*European Research Studies Journal*
*Volume XXIV, Issue 3B, 2021*

*pp. 1012-1026*

# Leading Research by Institutions and Authors: A Modern Research Analysis

*Submitted 04/07/21, 1st revision 14/08/21, 2nd revision 14/09/21, accepted 30/09/21*

Wieslawa Gryncewicz[1], Monika Sitarska-Buba[2]

***Abstract:***

***Purpose:*** *The article concludes on the importance of scientometric analysis in the process of finding partners for scientific cooperation.*
***Design/Methodology/Approach:*** *Scientometric analysis.*
***Findings:*** *The paper quantitatively characterizes the literature thematically related to the area of data science. The scientometric method was used and the Scopus database was the source of data for analysis. The study made it possible to determine when the first publications on data science appeared and how the dynamics of the growth of the literature in this range developed in years 2006-2021. The authors identified leading researchers and their area of interest, along with having reported statistics that describe top countries, institutions, subject areas, and founding sponsors.*
***Practical Implications:*** *Executing scientific analysis can help to identify institutions and researchers to build cooperation partneship.*
***Originality/value:*** *Our study has shown developing areas of data science methodology to identify leading institutions and researchers.*

***Keywords:*** *Data science, bibliometric analysis, visualization map.*

***JEL Code:*** *D83, C13.*

***Paper type:*** *Research article.*

***Acknowlegdement:*** *The project is financed by the Ministry of Science and Higher Education in Poland under the programme "Regional Initiative of Excellence" 2019 - 2022 project number 015/RID/2018/19 total funding amount 10 721 040,00 PLN.*

[1]*Ph.D., Wroclaw University of Economics and Business, Wrocław, Poland,*
*e-mail:* wieslawa.gryncewicz@ue.wroc.pl*;*
[2]*Ph.D. Wroclaw University of Economics and Business, Wrocław, Poland,*
*e-mail:* monika.sitarska-buba@ue.wroc.pl*;*

## 1. Introduction

In the modern digital world, business success and the ability to gain an edge in a competitive marketplace increasingly depend on access to data and on using it well. The amount of data in the world continues to grow. According to a March report by IDC, in 2025 there will already be 180 zettabytes of data (Reinsel *et al.*, 2021). On average, the amount of data collected grows by 23 percent each year, with companies producing it at a faster rate (28 percent) than consumers. The amount of data processed through IoT devices and sensors is also growing rapidly (by 33 percent). Data is coming in from all directions, in many sizes and formats. The challenge for enterprises is therefore to store it and ensure its availability and security. In this article, we will focus on yet another related aspect, namely, the appropriate processing of these data to enable their analysis and drawing conclusions. In fact, data itself usually means little and in order for it to carry any value, it has to be transformed into information and then into knowledge. By implementing appropriate data science (DS) tools and mechanisms, it is possible to extract information and knowledge from data to support business processes (Vermeulen, 2018). It is undoubtedly a key element in building competitive advantage on the market.

The subject of this article will be scientometric analysis of scientific publications devoted to the area of DS. The authors have conducted a comprehensive literature research in this field. Their result is a series of two articles, which managed to collect key information on these publications. This article is the first in this series and its aim is to identify key institutions and authors dealing with DS. In the second article, the authors focused on the analysis of subject areas and keywords within the discussed topic. Scientometric methods and indicators have made it possible not only to identify various phenomena and facts, but also to present the changes that have occurred over the years in the field in different contexts.

The motivation for writing this article was the authors' research interests, which oscillate around the issues of big data, machine learning and DS. The authors are concerned with the effectiveness of decision-making processes and search for tools and methods to improve them. The exploration of scientists and research centers that also deal with such topics has resulted in the collection of extensive research material, which inspired the creation of this article.

The remainder of the paper is structured as follows: Section two provides a short introduction to scientometrics, as a sub-field of bibliometric analysis that deals with scientific publications and a brief review of related work on scientometric analysis of DS publications. Next, section three presents the methodology and used tools applied in this paper. Furthermore, an exhaustive discussion on the bibliometric analyses of a selected dataset related to authors and research institutions is discoursed in section five. The article concludes on the importance of scientometric analysis in the process of finding partners for scientific cooperation.

## 2. Scientometric Analyses and Related Work

Bibliometrics is a set of statistical and mathematical methods used to analyse books, articles, and other publications. The subfield of bibliometrics that deals with the analysis of scientific publications is called scientometrics (Chellappandi and Vijayakumar, 2018). It enables assessment of the volume of "scientific production" reflected in scientific literature. This activity is a much more complex and complicated phenomenon, and there are also fields in which the results of research work are generally not published, such as military research or most research in industry that is confidential. Nonetheless, the analysis of the scientometric picture concerning publications on a specific topic allows for a quantitative assessment of this scientific output and makes it possible to draw conclusions about the dynamics of change in this area (Leydesdorff and Milojević, 2015).

Derek John de Solla Price is considered the father of scientometrics. He was a British physicist, historian of science and information scientist. The second historically important figure was Eugene Eli Garfield, an American linguist and businessman. He created the Science Citation Index and founded the Institute for Scientific Information which is widely used for scientometric analysis. The first Science Citation Index appeared in 1963 and covered 102,000 articles published in 1961 in 613 selected journals. The Science Citation Index was later transformed into the Web of Knowledge database and then into Web of Science. This was done initially manually until large scale electronic databases and associated computer algorithms were able to cope with the vast numbers of documents in most scientometric collections. The first automatic citation indexing was done by CiteSeer in 1997.

In the international area, in the context of all sciences, the most important bibliographic databases for documenting publications are the Web of Science and the Scopus databases. In our research we have selected the Scopus database. More information about research methodology is presented in section 3. As mentioned before, scientometrics deals with the measurement of the impact of research papers and academic journals, as well as the identification, analysis and statistical evaluation of scientific citations. In general, we can distinguish the following roles played by impact factors and other scientometric rankings (Lane, 2009):

- they allow to evaluate the dynamics of the growth of the literature in a research area,
- they are an indicator of the "productivity" of researchers,
- a measure of the impact of the researcher's work in a particular research area,
- a tool for choosing among different journals,
- a leading indicator of the status of universities,
- they are used to assess the impact of funding on scientific development,

– they are used to measure the impact of science on the economy.

Other applications of scientometric analysis may also be indicated, e.g., cooperation partners may be sought on the basis of rankings of authors or scientific institutions working on a given topic. A textual analysis approach can be used to identify the main disciplines in which research is conducted and highlight important keywords in the area. Scientometric analyses also enable the identification of new directions for future research (Donthu *et al.*, 2020; Leung *et al.*, 2017; Xu *et al.*, 2018). It is also worth noting that, whatever the purpose for which we will be using scientometric analysis is, it allows for independent evaluation and the avoidance of subjective bias.

There are many scientific works devoted to the scientometric analysis of research papers in a discipline. In the following section, we will present the state of the art of research on scientometric mapping of DS. In particular we will refer to those applications of scientometrics that relate to authors and research institutions. And this makes the present study unique in that we have not encountered such a deep analysis focused on the criteria indicated before.

Sarkar and Pal (2019), on the basis of the authors' affiliations, created a ranging of countries with the most authors publishing on DS. Other authors have focused a priori their studies on DS only on a particular country/region; for example, Gupta *et al.* (2011) – India, Singhal *et al.* (2014) - South Asia, Uddin and Singh (2014) - SAARC Countries. Singh *et al.* (2015) made a comparative study of research competitiveness of institutions/countries in one or more disciplines, including DS. Raban and Gordon (2020) focused on the evolution of bid data and DS literature, and author commitment was only one of the bibliometric indicators they used.

In our work we conducted analytical outcomes on authorship, the major contributors (institutions and individuals), and created collaboration maps presenting cooperation between research teams and individual academics.

## 3. Research Method and Tools Used

The authors decided to implement a two-stage approach to identify the list of articles to be put in the scope of the research. The first step was to determine which research database would be the best and most accurate source for the relevant article's selection. The authors verified the list of available research databases that covered, among others, EBSCO Publishing, Oxford Journals, Emerald, SAGE Journals, Science, Springer, Scopus and Web of Science. The first selection process was to identify databases that are not limited to one publisher or focus on selected science categories like medicine or management.

Comprehensive coverage in reputable journals and pervasive presence in the academic community of both databases convinced us to select Scopus and Web of Science for further research (Caviggioli and Ughetto 2019; Donthu *et al.*, 2020).

In the second step authors verified available selection's criteria and results analysis in both databases to determine which one will be selected for the bibliometric and content analysis in DS area.

Initial search criteria in Web of Science are topics that includes title, abstract, author keywords, and Keywords Plus, title, authors, year and few others that were not valid in our case. Initial search criteria in Scopus are very similar however deep-dive analysis of first results is more flexible. Scopus offers further limitations based on subject area and keyword, where in Web of Science only categories analysis is available. In both databases results analysis is very similar, reports presenting number of articles grouped by year, author, category, affiliation, country or organizations.

Scopus was chosen due to significant features i.e. deeper keywords limitation in search results, which allowed us to identify most relevant papers. The keyword selection was based on an initial review of available literature and panel discussion with experts from Wroclaw University of Economics, where they conduct post-graduate study in DS area. The first search was conducted with the keywords 'data science' and the first 30 papers were reviewed to identify synonyms that should be included in search criteria. 'Big data', 'machine learning', 'artificial intelligence' and 'data mining' keywords were identified as the review results.

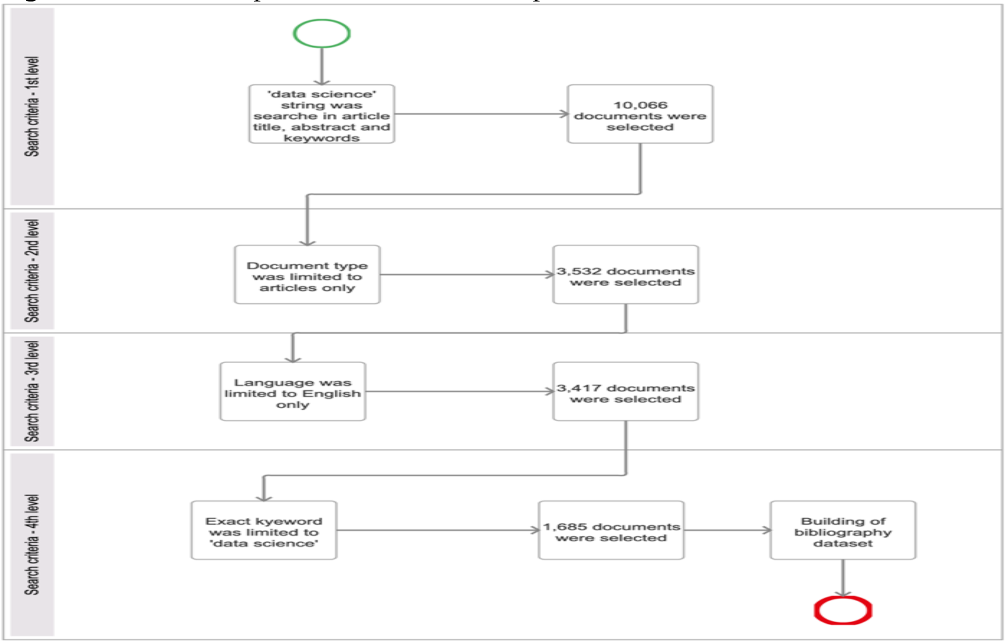
We asked our experts again to verify if these terms have been synonymously or prolifically applied in the extant literature. The outcome of the second panel stated that terms like 'big data', 'machine learning', 'artificial intelligence' or 'data mining' are part of data science category but shouldn't be considered as synonym of 'data science'.

Therefore, keyword used to scan Scopus databases for collecting research papers was limited to 'data science'. In a search query 'data science' string was searched in article title, abstract and keywords. The initial results were limited to documents that were categorized as articles and excluded all other document types, English was selected as publication language to consider only those articles that were published in this language. Lastly, we limited our search of Scopus to include only papers with exact-keyword: 'data science'.

No range was specified in terms of publication year to observe research category development over the time. The search algorithm was defined as follows: (TITLE-ABS-KEY "data science") AND (LIMIT-TO DOCTYPE, "ar") AND (LIMIT-TO LANGUAGE,"English") AND (LIMIT-TO EXACTKEYWORD, "Data Science")).

These selection criteria were utilized to create a comprehensive dataset of 1685 documents. The results were stored as comma-separated value (CSV) formats for further analysis. The search process is presented in Figure 1.

***Figure 1.*** *The search process executed in Scopus database*

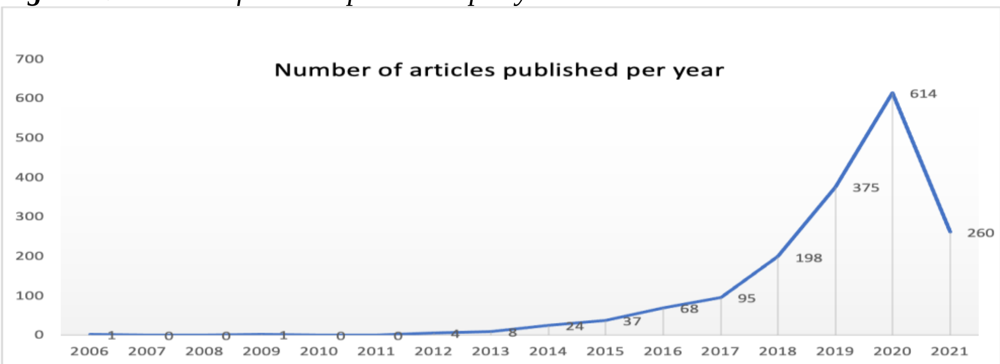



***Source:*** *Own diagram presenting search process.*

### 3.1 Initial Dataset Analysis

The 1685 articles that were included in the final dataset were published between 2006 and 2021, across 954 journals. The exponential growth of the publication number is visible in 2017, which indicates that within the last 4 years DS research has been doubled and explored by more research institutions (Figure 2). Based on the current trend it is very probable that by the end of 2021, around 1000 papers will have been published (since 2017 the number of articles has been doubled year by year).

***Figure 2.*** *Number of articles published per year*




***Source:*** *Own diagram created in MS Excel.*

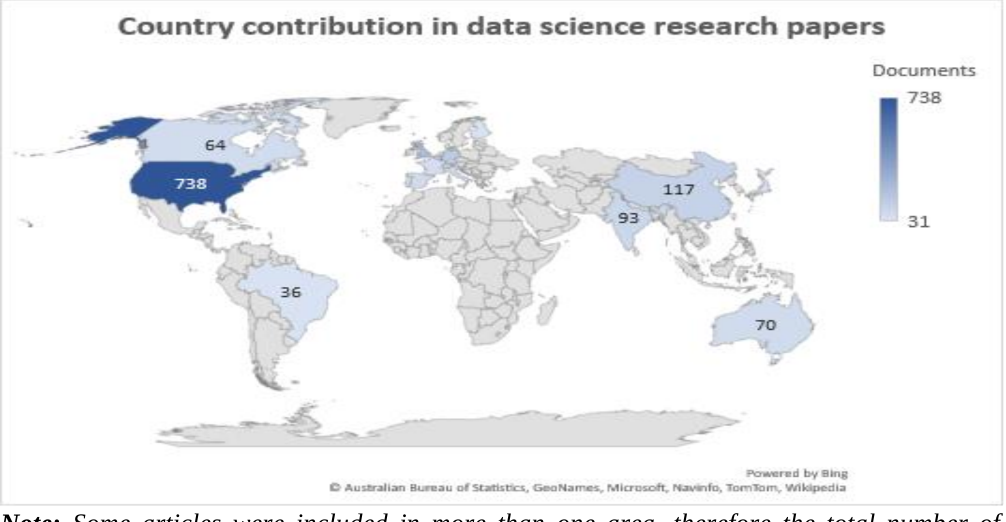
Country and region distribution of research papers prepared on the basis of the Scopus descriptive analytics shows that USA is the word leader in DS area with 738 articles (29% of total number of documents published) (Figure 3 and Table 1). The second and third place in the ranking are UK (219) and Germany (155). Authors affiliated with the USA published 3 times more than their colleagues from other regions, however two authors identified as leading contributors come from Germany, Joern Loetsch (28) from Goethe-Universität Frankfurt am Main and Amir Mosavi (23) from Bauhaus-Universität Weimar, however Mosavi provides few affiliations within one publication which means that one article is assigned to more than one country.

***Table 1.*** *Top 10 Countries by contribution.*

| Country/Territory | Documents | [%] |
|---|---|---|
| United States | 738 | 29% |
| United Kingdom | 219 | 9% |
| Germany | 155 | 6% |
| China | 117 | 5% |
| India | 93 | 4% |
| Italy | 75 | 3% |
| Australia | 70 | 3% |
| Spain | 70 | 3% |
| Canada | 64 | 3% |
| Netherlands | 60 | 2% |

***Source:*** *Own table created based on the dataset.*
***Note:*** *Some articles were included in more than one area, therefore the total number of articles in this table exceeds the total number of articles included in the final dataset.*

***Figure 3.*** *Contribution map by affiliation territory.*




***Note:*** *Some articles were included in more than one area, therefore the total number of articles in this figure exceeds the total number of articles included in the final dataset*
***Source:*** *Own diagram created in MS Excel.*

### 3.2 Leading Institutions and Authors in Data Science Area

Although the USA and the UK are leading countries by contribution, none of the top 10 authors has their affiliation placed in these countries. Table 2 presents the list of top 10 authors that published in DS area. In majority of cases, their institutions are placed in Europe, partial affiliation comes from China and one from Mexico. That conclusion forced us to run deeper analysis related to authors to verify what kind of dependencies are in place.

***Table 2.*** *Top 10 Authors by number of published articles.*

| Author | Affiliations | Documents |
|---|---|---|
| Lötsch, J. | Goethe-Universität Frankfurt am Main, Frankfurt am Main, Germany<br>Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Aachen, Germany | 28 |
| Mosavi, A. | Obuda University, Budapest, Hungary<br>Bauhaus-Universität Weimar, Weimar, Germany<br>Oxford Brookes University, Oxford, United Kingdom | 23 |
| Dehmer, M. | Private Universitat fur Gesundheitswissenschaften, Medizinische Informatik und Technik, Hall/Tyrol, Austria<br>Nankai University, Tianjin, China | 14 |
| Emmert-Streib, F. | BioMediTech Finland, Tampere, Finland<br>Tampereen Yliopisto, Tampere, Finland | 14 |
| Salas-Rueda, R.A. | Universidad Nacional Autónoma de México, Ciudad de México, Mexico | 11 |
| Ultsch, A. | Philipps-Universität Marburg, Marburg, Germany | 9 |
| Hummel, T. | Technische Universität Dresden, Dresden, Germany<br>Universitätsklinikum Carl Gustav Carus Dresden, Dresden, Germany<br>Bereich Medizin, Dresden, Germany | 8 |
| Kalso, E. | Helsinki University Hospital, Helsinki, Finland<br>Helsingin Yliopisto, Helsinki, Finland<br>University of Helsinki Faculty of Medicine, Helsinki, Finland | 7 |
| Kringel, D. | Goethe-Universität Frankfurt am Main, Frankfurt am Main, Germany | 7 |
| Tripathi, S. | University of Applied Sciences Upper Austria, School of Management, Steyr, Austria<br>University of Applied Sciences Upper Austria, Wels, Austria<br>Tampereen Yliopisto, Tampere, Finland | 7 |

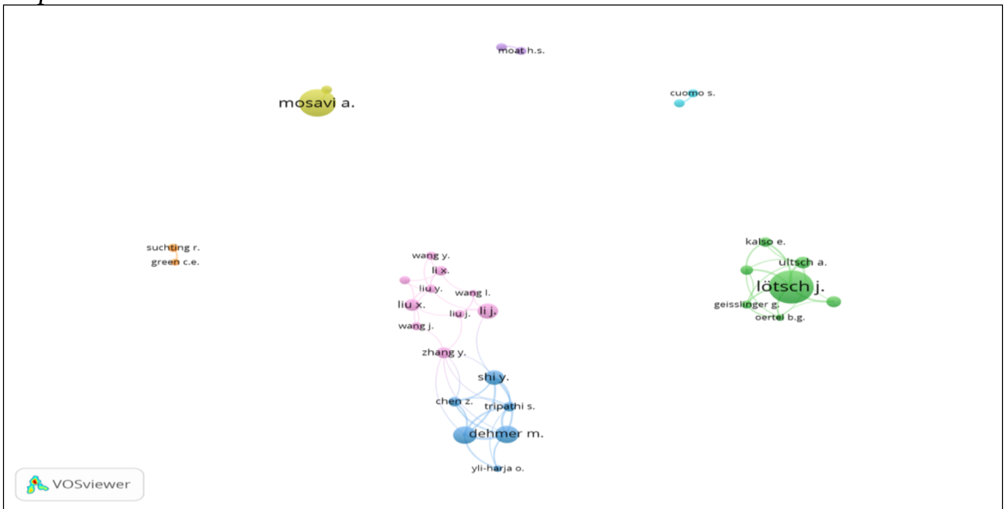
***Source:*** *Own table created based on the dataset.*

Authors performed network visualisation to better understand if there is any relation between countries and research center in context of co-authors metrics. To run network visualisation, authors utilised VOSviewer (Caviggioli and Ughetto, 2019) as one of most popular software used in other bibliometric research papers (Khanra *et al.*, 2020; 2021; Tandon *et al.*, 2021). VOSviewer enables very efficient and innovative visualisation and analysis methods of large datasets. It creates clusters of items automatically based on the selected criteria. Items which are the objects of interest in our case are authors of the publication.

Between any pair of items there can be a link. A link is a connection or a relation between two items. Items are grouped into clusters. A cluster is a set of items included in a network map (Waltman, 2020). To review relations between authors we run co-authorship analysis, with authors as an analysis unit. Minimum number of documents per author should be at least 5. Because we were interested to identify relation and connections between authors, we excluded items whose links were 0 and 1. With all these criteria we finally got a visualisation map of 7 clusters that group 31 authors (Figure 4).

In the network visualization, items are represented by their label and by a circle. The size of the label and the circle of an item is determined by the weight of the item. The higher the weight of an item, the larger the label and the circle of the item (Van Eck and Waltman, 2020).

***Figure 4.*** *Network visualisation of co-authors. VOSviewer network visualisation map.*

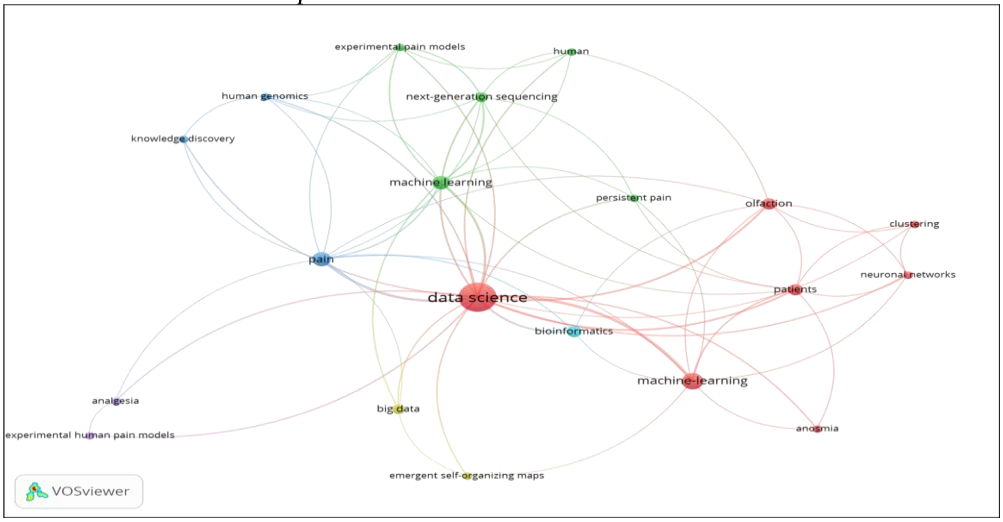



***Note:*** *Threshold criteria—minimum of five citations; resolution—1.0.*
***Source:*** *Own map created in VOSViewer.*

The largest labels have been created for J. Loetsch, A. Mosavi and for M. Dehmer and F. Emmert-Streib. This was an expected result because these authors provided the largest number of articles. The largest total link strength has been created for J. Loetsch (41) and for M. Dehmer and F. Emmert-Streib (37). Link's attribute indicates the number of co-authorship links of a given researcher with other researchers. The Total Link Strength (TLS) attribute indicates the total strength of the co-authorship links of a given researcher with other researchers (Van Eck and Waltman, 2020).

Almost all clusters are independent, which would suggest a thesis that authors are specializing in a different subject area of DS and build research consortium for the selected implementation area. The green cluster groups 7 authors: J. Loetsch, G. Geisslinger, T. Hummel, E. Kalso, D. Kringel, B.G. Oertel and A. Ultsch. The leading author in this cluster is J. Loetsch. They focused on DS and machine learning implementation within biomedical data used to build pain sensitivity phenotypes, pain models, pain interface, predictors of persistent pain for patients after different treatments like cancer, dementia, or rheumatoid arthritis-related illnesses. Keyword's network visualisation (Figure 5) created for this cluster confirms human pain-oriented research focus and data-driven decision processes. Their research is published as a part of Transnational Medicine and Pharmacology Project Group in Germany.

***Figure 5.*** *Network visualisation of author keywords for green cluster. VOSviewer network visualisation map*



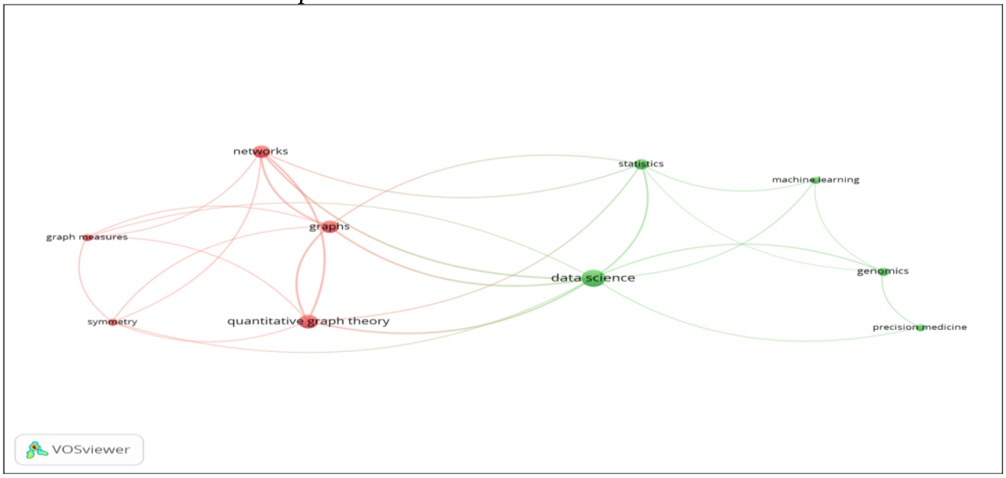

***Note:*** *Threshold criteria—minimum of five citations; resolution—1.0.*
***Source:*** *Own map created in VOSViewer.*

The blue cluster consists of 6 authors: M. Dehmer, Z. Chen, F. Emmert-Streib, F., Y. Shi, O. Yli-Harja, and S. Tripathi. Leading authors in this cluster are M. Dehmer and F. Emmert-Streib, with 14 articles. The majority of their papers focused on the improvement and development of theoretical models based on a random polynomial approach, graph theory, network science, search engine or Randić entropy.

The presented author keywords network visualisation (Figure 6) also highlighted that precision medicine and genomics are the implementation area of algorithms workout by researchers cooperating within the blue cluster. This cluster is an example of consortium of different institutions from Germany (Private Universitat fur Gesundheitswissenschaften, Medizinische Informatik und Technik), Finland (Tampere University of Technology) and China (Nankai University). It is also connected to the next cluster (pink) via Y. Zhang from School of Mathematics and

Statistics, Shandong University at Weihai, China. In this cluster it is hard to identify common research areas and the leading author.

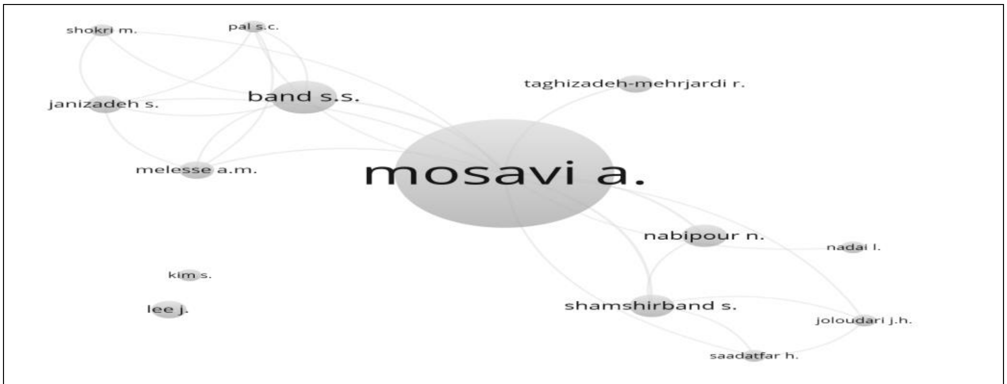
***Figure 6.*** *Network visualisation of author keywords for the blue cluster. VOSviewer network visualisation map.*




***Source:*** *Own map created in VOSViewer.*

The last cluster, with a significant number of publications inside, is the yellow one (bubble size visualises the number of items reaching analysis criteria). It includes only two authors: A. Mosavi and S.S. Band. Most probably the majority of papers were published without co-authors, or cooperation was only on a single paper. Deeper co-author analysis performed for this cluster yielded that cooperation focused on single papers. We changed visualization map criteria to display more documents and the minimum number of documents per author was set for 1. With those criteria, the following network visualization was created (Figure 7).

***Figure 7.*** *Network visualization of yellow cluster.*

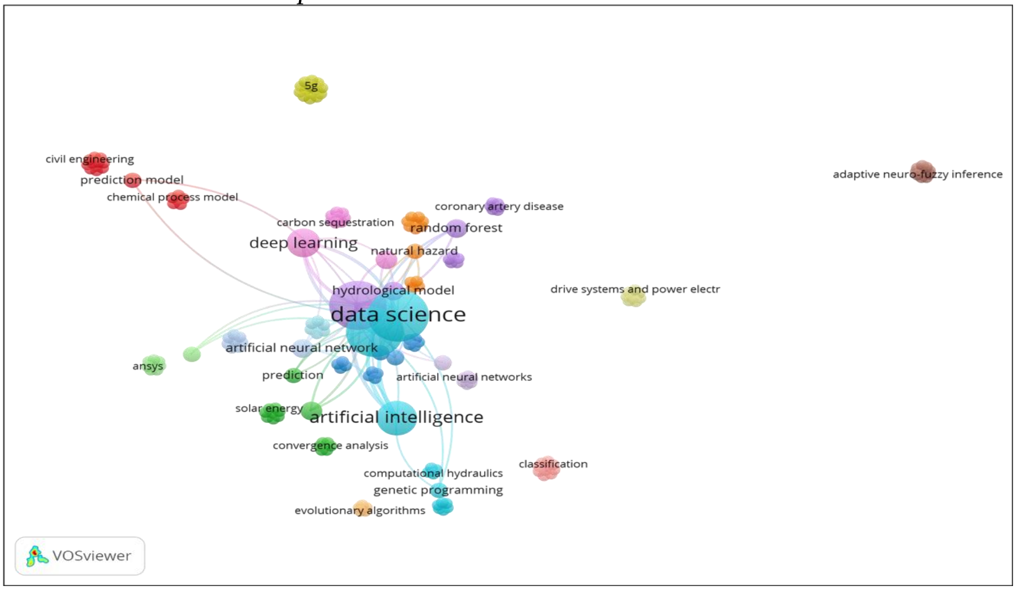



***Note:*** *Minimum number of documents has been set up to 1.*
***Source:*** *Own map created in VOSViewer.*

In contrast to J. Loetsch, who has built a permanent research team, A. Mosavi publishes his papers with individual researchers from different academic centers and research institutions. This is also likely to contribute to the broad research area covered in his publications. In some papers, the focus is on the implementation of deep learning and machine learning algorithms to predict and estimate groundwater nitrate concentration, the bank profile shape and dimensions of stable channels using the maximum entropy principle, the transverse shear stress distribution in a rectangular channel or mapping the spatial distribution of soil texture fractions. His last publications from 2021 describe algorithms implemented to model wave energy converter using long short-term memory and to model renewable energy systems.

One-time cooperation is also visible within author keywords visualization map (Figure 8). There are a lot of keywords that were specified only once and are not linked to other keywords defined.

***Figure 8.*** *Network visualisation of author keywords for yellow cluster. VOSviewer network visualisation map*




***Source:*** *Own map created in VOSViewer.*

The analysis of top 10 researchers should bring us to the conclusion that they are corelated with top 10 institutions (Table 3). The relation is not so obvious. University of Washington is in the first place in the ranking, with 35 documents. None of the researchers come from this institution. Moreover, institutions all other the USA are not connected to researchers that were listed as top 10. We do see synergy in Germany where Goethe-Universität Frankfurt am Main and Fraunhofer Institute for Molecular Biology and Applied Ecology IME are affiliations of J. Loetsch and Technische Universität Dresden is an affiliation of T. Hummel.

***Table 3.*** *Top 10 affiliations by number of documents.*

| Affiliation | Documents |
|---|---|
| University of Washington | 35 |
| Goethe-Universität Frankfurt am Main | 31 |
| University of Michigan, Ann Arbor | 27 |
| Fraunhofer Institute for Molecular Biology and Applied Ecology IME | 25 |
| University of Pennsylvania | 25 |
| Technische Universität Dresden | 24 |
| Harvard Medical School | 22 |
| Stanford University | 22 |
| Massachusetts Institute of Technology | 22 |
| University of California, Berkeley | 22 |

***Source:*** *Own map created in VOSViewer.*

Nevertheless, the top 10 affiliation ranking confirms that American institutions play a leading role as it was classified from the perspective of country and region as well. To better understand the phenomena here, we tried to create network visualization map of organizations and of authors whose documents were limited to the American region only. There are 3132 authors who represent 2412 organizations based in the USA, only 13 authors published more than 5 documents and only 5 institutions published more than 4 documents. In that situation VOSViewer would create the map of thousands of unconnected nodes, which is not a goal of the co-author analysis. That conclusion points to the thesis that the strength of the USA is built on a scale of publications, institutions and researchers whose affiliations are assigned to this country.

From that perspective (number of documents), there is no leading institution or author in the American region. The scale of institutions and authors ranked the USA as the leading region in the world.

## 4. Conclusion

In the presented dataset, the first publication related to the topic of DS appeared in 2006, while the highest growth rate in their number occurred in 2017. An exponential increase in the number of publications can be seen in 2017, indicating that in the last 4 years DS research has been intensified and is being explored by more and more research institutions. The scientometric methods and indicators made it possible to uncover the numerous connections, dependencies, and centers that play a key role in research during this period, as well as the dynamics of change over the years.

Through network visualisation, we indicated authors with the highest number of articles. An interesting finding was that although the USA and UK are leading countries in terms of research contributions, none of the top 10 authors have affiliations in these countries. In most cases, their institutions were located in Europe, some affiliations were from China and one from Mexico. On the network visualisation map we obtained independent clusters of scientific collaboration. This may indicate that the authors specialise in different areas of DS applications and build research consortia for a selected area. The keyword network visualisation allowed an in-depth investigation of the research topics of these clusters. However, we noticed that some authors work in teams and are focused on a specific area of DS implementation, while others publish mostly with different authors, from very diverse research centers in different countries, and the collaboration is usually only on one article.

This is a very interesting analysis. Based on the authors' affiliations, it was possible to identify the research centers that play a key role in DS research. It was also possible to identify the authors or even consortia that publish the most in this area. In conclusion, we can see that scientometric analysis can be a tool with which we can search for partners and research centers for scientific cooperation, look for top research directions, including those that are most supported by sponsoring institutions.

It is worth noting the unique character of this article, because such a deep scientometric investigation of scientific papers in the field of DS, related to academics and research institutions, has not been encountered by the authors so far. As mentioned in the introduction, this paper is the first of a series on scientometric analysis of DS-related scientific papers. In the second one, based on the identified main contributors, an analysis of the subject areas and keywords that appeared in their publications was conducted.

## References:

Almenberg, J., et al. 2020. Attitudes toward debt and debt behavior. The Scandinavian Journal of Economics. https://doi.org/10.1111/sjoe.12419.

Camacho, D., Panizo-LLedot, Á., Bello-Orgaz, G., Gonzalez-Pardo, A., Cambria, E. 2020. The four dimensions of social network analysis: An overview of research methods, applications, and software tools. Information Fusion, 63, 88-120.

Caviggioli, F., Ughetto, E. 2019. A bibliometric analysis of the research dealing with the impact of additive manufacturing on industry, business and society. Int. J. Prod. Econ., 208.

Chellappandi, P., Vijayakumar, C.S. 2018. Bibliometrics, Scientometrics, Webometrics / Cybermetrics, Informetrics and Altmetrics - An Emerging Field in Library and Information Science Research. Shanlax International Journal of Education, 7(1).

Donthu, N., Kumar, S., Pattnaik, D. 2020. Forty-five years of Journal of Business Research: a bibliometric analysis. J. Bus. Res., 109.

Du, Y., Khan, H.R. 2020. Data Science for Librarians. Libraries Unlimited.

Global Data Protection Index 2020 Snapshot. https://www.delltechnologies.com/en-us/data-protection/gdpi/index.htm#.

Global Data Sphere. https://www.idc.com/getdoc.jsp?containerId=IDC_P38353.

Gryncewicz, W., Sitarska-Buba, M. 2021. Data science in decision-making processes: A scientometric analysis. European Research Studies Journal, 24(3B), 1061-1074.
Gupta, B.M., Kshitij, A., Verma, C. 2011. Mapping of Indian computer science research output, 1999-2008. Scientometrics, 86(2), 261-283.
Khanra, S., Dhir, A., Mäntymäki, M. 2020. Big data analytics and enterprises: a bibliometric synthesis of the literature. Enterp. Inf. Syst., 14(6), 737-768.
Khanra, S., Dhir, A., Kaur, P., Mäntymäki, M. 2021. Bibliometric analysis and literature review of ecotourism: toward sustainable development. Tour. Manag. Perspectives, 37, 100777.
Kumar, S., Garg, K.C. 2005. Scientometrics of computer science research in India and China. Scientometrics, 64(2), 121-132.
Lane, J. 2009. Assessing the Impact of Science Funding. Science, vol. 324.
Leung, X.Y., Sun, J., Bai, B. 2017. Bibliometrics of social media research: a co-citation and co-word analysis. Int. J. Hosp. Management, 66.
Leydesdorff, L., Milojević, S. 2015. Scientometrics. International Encyclopedia of Social and Behavioral Sciences, Section 8.5: Science and Technology Studies, Subsection 85030, Elsevier.
Perron, B.E., Victor, B.G., Hodge, D.R., Salas-Wright, C.P., Vaughn, M.G., Taylor, R.J. 2017. Laying the Foundations for Scientometric Research: A Data Science Approach. Research on Social Work Practice, 27(7), 802-812.
Raban, D.R., Gordon, A. 2020. The evolution of data science and big data research: A bibliometric analysis. Scientometrics, 122, 1563-1581.
Reinsel, D., Rydning, J., Gantz, J. 2021. Worldwide Global Data Sphere Forecast, 2021-2025: The World Keeps Creating More Data — Now, What Do We Do with It All? https://www.idc.com/getdoc.jsp?containerId=US46410421.
Sarkar, A., Pal, A. 2019. Where does data science research stand in the 21st century: Observation from the standpoint of a scientometric analysis, 2561, 1-9.
Sarkar, A., Pal, A. 2020. Authorship Pattern of 21st Century Data Science Research: A Scientometric Evaluation, 4263, 6-15.
Singh, V.K., Banshal, S.K., Singhal, K., Uddin, A. 2015. Scientometric Mapping of Research on Big Data. Scientometrics, 105(2).
Singh, V.K., Uddin, A., Pinto, D. 2015. Computer Science Research: The Top 100 Institutions in India and in the World. Scientometrics, 104(2), 529-553.
Singhal, K., Banshal, S.K., Uddin, A., Singh, V.K. 2014. The information technology knowledge infrastructure and research in South Asia. Journal of Scientometric Research, 3(4), 134-142.
Tandon, A., Kaur, P., Mäntymäki, M., Dhir, A. 2021. Blockchain applications in management: A bibliometric analysis and literature review. Technological Forecasting & Social Change, 166, 120649.
Uddin, A., Singh, V.K. 2014. Mapping the Computer Science Research in SAARC Countries. IETE Technical Review, 31(4), 287-296.
Van Eck, N.J., Waltman, L. 2014. Visualizing bibliometric networks. Measuring Scholarly Impact. Springer, 285-320.
Van Eck, N.J., Waltman, L. 2020. VOSviewer Manual: Manual for VOSviewer Version 1.6.16. Univeristeit Leiden.
Vermeulen, A.F. 2018. Practical Data Science: A Guide to Building the Technology Stack for Turning Data Lakes into Business Assets. Apress.