Scholar Publishing Group
International Journal of Educational Innovation and Science
https://doi.org/10.38007/IJEIS.2022.030405
ISSN 2789-7184 Vol. 3, Issue 4: 55-68

Visual Analysis of Plant Science Knowledge Map Based on CiteSpace Ⅱ in Library Collection

Norizan Diah*

University of New South Wales Sydney, Australia

**corresponding author*

Keywords: Knowledge Atlas, CiteSpace II, Plant Science, Visual Analysis

Abstract: The research on the subject identification method based on knowledge map is a method research that combines a variety of metrological methods and scientific knowledge map technology to deeply analyze and analyze the structural relationship of the subject knowledge system, identify and detect research hot topics and their changing trends in the subject area, with a view to To better help scientific researchers better grasp the subject structure and hot topics from large-scale scientific and technological literature, and become a new method and new way for scientific and technological decision-makers to effectively carry out scientific and technological management in the new environment. This article first summarizes, combs, and refines the knowledge graph theory and its research methods; secondly, it explains the co-word analysis theory and co-occurrence analysis theory in detail, including the construction of analysis models and analysis processes, and the key technologies involved in journal relationship analysis. Research; Finally, in the experimental part, this article uses the literature of the molecular plant in the Web of Science database from the beginning of 2008 to March 2020 as the research object, using CiteSpace II software for literature co-citation analysis and co-word analysis, and knowledge visualization map method Demonstrate research institutions, knowledge bases, research hotspots, and research fronts in the field of plant science for more than a decade. Visualization results show that seed development is the most used burst word in molecular plant journals in the past three years, with 44 times of use. Compared with other burst words, the frequency of use is more than 30%, indicating that seed development is a plant science in recent years. Research fronts and research hotspots in the field.

1. Introduction

Scientific and technological literature, as an important soft carrier of science and technology, contains a large amount of scientific and technological knowledge. It is also an important source for the publication, tracing, citation, and search of scientific and technological achievements. It is the

most important way for researchers to obtain relevant information and knowledge. One. But nowadays, the rapid development of modern science and technology, and the rapid spread in the world along with the development of the Internet, has led to the explosive growth of global knowledge, which has also brought about difficulties in the choice of knowledge and information.

Faced with the explosive growth of huge amounts of document data and information knowledge, scientific researchers have used traditional classic document retrieval methods to query important documents to obtain professional knowledge. It is difficult to ensure that key documents and cutting-edge knowledge can be found. Scientific researchers It is impossible to quickly acquire knowledge and grasp scientific research trends from a large collection of documents. Therefore, for scientific researchers, how to quickly and accurately grasp the development of subject areas, understand research hot topics, and find breakthroughs in scientific and technological innovation has become their urgent needs. Similarly, traditional science and technology management theories and methods also encountered many problems in the new era. Sci-tech managers and sci-tech policy makers need to stand on a global level, comprehensively understand the current status and latest trends in the field of science and technology from a macro, holistic, and top-level perspective, keep up with the pace of science and technology development, and make accurate predictions of future technology development trends in various fields Sentence. However, the rapidly changing pace of knowledge update and the huge amount of scientific and technological information data make it difficult for scientific and technological decision-makers to identify and obtain valuable information and knowledge in a timely manner. Therefore, new scientific and technological evaluation methods are urgently needed to help scientific and technological managers break through the bottleneck of the existing scientific and technological evaluation mode and improve the efficiency of scientific and technological decision-making.

At present, Mapping Knowledge Domains is emerging as a new means to track the development of disciplines, detect the structure of knowledge in disciplines, and identify hot topics in disciplines [1-2]. It shows the complex fields of modern scientific and technological knowledge through data mining, information processing, knowledge measurement and graphic drawing. Using the theory and method of scientific knowledge atlas, it can target the development trend of science and technology in a certain industry or discipline and its related fields and its related In-depth research on the knowledge structure, and then discover potential general and special laws in scientific and technological activities, so that researchers are no longer difficult to choose new areas of interest, and also enable scientific and technological managers to effectively monitor the development of disciplines, thereby promoting the overall discipline development of. Jun Q I et al[3] expanded the field of green infrastructure (GI) research by introducing multidisciplinary research. The scientific knowledge atlas analysis software Cite Space was used to systematically study the results of geographical indication research, reveal the main areas, development trends and research hotspots of geographical indication research, and provide theoretical basis for further research. Xiuling L et al [4], in order to explore the hot technologies and development trends of the textile industry, 18 core journals of the textile industry were selected as research samples in the CNKI database. Use common word analysis and Cite Space to draw a map of scientific knowledge. Finally, according to the dynamic clustering chart, the development path of the textile industry in the past ten years can be analyzed, so that the research status of the textile industry is vividly displayed, and it provides a reference for the further research of the textile industry in China. Zhang J et al [5] embedded a large-scale knowledge graph composed of entities and relationships into a continuous vector space, which made up for the shortcomings of the knowledge graph's incomplete relationship mapping. The improved knowledge graph is accurate in analyzing the latest technology and other aspects.

The rate has improved. Song Q et al[6] proposed a graph summary framework for querying heterogeneous and large-scale knowledge graphs to facilitate the search of knowledge graphs.

Based on the above research background, this article uses the literature of "molecular plant" in the Web of Science database as the research object from the 2008 issue to March 2020. The paper uses CiteSpace II software to perform literature co-citation analysis and co-word analysis, and uses knowledge visualization maps. The methods show research institutions, knowledge base, research hotspots and research frontiers in the field of plant science for more than ten years.

2. Proposed Method

2.1. Introduction to Knowledge Graph

(1) Connotation and classification of knowledge map

With the continuous development of bibliometric theory and information visualization technology, the concept of knowledge map has gradually evolved into a new type of research field, which has become an important research method and tool in the field of scientific metrology. Through information technology such as data mining, knowledge measurement, and map construction, the evolution and structure of scientific knowledge in a certain field can be visualized, and the relationship between the activities and structure of scientific knowledge can be observed, and the evolution law in a certain field can be displayed. The knowledge map is constructed by visualizing the relationship between scientific knowledge. It contains many advantages, such as intuitive, easy to understand, quantitative, and objective. Therefore, as an efficient and comprehensive visual analysis tool, the knowledge map is widely used in the research of more disciplines, such as management, metrology, etc [7-8].

Knowledge maps are usually divided into six types: traditional scientometric maps, triangular configuration maps, multidimensional scale maps, social network analysis maps, self-organizing maps, and path-finding network maps. Figure 1 shows the different types of knowledge graphs.

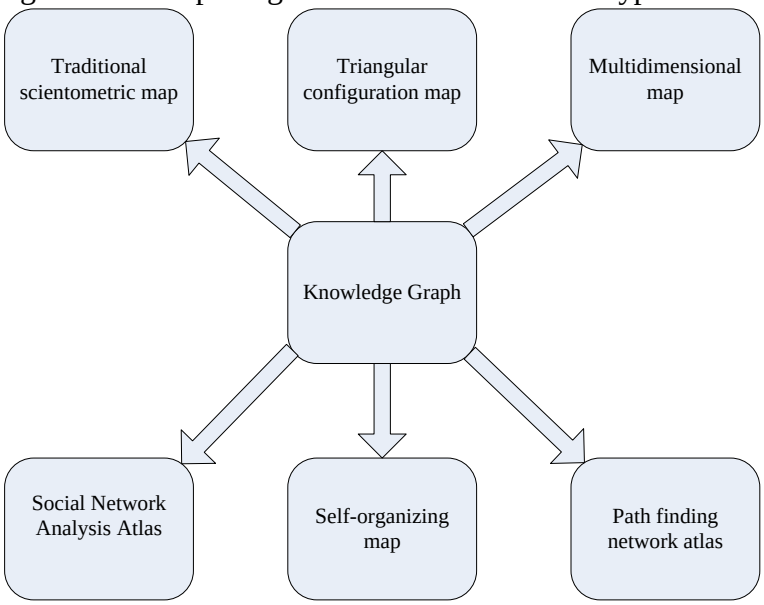

Figure 1. Different types of knowledge graphs

(2) Main functions of knowledge map

As an important method and tool for studying the relationship between the development process

and the structure of knowledge in the field of scientific knowledge, the main function of knowledge map research is to analyze the evolution of scientific knowledge, and to predict the research front and future development direction while analyzing the basic theory of scientific knowledge . The main functions of the knowledge map include tracking scientific frontiers, detecting scientific research hotspots, analyzing scientific evolutionary history, investigating scientific cooperation networks, evaluating the status of scientists, constructing scientific discovery theories, and assisting scientific work decisions. Figure 2 shows the main functions of the knowledge graph.

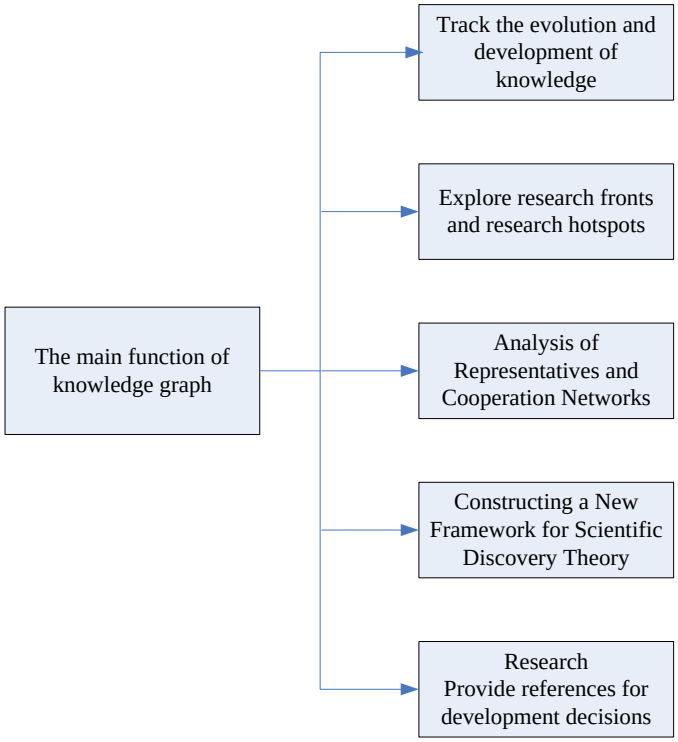

Figure 2. The main function of knowledge graph

1) Track the evolution and development of knowledge. Emerging knowledge map technology is widely used in the tracking of the evolution and development of knowledge under the support of visualization technology. The main working principle is to visually express the citations and citation relationships between documents using a visual map, and at the same time analyze the literature citation relationship based on time-sharing, dynamics, and multiple perspectives.

2) Exploring research fronts and research hotspots. The literature co-citation knowledge network provides a large number of data indicators for analysis in the field of scientific knowledge. By studying intuitive network key nodes and secondary document retrieval analysis based on emergent node literature, you can explore the frontiers of data source fields and Visually predict future trends.

3) Analyze representative people and cooperation networks. An important point about knowledge map research is the simple and direct presentation of author co-citation analysis and author co-occurrence analysis using visualization techniques. Researchers can analyze the in-depth content of key authors, author cooperation networks, and geographical distribution of author cooperation by constructing author public citation networks in the research field.

4) Construct a new framework for scientific discovery theory research. By analyzing the network structure and data indicators in the knowledge map, the structural holes in the map are studied, and the network nodes provided in the knowledge structure are used to comprehensively analyze the literature in the field of scientific knowledge to build a knowledge theoretical framework.

5) Provide references for development decisions. Use literature measurement tools to construct various types of knowledge maps, and visually analyze the development law, research hotspots, and development directions of the research field in order to scientifically provide decision-making references for the development planning of the research field and the overall research direction.

(3) Construction method of knowledge map

The knowledge graph construction process can be divided into eight steps, as shown in Figure 3 below:

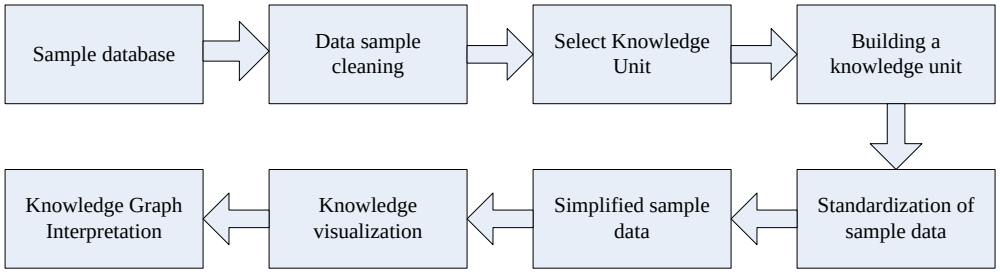

Figure 3. Knowledge Graph Construction Process

The acquisition of sample data is the premise and basis for drawing a knowledge map. Researchers can use literature and network knowledge as data sources. After obtaining the data, the sample data should be cleaned. The main tasks include deleting the duplicated data, checking for errors, and filtering. As the basic unit of data processing and analysis, the knowledge unit will have different structures for the construction of the knowledge map. One step that cannot be ignored in the construction of the knowledge graph is to standardize the sample data. Specifically, the set theory method or the probability theory method can be used for standardization. Next, you can use factor analysis, cluster analysis, and other methods to further simplify the data, so that you can better explore the relationship between each knowledge unit. The data processed through the above steps can be used to construct a knowledge map of knowledge visualization in related metrology software. Finally, the in-depth interpretation of the generated knowledge map is carried out. The methods used are: time analysis, space analysis, topic analysis, network analysis, etc.

2.2. Common Knowledge Map Research Methods and Their Theories

(1) Common knowledge map research methods

Common knowledge map research methods include bibliometric analysis, co-citation analysis, word frequency analysis, co-word analysis, and multivariate statistical analysis.

Bibliometric analysis is a kind of information statistics method with a long history. It mainly has the following types: bibliometrics by time, bibliometrics by journals, and bibliometrics by authors. In China, researchers are involved in the above types of analysis in the process of knowledge map research, but it is more common to use journals for research. The co-citation analysis method is specifically a process of displaying co-citation network views among many documents in the form of clustering groups through clustering analysis, multi-dimensional scaling and other measurement methods. This method can be reduced by computer devices Difficulty of analysis, visually shows a smaller number of groups. The currently widely used co-citation analysis methods are: author co-citation, journal co-citation, and subject co-citation. Author co-citation analysis is the construction and analysis of co-citation networks for authors, and then analyzes and summarizes the connections and characteristics of researchers in the field, in order to predict the evolutionary trend of research angle relationships. Journal co-citation analysis is based on the structure of core journals in the research field of the construction of journal co-citation networks, and analyzes professional

journals focusing on different professional fields. Disciplinary co-citation analysis is based on the construction of disciplinary co-citation network, showing the cross-relationships and dependencies between disciplines, and then reflecting the discipline composition and structural characteristics of the discipline system. The term frequency analysis method mainly counts the frequency of occurrence of keywords with intelligence significance and its changing rules, and then analyzes the hotspots in the research field and the cutting-edge trends. The co-word analysis method uses keywords, authors, etc. as the node types in the statistical document information to construct a knowledge network, and then it can visually display the topic structure, author cooperation network, and subject knowledge structure sources in the field. In case studies, domestic researchers use the method of keyword co-occurrence analysis. Multivariate statistical analysis means multivariate statistical analysis using computer tools to analyze random variables related to the subject. It is characterized by dimensionality reduction techniques, including factor analysis, cluster analysis, and multidimensional scale analysis.

(2) Co-occurrence analysis theory

Co-occurrence is the correlation between the appearance of one thing and another or more things [9-10]. It is common in nature and our daily life, such as the co-occurrence of orchids and orchid peaks, the co-occurrence of traditional Chinese medicine prescriptions, the co-occurrence of added ingredients in different foods, the co-occurrence of different netizens in the same comment on Weibo, etc. . The co-occurrence relationship of food ingredients can well reflect the correlation of food ingredients and provide a basis for understanding the composition of food groups. The co-occurrence phenomenon discussed in this article is a co-occurrence phenomenon between knowledge units of plant science literature. The co-occurrence analysis of knowledge units of plant science literature is as follows:

Co-occurrence of plant science literature mainly refers to the phenomenon that the same or different types of knowledge units co-occur. Such as co-occurrence of different authors in the same paper (author's cooperation), co-occurrence between different keywords (co-word of keywords), co-occurrence between different references (co-citation of references), etc. . In the plant science literature collection, some knowledge units are potentially related, and co-occurrence frequencies are often used to measure such associations. Different measurement methods also have different names, such as co-occurrence of keywords, co-citation of documents, coupling of documents, and cooperation of authors. There are various methods for co-occurrence classification of knowledge units in plant science literature. Some scholars have proposed to divide co-occurrence into intra-co-occurrence and inter-co-occurrence based on whether the co-occurrence unit appears in the same paper. In-text co-occurrence refers to the co-occurrence between plural entries in the same field of the same article, as if cited; inter-co-occurrence is the co-occurrence of the same entry in the same field in different articles, such as citation coupling. According to the relationship between the knowledge units of plant science literature, this paper divides the co-occurrence types into co-occurrence and hetero-co-occurrence: Different types of knowledge units are directly or indirectly related. For example, the document-keyword matrix established by document-indexed keywords measures the similarity of documents; this co-occurrence phenomenon is translated into "Occurrence" in English to express a kind of affiliation . If a keyword has only two cases in the document, that is, it appears or does not appear, so the matrix of this relationship is a 0-1 matrix. This type of knowledge unit co-occurrence also includes document-author, document-keyword, document-classification, etc. This type of co-occurrence relationship is mainly established directly through different types of knowledge units. In addition, there is a co-occurrence network of indirect relationships, such as author-keyword. Co-occurrence relationships established between

author-institution fall into this category. This co-occurrence is called hetero-co-occurrence, and it is the relationship established between different knowledge units.

Indirect association of the same unit of knowledge. The indexing items in plant science literature contain information such as author, organization, keywords, and text classification. These single knowledge units usually contain more than two items. If the number of keywords in a document is usually around 5, then co-word analysis can be performed on these keywords. Similar co-occurrence analysis of knowledge units established by the indirect relationship of the same knowledge unit also includes Co-authorship (author, institution, country / region), Co-citation (document, author, journal), etc., such co-occurrence networks It is often called "Co-Occurrence" in English.

(3) Co-word analysis theory

Content word analysis is based on vocabulary (content words) that can reflect the subject matter of a document as a research object. It conducts quantitative analysis of similarity or dissimilarity, thereby researching and revealing the internal relationship between documents, scientific structure and its evolution. Content words include two types, controlled words and uncontrolled words. Research topics usually consist of a set of interrelated vocabularies, and various combinations can be made between words to express complex concepts in all aspects of the subject content. The distribution of vocabulary is closely related to the development of the subject area. At the embryonic stage, new vocabulary began to appear. With the gradual development of the discipline, the number of vocabulary and the frequency of use increased rapidly. The emergence, merger, differentiation and demise of vocabulary reflect the development and change of knowledge information from the depth of the literature content. By tracking the changes of these words in different periods, we can reveal the evolution of scientific development and its dynamic structure. Compared with the citation analysis method based on the whole document, the content word analysis method is more effective in subject identification research, and its pertinence and accuracy have been further improved. Word frequency analysis and co-word analysis both belong to the content word analysis method [11-12].

The theory of the word frequency analysis method is that in any article, the frequency of words appears according to the following rule: if you count the frequency of each word in a longer article, and arrange them in descending order from high to low, use Natural numbers assign rank numbers to these words, that is, the most frequent word rank is 1, the second most frequent word rank is 2, and the least frequent word rank is D (or L). If f is used for frequency and r is used for rank number, then: $f \cdot r = c$. Where c is a constant. But the constant here is not an absolute constant, but fluctuates around a central value. The above formula is consistent with Zipf's previously verified quantitative form, called Zipf's law (or Zipf's first law).

The co-word analysis method can be used to reveal hot topics in the research field, analyze the development process, characteristics of the field disciplines and the relationship between fields or disciplines from different perspectives horizontally and vertically, monitor the level of scientific research, and track the dynamics of discipline development history Static structure, etc. The common word analysis method is mainly based on counting the number of times a group of words appear in the same document. If the number of times a group of words appears in a document, the group of words is considered to have a closer topic relationship. Then through further cluster analysis, the subject and theme changes represented by these words are studied. Co-word cluster analysis is a commonly used method in co-word analysis. The co-word frequency is used as the analysis object, and the statistical method of clustering is used to simplify the intricate co-word network relationship between many analysis objects into a relative number. There are fewer

relationships between several groups and the process of clustering is visually expressed. The clusters formed by these closely related subject words can express the composition of a branch of a certain field. Studying the composition, evolution, and disappearance of clusters is the focus of co-word cluster analysis.

2.3. Information Visualization Tools

With the vigorous development of information visualization technology, the development of visualization tools has received increasing attention from researchers. At present, the more common visualization tools include three types: visualization development tools, analysis tools, and toolkits. Specific commonly used information visualization tools are shown in Table 1.

Table 1. Information visualization tool sheet

Name of Software	Free	Preprocessing	Standards and Simplifications	Atlas interpretation
CiteSpace II	Yes	Time period / data simplification / network simplification	Consine/Jaccard/ Dice	Catastrophe monitoring, diachronic analysis
VOSviewer	No	Deduplication / Time Division / Data Simplification	Equivalence index	diachronic analysis
HistCite	Yes	Deduplication / Time period / data simplification / network simplification	User-defined	Catastrophe monitoring
SCI[2]	Yes	Deduplication /Time period / data simplification / network simplification	User-defined	Catastrophe monitoring

At present, there are more than ten kinds of literature mapping software, and each software has its different advantages. For example, VOSviewer is clear and detailed in topic clustering; SCI2 is more flexible in the analysis of topic words; HistCite presents the cross-references of the documents in the downloaded data set in the form of a network in time; BibExcel is based on the tag of the original data of the data set Provides a variety of literature analysis functions (requires external software to assist in visualization, such as Gephi, VOSviewer, Pajek, etc.); CiteSpace II is well known for its powerful literature co-citation analysis, and with continuous development of algorithms and functions Continuous optimization. CiteSpace has been widely used in more than 60 fields such as computer science, information science and medicine.

3. Experiments

3.1. Experimental Data Set

According to the Journal Citation Report (JCR) published by the American Institute of Scientific

Information (ISI) on June 18, 2010, the Institute of Plant Physiology and Ecology (IPPE) of the Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences The Molecular Plant's 2010 Impact Factor (IF) for the academic period sponsored by the Society of Science (CSPP) and sponsored by the Life Science Journal of Shanghai Life Science Information Center of the Chinese Academy of Sciences is 4.296. This impact factor ranks fourth among 112 Chinese SCI journals, and ranks first among domestic journals in the field of plant science. Ranked 14th (top 8%) in 190 international plant science journals included in the JCR report. "Molecular Plant" was founded in 2008. In recent years, articles have focused on plant hormones, plant stress resistance, plant development, and signal transduction, leading the research direction of plant molecular biology. In this paper, the scientific visual map analysis software Cite SpaceII is used to process and analyze the literature published in "Molecular Plant" over the years, to draw up a knowledge map, and to interpret the evolutionary path and research frontier displayed by the knowledge map, in order to fully grasp the plant The dynamic process, characteristics and laws of development in the scientific field.

Before constructing a plant science knowledge map, first collect and process data. The data source is Web of Science (http://webofknowl-edge.com/WOS), select the Science Citation Index Expanded (SCI-EXPANDED) database in the advanced search, select the year 2008 to 2020, and enter the search formula "SO = (Molecular Plant) ". A total of 1752 articles published in the journal from 2008 to 2020 were retrieved. Select 500 records at a time for all records to output. The full record format includes the cited references. Save as a plain text file. The file name is download_1-500.txt ~ download _1500-1752.txt for future use. The downloaded data includes the following fields. : Author, title, source, cited references, key words, abstract, etc.

3.2. Experimental Platform

First, use EXCEL 2013 to conduct a simple bibliometric analysis of the research data. Before constructing a knowledge map of the research in the field of mobile libraries, perform a simple statistical analysis on the annual volume of publications. Use the statistical function of EXCEL to roughly analyze the basic situation of domestic and foreign research. Understanding. Secondly, use information visualization tools to construct knowledge maps for mobile library research.

CiteSpace II is a knowledge graph analysis tool developed by Dr. Chaomei Chen, a professor at the School of Information Science and Technology of Lexel University in the United States. It is a free visual knowledge graph tool suitable for multivariate, time-sharing and dynamic complex network analysis. The main users of CiteSpace II are scientific researchers, medical circles, scientific policy researchers, and medical librarians. It integrates information visualization methods, bibliometric methods, and data integration methods. It draws a map and establishes the basis for the relationship between nodes. Are "co-citations" and "quotations." In summary, this is a software used in scientific literature to identify and display new trends and new developments in science. CiteSpace II can be used to find research advances and current research frontiers in a certain subject area.

4. Discussion

4.1. Analysis of Papers Published by Journals

(1) Paper load

Journal load is one of the basic indicators describing the production capacity of a journal. It is

defined as the total number of papers published by the journal in a given time. Molecular Plant was founded in 2008. In recent years, articles have focused on plant hormones, plant stress resistance, plant development, and signal transduction, leading the research direction of plant molecular biology. Figure 4 shows that the overall trend in the number of plant science articles published by Molecular Plant since 2008 has continued to increase. From 93 articles in 2008 to 185 articles in 2013, the number of articles has almost doubled compared to 2008. . The number of articles published in the journal has declined in the past two years, with 163 articles published in 2019.

Table 2 shows the impact factors for Molecular Plant from 2014 to 2018.

Table 2. Molecular plant impact factors 2014-2019

Journal Title	Impact factor	Year
Molecular Plant	7.012	2014
	7.826	2015
	8.359	2016
	8.973	2017
	9.326	2018

Table 3 shows the comparison of the impact factors of many plant journals in 2018.

Table 3. Impact factors of plant journals

Journal Title	Impact factor
TRENDS IN PLANT SCIENCE	12.149
Nature Plants	11.471
Molecular Plant	9.326
PLANT CELL	8.228
PLANT BIOTECHNOLOGY JOURNAL	6.305

Figure 4 shows the trend of the amount of plant science articles in the Molecular Plant journal from 2008 to mid-March 2020.

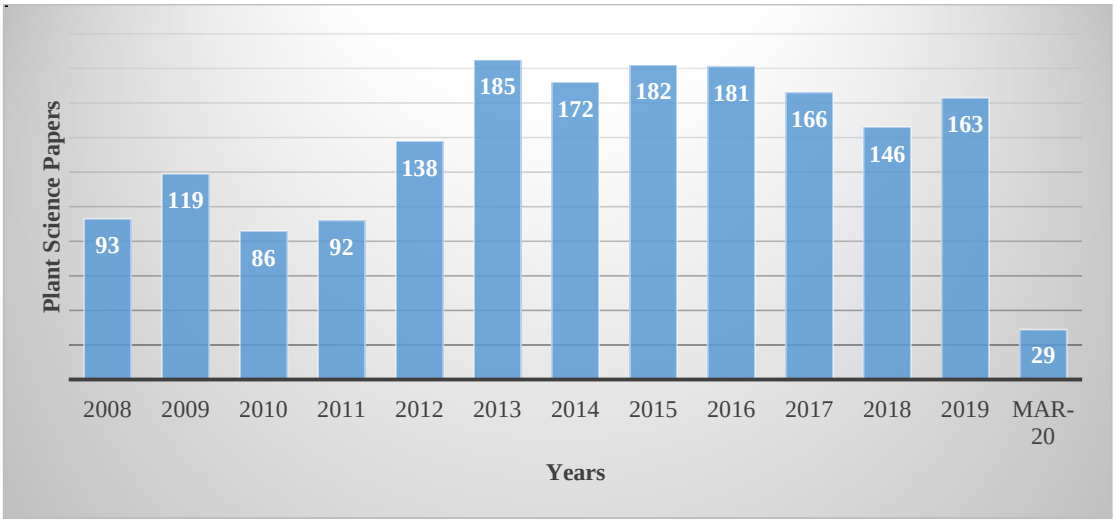

Figure 4. Statistics of journal articles from 1999 to 2019

(2) Analysis of the volume of papers issued by the state and research institutions

Install the Java runtime environment, start and run the Cite Space II software (version number v.5.6.R4), and set the corresponding data source directory and project directory. Cite Space II can visually display the number and time of papers published in various countries by the size and color of the circle. In the CiteSpace II software interface, the network node selects the country and institution, and the source of the subject word selects the title, abstract, abstract, and identifier, and

The data extraction object is Top30. The value of the Years per slice is set to 1, which will be divided into 12 periods for processing from 2008 to 2020. Run CiteSpace II to get the country distribution map of Molecular Plant magazine. Click the "find cluster" button to perform cluster analysis on countries and research institutions (Figure 5). Among them, a circle represents a node, that is, a country or institution, the area of the circle represents the volume of posts, and the thin line represents the cooperative relationship between countries or institutions. The color blocks of different colors shown above represent the corresponding year, and 1 year is a time slice.

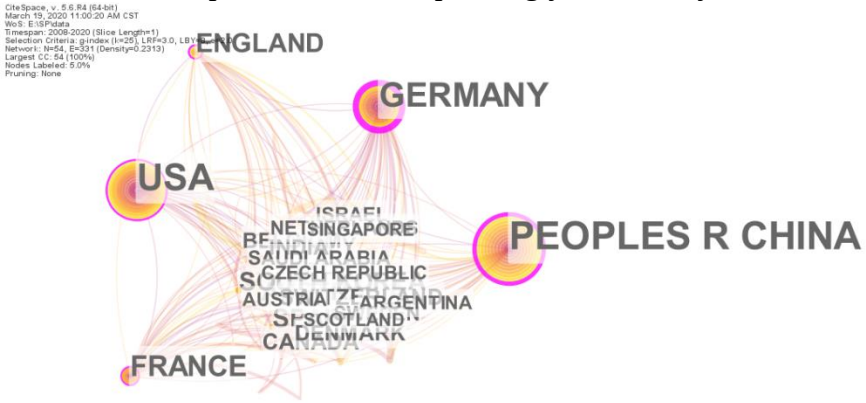

Figure 5. Cluster map of countries and institutions

From the perspective of the total number of posts in each node in Figure 5, China has the largest contribution rate of documents, far higher than other countries, accounting for 20% of the total number of posts, followed by the United States, Germany, France, and the United Kingdom. In Figure 5, there are 100 nodes in total, and only 64 connections. Although the cooperation and exchange of biological research has been very frequent in recent years, from the perspective of the entire network and the entire time span, there are fewer connections between countries or research institutions, indicating that most of the Top30 articles are between authors from the same country. Cooperation is done. Judging from the suddenness of publication, the United States and China are very prominent. The sudden increase in China's publications occurred in 2014-2017, indicating that China's publications in top magazines have increased rapidly in the past five years. After a period of imitation, learning and accumulation in the field of plant science in China, it has entered a rapid development Period.

4.2. Research Hotspots and Frontier Literature Analysis

Use the keywords in the bibliography and use the Cite Space II software to identify hotspots in bibliometrics research. Keywords are the author's summary and refinement of the core of the article, and the essence of an article. Therefore, the keywords of the article are analyzed. Frequently used keywords are often used to determine a hot topic in a research field. The network node only selects the keyword (Keyword), other strategies remain unchanged. Cite SpaceII can analyze the change of research hotspots by showing the change of keywords over time through the time zone view. Figure 6 shows the changes in the hotspots of plant science research in the journal Molecular Plant in the past three years.

Figure 6. Changes in research hotspots

Combining statistics, it can be seen that the most used keyword in Molecular Plant magazine is "Arabidopsis thaliana", which appears 350 times in the entire network. This shows that Arabidopsis as a "model plant" has contributed greatly to the entire plant biology. . Followed by "gene expression" and "protein" were used 205 times and 157 times, respectively. The study of gene and protein level is the core of modern botany (that is, plant molecular biology). The function of genes and proteins reveals the essence of plant development. The use of "transcription factor" as a keyword is also very frequent, and the suddenness is relatively high. The transcription factor is related to the expression of protein-regulated genes, which is a continuation of the previous two keywords. "Signal transduction" involves various stages of plant growth and development, and in recent years it has gradually become a research hotspot. As a pair of closely related keywords, transgenic plants and wild type are the most commonly used research methods in plant biology to discover the function of genes through the comparison of transgenic plants and wild type.

Figure 7 shows the top 10 keywords used by Molecular Plant magazine from 2016 to 2020. As can be seen from Figure 7, many keywords in Molecular Plan became prominent after 2016, such as plant development, transcription factors, gene families, functional analysis, in vitro, differentiation, etc., indicating the research areas related to these keywords It is becoming more and more popular, and it is also the forefront of plant biology research.

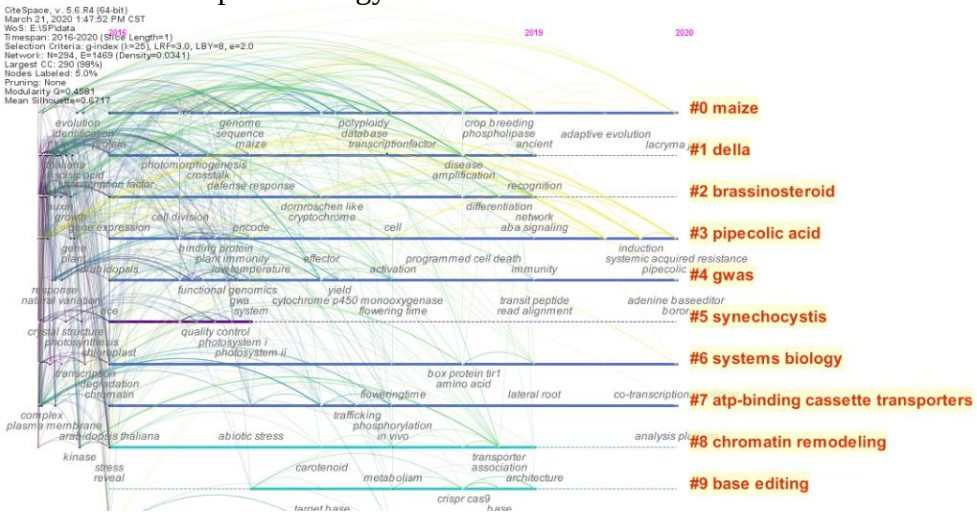

Figure 7. Unexpected words

Figure 7 shows that the most frequently used burst word is maize, which reached 44 times, and is concentrated in the area after 2016. This shows that plant seeds have become a popular research object in the field of botany. The emergent words that are used collectively also include transcription factors, plant growth, cell division, protein-RNA interactions, transgenic plants, root growth, green fluorescent protein (GFP), etc., indicating that the fields related to these emergent words are also recent years Research hotspots.

5. Conclusion

Traditional science and technology management theories and methods have encountered many problems in the new era. Science and technology managers and science and technology policy makers need to stand on a global level, comprehensively understand the current state of science and technology in the field from the macro, overall, and top-level perspectives, and keep up with technological development To make accurate predictions of future technological development trends in various fields. However, the rapidly changing pace of knowledge update and the huge amount of scientific and technological information data make it difficult for scientific and technological decision-makers to identify and obtain valuable information and knowledge in a timely manner. Therefore, new scientific and technological evaluation methods are urgently needed to help scientific and technological managers break through the bottleneck of the existing scientific and technological evaluation mode and improve the efficiency of scientific and technological decision-making.

The combination of scientific knowledge maps and other methods has been widely used in the research of subject identification in subject areas. However, there are still some shortcomings in the research of these methods themselves, and the method of strategy combination and how to use these methods in combination for subject recognition are still Further optimization and standardization are needed, and some key links involved also have room for further improvement.

Based on the above background, the article takes all the literature contained in the journal "Molecular Plant", which has the highest impact factor in the field of plant science, as the research object in the past 12 years. Through the data processing and analysis of Cite Space II software, it displays plant science in the form of knowledge maps. The distribution of research power in the field, and related important academic literatures, hot keywords, etc., analyzes the knowledge structure and cutting-edge development trends on which the current development of plant science depends.

References

[1] Wang Ruojia, Li Ying.(2016). "Comparative Study on Competitive Intelligence Based on Mapping Knowledge Domains at Home and Abroad.", Journal of Intelligence, 20(1),pp.12-19.

[2] Ludolph R , Allam A , Schulz P J.(2016). "Manipulating Google ' s Knowledge Graph Box to Counter Biased Information Processing During an Online Search on Vaccination: Application of a Technological Debiasing Strategy.", Journal of Medical Internet Research, 18(6),pp.137-144.https://doi.org/10.2196/jmir.5430

[3] Jun Q I , Ping-Ting H , Xue-Qiong T.(2015). "Review of Green Infrastructure Research from the Perspective of Scientific Knowledge Map." , Journal of Beijing Forestry University(Social Sciences), 14(2),pp.88-95.

[4] Xiuling L , Guangchun R.(2016). "Knowledge Map and Visualization Analysis in Textile Industry Based on Cite Space." , Journal of Silk, 16(12),pp.4-12.

[5] Zhang J.(2015). "Knowledge Graph Embedding by Translating on Hyperplanes.", AAAI - Association for the Advancement of Artificial Intelligence, 24(13),pp.237-245.
[6] Song Q , Wu Y , Lin P , et al.(2018). "Mining Summaries for Knowledge Graph Search.", IEEE Transactions on Knowledge & Data Engineering, 33(9),pp.125-136.https://doi.org/10.1109/TKDE.2018.2807442
[7] HOU Dongyang.(2018). "Mapping Knowledge Domains of European Communication Research:An Analysis of Two Journals of European Communication(2010-2014).", Jinan Journal(Philosophy & Social Science Edition), 14(19),pp.89-97.
[8] Sun Y.(2016). "Mapping Knowledge Analysis of Information Classification and Coding Based on CSSCI.", Journal of Computational & Theoretical Nanoscience, 13(12),pp.9244-9250.https://doi.org/10.1166/jctn.2016.6311
[9] Hoo W L , Chan C S.(2017). "Zero-Shot Object Recognition System based on Topic Model.", IEEE Transactions on Human-Machine Systems, 45(4),pp.518-525.https://doi.org/10.1109/THMS.2014.2358649
[10] Shi Y , Liu J , Shi L , et al.(2019). "An Improved Real-Time Recommendation for Microblogs Based on Topic.", International Journal of Pattern Recognition and Artificial Intelligence, 23(12),pp.613-620.
[11] Liu, Yanping, Huang, Ren, Li, Yugang.(2017). "The Word Frequency Effect on Saccade Targeting during Chinese Reading: Evidence from a Survival Analysis of Saccade Length.", Frontiers in Psychology, 8(11),pp.1-10.https://doi.org/10.3389/fpsyg.2017.00116
[12] Chunmei Gan, Weijun Wang.(2015). "Research Characteristics and Ctatus on Social Media in China: A Bibliometric and Co-word Analysis.", Scientometrics, 105(2),pp.1167-1182.https://doi.org/10.1007/s11192-015-1723-2