Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Chaos, Solitons and Fractals 139 (2020) 110052

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

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

Bibliometric method for mapping the state of the art of scientific production in Covid-19

Mohamed El Mohadab*, Belaid Bouikhalene, Said Safi

Laboratory of Mathematics Innovation and Information Technology (LIMATI), Department of Mathematics and computers Sciences, Polydisciplinary faculty Beni Mellal, Sultan Moulay Slimane University, Morocco

ARTICLE INFO

Article history:
Received 3 June 2020
Accepted 23 June 2020
Available online 30 June 2020

Keywords:
Covid-19
Scientific production
Bibliometric method
Bibliometric analysis
Scientific research

ABSTRACT

Global scientific production around the Covid-19 pandemic, in the various disciplines on the various international scientific bibliographic databases, has grown exponentially. The latter builds a source of scientific enrichment and an important lever for most researchers around the world, each of its field and its position with an ultimate aim of overcoming this pandemic. In this direction, bibliometric data constitute a fundamental source in the process of evaluation of scientific production in the academic world; bibliometrics provides researchers and institutions with crucial strategic information for the enhancement of their research results with the local and international scientific community, especially in this international pandemic.

1. Introduction

The latest statistics indicate that there has been an exponential increase in the number of publications since the discovery of the Covid-19 pandemic; the results provide a comprehensive view of interdisciplinary research in medicine, biology, finance and other fields.

The number of publications in international databases aims to disseminate and share the contributions and advances of academic research from different groups of researchers from different universities and countries in the thematic of Covid-19.

Bibliometrics [1] is a tool for mapping the state of the art in a field related to given scientific knowledge. So the use of bibliometric analysis [2] to identify and analyze the scientific performance of authors, articles, journals, institutions, countries through the analysis of keywords and the number of citations constitutes an essential element which provides researchers with the means to identify avenues and new directions in relation to a theme of scientific research.

2. Bibliometrics at the service of scientific research

Scientometrics [3] is considered as the science of measurement and the analysis of science which is based on an input set and an output set which uses bibliometrics in the field of study of publications. The latter is a meta-science which takes science as its object of study based on three elements of scientific activity: its inputs, its outputs and its impacts. Thus, it makes it possible to map and broaden knowledge on a research field, by clarifying the links between the authors, the publications, the institutions, and other characteristics of the studied field.

Scientific publications [4] represent all publications in newspapers or conferences, either chapters in scientific books or scientific patents. All these types of publications represent the work of a researcher who publishes these works with the aim of circulating these results in databases which have broad international visibility and scientific credibility such as Web of Science, Scopus... and renowned publishing houses such as Elsevier, Springer, Wiley, etc.; but with all the efforts made, the benefits that can be drawn remain limited if we cannot manage this large mass of publication which is added every day to the thousands or millions of existing scientific papers.

Bibliometric data is used for:

- Measure and compare the scientific output of the researcher, research groups, institutions, regions or countries using indicators based on:
  - The number of publications.
  - The quotes received.
  - The collaborations.
- Identify the most important or influential journals in a given field.
- Monitor the evolution over time of a discipline or research subject.

* Corresponding author.
E-mail address: m.elmohadab@gmail.com (M.E. Mohadab).

https://doi.org/10.1016/j.chaos.2020.110052
0960-0779/

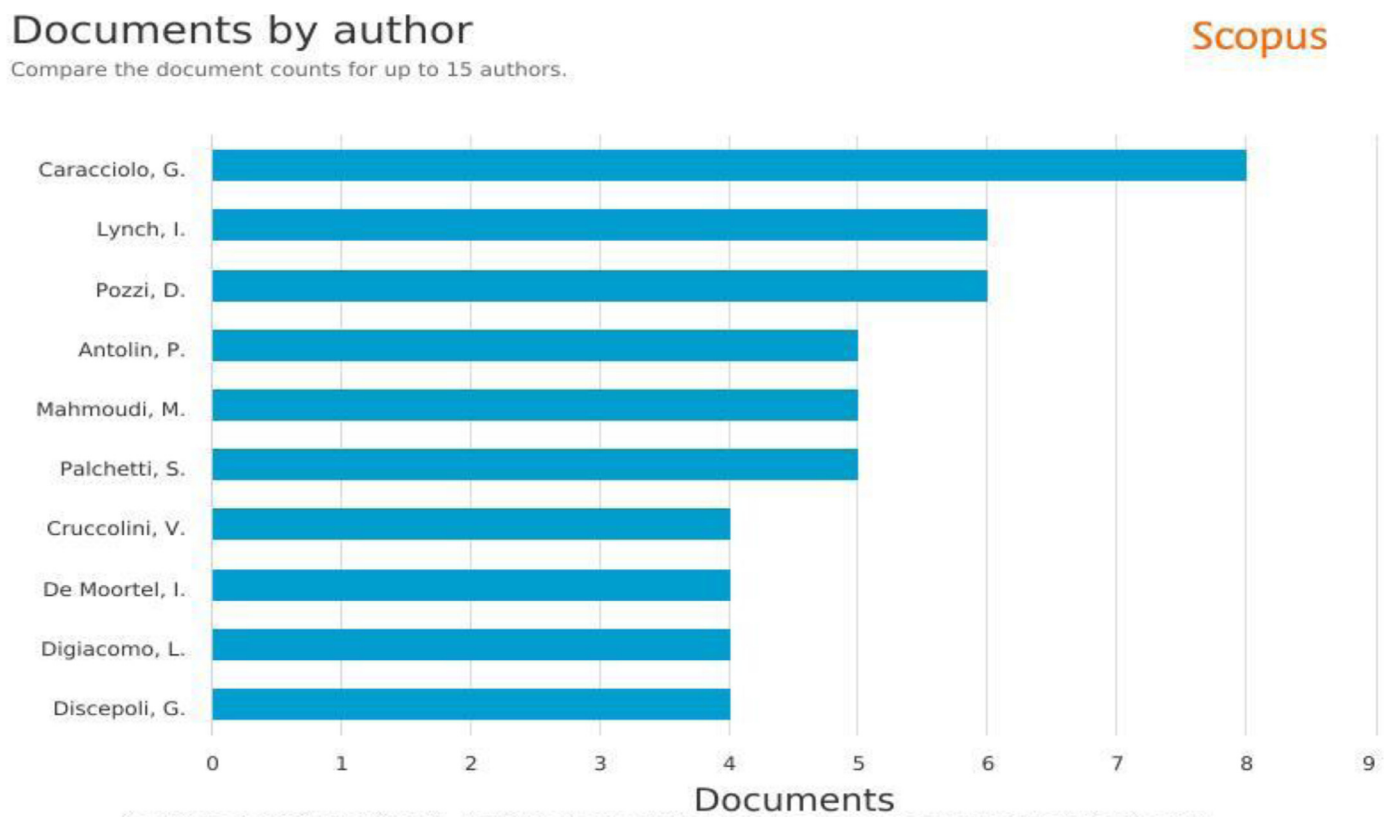

Fig. 1. Statistics of the best author published for Covid-19 on Scopus.

These data represent the main part of the data provided for each paper by the databases which allow bibliometrics to carry out statistical processing, and bibliometric analysis.

3. Statistical overview on Covid-19

3.1. The international context

According to statistics provided by Johns Hopkins University [5] until May 23, 2020, the death of more than 339,949 people worldwide, was the infection of 5,267,452, considerable efforts were made in the various disciplines relating to the treatment of this pandemic either from near or far.

Since the beginning of the year, Covid-19 represents an increasing interest for researchers from all over the world, in response to this crisis, a lot of research was carried out in many fields of research (medical, biology, financial, ...) by several Institutions and organizations, either public or private worldwide, each with their own means available.

By reviewing most of the scientific databases, the search to identify the scientific output related to the subject of Covid-19 [6] was carried out using a set of terms as search criteria, the language of the documents is the English because it is the universal language of research, all disciplines are authorized in order to provide a global view of Covid-19 research in the various disciplines, research is limited to the period from early 2020 (Beginning of the pandemic a been listed) so far Figs. 1–17.

❖ SCOPUS [7]:

Using the Scopus search engine to search for the word “covid-19” and “coronavirus” from 01/01/2020 until 23/05/2020, we find 10,228 documents:

- According to the authors:
- According to the institutions:
- According to the country:
- According to the type of documents:
- According to the domains:

❖ Web of Science [8]:

Using the search engine of Web of Science to search for the word “covid-19” and “coronavirus” from 01/01/2020 until 23/05/2020 results in 5,161 documents:

- According to the authors:
- According to the institutions:
- According to the country:
- According to the type of documents:
- According to the domains:

3.2. The African and Arab context

❖ Scopus:
■ Africa:
■ Arab:
❖ World of Science:
■ Africa:
■ Arab:

4. Methodology of the analysis of bibliometric data

The exploitation of the bibliometric parameters available on the scientific data base on multiple field and discipline makes it possible to release relevant information which can meet the expectations of researchers, research teams and research institutes. The bibliometric analysis reveals to the researcher exact information for the construction of new research as in the case of our study on Covid-19.

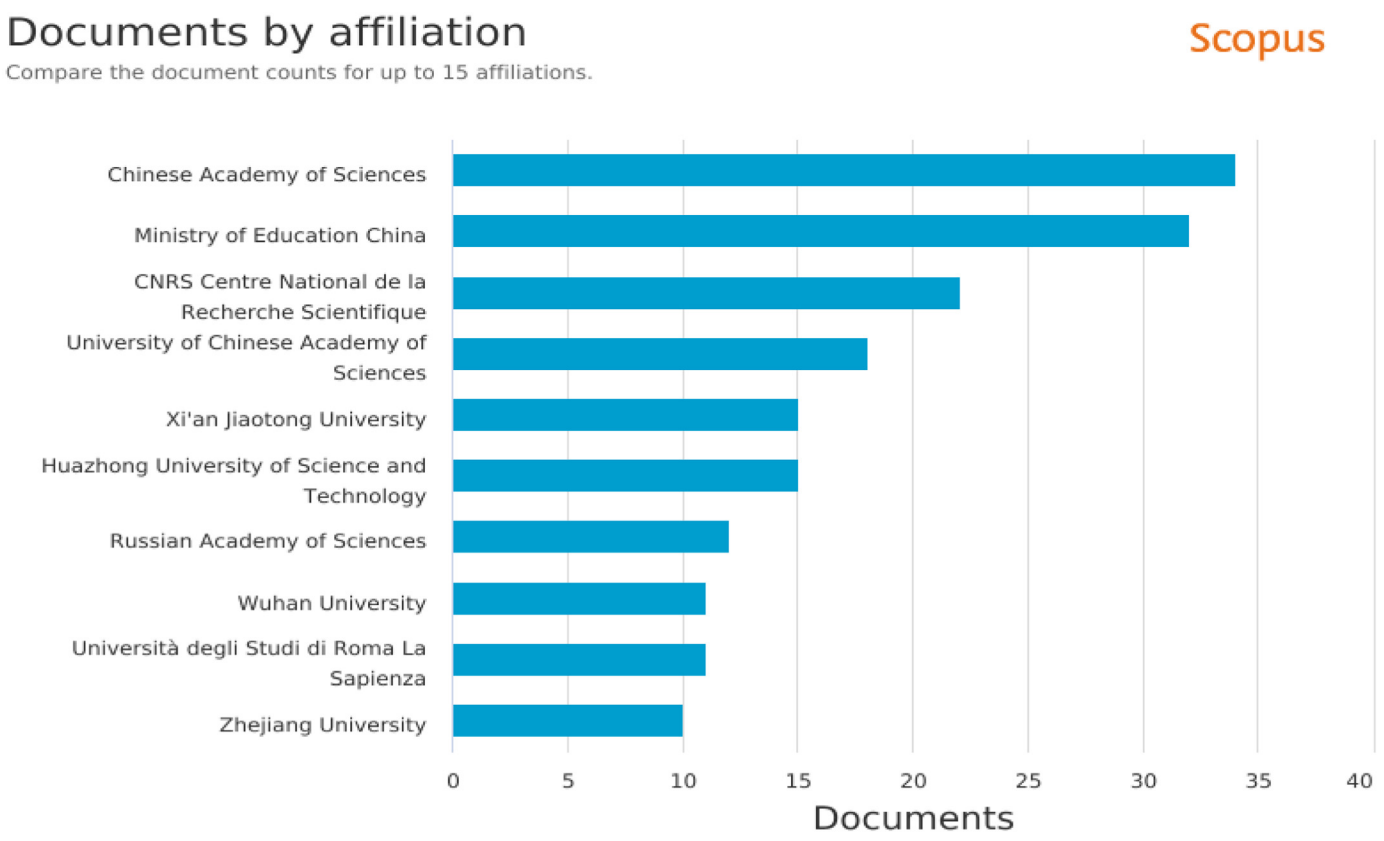

Fig. 2. Statistics of the best 10 institutions published for Covid-19 on Scopus.

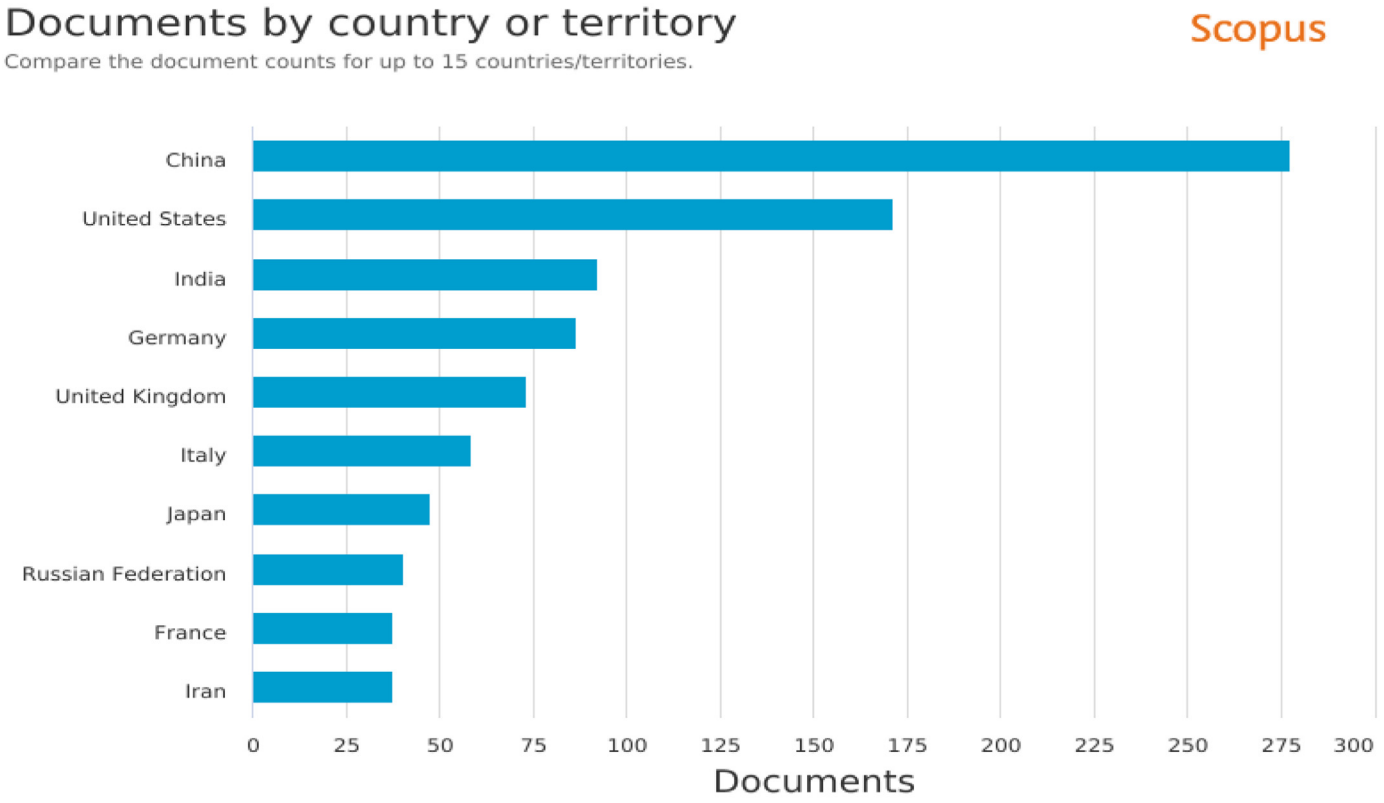

Fig. 3. Statistics of the best 10 countries published for Covid-19 on Scopus.

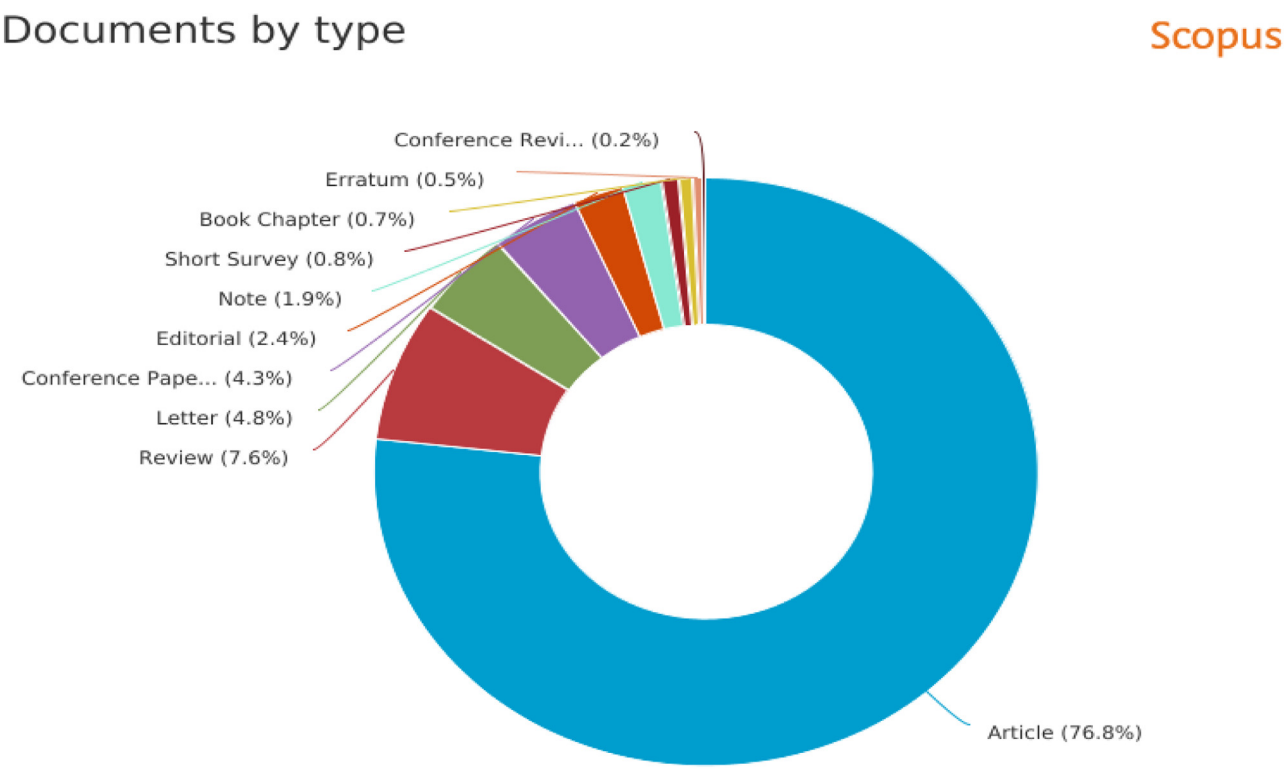

Fig. 4. Statistics of the type of document published for Covid-19 on Scopus.

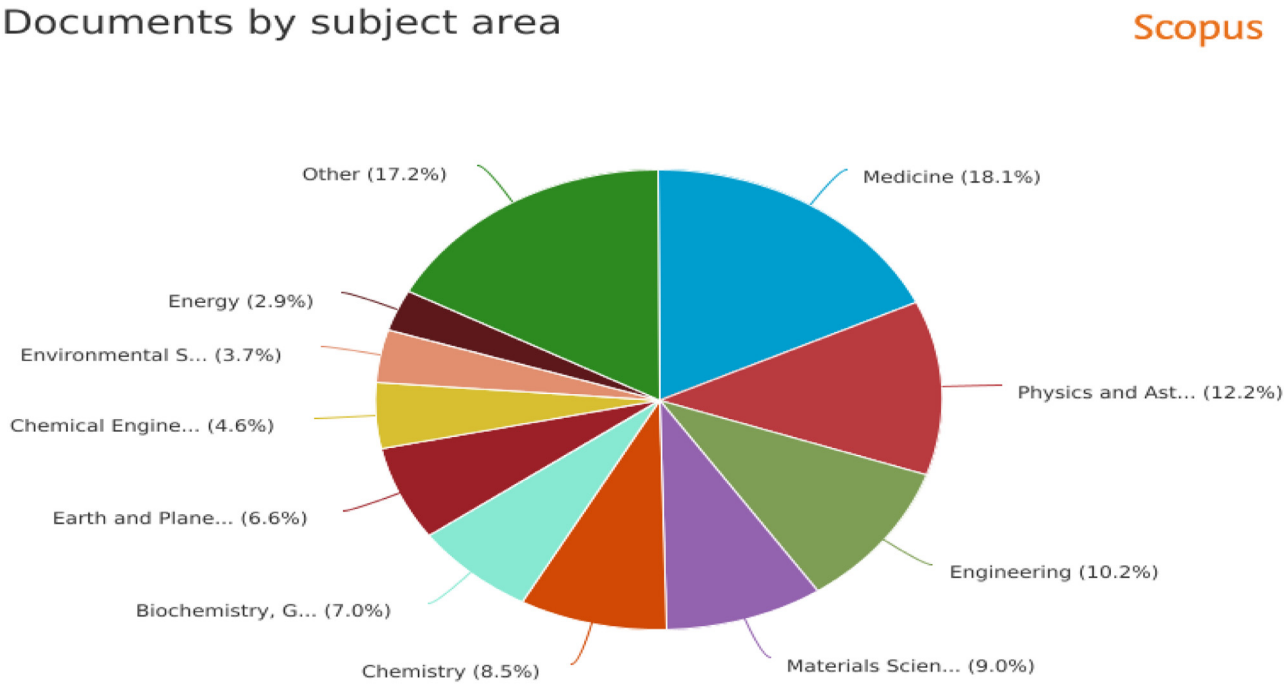

Fig. 5. Domain statistics published for Covid-19 on Scopus.

This study was carried out on the basis of specific research using the three databases (Scopus, Web of Science, Pubmed) from the beginning of 2020 until 23/05/2020. The sample consists of 5,161 academic publications (Web of Science), 10,228 academic publications (Scopus) and 7,991 academic publications (Pubmed). The use of bibliometrics will contribute to the exploration and description of the existing scientific literature on the theme of Covid-19.

The steps taken to achieve the desired results are manifested as:

The use of bibliometric tools plays an important role in guiding a particular field of study by collecting scientific data and synthesizing the results obtained.

Statistics from different bibliographic databases which differ either in terms of data volume or coverage constitutes a reliable source for bibliometric indicators [9].

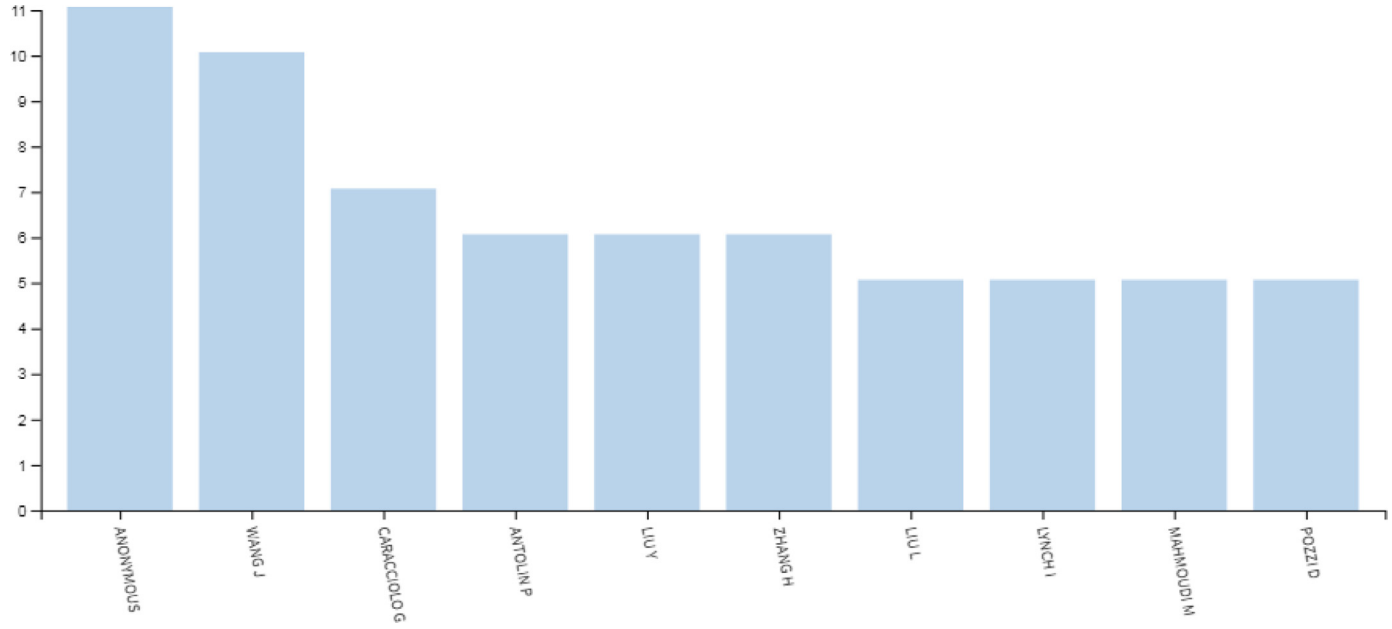

Fig. 6. Statistics of the best 10 author published for Covid-19 on Web of Science.

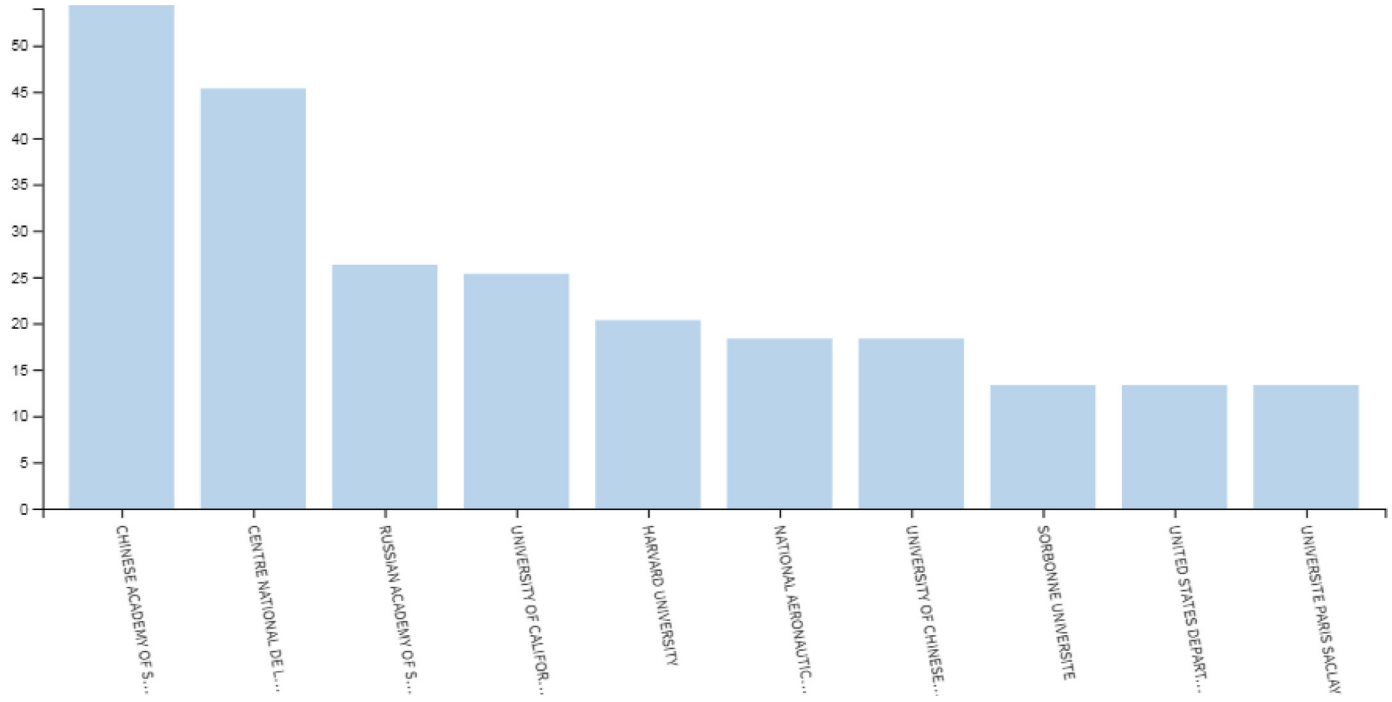

Fig. 7. Statistics of the best 10 institutions published about Covid-19 on Web of Science.

Choosing the right database, the right keywords and applying the filters that reflect the research objectives is a crucial step to have reliable results.

Among the credible scientific database which brings together most of the publishing houses known as Elsevier, Taylor & Francis, Springer..., we find Scopus, web of Science and for the medical field Pubmed [10] equipped with different filters to refine the search and limit the results found.

Some researches try to analyze data coming from the various scientific databases, but there are structural differences between the platforms. Thus the differences in the classification of information adopted by each of them builds an obstacle for an exploitation of the common data.

For a good bibliometric analysis, we choose the following bibliometric data:

- Article title.
- Authors.
- Keywords.
- Number of citations.
- Year of publication.
- Journals.
- Type of documents.
- Institution.
- Country.
- Field of research.

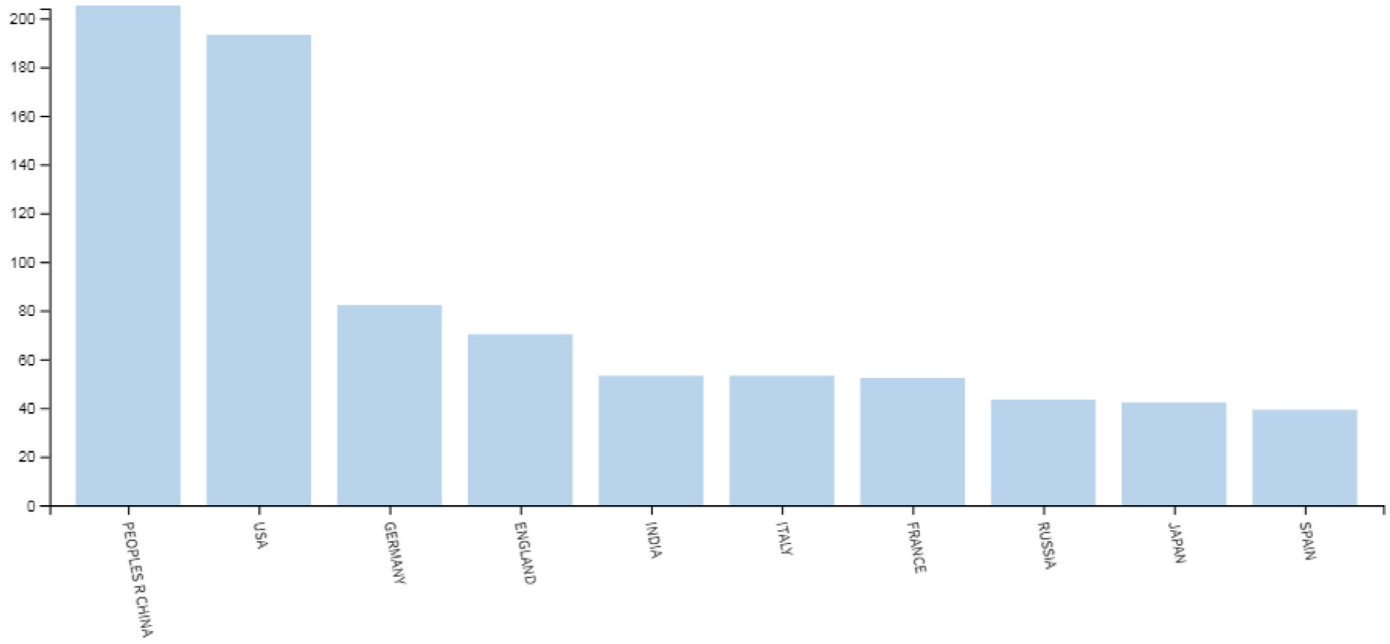

Fig. 8. Statistics of the best 10 countries published about Covid-19 on Web of Science.

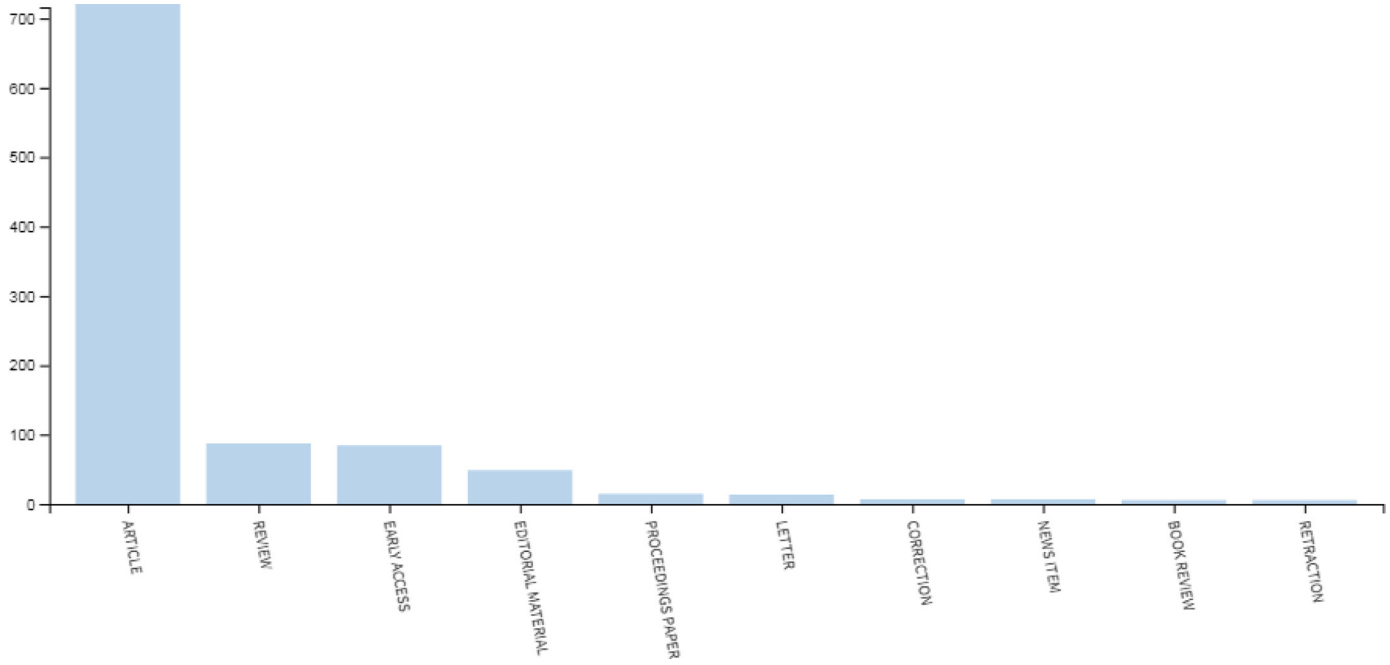

Fig. 9. Statistics of the type of document published about Covid-19 on Web of Science.

Regarding the indicators used by Scopus we find:

- H-index [11]: is based on the highest number of articles with at least the same number of citations.
- CiteScore: measures the average number of citations received per document published in the serial publication.
- SJR: measures the weighted citations received by the periodical, the weighting of the citations depends on the domain and the prestige of the citing series.
- SNIP: the standardized paper impact of the source which measures the actual citations received compared to the expected citations for the field of serial publication.

Regarding the indicators used by Web of Science we find:

- H-Index: the most used research indicator that measures both the productivity and the impact of an author's scientific production.
- The impact factor: measures the importance of a review according to the number of citations received in a year.
- Journal Citation Reports: Web of science product and an authoritative resource for impact factor data.

In the present case study, the keywords employed are "Covid-19" / "Coronavirus" from the beginning of 2020 (date of the start of the pandemic). The search should focus mainly on the

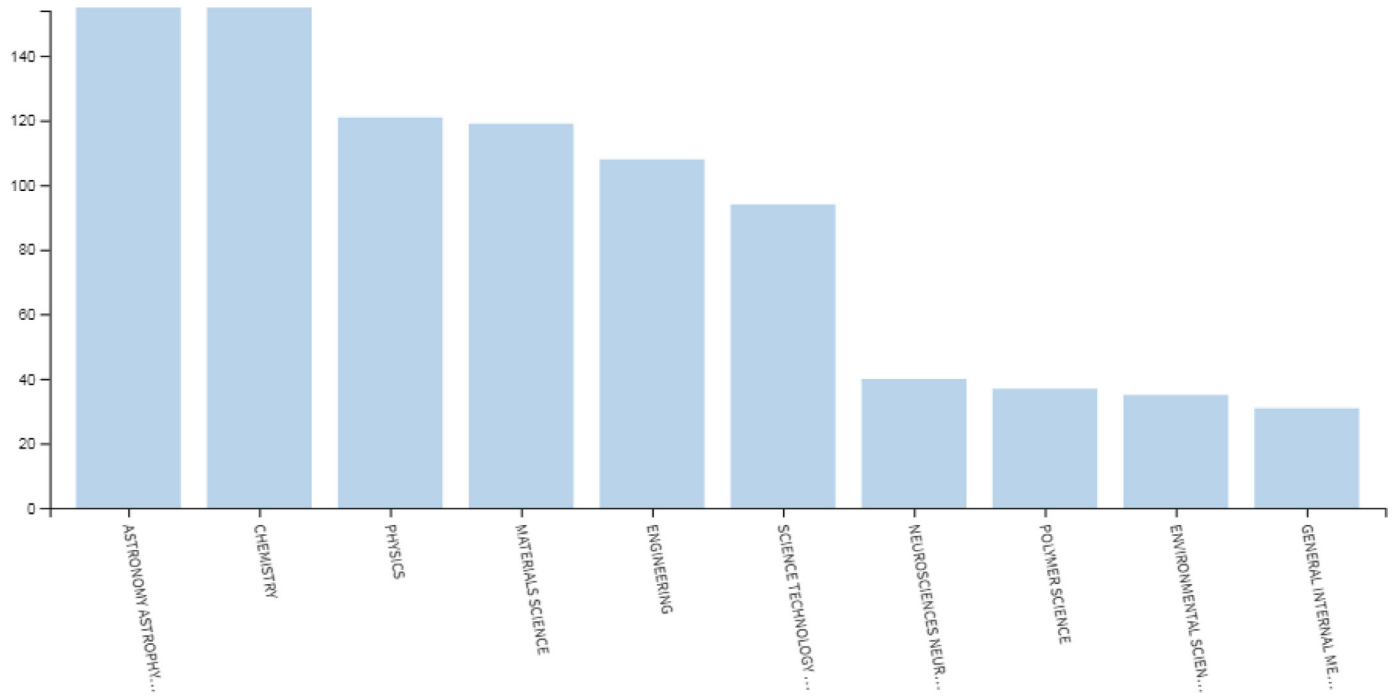

Fig. 10. Domain statistics published for Covid-19 on Web of Science.

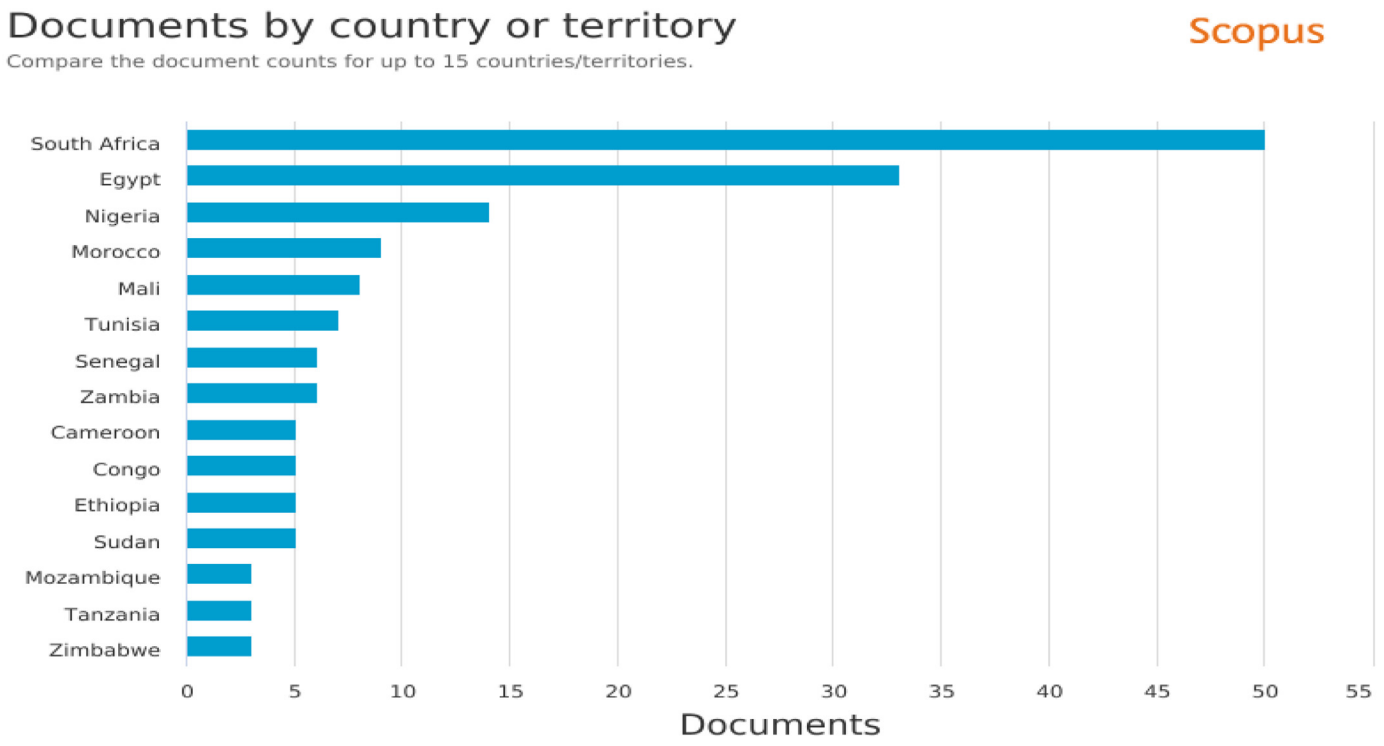

Fig. 11. Statistics of the best 10 African countries published for Covid-19 on Scopus.

titles, keywords and abstracts of articles in each of the databases. Then the results found for each of the three databases (Scopus, Web of science, Pubmed) builds our separate database on which our bibliometric analysis will be applied. We export the data from Scopus in format (.csv), Web of science, Pubmed in format (.txt).

Next, we use the VOSviewer software [12] which represents a high-performance solution with numerous viewing options with co-quotation, co-word, co-author network analysis.

4.1. Identification and analysis of research trends on Covid-19

Through bibliometric analyzes we try to get the trends of scientific research in the theme of Covid-19.

4.1.1. Analysis of authors, institutions and countries

In order to observe and evaluate the trends in publications in the thematic of Covid-19, the VOSviewer software was used to analyze the academic literature and examine the evolution of pub-

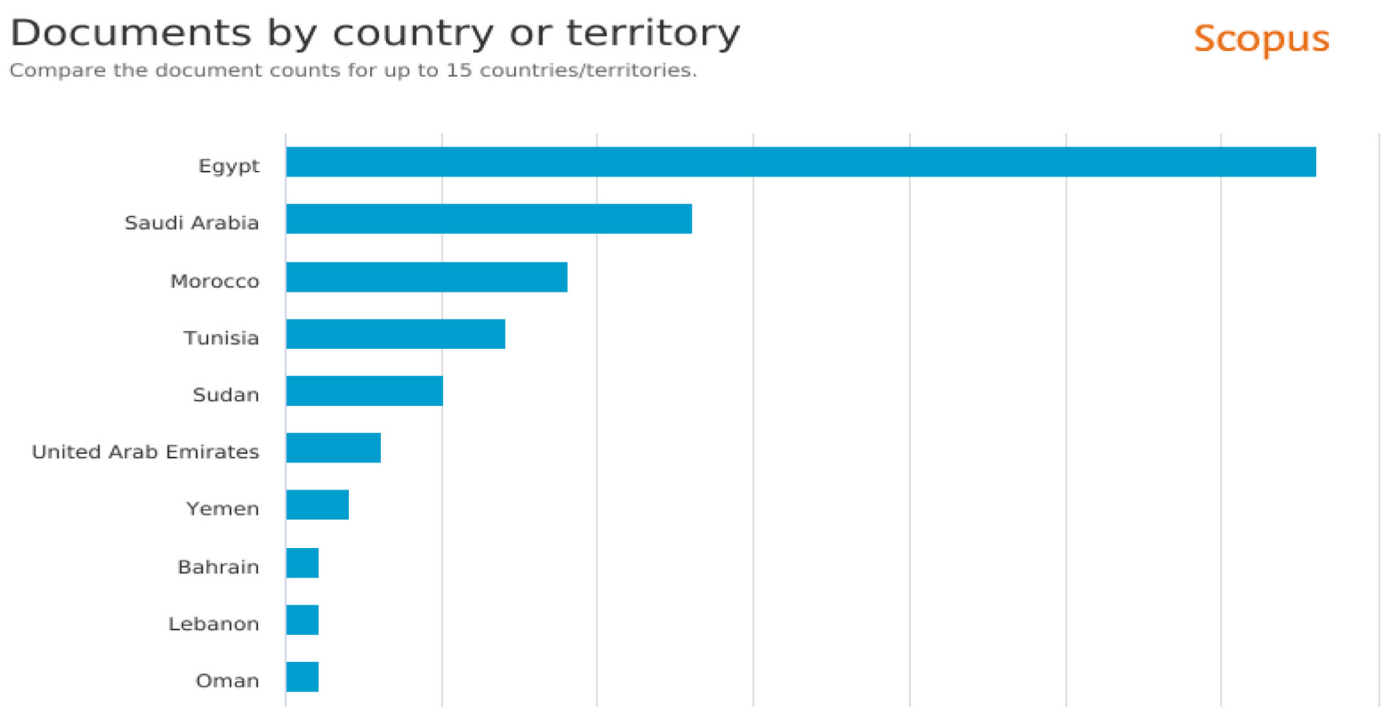

Fig. 12. Statistics of the best 10 Arab countries published about Covid-19 on Scopus.

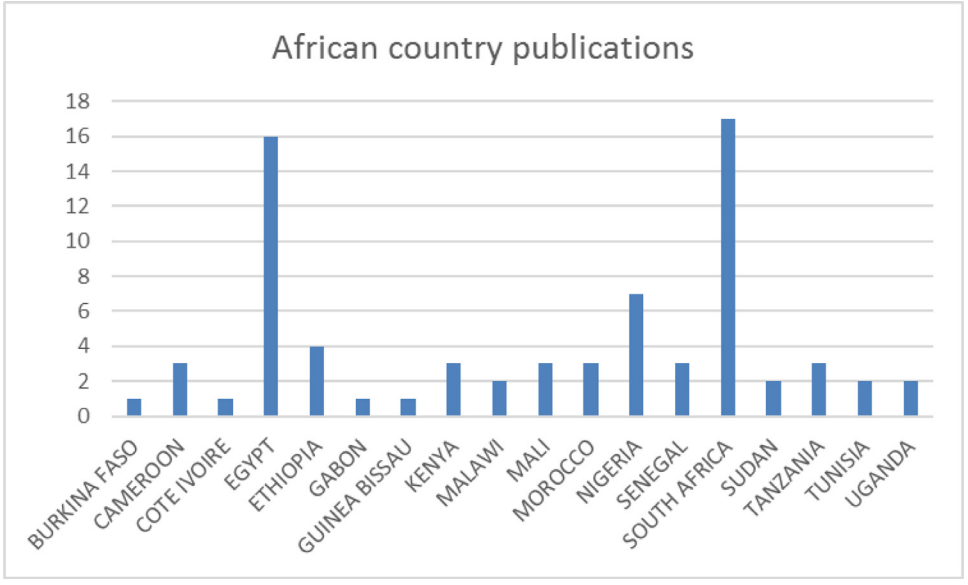

Fig. 13. Statistics of the best 10 African countries published about Covid-19 on Web of Science.

lished articles, co-authorship, geographic area (country) of authors, co-citation, co-occurrence.

The analysis of the authors belonging to the database allows to have a global view on the authors active in the thematic by offering the possibility to follow the work of these researchers by opening the door to achieve cooperation and partnerships.

Thus, the analyzes of research institutions and countries constitute an effective asset for finding the pillar institutions in each field, with the aim of seeking possible cooperation at the level of research institutions.

The software used for viewing and mapping the structure of a research are including Bibexcel, Histcite, Citespace, Gephi, and VOSviewer. For this work, we chose to work with VOSviewer because it allows us to easily display and interpret the display of large bibliometric maps.

In order to carry out the various analyzes previously cited and to examine the evolution of the articles published, we have for:

❖ Scopus:

- For authors:

We have 21 clusters distributed as follows:

Cluster 1-2: 42 items; Cluster 3: 29 items; Cluster 4-5: 27 items;
Cluster 6-7-8: 26 items; Cluster 9: 25 items; Cluster 10: 23 items;

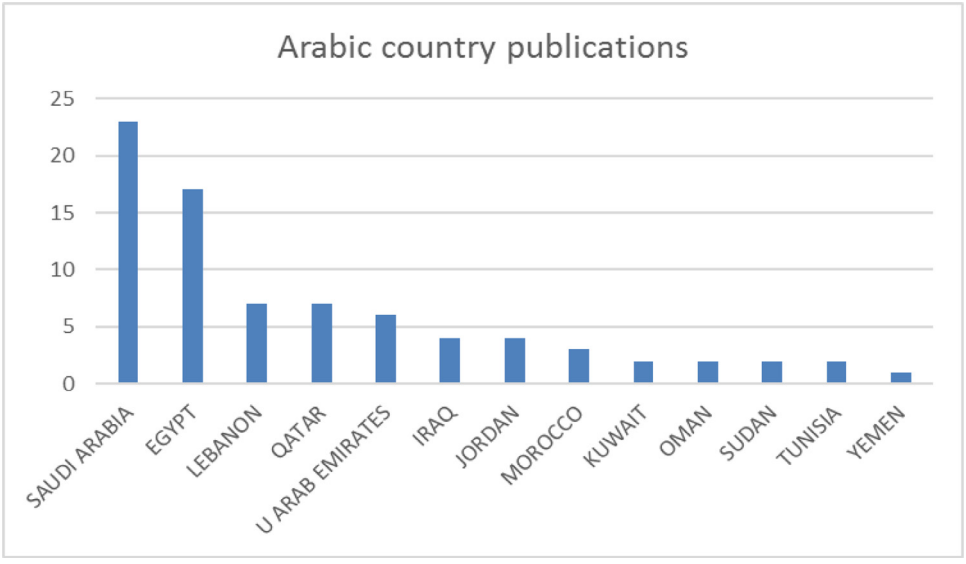

Fig. 14. Statistics of the best 10 Arab countries published about Covid-19 on Web of Science.

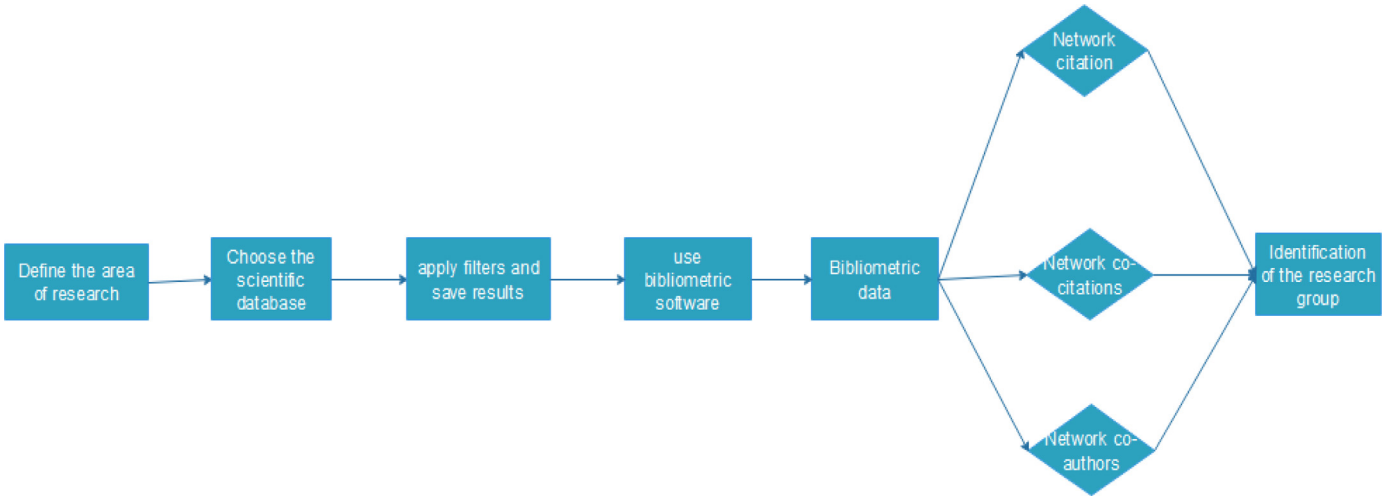

Fig. 15. Stage of the proposed bibliometric method.

Cluster 11: 21 items; Cluster 12-13: 19 items; Cluster 14-15: 16 items;
Cluster 16:15 items; Cluster 17: 14 items; Cluster 18: 13 items;
Cluster 19-20: 11 items; Cluster 21: 7 items.

The results clearly show that there are 21 groups of researchers collaborating with each other.

- For institutions:

We have 1 cluster which contains 12 items.
We deduce that most institutions collaborate with each other on an international scale and not at the regional or continental level.

- For countries:

We have 9 clusters distributed as follows:

Cluster 1-2-3: 5 items; Cluster 4-5-6: 4 items; Cluster 7-8-9: 3 items.

As we see in Fig. 21, the map indicates a large node representing China which means the great involvement of the Chinese giant through these researchers in the various research fields related to Covid-19.

❖ World of Science:

- For authors:

Bibliometric studies are used to identify networks of researchers or to map the structure of researchers in a given research area.
We have 9 clusters distributed as follows:

Cluster 1:46 items; Cluster2: 46 items; Cluster3: 20 items; Cluster 4:16 items;
Cluster 5:15 items; Cluster 6:11 items; Cluster 7: 11 items; Cluster 8: 10 items;
Cluster 9: 10 items.

The results clearly show that there are 9 groups of researchers who collaborate. Two groups have a significant number of researchers despite an exponential increase in the number of publications since the start of the pandemic, international collaboration between the authors remains low.

- For institutions:

The network analysis of research institutions with the highest number of links in this area are the institutions of the

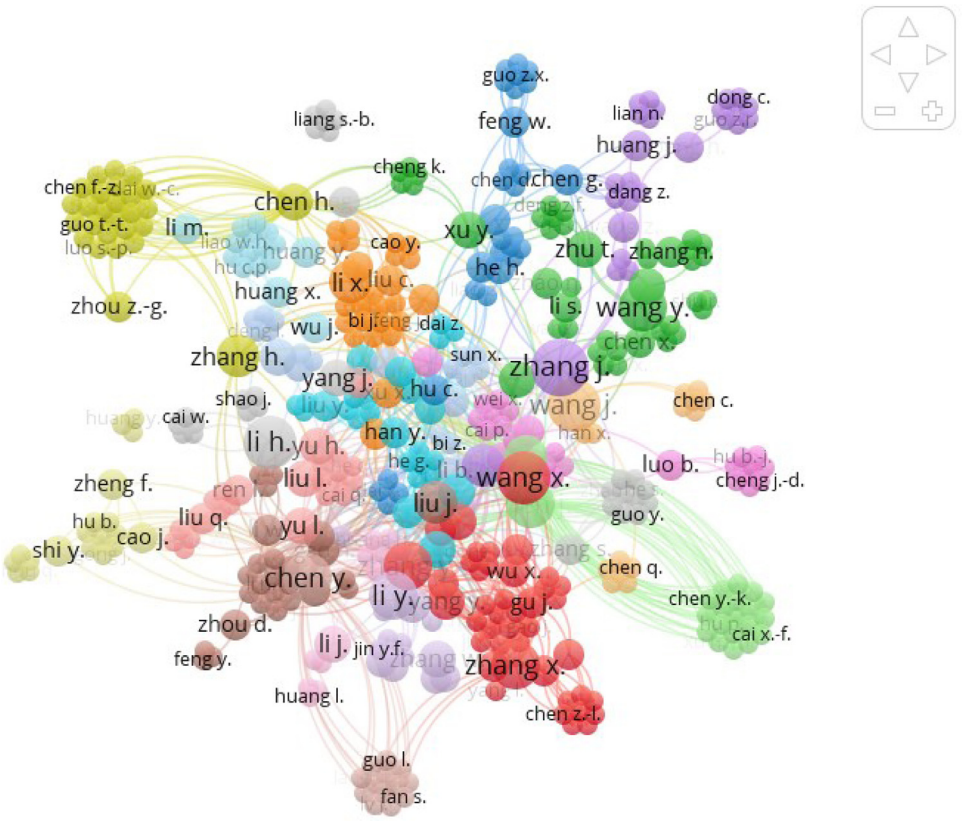

Fig. 16. Author co-authorship network in the "Network visualization" display mode.

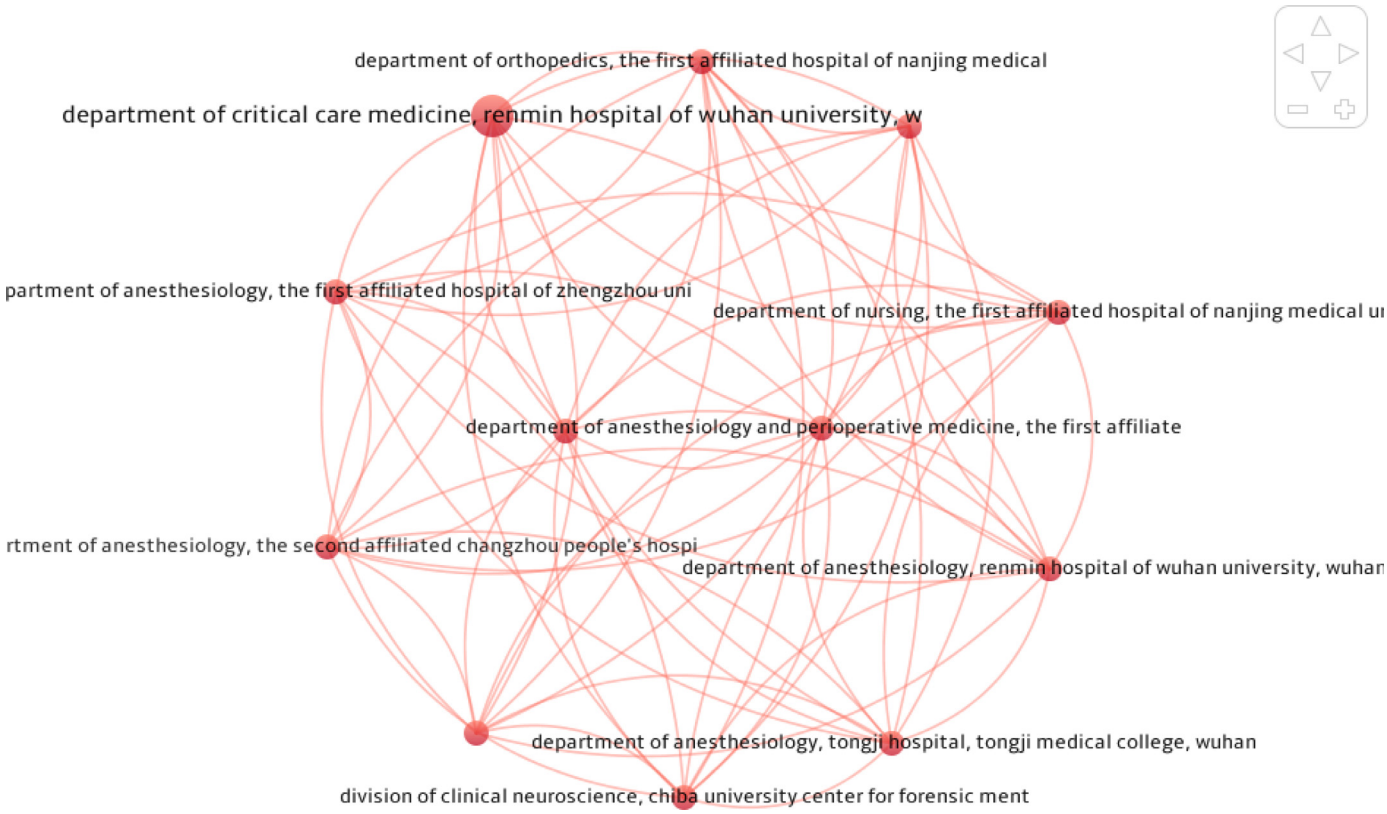

Fig. 17. Author organizations network in the "Network visualization" display mode.

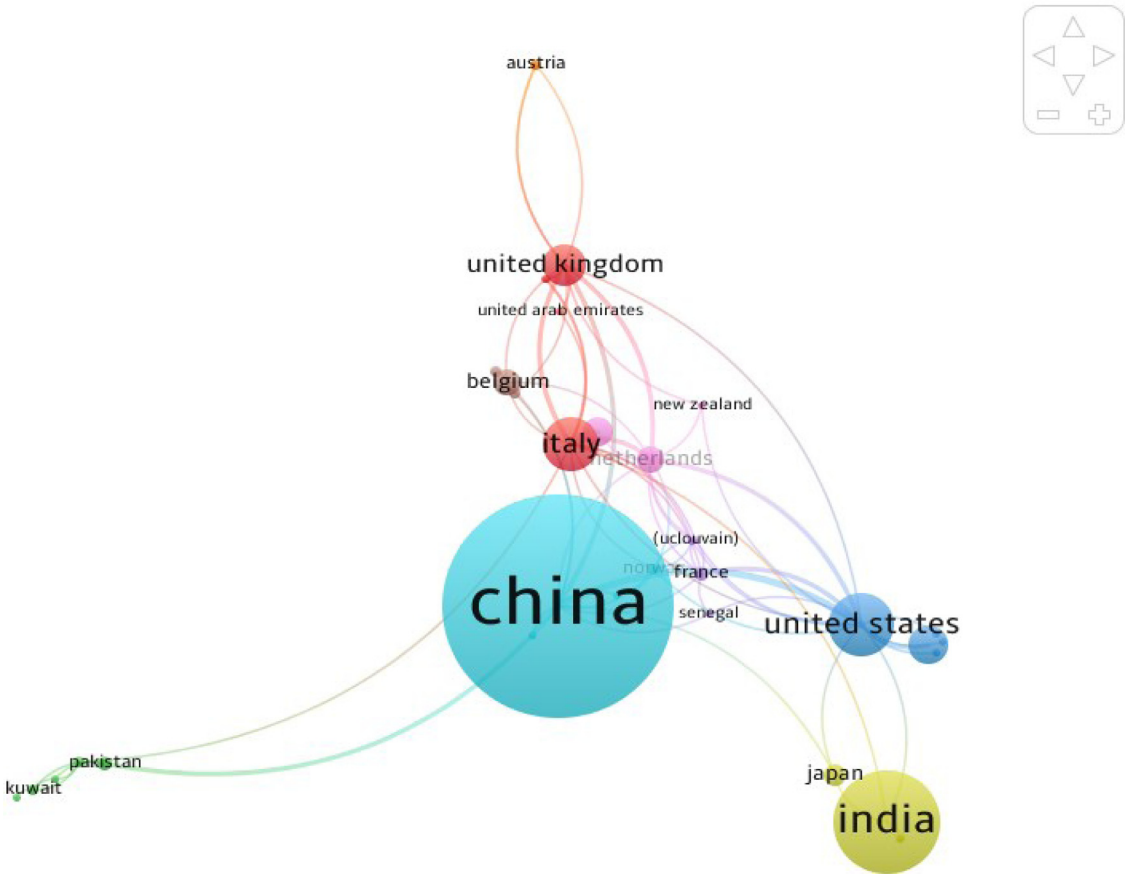

Fig. 18. Country organizations network in the "Network visualization" display mode.

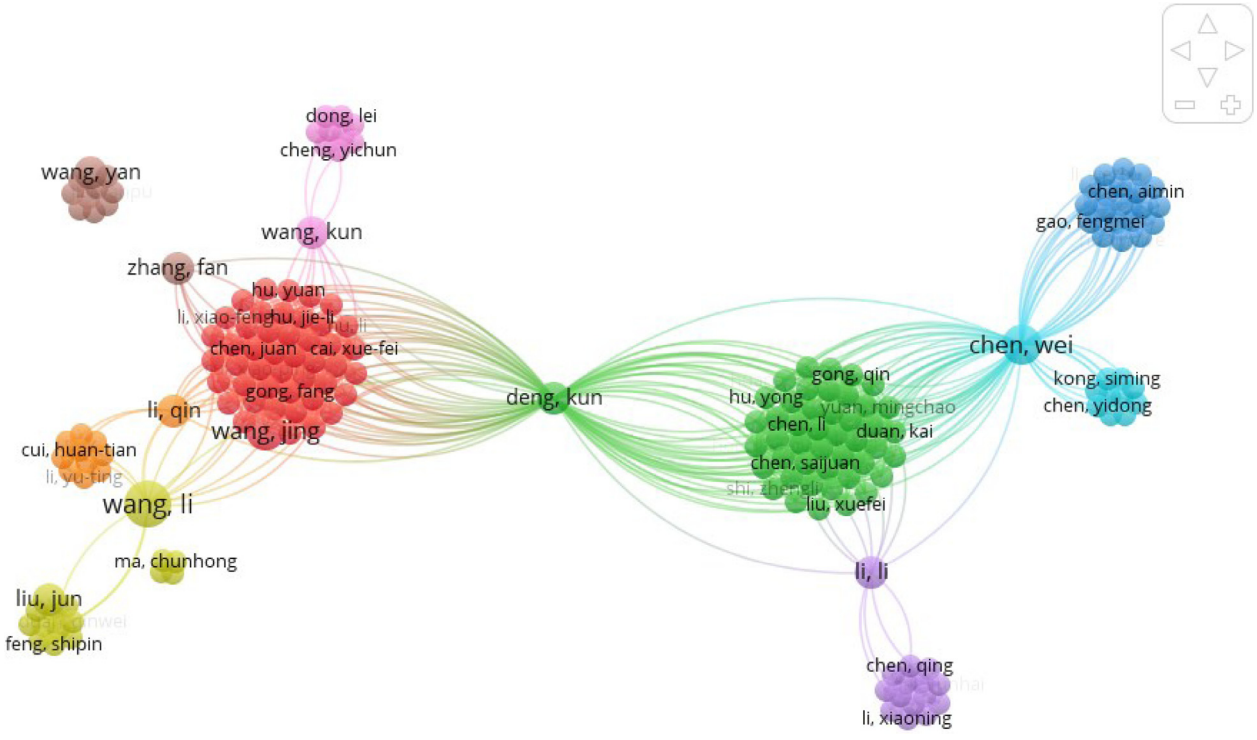

Fig. 19. Author co-authorship network in the "Network visualization" display mode.

Fig. 20. Author organizations network in the "Network visualization" display mode.

United States and China. In other words, the institutions in China and the United States have the highest total liaison force for collaboration with various institutions from different continents.

We have 31 clusters distributed as follows:

Cluster 1:33 items; Cluster2: 31 items; Cluster3: 30 items; Cluster 4:28 items;
Cluster 5-6: 27 items; Cluster 7-8: 25 items; Cluster 9-10: 24 items;
Cluster 11-12: 23 items; Cluster 13-14: 21 items; Cluster 15-16: 20 items;
Cluster 17-18: 16 items; Cluster 19-20: 15 items; Cluster 21: 14 items;
Cluster 22: 13 items; Cluster 23: 10 items; Cluster 24: 9 items;
Cluster 25-26: 7 items; Cluster 27: 6 items; Cluster 28-29-30-31: 5 items.

From the results found, it can be deduced that geographic proximity between institutions tends to strengthen the collaborative relationships of institutions. Thus, it warns of the need to expand cooperation in other regions, countries or continents.

- For countries:

The analysis of the network of countries is an important form of analysis which makes it possible to visualize the most influential countries in a given field of research, thus it exposes the degree of scientific cooperation between the countries.

We have 11 clusters distributed as follows:

Cluster 1: 7 items; Cluster 2-3: 6 items; Cluster 4: 5 items; Cluster 5: 4 items;
Cluster 6-7-8: 3 items; Cluster 9-10-11: 2 items.

As we can see in Fig. 18, the map shows a large node representing the countries and regions with the highest number of publications: China, United States, Italy, England, France and Spain Figs. 19 and 20.

❖ Pubmed:

- For authors:

We have 6 clusters distributed as follows:

Cluster 1:27 items; Cluster 2-3-4: 15 items; Cluster 5: 7 items; Cluster 6: 4 items.

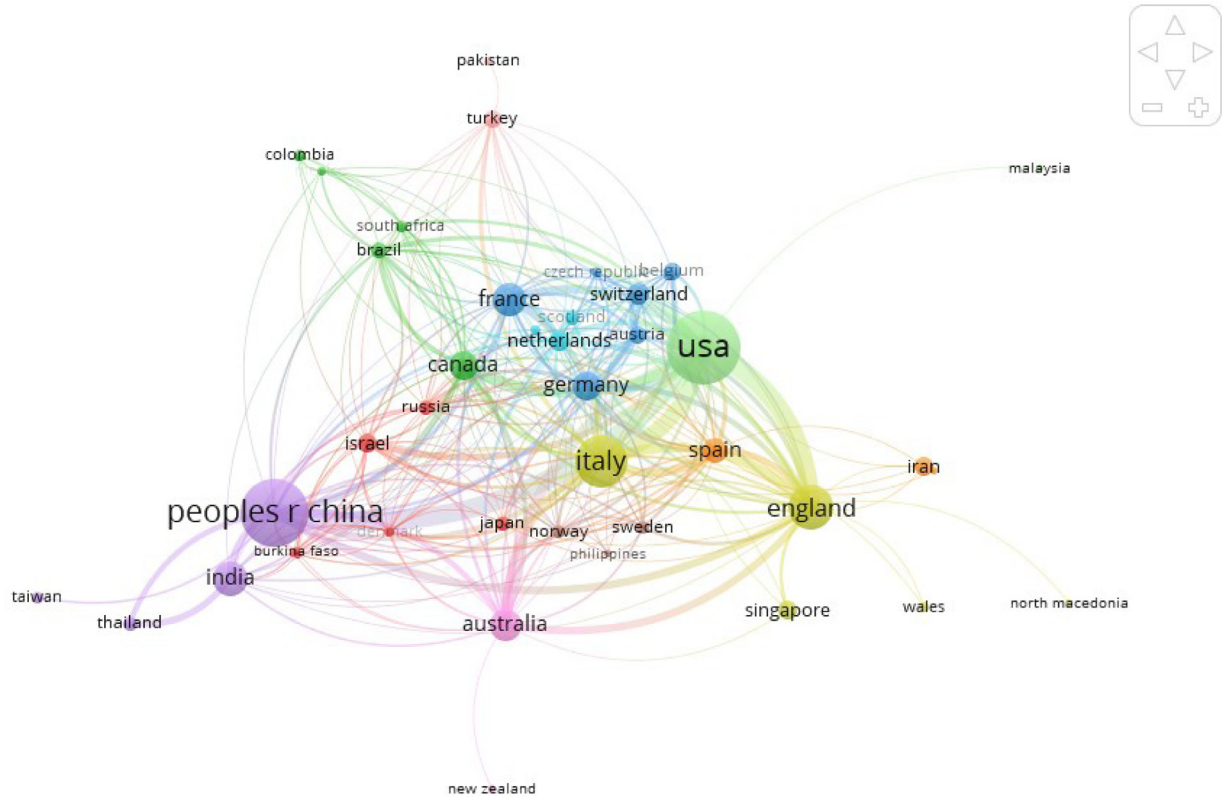

Fig. 21. Country organizations network in the "Network visualization" display mode.

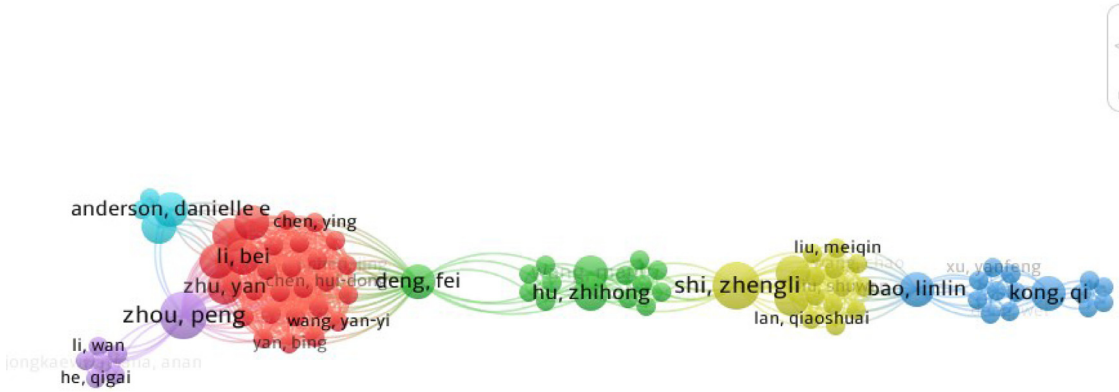

Fig. 22. Author co-authorship network in the "Network visualization" display mode.

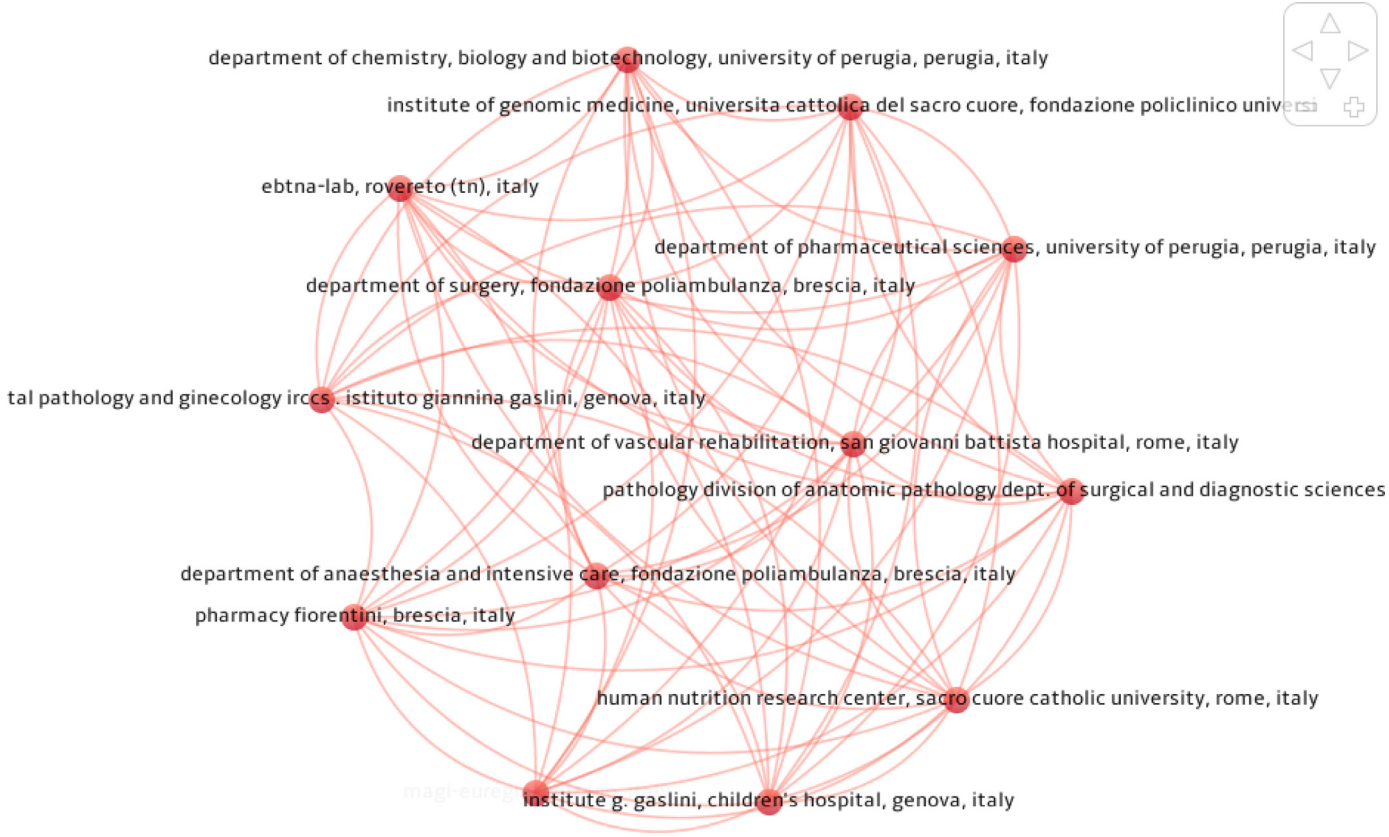

Fig. 23. Author organizations network in the "Network visualization" display mode.

The results clearly show that there are 6 groups of researchers who collaborate with each other, a group has a large number of researchers, followed by a group that is distinguished by the number of researchers who compose them.

- For institutions:

In 1 cluster with 13 items, we notice that there is a significant presence of Italian medical institutions, the analysis of data from Pubmed by VOSviewer does not offer the possibility of analyzing the network of countries.

4.1.2. Analysis of keywords

❖ VOSviewer:

We have 3 clusters distributed as follows:

Cluster 1: 6 items; cluster 2-3: 4 items.

The results found build a map dividing the keywords into three groups with the minimum number of occurrences of a keyword fixed at 6 elements for the first group and 4 elements for the second and third group. The keyword "Coronavirus" has the highest occurrence and total binding strength, other keywords with a high occurrence include "Sars-cov-2", "Covid-19" Figs. 22 and 23.

❖ Wordle [13]:

Among the existing display means, there is the word cloud which is a practical tool allowing to have a dimensional visualization of the keywords most used in the database. For our case, we use wordle which is an analysis tool which makes it possible to display a word cloud which gives greater importance to the words which appear more frequently in the source text, for the three scientific databases already mentioned, we find:

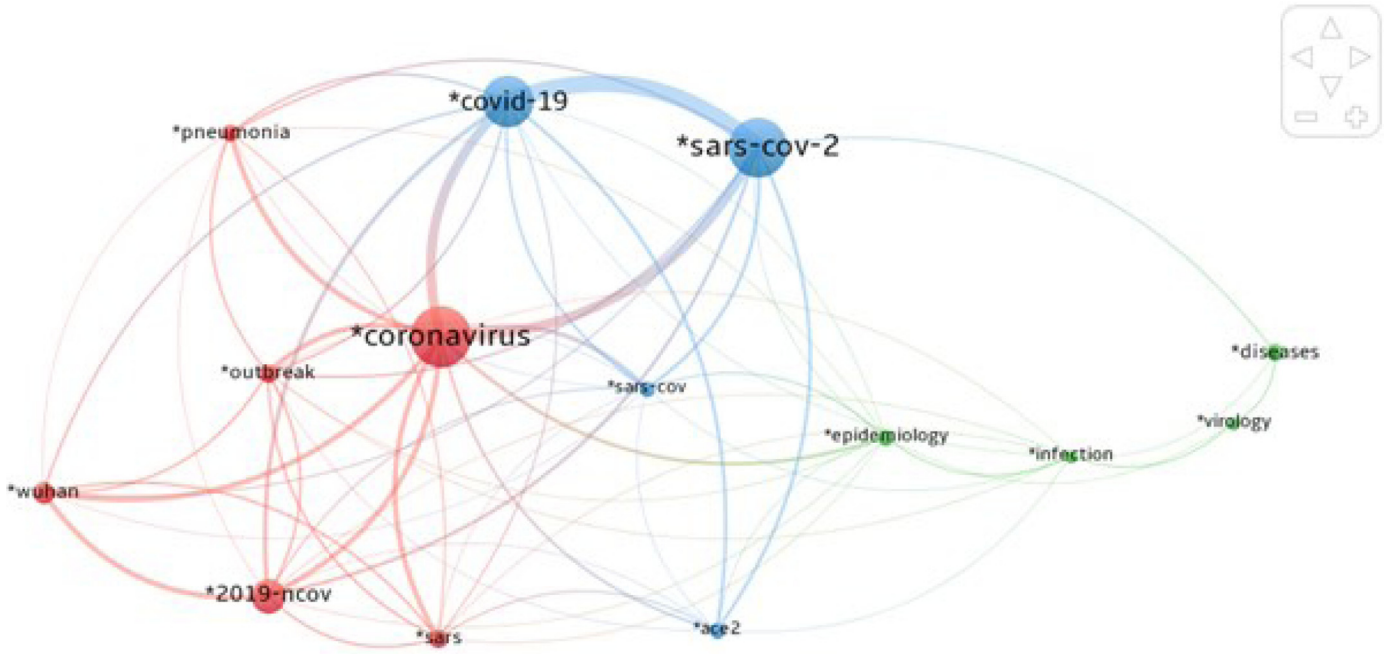

Fig. 24. Author keywords network in the "Network visualization" display mode.

Fig. 25. Wordle-generated image of the words Covid-19 for Scopus.

Fig. 26. Wordle-generated image of the words Covid-19 for Web of science.

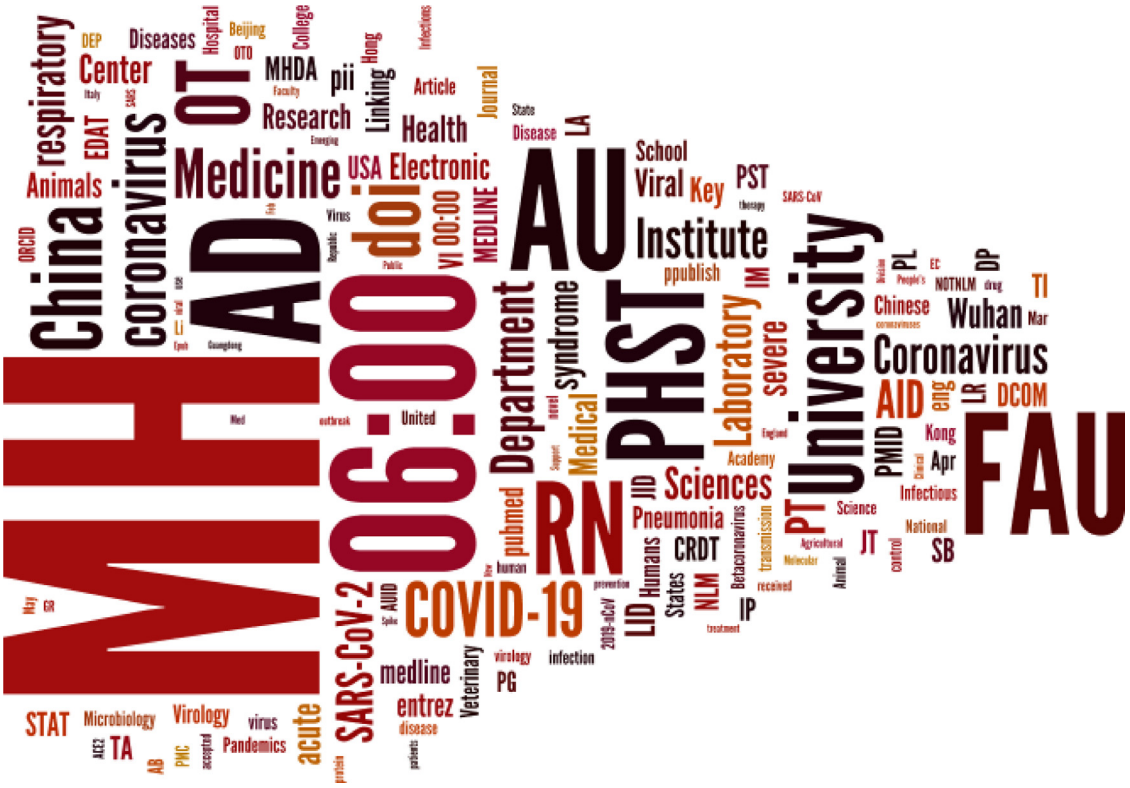

Fig. 27. Wordle-generated image of the words Covid-19 for Pubmed.

- Scopus:
- Web of Science:
- Pubmed:

From the Figs. 24–27 provided by VOSviewer and Wordle, a set of words related to the pandemic such as (Covid-19, Coronavirus, Sars-cov-2, 2019-ncon) as synonyms used in scientific literature, so the appearance of terms (China, Wuhan, USA) refers to the place of appearance of the pandemic and the countries that are conducting research to find the vaccine, too (Medical, Health, Hospital, virology) refers to the most concerned research area, (Zhang, Wang) for the most productive researchers in the topic of Covid-19 in scientific databases.

5. Conclusion

Since the onset of the pandemic, considerable effort has been invested by researchers worldwide depending on the fields and resources available, an exponential increase in scientific production has been recorded in the various databases around the Covid-19.

In this work, we opted for a statistical study for the data from the bibliographic databases Scopus, Web of Science for the theme of Covid-19. The scientific contribution of researchers from the USA and China shows a total involvement of institutions from these two countries, so for the African continent researchers from "South Africa and Egypt are the exception, while for the Arab region Saudi Arabia and Egypt are leading the efforts of the Arab countries for this pandemic.

Afterwards, a bibliometric analysis method was adopted in order to map the state of the art on the theme of Covid-19, so the three scientific databases (Scopus, Web of Science, Pubmed) were used. Thus, the search must be precise and planned by combining the precision of the terms to be used and adequate filters to refine the results found, in order to conduct a relevant bibliometric analysis by analyzing the contributions of the authors, institutions, countries and the words- keys.

Finally, it is well known that the method presented remains applicable for other scientific themes and not only for the Covid-19 theme, it should be noted that the results obtained with the application of the proposed method may vary depending on the basis of scientific data chosen and the appropriate filters in order to present the evolution of published articles, co-authors, geographic area of the authors, co-citation, co-occurrence analysis and keywords.

Declaration of Competing Interest

We wish to draw the attention of the Editor to the following facts which may be considered as potential conflicts of interest and to significant financial contributions to this work. [OR] We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome. We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us. We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

References

[1] Sengupta IN. Bibliometrics, informetrics, scientometrics and librametrics: an overview. Libri 2009;42(2):75–98.
[2] Tiernoa NR, Gonzalez-Cruz TF, Martinez JL. An overview of qualitative comparative analysis: a bibliometric analysis. J Innov Knowl 2017;2(1):15–23.

[3] Mingers J, Leydesdorff L. A review of theory and practice in scientometrics. Eur J Oper Res 2015;246(1):1–19.
[4] Kelleher C, Wagener T. Ten guidelines for effective data visualization in scientific publications. Environ Model Softw 2011;26(6):822–7.
[5] https://coronavirus.jhu.edu Accessed 23/05/2020.
[6] https://www.who.int Accessed 23/05/2020.
[7] https://www.scopus.com Accessed 23/05/2020.
[8] https://webofknowledge.com Accessed 23/05/2020.
[9] M. Franceschet, "A comparison of bibliometric indicators for computer science scholars and journals on Web of Science and Google Scholar", Vol 83(1), Pages 243–258, 2010.
[10] https://pubmed.ncbi.nlm.nih.gov Accessed 23/05/2020.
[11] Bornmann L, Daniel H-D. Does the h-index for ranking of scientists really work? Scientometrics 2005;65:391–2.
[12] van Eck N, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 2010;84(2):523–38.
[13] Viegas FB, Wattenberg M, Feinberg J. Participatory visualization with Wordle. IEEE Trans Vis Comput Graph 2009;15(6):1137–44.