ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/365300990

Conceptual structure and current trends in Artificial Intelligence, Machine Learning, and Deep Learning research in sports: A bibliometric review

Preprint · November 2022

DOI: 10.1101/2022.11.09.515813

CITATIONS
6

READS
150

4 authors:

Carlo Dindorf
Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau
71 PUBLICATIONS **334** CITATIONS
SEE PROFILE

Eva Bartaguiz
Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau
38 PUBLICATIONS **163** CITATIONS
SEE PROFILE

Freya Gassmann
Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau
116 PUBLICATIONS **412** CITATIONS
SEE PROFILE

Michael Fröhlich
Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau
450 PUBLICATIONS **3,237** CITATIONS
SEE PROFILE

All content following this page was uploaded by Carlo Dindorf on 11 November 2022.

The user has requested enhancement of the downloaded file.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Conceptual structure and current trends in Artificial Intelligence, Machine Learning, and Deep Learning research in sports: A bibliometric review

Carlo Dindorf[1,*], Eva Bartaguiz[1], Freya Gassmann[2], Michael Fröhlich[1]

[1]Department of Sport Science, Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau (RPTU), Kaiserslautern, Germany

[2]Department of Empirical Social Research, Rheinland-Pfälzische Technische Universität Kaiserslautern, Kaiserslautern, Germany

[*]Corresponding author

E-mail: carlo.dindorf@sowi.uni-kl.de

Abstract

Artificial intelligence and its subcategories of machine learning and deep learning are gaining increasing importance and attention in the context of sports research. This has also meant that the number of corresponding publications has become complex and unmanageably large in human terms. In the current state of the research field, there is a lack of bibliometric analysis, which would prove useful for obtaining insights into the large amounts of available literature. Therefore, the present work aims to identify important research issues, elucidate the conceptual structure of the research field, and unpack the evolutionary trends and the direction of hot topics regarding key themes in the research field of artificial intelligence in sports. Using the Scopus database, 1215 documents (reviews and articles) were selected. Bibliometric analysis was performed using VOSviewer and bibliometrix R package. The main findings are as follows: (a) the literature and research interest concerning AI and its subcategories is growing exponentially; (b) the top 20 most

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

cited works comprise 32.52% of the total citations; (c) the top 10 journals are responsible for 28.64% of all published documents; (d) strong collaborative relationships are present, as along with small, isolated collaboration networks of individual institutions; (e) the three most productive countries are China, the USA, and Germany; (f) different research themes can be characterized using author keywords with current trend topics, e.g., in the fields of biomechanics, injury prevention or prediction, new algorithms, and learning approaches. AI research activities in the fields of sports pedagogy, sports sociology, and sports economics seem to have played a subordinate role thus far. Overall, the findings of this study expand knowledge on the research situation as well as the development of research topics regarding the use of artificial intelligence in sports, and may guide researchers to identify currently relevant topics and gaps in the research.

Keywords: sports; artificial intelligence; machine learning; deep learning; training; exercise; big data; sports analytics; physical activity; athletes; technology; data mining; review; survey; meta-analysis

Introduction

In general, both artificial intelligence (AI)[1] and the subcategories of machine learning (ML) and deep learning have experienced a boom in recent years. AI is generally understood to be a type of human intelligence that is executed by machines. Machine learning is an approach to achieving AI, describing computer systems that are able to learn from examples or experiences without being explicitly

[1] Hereafter, the term artificial intelligence (AI) is used even when individual publications refer specifically to machine learning (ML) or deep learning.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

programmed. Deep learning has also emerged as an important technique for implementing ML [1].

AI has already established itself in many situations and is now indispensable. For example, AI has experienced great success in facial recognition on smartphones [2], as well as speech processing, synthesis, and translation (e.g., Google Translate, Alexa, Google Assistant [3]), and has long since become routine in everyday life. It has proven itself in numerous real-world applications, including facial recognition [4], speech recognition [5], object recognition [6], malware detection [7], and spam filtering [8]. In addition, the main areas of application can be found in safety-critical fields of application in which errors have a tremendous impact (e.g., autonomous driving [9] or healthcare and medicine [10–13]).

The AI boom in recent years, along with the progressive development of the associated approaches, has been driven by various factors—including the availability (i.e., quality and quantity) of data, which decisively influences the success of AI models [14]. There have often been limited data in the past, but today, modern technologies often generate vast amounts of large and complex datasets [15]. In contrast to the original problem with the availability of data, there are now difficulties in coping with the sheer volume of data [16]. Parallel to the increase in the amount of available data, immense developments have taken place in the field of modeling, which have contributed to models becoming more and more accurate [1]. In addition, the implementation of ML models is becoming easier and more accessible for a large number of people thanks to user-friendly open-source programming libraries such as scikit-learn [17] or PyTorch [18]. Furthermore, hardware developments, the associated increase in available computing power, and the consequent acceleration

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

of calculations play a decisive role [19]. These drivers have ensured that AI has opened up many new possibilities and gradually found its way into sports contexts and applications [20,21]. Hence, the present work is based on an understanding of the field of AI in sports as the application of methods from the general field of AI in the context of sports science and practical sports issues.

The USA is considered to be a pioneer in using AI in sports, both in research and in practice. AI has been used for game and player analysis for several years, especially in American football and baseball [22]. For example, studies have shown that the analysis of various types of data can help support coaches' tactical decision-making [23], as well as training and competition planning [24]. Furthermore, this kind of analysis has great potential for preventing injuries [20] which, in turn, is accompanied by a reduction in negative consequential costs and cancellations of competitions. Increasing numbers of major football clubs are also using AI and data analytics. First and foremost, FC Barcelona and TSG Hoffenheim have their own research labs (Barça Innovation Hub and TSG Research Lab, respectively). Liverpool are also implementing data analytics solutions and AI to support managerial decisions [25]. Even fan interactions can be monitored and analyzed for better performances [26]. On the topic of available data for analysis purposes, players are monitored via numerous different sensors during matches, competitions, and training. For example, with regard to the sub-disciplines of sports science, applications can be found in the fields of training science [24], biomechanics [27], and sports medicine [28]. Against the backdrop of the ever-increasing amounts of data, AI has become a key technology.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The research situation with respect to AI technologies in sports has also led to a large bibliographic dataset, which is impractical for manual investigation. There is a need to obtain optimal insights into the available publication data using appropriate analysis techniques, since it is crucial to analyze the relevant research and structures in depth to uncover existing knowledge, challenges, and research deficits in order to derive and justify new research necessities. A promising approach, which is becoming increasingly important in many fields of research, is bibliometric analysis. For example, bibliometric analyses are used for uncovering emerging trends, collaboration characteristics, and research constituents, as well as for exploring the intellectual structures of specific domains based on their respective literature base [29,30]. Bibliometric techniques (as well as meta-analyses) seem especially promising and suitable for making sense of the large volumes of literature, as classic review methods need to focus on a manageable amount of literature [31]. Such techniques are superior to meta-analysis if the publications in question are very heterogeneous. Furthermore, unlike classic systematic literature reviews, which tend to be affected by the interpretation bias of scholars, bibliometric analysis (or meta-analysis) relies on more quantitative methods; thus, it can avoid or at least reduce such bias [29].

Regarding the basic review situation, several works have been published to date in the context of AI and sport-specific fields of its application [21,32–35]. In addition to AI topics, numerous general works can be found in the context of sports using bibliometric methods (e.g., football [36–38]; sports tourism [39,40]; sports sustainability [41]; sports management [42]; sports nutrition [43]; sports science in general [44]). However, looking at different literature analysis methods applied in the context of sports-related AI research shows that only a few works have been

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

published to date using bibliometric techniques. The authors of [45] analyzed literature regarding big data and artificial intelligence in sports using a bibliometric approach; however, their focus was restricted to articles in the CNKI database. The authors of [46] analyzed trends of AI applications in sports, with a focus on patents. Moreover, some works have used bibliometric techniques in the context of sports, focusing on topics that overlap with AI. For example, [47] analyzed publicly available baseball data in the context of big data characteristics. Furthermore, [48] investigated applications of wearable devices in sports using bibliometric techniques.

To the best of our knowledge, there is still a research deficit in the bibliometric analysis of available literature in the field of AI in sports using a broad literature base. Thus, it can be concluded that the overall scientific landscape has not been thoroughly analyzed thus far, and great potential remains.

Therefore, the present study aims to identify important research issues, reveal the conceptual structure, and unpack evolutionary trends and the directions of hot topics regarding key themes in the research field of AI in sports. Due to the presence of large bibliographic datasets, as well as the broad and heterogeneous literature in the context of machine learning research in sports, the application of a bibliometric analysis appears to be excellently suited to the intended object of investigation.

The following contributions of research constituents were analyzed using bibliometric methods through general performance analysis, and their relationships were identified using science mapping. Longitudinal development and current trends were further analyzed by looking at the occurrence of keywords over time and in the latest research publications. For in-depth background information on the applied analysis techniques, see [29], for example.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Methods

Fig 1. General workflow of the study.

For this study, data from the Scopus database were used (accessed on 6 September 2022). To map the most prominent subcategories of artificial intelligence, the terms “machine learning” and “deep learning” were included in the search, and “artificial intelligence” was the general term. It should be noted that the Scopus search engine is not case-sensitive and considers hyphens as spaces.

The document types selected were articles and reviews, as commonly performed in literature searches [40,49], with a focus on works of the highest quality. As usual in bibliometric analyses, we focused on one language only (i.e., English). Due to the broad range of different subject areas addressed by publications in the context of sports, the exclusion of inappropriate literature based on subject areas of the database was not suitable for the aims of this work. Inappropriate literature in this context can be defined as publications that use the term “sport*” in the sense of a conceivable application of their research but do not carry out the research in the direct context of sports. However, in the present work, the focus was on works that are closely related to sports. We dealt with this problem using a literature search focusing on titles and the keywords, searching only for publications that included the term “sport*” in their title or keywords. Overall, this procedure allowed us to ensure that sports were the direct focus of the research, rather than a side issue. However, there was also the potential for inclusion of papers whose association with sports is debatable, depending on the individual and cultural understanding of the nature of sport. The final search procedure is shown in detail in Figure 1, resulting in 1215 documents from 428 journals (1157 articles, 58 reviews).

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Based on the final search, a csv file was exported. Erroneous or incomplete entries, as along with duplicates, were removed. Furthermore, abbreviations were replaced with their full designations (e.g., AI → artificial intelligence; ML → machine learning), while grammatical numbers were unified in the keywords (e.g., sport → sports) in the exported csv file. Incorrect and inappropriate keywords were manually removed.

Bibliometric analysis was performed using VOSviewer [50] and the R package bibliometrix, as well as the corresponding web interface biblioshiny [51]. VOSviewer and bibliometrix are commonly used in bibliometric research [49,52,53]. For a detailed overview of the software tools used, see [54]. Accordingly, the selection of the used software tools can be justified, since bibliometrix stands out due to the large number of different analysis methods available, while VOSviewer provides outstanding network visualizations.

In addition to the general performance analysis of the contributions of research constituents (e.g., articles, journals, institutions, authors, countries) to the considered field—using publication counts reflecting productivity and citations for measuring impact and influence [29]—the following science mapping methods were applied to identify the relationships between the research constituents: Co-authorship analysis focuses on revealing social interactions and relationships between authors, institutions (co-authorship institutions), and/or countries (co-authorship countries). Co-citation analysis identifies relationships among cited publications to reveal underlying themes in the research field. The co-occurrence of two publications in another publication´s reference list forms a connection between them [29]. As co-citation analysis focuses only on highly cited publications,

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

bibliographic coupling was further applied to increase the visibility of recent and niche publications [29]. The analysis of bibliographic coupling shows which articles and authors are correlated with one another through multiple citations. A bibliographic coupling is present if two documents cite the same reference [55]. As it is recommended to use bibliographic coupling for specific timeframes [56], our analysis focused on the two recent periods 2019-2020 and 2021-2022.

Scopus uses the indexed keywords for the search; however, our keyword results are presented for the author keywords. This decision was based on our objective of starting from a broad, exploratory point of view in order to make the best use of any potentially relevant information and then focusing on aspects seen as relevant from the perspective of the scientific community. In addition to the general analysis of the total keyword occurrence, a longitudinal analysis of the keywords' occurrence per year was also performed. This analysis was intended to reveal trending topics of AI research in sports based on the author keywords, which are generally sufficient to conclude topical aspects in a research field, as they are connected to the content of publication [57].

Furthermore, based on the author keywords, a thematic map of AI in sports was constructed. In general, the aim was to obtain insights into the current status of the field as well as to provide knowledge regarding the possible future development of research themes within the field. Therefore, the author keywords and their interconnections were clustered into themes. Characterization was performed through the properties density (i.e., measurement of the cohesiveness among the nodes, or internal cohesion; vertical axis) and centrality (i.e., measurement of correlation between different topics, or external cohesion; horizontal axis) [58]. The

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

more relationships a node shows with others in the network, the greater its centrality and importance. The density representing the cohesiveness among the nodes gives clues as to the capability of a research field to develop and sustain itself [59].

Results

General performance analysis

In terms of the various Scopus subject areas, “Computer Science” (n = 741) ranked in first place, followed by “Engineering” (n = 517) and “Mathematics” (n = 242). The first publications were recorded in the year 1988 [60]. There was a steady increase in the annual number of publications, with the first major increase in 2012 and a second sharp and continuous increase from 2017 onwards (see Figure 2). Therefore, this research field can be thought of as being relatively young, especially in relation to the recent increase in publications. The context of the COVID-19 pandemic in the years 2020–2022 seems to have had little impact, with no visible drop in publications compared to other research fields, e.g., life sciences [61]. However, due in part to long publication times, it remains unclear whether the pandemic´s effects will become apparent with a shift in the next few years.

Article citation analysis is the most widely used method to study the impact of authors, journals, and articles, as it identifies the key works in a research area [62]. Overall, on average, every document had 9.53 citations based on the global citation counts using Scopus data. Along with the total increase in the number of published documents over time, a parallel increase in the average number of citations per year was observable (see Figure 2).

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

An anomaly was detected for the year 2007 in terms of the total citations per document. This can be explained by the appearance of the most cited article "Checkers Is Solved" (Schaeffer et al., 2007) (see also: Table 1), which caused an atypical total citation count in relation to the early development of AI research in sports during that time. Checkers is a strategy board game that can be understood as a sport, similar to chess. In terms of citations per document, an unequal distribution was present, where the majority of works were only cited a few times. The proportion of papers with two or fewer citations was 53.33%. There are two potential reasons for this: the articles were published very recently, or they do not generate sufficient academic interest to be cited by the research community.

Fig 2. Evolution of the annual scientific production as the number of articles per year (left y-axis) and mean total citations per year (right y-axis). TC = total citations

Overall, the reviewed documents were published in 428 different journals. The 15 most productive journals in terms of the number of published documents are presented in Figure 3. The top 10 journals are responsible for 28.64% of all published documents. The most productive journal was *Computational Intelligence and Neuroscience* (h-index: 61; 2021 impact factor: 3.64), followed by *Sensors* (h-index: 196; 2021 impact factor: 4.35) and *IEEE Access* (h-index: 158; 2021 impact factor: 4.34). It is notable that the top-ranked journals include not only journals with specific aims and scopes, but also those with a broader possible range (e.g., *PLOS One*). However, journals that focus on information technology and engineering comprise the majority. This indicates that AI in sports is interdisciplinary, because most journals do not focus on sports. The *Journal of Sport Science* is only ranked 10th, which can be taken as a sign that AI is not the main topic in sports science at all.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig 3. The top 15 most productive journals for research articles in the context of sports and artificial intelligence.

Top 20 overall most cited works

The top 20 most cited works (see Table 1) account for 32.52% of the total citations. The most common theme in the top 20 most cited works is human activity or action recognition (e.g., [32,63–67]), followed by predictive systems (e.g., [68,69]) and systems for object or player tracking (e.g., [70–72]). Only one study investigated AI in the direct context of motion analysis [73]. Most studies developed an algorithm or a new methodological approach (e.g., [74–76]). The other studies focused on reviewing the literature [32,63,70,73,77]. The focus of the data was mainly on wearables (e.g., [63,65,67,76]), followed by visual data (e.g., [71–73,78]). Team sports were analyzed more frequently compared to other sports (e.g., [69,70,77]).

The publications that were a maximum of four years old and featured among the top 20 most cited publications in this context (referred to here as “hot papers”; see also: Table 1) could take on special importance in future research if they continue to be cited over the next few years.

279 **Table 1. The top 20 most cited works.** TC = total citations; normal. = normalized; = “hot papers” published in the last four years.

Title	Journal	First Author	Publication Year DOI	TC	TC per year
Deep Learning Algorithms for Human Activity Recognition Using Mobile and Wearable Sensor Networks: State of the Art and Research Challenges	*Expert Systems with Applications*	H. F. Nweke	2018 10.1016/j.eswa.2018.03.056	404	80.8
Using Text Mining and Sentiment Analysis for Online Forums Hotspot Detection and Forecast	*Decision Support Systems*	N. Li	2010 10.1016/j.dss.2009.09.003	365	28.08
Learning Sparse Representations for Human Action Recognition	*IEEE Transactions on Pattern Analysis and Machine Intelligence*	T. Guha	2011 10.1109/TPAMI.2011.253	296	26.91
Checkers Is Solved	*Science*	J. Schaeffer	2007 10.1126/science.1144079	277	17.31
A Deep Learning Approach to on-Node Sensor Data Analytics for Mobile or Wearable Devices	*IEEE Journal of Biomedical and Health Informatics*	D. Ravi	2016 10.1109/JBHI.2016.2633287	254	42.33
Learning to Select and Generalize Striking Movements in Robot Table Tennis	*The International Journal of Robotics Research*	K. Mülling	2013 10.1177/0278364912472380	241	24.1
Big Data and Tactical Analysis in Elite Soccer: Future Challenges and Opportunities for Sports Science	*SpringerPlus*	R. Rein	2016 10.1186/s40064-016-3108-2	205	29.29
Haptic-Feedback Smart Glove as a Creative Human–Machine Interface (HMI) for Virtual/Augmented Reality Applications	*Science Advances*	M. Zhu	2020 10.1126/sciadv.aaz8693	184	61.33

Recognizing Daily and Sports Activities in Two Open Source Machine Learning Environments Using Body-Worn Sensor Units	*The Computer Journal*	B. Barshan	2014 10.1093/comjnl/bxt075		179	17.9
Recognizing Human–Object Interactions in Still Images by Modeling the Mutual Context of Objects and Human Poses	*IEEE Transactions on Pattern Analysis and Machine Intelligence*	B. Yao	2012 10.1109/TPAMI.2012.67		174	15.82
A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System	*Sports Medicine - Open*	S. L. Colyer	2018 10.1186/s40798-018-0139-y		154	30.8
Deep Learning for Fall Detection: Three-Dimensional CNN Combined With LSTM on Video Kinematic Data	*IEEE Journal of Biomedical and Health Informatics*	N. Lu	2019 10.1109/JBHI.2018.2808281		138	34.5
Spatio-Temporal Analysis of Team Sports	*ACM Computing Surveys*	J. Gudmundsson	2018 10.1145/3054132		133	22.17
Learning to Track and Identify Players from Broadcast Sports Videos	*IEEE Transactions on Pattern Analysis and Machine Intelligence*	W.-L. Lu	2013 10.1109/TPAMI.2012.242		132	13.2
Machine Learning for Activity Recognition: Hip versus Wrist Data	*Physiological Measurement*	S. G. Trost	2014 10.1088/0967-3334/35/11/2183		127	14.11
Real-Time Posture Reconstruction for Microsoft Kinect	*IEEE Transactions on Cybernetics*	H. P. H. Shum	2013 10.1109/TCYB.2013.2275945		109	10.9
Machine and Deep Learning for Sport-Specific Movement Recognition: A Systematic Review of Model Development and Performance	*Journal of Sports Sciences*	E. E. Cust	2019 10.1080/02640414.2018.1521769		104	26
A Machine Learning Framework for Sport Result Prediction	*Applied Computing and Informatics*	R. P. Bunker	2019 10.1016/j.aci.2017.09.005		97	24.25

Multiple Player Tracking in Sports Video: A Dual-Mode Two-Way Bayesian Inference Approach With Progressive Observation Modeling	*IEEE Transactions on Image Processing*	J. Xing	2010 10.1109/TIP.2010.2102045	97	8.08
Relationships Between Internal and External Training Load in Team-Sport Athletes: Evidence for an Individualized Approach	*International Journal of Sports Physiology and Performance*	J. D. Bartlett	2016 10.1123/ijspp.2015-0791	94	15.67

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Countries, organizations, and global structures

AI in sports is a global research topic. Analysis of the total authorship data of each document showed that 128 countries contribute to the research on this topic. However, China leads in terms of the total number of publications by a considerable margin, followed by the USA and Germany (see Table 2). On average, each publication has 3.65 co-authors. Regarding the mean numbers of authors from the most productive countries, China has the fewest authors per document. This can be explained by the fact that 27.33% of the publications with at least one Chinese author had only one author; in comparison to the USA and Germany—with 7.38% and 2.99% single authorship, respectively—this is a relatively high proportion. It should be noted that 21.15% international co-authorship is also present; therefore, a single publication can be counted in several countries.

Table 2. Top 10 most productive countries based on the author´s countries, total citations (TC) per country based on document citations, and mean number of authors per country based on the authors´ countries.

Countries	N publications	TC	Mean number of authors
China	1244	1723	2.72
USA	591	1031	5.27
Germany	204	842	4.40
UK	170	671	4.39
Australia	169	815	4.54
South Korea	149	395	3.56
Japan	134	139	5.39
Spain	134	159	5.02
Canada	125	796	4.23
India	112	107	3.76

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

With 38 articles, the Hospital for Special Surgery (New York City, New York) takes first place among the most productive organizations, followed by Victoria University (Melbourne, Australia) and Ghent University (Ghent, Belgium) (Figure 4a). Surprisingly, although China is the most productive country in terms of publication output, the leading Chinese university only ranks in 6th place (Beijing Sport University). This indicates that many institutions in China seem to be involved in the publications without any individual institution being particularly active. The aforementioned high proportion of single authorship in Chinese publications, without collaborations between institutions, also appears to be related to this.

It is notable that the trend toward multi-university publications has continued to increase, especially in recent years (Figure 4b). This could be interpreted as a sign that AI in sports has moved out of its initial niche status. An examination of the collaboration between organizations (Figure 5) shows that publications are often produced through cooperation between individual universities. There are larger associations of collaborating universities as well as islands formed through small collaboration networks. The resulting networks may be due to the need to bundle expertise, since individual universities mainly provide sporting or IT expertise, which must be brought together to process the research object in the sense of interdisciplinary or transdisciplinary research. According to the existing small, unconnected collaborative networks, networking still has great potential insofar as there is a fit in terms of the content and methodology of the research institutes.

Fig 4. (a) Top 10 most productive organizations based on authors' affiliations. (b) Distribution of publications according to the involvement of one vs. multiple universities.

Fig 5. Collaboration network organizations using the bibliometrix package (parameters: Walktrap clustering algorithm, 500 nodes displayed). Exemplary subnetworks are also presented, zoomed in with the names of exemplary organizations for visualization reasons.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Examination of the international collaboration, which indirectly depicts the connections of the authors from different countries, shows that five clusters of collaboration groups are present (Figure 6). The top three most productive countries (i.e., China, the USA, and Germany) are in different clusters. In the blue cluster, the most productive countries are China, South Korea, and India. In the yellow cluster, the USA is the most active country, followed by Spain and Canada. In the red cluster, the most productive countries are Germany, Turkey, and Japan. In the green cluster, the United Kingdom is the leader, followed by Australia and Slovenia. It is striking that geographic proximity is a strong explanatory factor but is not necessarily related to cluster affiliation. For example, Germany collaborates strongly with the nearby countries of the Netherlands and Belgium, but also with Japan. The same applies to the cooperation between the USA and Spain.

Substantial collaboration is visible between the different clusters. However, the collaboration between China and the other main clusters is lower than the collaboration of the remaining clusters. Using the terms of the World Bank for the classification of the countries [79], it seems that high-income countries and middle-income countries collaborate more within each cluster.

Fig 6. Collaboration world map/co-authorship countries: The minimum number of documents per country was set to 10, resulting in 31 countries. Cluster memberships are distinguished by color. The node size reflects the productivity of the country based on the number of publications in collaboration with other countries; the shorter the distance between the nodes, the more closely connected the countries are (based on co-authored articles).

Authors and social structures

The research field of AI in sports is characterized by a large number of authors. In total, 3411 different authors published their works during the period of time investigated in this review. Only 188 authors of single-authored publications

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

were present. On average, each publication shows 3.65 co-authors, and 21.15% of the papers are written by international author groups. The majority of the most productive authors are located in China. Due to the relatively high level of correspondence between names according to the country-specific characteristics in China [80], the most productive authors are not presented here, so as to rule out any possible errors or injustice, since authors were not identified on the basis of a clearly identifiable ID (e.g., ORCID ID) but on the basis of their first and last names which, due to the abovementioned situation, may have led to an incorrect merging of authors with the same first and last names. However, as it is reported that co-citation analysis is relatively immune to author name ambiguity issues [81], the corresponding mapping was performed to reveal the social structures of the authors (Figure 7).

In the center of Figure 7, we can see a large number of strongly networked and collaborating authors who come from China in particular. In the periphery, however, some authors do not show any networking, or small groups are formed that are not further connected to other authors. Unused potential in networks could be the reason for these islands' formation. On the other hand, it also seems conceivable that the small, separate groups are pursuing very specific research areas and, therefore, there are no networking opportunities with researchers from other subject areas with methodological or content-related similarity.

Fig 7. Co-authorship networks using VOSviewer: The minimum number of documents per author was set to 3, resulting in 172 authors. Cluster memberships are distinguished by color. The node size (author) reflects the number of received citations; the shorter the distance between the nodes, the more closely connected they are (based on co-citations).

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Co-citation analysis and Bibliographic coupling

The co-citation analysis revealed the presence of six clusters (see Figure 8). Two clusters were not connected with other clusters (shown in green and blue). The overall most cited publications in the clusters were those of Breiman (2001) with the random forest algorithm [82] and Hochreiter and Schmidhuber (1997) with long short-term memory (LSTM) [83].

Focusing on the main topics of each cluster, it is hard to find significant differences. The first cluster, shown in red, could be grouped under primary algorithmic and machine learning literature—most notably the paper of Breiman (2001) [82].

Many papers in the purple cluster deal with deep convolutional neural networks as well as human action or image recognition. Furthermore, the works of Zhang et al. (2021) [84] and Cai et al. (2021) [85] (green cluster), which are not connected to other clusters, can be attributed to image segmentation. The publications in the orange cluster (e.g., LeCun et al., 2015) [86] mainly deal with the themes of wearables and deep learning.

Fig 8. Co-citation analysis of documents using the bibliometrix R package: Leading eigenvalues were used as a clustering algorithm. Isolated nodes were removed. Fifty nodes are displayed. Cluster memberships are distinguished by color. The node size (article) reflects the number of received citations; the shorter the distance between the nodes, the more closely connected they are (based on co-citations).

The results of the bibliographic coupling are presented in Figures 9 and 10. A look at the most frequently cited publications in the recent period (Figure 9) shows that only a few thematically distinguishable clusters can be formed for the years 2021 and 2022. With a minimum of three citations, the bibliometric analysis obtained 14 clusters. The first cluster, shown in red, contains 17 items, mainly concerning the themes of machine learning applications in the context of sports and health, as well

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

as decision-making; in this cluster, the most representative paper was published by Cao et al. (2021) [87]. The green cluster includes 16 items and could be summarized under the topic of motion recognition; Cui et al. (2021) [88] published the most cited paper in this cluster. The blue cluster consists of 14 items; most of these publications address the topic of AI in medicine; the study of Ramkumar et al. (2021) [89] was cited the most, with 20 citations.

Thirteen clusters were obtained via bibliometric analysis for 2019 and 2020 (see Figure 10). The red cluster contains 21 items, most of which can be assigned to the topic of team sports; the most cited paper in this cluster was published by Bunker et al. (2019) [68]. The cluster shown in green, with 16 items, can be summarized under the topic of sports recognition or recognition methods; the most representative paper in this cluster was published by Cust et al. (2019) [32]. The blue and purple clusters pertain to research with detections/tracking and big data, as well as machine learning; the most cited papers in the blue cluster were authored by Yoon et al. (2019) [90] and Minhas et al. (2019) [91]; in the purple cluster, the review by Claudino et al. (2019) [20] was cited the most.

Fig 9. Bibliographic coupling for the years 2021 and 2022 using VOSviewer: The minimum number of citations per document was set to 3 (180 out of 714 selected). Only the largest set of connected items is displayed (129 items). Cluster memberships are distinguished by color. The node size (article) reflects the number of citations received; the shorter the distance between the nodes, the more closely connected they are (based on coupling).

Fig 10. Bibliographic coupling for the years 2019 and 2020 using VOSviewer: The minimum number of citations per document was set to 5 (163 out of 257 selected). Only the largest set of connected items is displayed (135 items). Cluster memberships are distinguished by color. The node size (article) reflects the number of received citations; the shorter the distance between the nodes, the more closely connected they are (based on coupling).

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Keywords / Themes: general trends and longitudinal evolution

The overall occurrence of the 25 most frequently occurring author keywords is shown in Figure 11. The top four keywords map the search terms used for the Scopus search. It is notable that “machine learning” is the most frequently used term; the subarea “deep learning” is second, followed by “artificial intelligence”. A possible reason for this order is the more precise and specific scientific language [92] that specifically names the subfields instead of speaking of AI in a general way. The fact that the keyword “sports” only came in 4th place is a sign that the authors do not use sport as a keyword, in contrast to Scopus. This could be explained by two possible mechanisms: On the one hand, it is possible that authors do not use sport itself as a key term, referring instead to football, biomechanics, injury prevention, etc., as subsets of sport. Thus, sport(s) would be too general or too obvious as a keyword. On the other hand, it is also possible that authors do not see to the relevance of sport in their research compared to the AI topics.

The keywords can be manually grouped according to their meaning. The keywords “artificial neural networks”, “convolutional neural networks”, and “support vector machine” can be subsumed under the term *algorithms* (a). The keywords “inertial measuring units”, “wearable sensors”, “accelerometer”, “wearables”, and “sensors” can be assigned to *measuring methods or tools* (b). It is striking that all of these methods can also be used outside of laboratory settings in the field and make huge amounts of data relatively easily accessible, which favors the emergence of “big data” characteristics [93]. Specific *AI tasks* (c) are mapped with the terms “classification”, “activity recognition”, “action recognition”, and “prediction”. Different *research fields* (d) are mapped with the terms “sports analytics”, “computer vision”, “data mining”, "internet of things", and “biomechanics”. Of the 25 keywords shown,

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

only “soccer” and “football” name specific *sports categories* (e). A closer analysis shows that football is used synonymously with soccer in most of the publications included, in accordance with country-specific conventions. Other sports were also mentioned as specific keywords; however, their frequency was significantly lower (see Figures 12 and 13). Thus, it can be seen that the sport of soccer/football is very dominant, which can also be explained by the significant financial interests behind this sport [94].

Overall, according to the meaning of the dominant keywords, a high overlap with the contents of the most cited publications listed in Table 1 could be observed. It is notable that when the topics that are covered by the keywords or the most frequently cited publications are assigned to sub-disciplines of sports science, their primary focus is on sports informatics, training science, biomechanics, and sports medicine. Research activities in the fields of sports pedagogy, sports sociology, and sports economics were not evident from our analyses.

Fig 11. TreeMap of the 25 most frequently occurring author keywords

Figure 12 presents the corresponding thematic map of AI research in sports. The themes are divided into four quadrants by the horizontal and vertical dashed lines: motor/driving themes (upper right; Q2), basic/underlying themes (lower right; Q4), niche/very specialized themes (upper left; Q1), and disappearing or emerging themes (lower left; Q3). Themes in Q4—such as in general AI (inclusive of its subareas), different specific sports, and subfields of the AI pipeline—are basic, foundational themes that are very important for the field's general development, characterized by high centrality but low density. Various measuring technologies (e.g., accelerometers, inertial sensors, gyroscopes) can be subsumed under the driving themes of Q2; this makes sense because they often provide complex data,

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

making AI techniques more and more relevant due to static inference methods being limited in those contexts [95]. Looking at Q3, including the subsequent results of trending topics, sports medicine can be identified as an emerging theme. The themes of Q1 (e.g., “simulation”, “hidden Markov model”) are highly specialized and rare, creating internal bonds, as indicated by their low centrality but high density. The theme “performance analysis” is located between all quadrants, indicating that some aspects are basic and necessary for file development, some are emerging or declining, some are fundamental drivers, and some are highly specialized niche themes.

Overall, it is notable that the manual grouping of the most frequently occurring keywords fits with the data-driven clustering of the thematic map to a high degree; this also speaks to the coherence of the obtained solution.

Fig 12. Thematic map using Walktrap for clustering: The size of the spheres maps the total number of documents to each theme. The top three most frequently occurring labels are displayed per cluster. The more relations a node shows with others in the network, the greater its centrality and importance. The density representing the cohesiveness among the nodes gives clues as to the capability of a research field to develop and sustain itself

Figure 13 shows the evolution of author keywords, which has attracted particular attention in recent years. In addition to the neural networks that have been in use for a long time, as well as the support vector machine (see Figure 9), the random forest algorithm has recently been given special attention. This is consistent with the results of the co-citation analysis, where a large cluster was found around the work of Breiman (2001) [82]. According to methodological aspects in the context of the AI workflow, attention has also recently been paid to “feature extraction” and “transfer learning”.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The biomechanical aspects of “pose estimation” and “kinematics” have only recently attracted particular attention. Furthermore, the field of “sports medicine” and the matching topic of injury prevention are mentioned frequently—especially recently. The sport of running and e-sports have also received attention in connection with AI in recent years.

Unsurprisingly, “COVID-19” also attracted attention in 2021 and 2022 during the period when the world was still struggling deeply with the pandemic. The future infection situation will probably be decisive for the progression or decline of this topic. Another aspect that has emerged—especially in 2022—is the ethical discussion of AI in sports. Please refer to Figure 11 for information on further trend topics.

At this point, it should be noted that our analysis probably does not depict topics that have arisen very recently. For example, few recent publications have addressed the black-box character of AI models [96,97], which does not comply with the General Data Protection Regulation (GDPR) [98], using explainable AI. However, this theme does not appear as a trend topic using the author keywords at present, although it may be visible in a bibliometric analysis in the next few years.

Fig 13. Trend topics using author keywords: Keywords are displayed that obtained more than 80% of their total count in the years 2020–2022. For comparison, the keywords “machine learning” and “sports” are additionally displayed on top. Each y-axis maps the total number of occurrences.

Limitations and further research

In addition to the aforementioned advantages for our research aims, the bibliometric methodology is also associated with some inherent limitations. The analysis is based on technical decisions, e.g., the selection of English as the search

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

526 language or the focus on the highest-quality papers (i.e., articles and reviews), which
527 may have resulted in the exclusion of other research papers. Therefore, future
528 research should also consider the inclusion of other document types, e.g.,
529 conference papers or preprints, as these may provide better insights into the very
530 latest research themes because their publication times are often shorter. Articles in
531 anthologies or monographs could also be interesting, since social science sub-
532 disciplines of sport use them as publication opportunities. Furthermore, it has been
533 shown that different databases (e.g., Google Scholar, Scopus, Web of Science) have
534 certain differences in their search results [99]. Therefore, a multisource comparison
535 and mutual supplementation of the different databases could result in a broader
536 literature overview. Reported issues with the ambiguity of author names could be
537 addressed in the future by using ORCID IDs to prevent potential incorrect merging of
538 authors with the same first and last names [81].

539 In addition, the used search terms could be expanded to include further
540 aspects that are associated with the term “sports”, in order to identify publications
541 that deal with topics that can be assigned to sports without mentioning the term itself
542 (e.g., physical activity, exercise, training). At this point, the problem of defining the
543 term sports should not go unmentioned. One should be aware that, especially in
544 everyday life, there are a wide variety of (individual and cultural) ideas and
545 definitions of sports—and this applies not only to the term “sports”, but also to terms
546 that are often associated with sports, such as the term “training” [100,101]. For
547 example, a central point of discussion is whether e-sports constitute sports [102].
548 This problem was dealt with in this study in such a way that, in simplified terms, it
549 was assumed that publications containing the term “sport*” could be assigned to the
550 research interest of sports science. Therefore, publications from the field of e-sports

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

or—in some cases—from the field of board games (e.g., “Checkers is solved” [74]) were also included. It should be noted that, depending on the individual or context-specific understanding of the definition of sports, these publications might also be classified as unrelated to sports.

Conclusion

The present study provides a detailed analysis of the leading journals, authors, institutions, keywords, and research themes associated with interconnections between the fields of AI and sports. However, it seems that AI has not yet taken a leading role in sports science but has been gaining more and more importance recently, and sport is an area of application for AI. Our key findings are as follows: (a) The literature and research interest concerning AI and its subcategories is growing exponentially, with considerable potential for further growth in the coming years. (b) The top 20 most cited works account for 32.52% of the total citations, including the presence of “hot papers”, which are still relatively young and have already received considerable numbers of citations. (c) The top 10 journals are responsible for 28.64% of all published documents and indicate a trend of interdisciplinary research. (d) In addition to strong collaborative relationships, delimited small collaboration networks of individual institutions are identifiable. (e) The three most productive countries are China, the USA, and Germany; collaborations appear to be taking place more within the clusters of high- and middle-income countries. (f) Different research themes can be characterized using author keywords, with potential current trend topics including the fields of biomechanics, injury prevention or prediction, new algorithms, and learning approaches. AI

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

research activities in the fields of sports pedagogy, sports sociology, and sports economics seem to have played a subordinate role so far.

These findings provide a better understanding of the research situation regarding the use of AI in sports and may aid researchers in identifying currently relevant research topics, as well as research gaps, through revealing fields of research with much evidence and those that have only been considered marginally thus far. Furthermore, they may be helpful in improving the obviously necessary networking in the context of interdisciplinary and transdisciplinary research, in order to combine competencies from different disciplinary approaches. Hence, this work could be a starting point for systematically initiating stronger networking between authors and their countries/institutions, which might offer yet-untapped potential in their interconnection according to the present study's results.

Today, AI seems to play an important role not only in sports research, but also in sports practice. Numerous companies (e.g., Apple, Microsoft, Enduco, iDIERS, Aaptiv) are advertising commercial offers that use AI technologies and, thus, have partially opened up areas that research has not yet sufficiently investigated. Therefore, it seems important that research does not lag behind in questioning these economic interests with scientific evidence.

Funding: This work was supported by the BMBF under grant numbers 16SV7115 and 03IHS075B.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

1. Nguyen G, Dlugolinsky S, Bobák M, Tran V, López García Á, Heredia I, et al. Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey. Artif Intell Rev. 2019; 52:77–124. doi: 10.1007/s10462-018-09679-z.
2. Wasnik P, Raja KB, Ramachandra R, Busch C. Assessing face image quality for smartphone based face recognition system. 5th International Workshop on Biometrics and Forensics (IWBF). Coventry, UK: IEEE; 2017. pp. 1–6.
3. Hoy MB. Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. Med Ref Serv Q. 2018; 37:81–8. doi: 10.1080/02763869.2018.1404391 PMID: 29327988.
4. Hu G, Yang Y, Yi D, Kittler J, Christmas W, Li SZ, et al. When Face Recognition Meets with Deep Learning: an Evaluation of Convolutional Neural Networks for Face Recognition [updated 9 Apr 2015]. Available from: https://arxiv.org/pdf/1504.02351.
5. Hinton G, Deng L, Yu D, Dahl G, Mohamed A, Jaitly N, et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. IEEE Signal Process Mag. 2012; 29:82–97. doi: 10.1109/MSP.2012.2205597.
6. Fink M, Liu Y, Engstle A, Schneider S-A. Deep Learning-Based Multi-scale Multi-object Detection and Classification for Autonomous Driving. In: Bertram T, editor. Fahrerassistenzsysteme 2018. Von der Assistenz zum automatisierten Fahren. 4. Internationale ATZ-Fachtagung Automatisiertes Fahren. Wiesbaden, Germany: Springer; 2019. pp. 233–42.
7. Naway A, LI Y. A Review on The Use of Deep Learning in Android Malware Detection. International Journal of Computer Science and Mobile Computing. 2018; 7. doi: 10.48550/arXiv.1812.10360.
8. Guzella TS, Caminhas WM. A review of machine learning approaches to Spam filtering. Expert Systems with Applications. 2009; 36:10206–22. doi: 10.1016/j.eswa.2009.02.037.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

624 **9**. Bojarski M, Del Testa D, Dworakowski D, Firner B, Flepp B, Goyal P, et al. End to End
625 Learning for Self-Driving Cars [updated 25 Apr 2016; cited 14 Sep 2021]. Available
626 from: https://arxiv.org/pdf/1604.07316.
627 **10**. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide
628 to deep learning in healthcare. Nat Med. 2019; 25:24–9. Epub 2019/01/07.
629 doi: 10.1038/s41591-018-0316-z PMID: 30617335.
630 **11**. Amato F, López A, Peña-Méndez EM, Vaňhara P, Hampl A, Havel J. Artificial neural
631 networks in medical diagnosis. J Appl Biomed. 2013; 11:47–58. doi: 10.2478/v10136-
632 012-0031-x.
633 **12**. Buch VH, Ahmed I, Maruthappu M. Artificial intelligence in medicine: current trends and
634 future possibilities. Br J Gen Pract. 2018; 68:143–4. doi: 10.3399/bjgp18X695213
635 PMID: 29472224.
636 **13**. Topol EJ. High-performance medicine: the convergence of human and artificial
637 intelligence. Nat Med. 2019; 25:44–56. Epub 2019/01/07. doi: 10.1038/s41591-018-
638 0300-7 PMID: 30617339.
639 **14**. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation
640 with a limited sample size. PLoS ONE. 2019; 14:1-20. Epub 2019/11/07.
641 doi: 10.1371/journal.pone.0224365 PMID: 31697686.
642 **15**. Helbing D, Frey BS, Gigerenzer G, Hafen E, Hagner M, Hofstetter Y, et al. Will
643 Democracy Survive Big Data and Artificial Intelligence. In: Helbing D, editor. Towards
644 digital enlightenment. Essays on the dark and light sides of the digital revolution. Cham:
645 Springer; 2018. pp. 73–98.
646 **16**. Jun Lee S, Siau K. A review of data mining techniques. Industr Mngmnt & Data
647 Systems. 2001; 101:41–6. doi: 10.1108/02635570110365989.
648 **17**. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-
649 learn: Machine Learning in Python. J Mach Learn Res. 2011; 12:2825–30.
650 **18**. Paszke A, Gross S, Chintala S, Chanan G, Yang E, DeVito Z, et al. Automatic
651 differentiation in PyTorch. In: Luxburg U von, Guyon I, Bengio S, Wallach H, Fergus R,

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

editors. 31st Conference on Neural Information Processing Systems (NIPS 2017). Red Hook, NY, United States: Curran Associates Inc.; 2017. pp. 1–4.

19. Al-Jarrah OY, Yoo PD, Muhaidat S, Karagiannidis GK, Taha K. Efficient Machine Learning for Big Data: A Review. Big Data Research. 2015; 2:87–93. doi: 10.1016/j.bdr.2015.04.001.

20. Claudino JG, Capanema DdO, Souza TV de, Serrão JC, Machado Pereira AC, Nassis GP. Current Approaches to the Use of Artificial Intelligence for Injury Risk Assessment and Performance Prediction in Team Sports: a Systematic Review. Sports Med - Open. 2019; 5:1–12. Epub 2019/07/03. doi: 10.1186/s40798-019-0202-3 PMID: 31270636.

21. Horvat T, Job J. The use of machine learning in sport outcome prediction: A review. WIREs Data Mining Knowl Discov. 2020; 10:1–28. doi: 10.1002/widm.1380.

22. Li B, Xu X. Application of Artificial Intelligence in Basketball Sport. J Educ Health Sport. 2021; 11:54–67. doi: 10.12775/JEHS.2021.11.07.005.

23. Herold M, Goes F, Nopp S, Bauer P, Thompson C, Meyer T. Machine learning in men's professional football: Current applications and future directions for improving attacking play. International Journal of Sports Science & Coaching. 2019; 14:798–817. doi: 10.1177/1747954119879350.

24. Me¸żyk E, Unold O. Machine learning approach to model sport training. Computers in Human Behavior. 2011; 27:1499–506. doi: 10.1016/j.chb.2010.10.014.

25. Lichtenthaler U. Mixing data analytics with intuition: Liverpool Football Club scores with integrated intelligence. JBS. 2022; 43:10–6. doi: 10.1108/JBS-06-2020-0144.

26. Hartig-Franc P. Big-Data, Surveillance, Sports and the Law. UNSWLJ Student Series. 2016; 16.

27. Dindorf C, Konradi J, Wolf C, Taetz B, Bleser G, Huthwelker J, et al. Machine learning techniques demonstrating individual movement patterns of the vertebral column: The fingerprint of spinal motion. Comput Methods Biomech Biomed Engin. 2021; 25:1–11. doi: 10.1080/10255842.2021.1981884.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

679 **28**. Rommers N, Rössler R, Verhagen E, Vandecasteele F, Verstockt S, Vaeyens R, et al.
680 A Machine Learning Approach to Assess Injury Risk in Elite Youth Football Players.
681 Med Sci Sports Exerc. 2020; 52:1745–51. doi: 10.1249/MSS.0000000000002305
682 PMID: 32079917.
683 **29**. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric
684 analysis: An overview and guidelines. Journal of Business Research. 2021; 133:285–
685 96. doi: 10.1016/j.jbusres.2021.04.070.
686 **30**. Verma S, Gustafsson A. Investigating the emerging COVID-19 research trends in the
687 field of business and management: A bibliometric analysis approach. Journal of
688 Business Research. 2020; 118:253–61. Epub 2020/07/02.
689 doi: 10.1016/j.jbusres.2020.06.057 PMID: 32834211.
690 **31**. Ramos-Rodríguez A-R, Ruíz-Navarro J. Changes in the intellectual structure of
691 strategic management research: a bibliometric study of theStrategic Management
692 Journal, 1980–2000. Strat Mgmt J. 2004; 25:981–1004. doi: 10.1002/smj.397.
693 **32**. Cust EE, Sweeting AJ, Ball K, Robertson S. Machine and deep learning for sport-
694 specific movement recognition: a systematic review of model development and
695 performance. J Sports Sci. 2019; 37:568–600. doi: 10.1080/02640414.2018.1521769
696 PMID: 30307362.
697 **33**. Farrahi V, Niemelä M, Kangas M, Korpelainen R, Jämsä T. Calibration and validation of
698 accelerometer-based activity monitors: A systematic review of machine-learning
699 approaches. Gait Posture. 2019; 68:285–99. Epub 2018/12/05.
700 doi: 10.1016/j.gaitpost.2018.12.003 PMID: 30579037.
701 **34**. Rahimian P, Toka L. Optical tracking in team sports. Journal of Quantitative Analysis in
702 Sports. 2022; 18:35–57. doi: 10.1515/jqas-2020-0088.
703 **35**. Ashfaq A, Cronin N, Müller P. Recent advances in machine learning for maximal
704 oxygen uptake (VO2 max) prediction: A review. Informatics in Medicine Unlocked.
705 2022; 28:100863. doi: 10.1016/j.imu.2022.100863.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

706 **36**. Cefis M. Football analytics: a bibliometric study about the last decade contributions.
707 University of Salento; 2022.
708 **37**. Escamilla-Fajardo P, Núñez-Pomar JM, Ratten V, Crespo J. Entrepreneurship and
709 Innovation in Soccer: Web of Science Bibliometric Analysis. Sustainability. 2020;
710 12:4499. doi: 10.3390/su12114499.
711 **38**. Pereira J, Rodrigues RG, Duarte P, Gouveia A. Social Media Marketing of Football
712 Clubs: A Study with Portuguese Football Clubs—SL Benfica, FC Porto, and Sporting
713 CP. In: Reis JL, Peter MK, Cayolla R, Bogdanović Z, editors. Marketing and Smart
714 Technologies. Singapore: Springer Singapore; 2022. pp. 555–68.
715 **39**. Oliveira JR de, Tobar FB, Capraro AM. Football tourism: a bibliometric analysis of
716 published works in the tourism-based journals (2003–2019). Journal of Sport &
717 Tourism. 2021; 25:317–35. doi: 10.1080/14775085.2021.1965010.
718 **40**. Jiménez-García M, Ruiz-Chico J, Peña-Sánchez AR, López-Sánchez JA. A
719 Bibliometric Analysis of Sports Tourism and Sustainability (2002–2019). Sustainability.
720 2020; 12:2840. doi: 10.3390/su12072840.
721 **41**. González-Serrano MH, Añó Sanz V, González-García RJ. Sustainable Sport
722 Entrepreneurship and Innovation: A Bibliometric Analysis of This Emerging Field of
723 Research. Sustainability. 2020; 12:5209. doi: 10.3390/su12125209.
724 **42**. Baier-Fuentes H, González-Serrano MH, Alonso-Dos Santos M, Inzunza-Mendoza W,
725 Pozo-Estrada V. Emotions and Sport Management: A Bibliometric Overview. Front
726 Psychol. 2020; 11:1512. Epub 2020/07/10. doi: 10.3389/fpsyg.2020.01512 PMID:
727 32754088.
728 **43**. Kiss A, Temesi Á, Tompa O, Lakner Z, Soós S. Structure and trends of international
729 sport nutrition research between 2000 and 2018: bibliometric mapping of sport nutrition
730 science. J Int Soc Sports Nutr. 2021; 18:12. Epub 2021/02/05. doi: 10.1186/s12970-
731 021-00409-5 PMID: 33546728.
732 **44**. Ma L, Wang Y, Wang Y, Li N, Fung S-F, Zhang L, et al. The Hotspots of Sports
733 Science and the Effects of Knowledge Network on Scientific Performance Based on

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

734 Bibliometrics and Social Network Analysis. Complexity. 2021; 2021:1–12.
735 doi: 10.1155/2021/9981202.
736 **45**. Xiang X, Wang Q. Metrology Atlas Research on the CNKI Literature Big Data of Sports
737 Artificial Intelligence Based on COOC/VOSviewer. 2021 2nd International Conference
738 on Information Science and Education. ICISE-IE 2021 : proceedings : 26-28 November
739 2021, Nanchang, China. Los Alamitos, CA: IEEE Computer Society, Conference
740 Publishing Services; 2021. pp. 720–4.
741 **46**. Hu T, Guo J, Zhang T, Liu J, Sun X, Chang Z. Analysis of the Development Trend of
742 Artificial Intelligence Technology Application in the Field of Sports-Based on Patent
743 Measurement. In: Soares MM, Rosenzweig E, Marcus A, editors. Design, user
744 experience, and usability. 11th International Conference, DUXU 2022, held as part of
745 the 24th HCI International Conference, HCII 2022, virtual event, June 26-July 1, 2022,
746 proceedings, pts. I-III / edited by Marcelo M. Soares, Elizabeth Rosenzweig, Aaron
747 Marcus. Cham, Switzerland: Springer; 2022. pp. 337–54.
748 **47**. Huang J-H, Hsu Y-C. A Multidisciplinary Perspective on Publicly Available Sports Data
749 in the Era of Big Data: A Scoping Review of the Literature on Major League Baseball.
750 SAGE Open. 2021; 11:215824402110615. doi: 10.1177/21582440211061566.
751 **48**. Yang Q, Sun J, Zhang ZY, Ding SC, Chen Z, Cui XY, et al. Application of wearable
752 devices in sports: behavior change and result effect. 2021 International Conference on
753 Health Big Data and Smart Sports (HBDSS). IEEE; 2021. pp. 137–47.
754 **49**. Wu H, Cheng K, Guo Q, Yang W, Tong L, Wang Y, et al. Mapping Knowledge
755 Structure and Themes Trends of Osteoporosis in Rheumatoid Arthritis: A Bibliometric
756 Analysis. Front Med (Lausanne). 2021; 8:787228. Epub 2021/11/23.
757 doi: 10.3389/fmed.2021.787228 PMID: 34888333.
758 **50**. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for
759 bibliometric mapping. Scientometrics. 2010; 84:523–38. Epub 2009/12/31.
760 doi: 10.1007/s11192-009-0146-3 PMID: 20585380.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

761 **51**. Aria M, Cuccurullo C. bibliometrix : An R-tool for comprehensive science mapping
762 analysis. Journal of Informetrics. 2017; 11:959–75. doi: 10.1016/j.joi.2017.08.007.
763 **52**. Taqi M, Rusydiana AS, Kustiningsih N, Firmansyah I. ENVIRONMENTAL
764 ACCOUNTING: A SCIENTOMETRIC USING BIBLIOSHINY. IJEEP. 2021; 11:369–80.
765 doi: 10.32479/ijeep.10986.
766 **53**. Rusydiana AS. Bibliometric analysis of journals, authors, and topics related to COVID-
767 19 and Islamic finance listed in the Dimensions database by Biblioshiny. Sci Ed. 2021;
768 8:72–8. doi: 10.6087/kcse.232.
769 **54**. Moral-Muñoz JA, Herrera-Viedma E, Santisteban-Espejo A, Cobo MJ. Software tools
770 for conducting bibliometric analysis in science: An up-to-date review. EPI. 2020; 29.
771 doi: 10.3145/epi.2020.ene.03.
772 **55**. Ma T-J, Lee G-G, Liu J, Lae R. Bibliographic coupling: a main path analysis from 1963
773 to 2020. IR. 2022; 27. doi: 10.47989/irpaper918.
774 **56**. Zupic I, Čater T. Bibliometric Methods in Management and Organization.
775 Organizational Research Methods. 2015; 18:429–72. doi: 10.1177/1094428114562629.
776 **57**. Song Y, Chen X, Hao T, Liu Z, Lan Z. Exploring two decades of research on classroom
777 dialogue by using bibliometric analysis. Computers & Education. 2019; 137:12–31.
778 doi: 10.1016/j.compedu.2019.04.002.
779 **58**. Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F. SciMAT: A new science
780 mapping analysis software tool. J Am Soc Inf Sci Tec. 2012; 63:1609–30.
781 doi: 10.1002/asi.22688.
782 **59**. Jelvehgaran Esfahani H, Tavasoli K, Jabbarzadeh A. Big data and social media: A
783 scientometrics analysis. 10.5267/j.ijdns. 2019:145–64. doi: 10.5267/j.ijdns.2019.2.007.
784 **60**. Chang C, Lin M-C, Leonard MS, Occena LG. Building an expert system for the design
785 of sport shoes. Computers & Industrial Engineering. 1988; 15:72–7. doi: 10.1016/0360-
786 8352(88)90066-6.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

787 **61**. Riccaboni M, Verginer L. The impact of the COVID-19 pandemic on scientific research
788 in the life sciences. PLoS ONE. 2022; 17:e0263001. Epub 2022/02/09.
789 doi: 10.1371/journal.pone.0263001 PMID: 35139089.
790 **62**. Mulet-Forteza C, Martorell-Cunill O, Merigó JM, Genovart-Balaguer J, Mauleon-
791 Mendez E. Twenty five years of the Journal of Travel & Tourism Marketing : a
792 bibliometric ranking. Journal of Travel & Tourism Marketing. 2018; 35:1201–21.
793 doi: 10.1080/10548408.2018.1487368.
794 **63**. Nweke HF, Teh YW, Al-garadi MA, Alo UR. Deep learning algorithms for human activity
795 recognition using mobile and wearable sensor networks: State of the art and research
796 challenges. Expert Systems with Applications. 2018; 105:233–61.
797 doi: 10.1016/j.eswa.2018.03.056.
798 **64**. Guha T, Ward RK. Learning sparse representations for human action recognition. IEEE
799 Transactions on Pattern Analysis and Machine Intelligence. 2012; 34:1576–88.
800 doi: 10.1109/TPAMI.2011.253 PMID: 22745001.
801 **65**. Barshan B, Yuksek MC. Recognizing Daily and Sports Activities in Two Open Source
802 Machine Learning Environments Using Body-Worn Sensor Units. The Computer
803 Journal. 2014; 57:1649–67. doi: 10.1093/comjnl/bxt075.
804 **66**. Yao B, Fei-Fei L. Recognizing human-object interactions in still images by modeling the
805 mutual context of objects and human poses. IEEE Transactions on Pattern Analysis
806 and Machine Intelligence. 2012; 34:1691–703. doi: 10.1109/TPAMI.2012.67 PMID:
807 22392710.
808 **67**. Trost SG, Zheng Y, Wong W-K. Machine learning for activity recognition: hip versus
809 wrist data. Physiol Meas. 2014; 35:2183–9. Epub 2014/10/23. doi: 10.1088/0967-
810 3334/35/11/2183 PMID: 25340887.
811 **68**. Bunker RP, Thabtah F. A machine learning framework for sport result prediction.
812 Applied Computing and Informatics. 2019; 15:27–33. doi: 10.1016/j.aci.2017.09.005.
813 **69**. Bartlett JD, O'Connor F, Pitchford N, Torres-Ronda L, Robertson SJ. Relationships
814 Between Internal and External Training Load in Team-Sport Athletes: Evidence for an

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Individualized Approach. International Journal of Sports Physiology and Performance. 2017; 12:230–4. Epub 2016/08/24. doi: 10.1123/ijspp.2015-0791 PMID: 27194668.

70. Gudmundsson J, Horton M. Spatio-Temporal Analysis of Team Sports. ACM Comput Surv. 2018; 50:1–34. doi: 10.1145/3054132.

71. Lu W-L, Ting J-A, Little JJ, Murphy KP. Learning to track and identify players from broadcast sports videos. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2013; 35:1704–16. doi: 10.1109/TPAMI.2012.242 PMID: 23681997.

72. Xing J, Ai H, Liu L, Lao S. Multiple player tracking in sports video: a dual-mode two-way bayesian inference approach with progressive observation modeling. IEEE Trans Image Process. 2011; 20:1652–67. Epub 2010/12/23. doi: 10.1109/TIP.2010.2102045 PMID: 21189238.

73. Colyer SL, Evans M, Cosker DP, Salo AIT. A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System. Sports Med - Open. 2018; 4:24. Epub 2018/06/05. doi: 10.1186/s40798-018-0139-y PMID: 29869300.

74. Schaeffer J, Burch N, Björnsson Y, Kishimoto A, Müller M, Lake R, et al. Checkers is solved. Science. 2007; 317:1518–22. Epub 2007/07/19. doi: 10.1126/science.1144079 PMID: 17641166.

75. Li N, Wu DD. Using text mining and sentiment analysis for online forums hotspot detection and forecast. Decision Support Systems. 2010; 48:354–68. doi: 10.1016/j.dss.2009.09.003.

76. Ravi D, Wong C, Lo B, Yang G-Z. A Deep Learning Approach to on-Node Sensor Data Analytics for Mobile or Wearable Devices. IEEE J Biomed Health Inform. 2017; 21:56–64. Epub 2016/12/23. doi: 10.1109/JBHI.2016.2633287 PMID: 28026792.

77. Rein R, Memmert D. Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. Springerplus. 2016; 5:1410. Epub 2016/08/24. doi: 10.1186/s40064-016-3108-2 PMID: 27610328.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

842 **78**. Lu N, Wu Y, Feng L, Song J. Deep Learning for Fall Detection: Three-Dimensional
843 CNN Combined With LSTM on Video Kinematic Data. IEEE J Biomed Health Inform.
844 2019; 23:314–23. Epub 2018/02/20. doi: 10.1109/JBHI.2018.2808281 PMID:
845 29994460.
846 **79**. The World Bank. How does the World Bank classify countries. n.y. Available from:
847 https://datahelpdesk.worldbank.org/knowledgebase/articles/378834-how-does-the-
848 world-bank-classify-countries.
849 **80**. Asia on the rise. Nature. 2007; 447:885. doi: 10.1038/447885a PMID: 17581541.
850 **81**. Strotmann A, Zhao D. Author name disambiguation: What difference does it make in
851 author-based citation analysis. J Am Soc Inf Sci Tec. 2012; 63:1820–33.
852 doi: 10.1002/asi.22695.
853 **82**. Breiman L. Random Forests. Machine Learning. 2001; 45:5–32.
854 **83**. Hochreiter S, Schmidhuber J. Long short-term memroy. Neural Computation. 1997;
855 9:1735–80.
856 **84**. Zhang X, Yang Y, Li Z, Ning X, Qin Y, Cai W. An Improved Encoder-Decoder Network
857 Based on Strip Pool Method Applied to Segmentation of Farmland Vacancy Field.
858 Entropy (Basel). 2021; 23. Epub 2021/04/08. doi: 10.3390/e23040435 PMID:
859 33917753.
860 **85**. Cai W, Zhai B, Liu Y, Liu R, Ning X. Quadratic polynomial guided fuzzy C-means and
861 dual attention mechanism for medical image segmentation. Displays. 2021; 70:102106.
862 doi: 10.1016/j.displa.2021.102106.
863 **86**. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521:436–44.
864 doi: 10.1038/nature14539 PMID: 26017442.
865 **87**. Cao Z, Liao T, Song W, Chen Z, Li C. Detecting the shuttlecock for a badminton robot:
866 A YOLO based approach. Expert Systems with Applications. 2021; 164:113833.
867 doi: 10.1016/j.eswa.2020.113833.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

868 **88**. Cui L, Su D, Yang S, Wang Z, Ming Z. TCLiVi: Transmission Control in Live Video
869 Streaming Based on Deep Reinforcement Learning. IEEE Trans Multimedia. 2021;
870 23:651–63. doi: 10.1109/TMM.2020.2985631.
871 **89**. Ramkumar PN, Kunze KN, Haeberle HS, Karnuta JM, Luu BC, Nwachukwu BU, et al.
872 Clinical and Research Medical Applications of Artificial Intelligence. Arthroscopy. 2021;
873 37:1694–7. Epub 2020/08/21. doi: 10.1016/j.arthro.2020.08.009 PMID: 32828936.
874 **90**. Yoon Y, Hwang H, Choi Y, Joo M, Oh H, Park I, et al. Analyzing Basketball Movements
875 and Pass Relationships Using Realtime Object Tracking Techniques Based on Deep
876 Learning. IEEE Access. 2019; 7:56564–76. doi: 10.1109/ACCESS.2019.2913953.
877 **91**. Minhas RA, Javed A, Irtaza A, Mahmood MT, Joo YB. Shot Classification of Field
878 Sports Videos Using AlexNet Convolutional Neural Network. Applied Sciences. 2019;
879 9:483. doi: 10.3390/app9030483.
880 **92**. Esser H, Klenovits K, Zehnpfennig H. Wissenschaftstheorie. Grundlagen und
881 Analytische Wissenschaftstheorie 1. Wiesbaden: Vieweg+Teubner Verlag; Imprint;
882 1977.
883 **93**. Phinyomark A, Petri G, Ibáñez-Marcelo E, Osis ST, Ferber R. Analysis of Big Data in
884 Gait Biomechanics: Current Trends and Future Directions. J Med Biol Eng. 2018;
885 38:244–60. doi: 10.1007/s40846-017-0297-2.
886 **94**. Dressler S, Dressler S, Rachfall T, Foerster-Trallo D. The Financial Impact. In: Faganel
887 A, Riznar I, Baruca A, editors. Impacts and implications for the sports industry in the
888 post-Covid-19 era. Hershey PA: Business Science Reference; 2021. pp. 40–59.
889 **95**. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nature Methods.
890 2018; 15:233–4. doi: 10.1038/nmeth.4642.
891 **96**. Teufl W, Taetz B, Miezal M, Dindorf C, Fröhlich M, Trinler U, et al. Automated detection
892 of pathological gait patterns using a one-class support vector machine trained on
893 discrete parameters of IMU based gait data. Clinical Biomechanics. 2021; 89:1–7.
894 doi: 10.1016/j.clinbiomech.2021.105452.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

895 **97**. Horst F, Lapuschkin S, Samek W, Müller K-R, Schöllhorn WI. Explaining the unique
896 nature of individual gait patterns with deep learning. Sci Rep. 2019; 9:1–13.
897 doi: 10.1038/s41598-019-38748-8 PMID: 30787319.
898 **98**. European Union. Regulation (EU) 2016/679 of the european parliament and of the
899 council of 27 april 2016 on the protection of natural persons with regard to the
900 processing of personal data and on the free movement of such data, and repealing
901 directive 95/46/ec (General Data Protection Regulation). Official Journal of the
902 European Union. 2016; L 119:1–88.
903 **99**. Harzing A-W, Alakangas S. Google Scholar, Scopus and the Web of Science: a
904 longitudinal and cross-disciplinary comparison. Scientometrics. 2016; 106:787–804.
905 doi: 10.1007/s11192-015-1798-9.
906 **100**. Dindorf C, Fröhlich M. Zur Konnotation und Denotation des Trainingsbegriffs in Theorie
907 und Praxis des Sports. Ger J Exerc Sport Res. 2020; 50:297–307.
908 doi: 10.1007/s12662-019-00641-x.
909 **101**. Willimczik K. Die Vielfalt des Sports : kognitive Konzepte der Gegenwart zur
910 Binnendifferenzierung des Sports. Sportwissenschaft. 2007; 37:19–37.
911 **102**. Willimczik K. eSport „ist“ nicht Sport – eSport und Sport haben Bedeutungen. Ger J
912 Exerc Sport Res. 2019; 49:78–90. doi: 10.1007/s12662-019-00569-2.

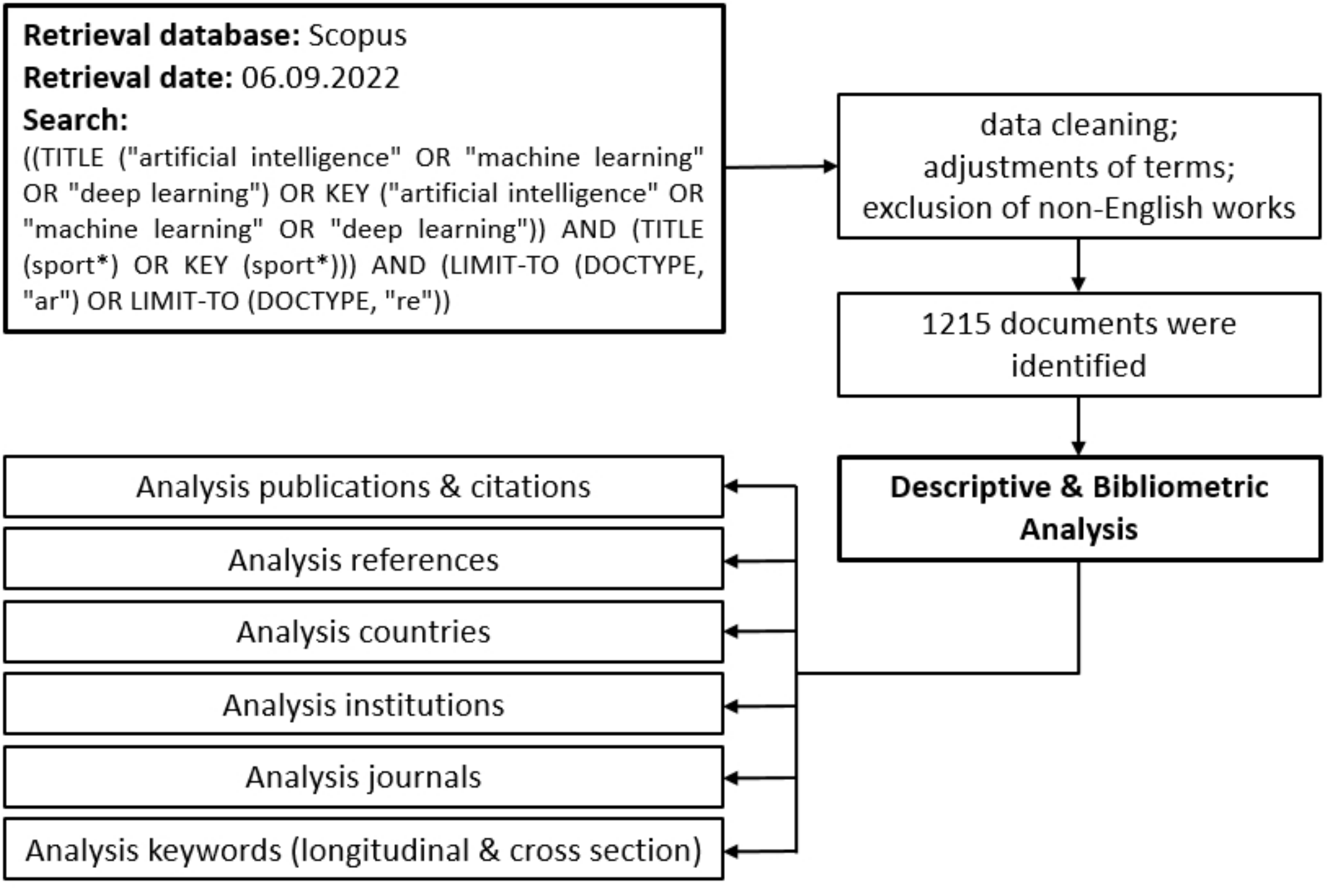
Retrieval database: Scopus
Retrieval date: 06.09.2022
Search:
((TITLE ("artificial intelligence" OR "machine learning" OR "deep learning") OR KEY ("artificial intelligence" OR "machine learning" OR "deep learning")) AND (TITLE (sport*) OR KEY (sport*))) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "re"))
data cleaning;
adjustments of terms;
exclusion of non-English works
1215 documents were identified
Descriptive & Bibliometric Analysis
Analysis publications & citations
Analysis references
Analysis countries
Analysis institutions
Analysis journals
Analysis keywords (longitudinal & cross section)

Figure 1

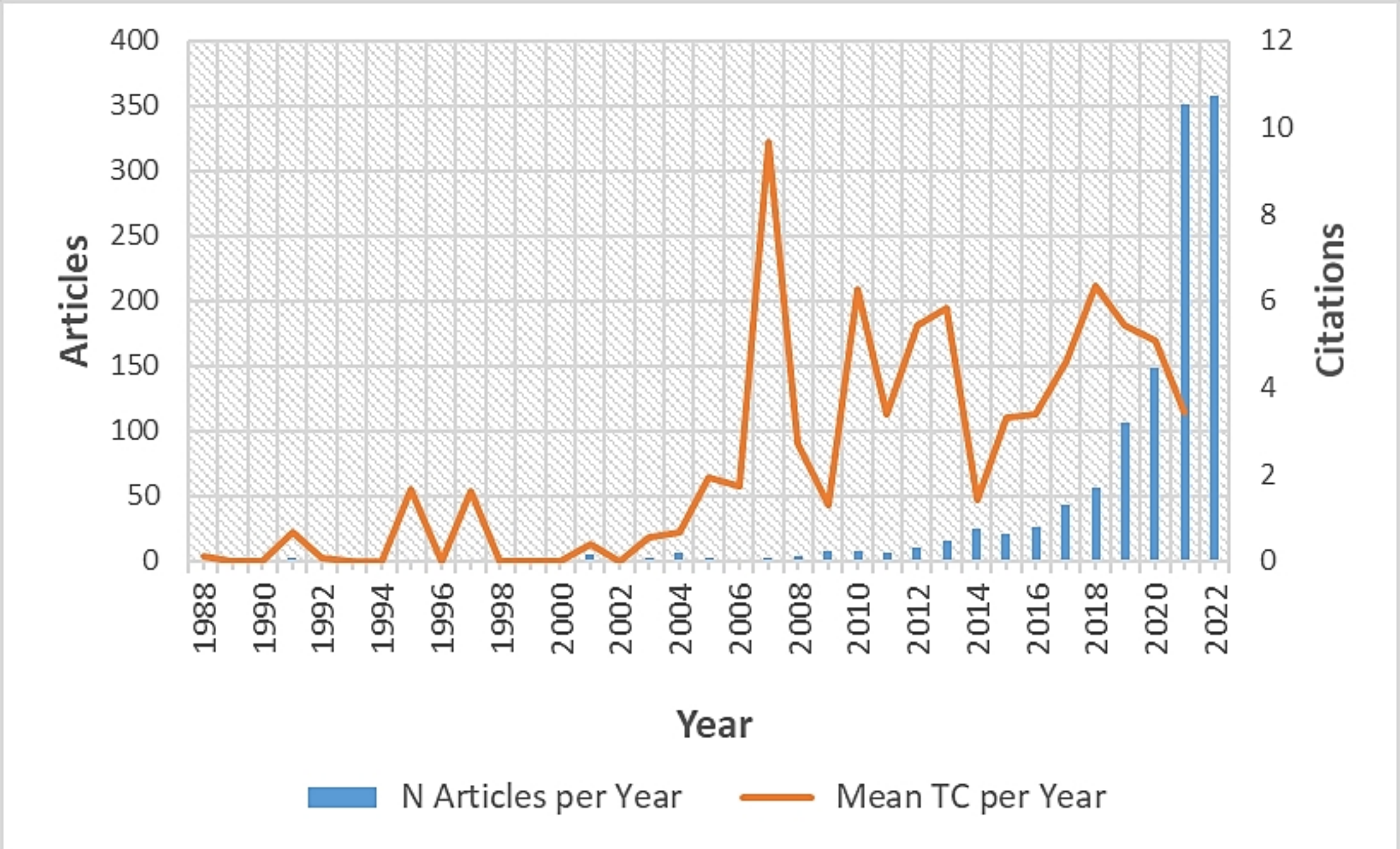
Articles
400
350
300
250
200
150
100
50
0
Citations
12
10
8
6
4
2
0
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
Year
N Articles per Year
Mean TC per Year

Figure 2

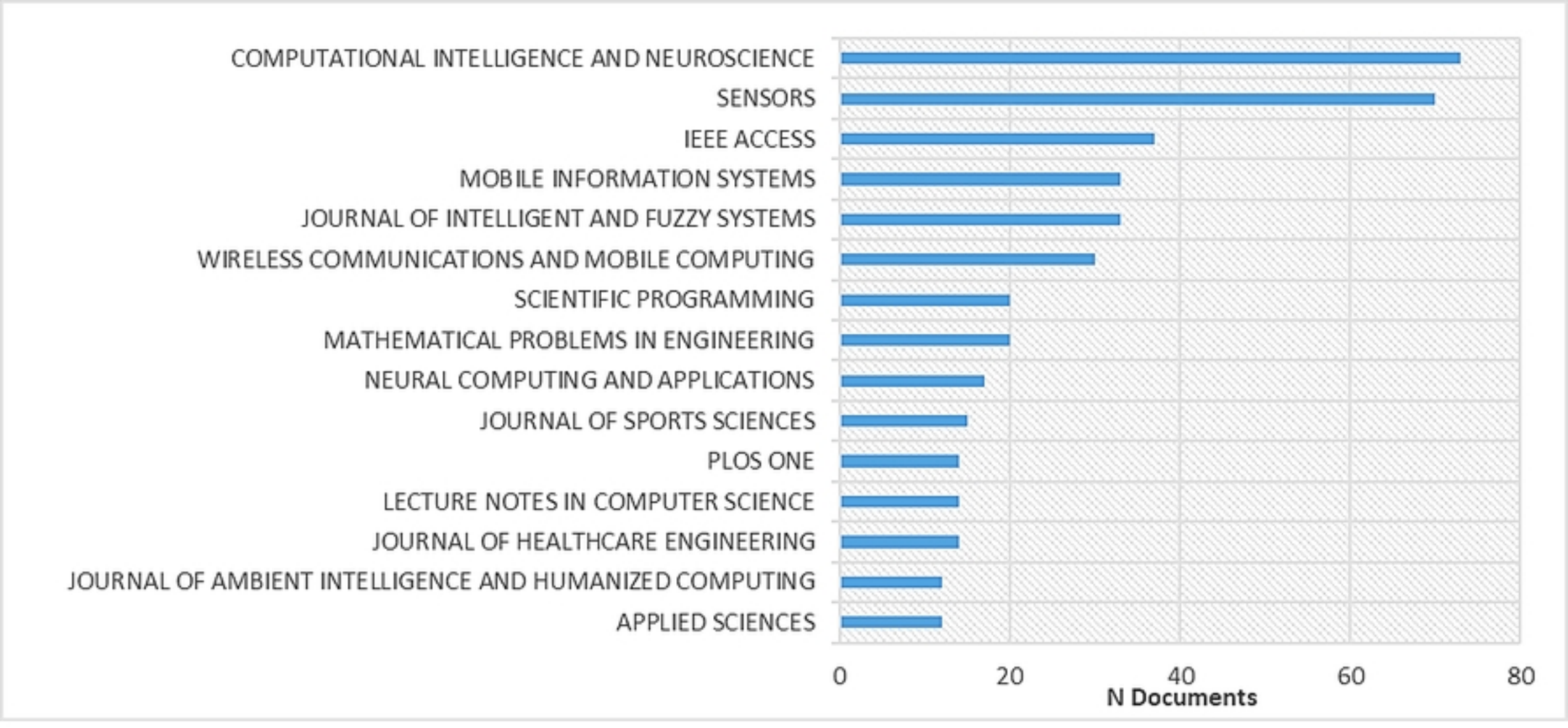
COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE
SENSORS
IEEE ACCESS
MOBILE INFORMATION SYSTEMS
JOURNAL OF INTELLIGENT AND FUZZY SYSTEMS
WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
SCIENTIFIC PROGRAMMING
MATHEMATICAL PROBLEMS IN ENGINEERING
NEURAL COMPUTING AND APPLICATIONS
JOURNAL OF SPORTS SCIENCES
PLOS ONE
LECTURE NOTES IN COMPUTER SCIENCE
JOURNAL OF HEALTHCARE ENGINEERING
JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING
APPLIED SCIENCES
0
20
40
60
80
N Documents

Figure 3

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

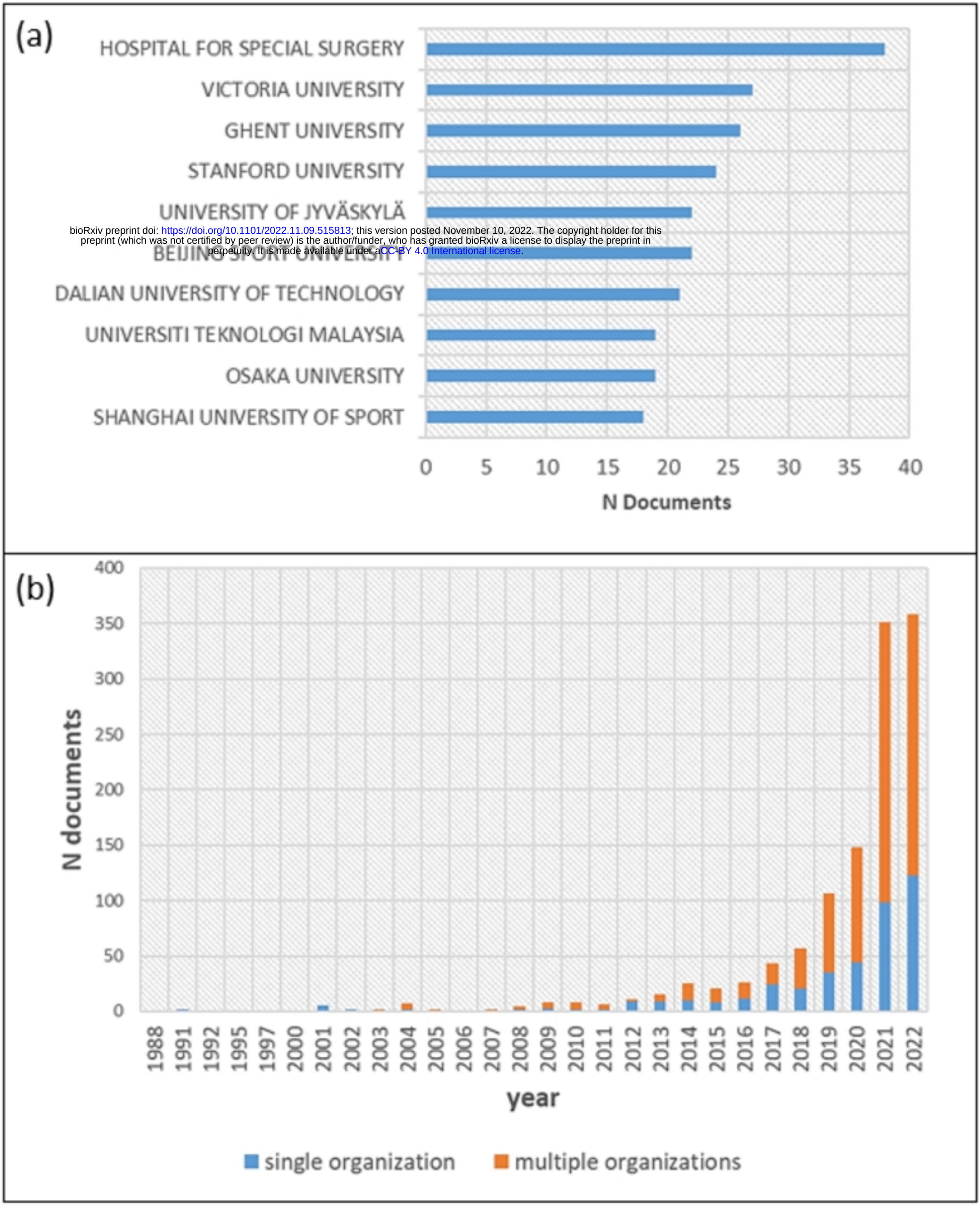

Figure 4

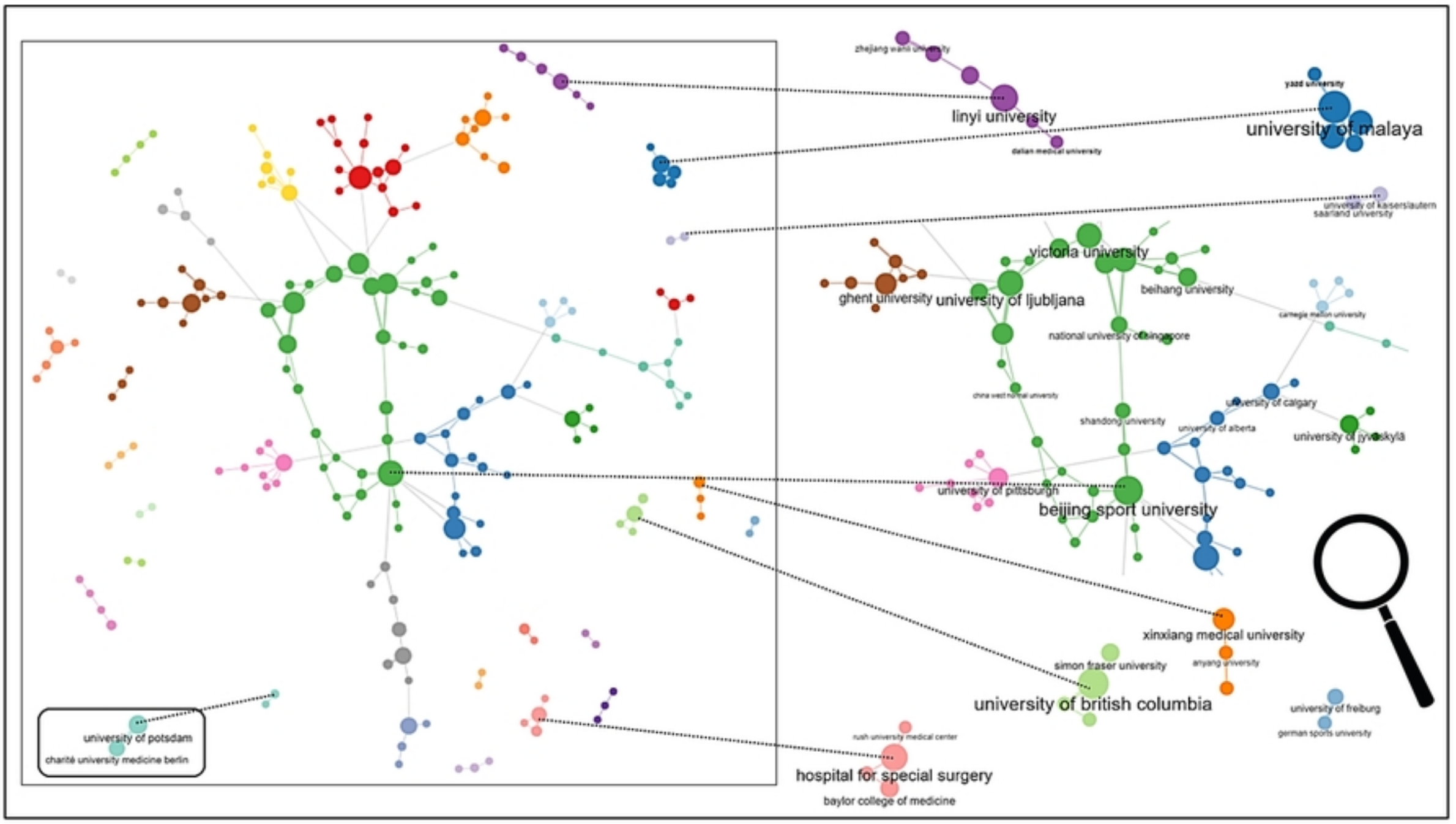
zhejiang wanli university
linyi university
dalian medical university
yazd university
university of malaya
university of kaiserslautern
saarland university
victoria university
ghent university
university of ljubljana
beihang university
carnegie mellon university
national university of singapore
china west normal university
university of calgary
shandong university
university of alberta
university of jyväskylä
university of pittsburgh
beijing sport university
xinxiang medical university
anyang university
simon fraser university
university of british columbia
university of freiburg
german sports university
rush university medical center
hospital for special surgery
baylor college of medicine
university of potsdam
charité university medicine berlin

Figure 5

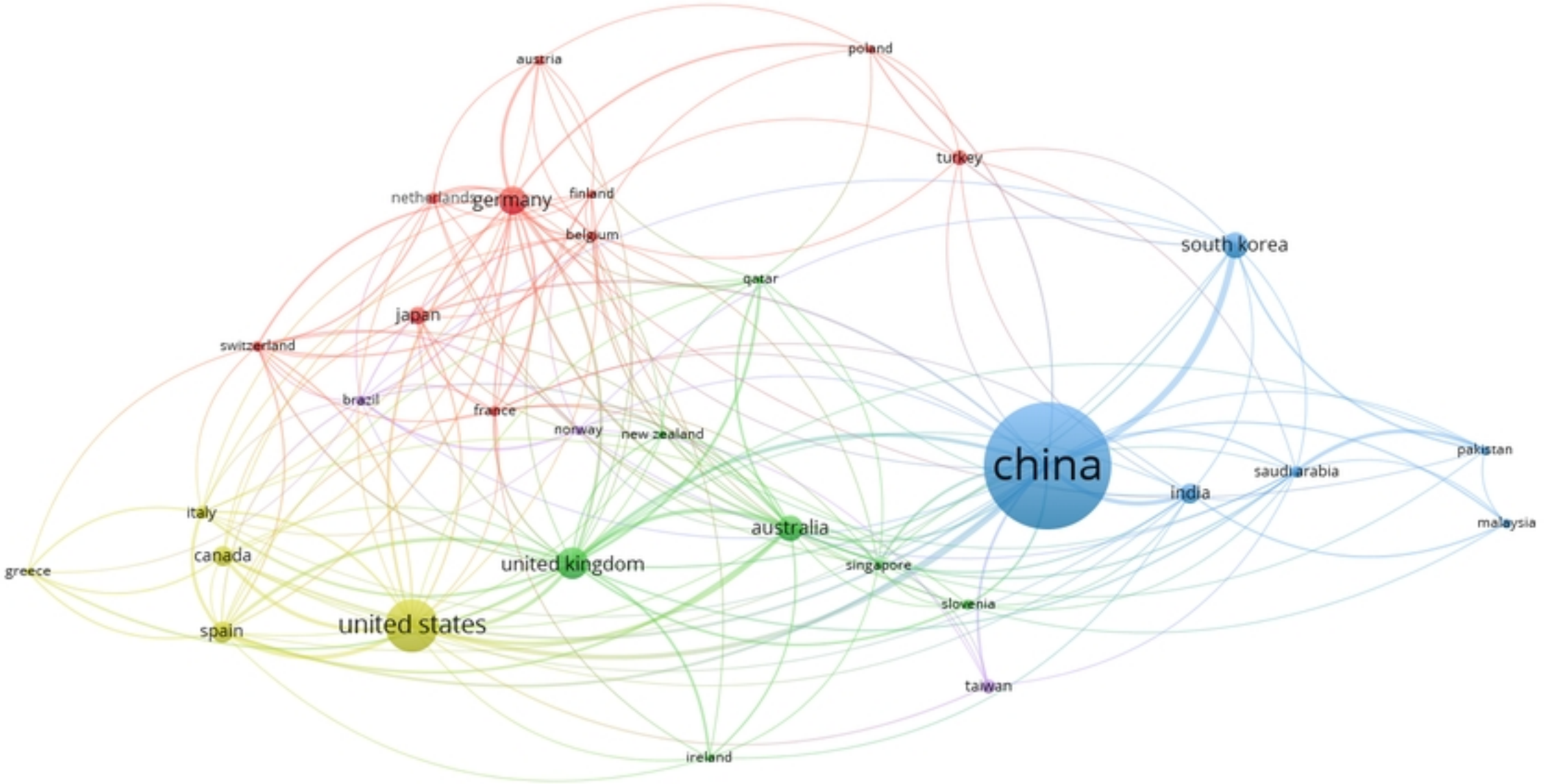
austria
poland
turkey
netherlands
germany
finland
belgium
south korea
qatar
japan
switzerland
brazil
france
norway
new zealand
china
pakistan
saudi arabia
india
italy
australia
malaysia
canada
united kingdom
singapore
greece
slovenia
spain
united states
taiwan
ireland

Figure 6

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515813; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

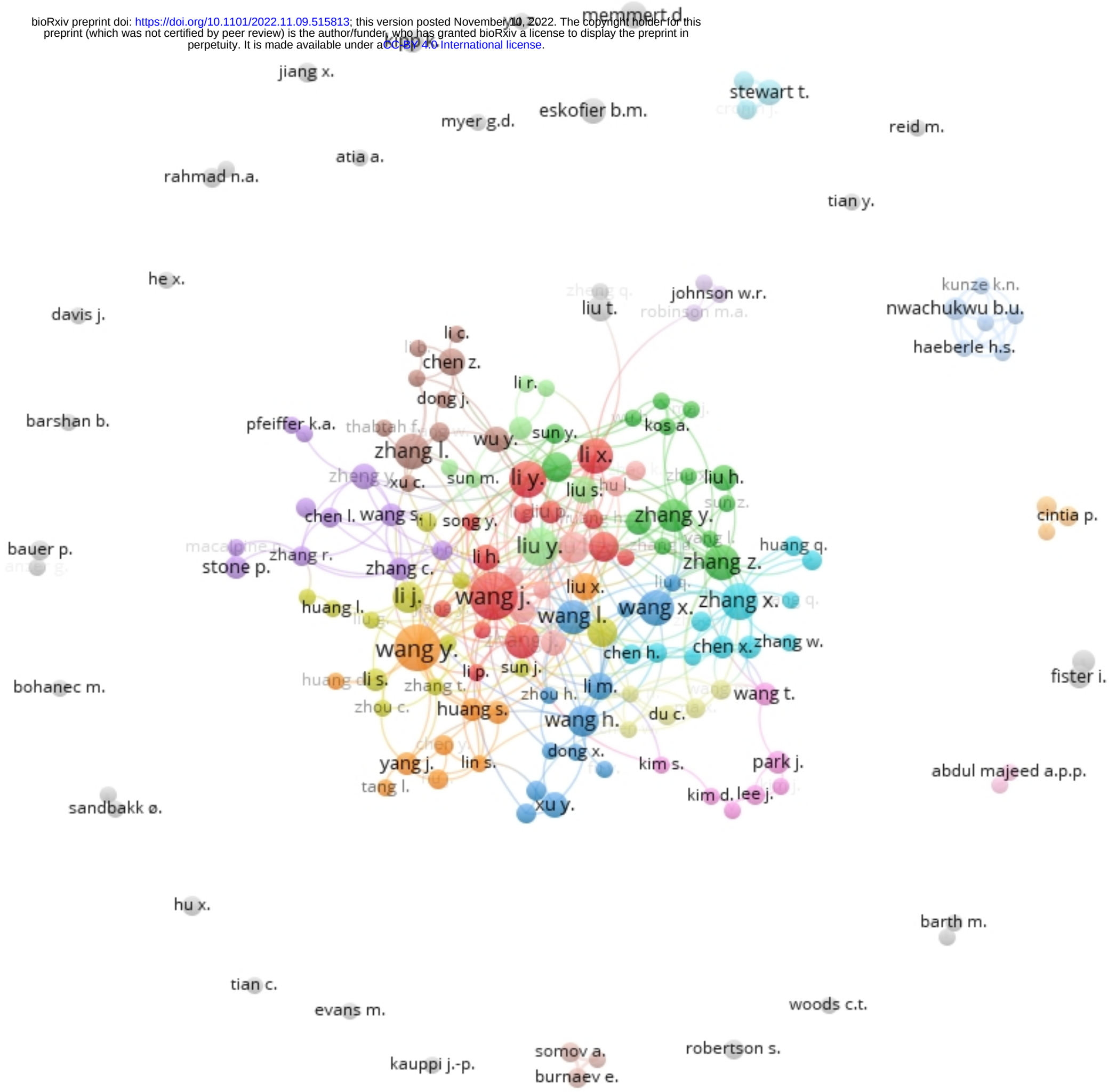

Figure 7

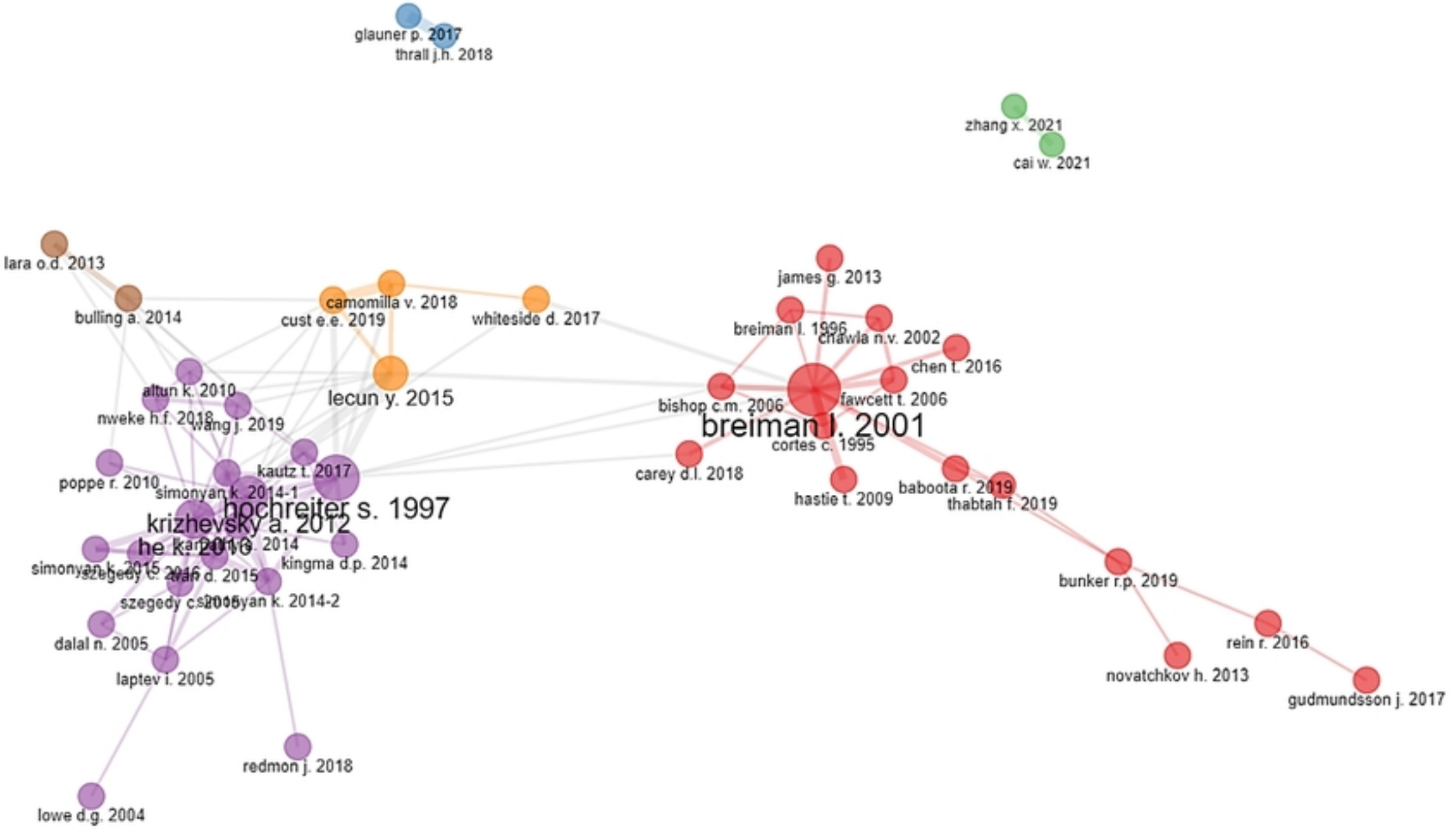
glauner p. 2017
thrall j.h. 2018
zhang x. 2021
cai w. 2021
lara o.d. 2013
bulling a. 2014
camomilla v. 2018
cust e.e. 2019
whiteside d. 2017
lecun y. 2015
altun k. 2010
nweke h.f. 2018
wang j. 2019
poppe r. 2010
kautz t. 2017
simonyan k. 2014-1
hochreiter s. 1997
krizhevsky a. 2012
he k. 2016
kingma d.p. 2014
szegedy c. 2015
ioffe s. 2015
szegedy c. 2016
simonyan k. 2014-2
dalal n. 2005
laptev i. 2005
redmon j. 2018
lowe d.g. 2004
james g. 2013
breiman l. 1996
chawla n.v. 2002
chen t. 2016
bishop c.m. 2006
fawcett t. 2006
breiman l. 2001
cortes c. 1995
carey d.l. 2018
hastie t. 2009
baboota r. 2019
thabtah f. 2019
bunker r.p. 2019
rein r. 2016
novatchkov h. 2013
gudmundsson j. 2017

Figure 8

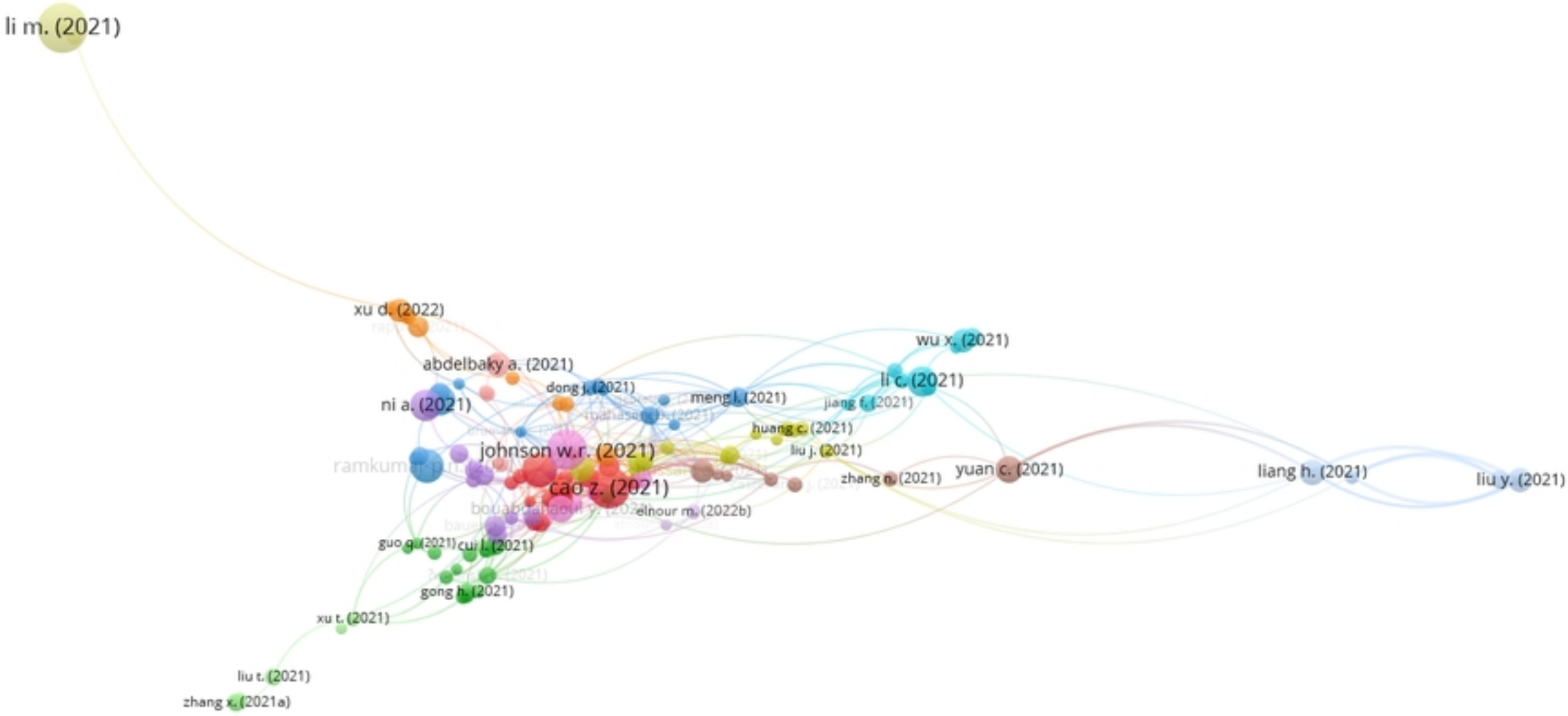
li m. (2021)
xu d. (2022)
abdelbaky a. (2021)
ni a. (2021)
dong j. (2021)
meng l. (2021)
wu x. (2021)
li c. (2021)
jiang f. (2021)
huang c. (2021)
liu j. (2021)
johnson w.r. (2021)
cao z. (2021)
zhang n. (2021)
yuan c. (2021)
liang h. (2021)
liu y. (2021)
elnour m. (2022b)
guo q. (2021)
cui l. (2021)
gong h. (2021)
xu t. (2021)
liu t. (2021)
zhang x. (2021a)

Figure 9

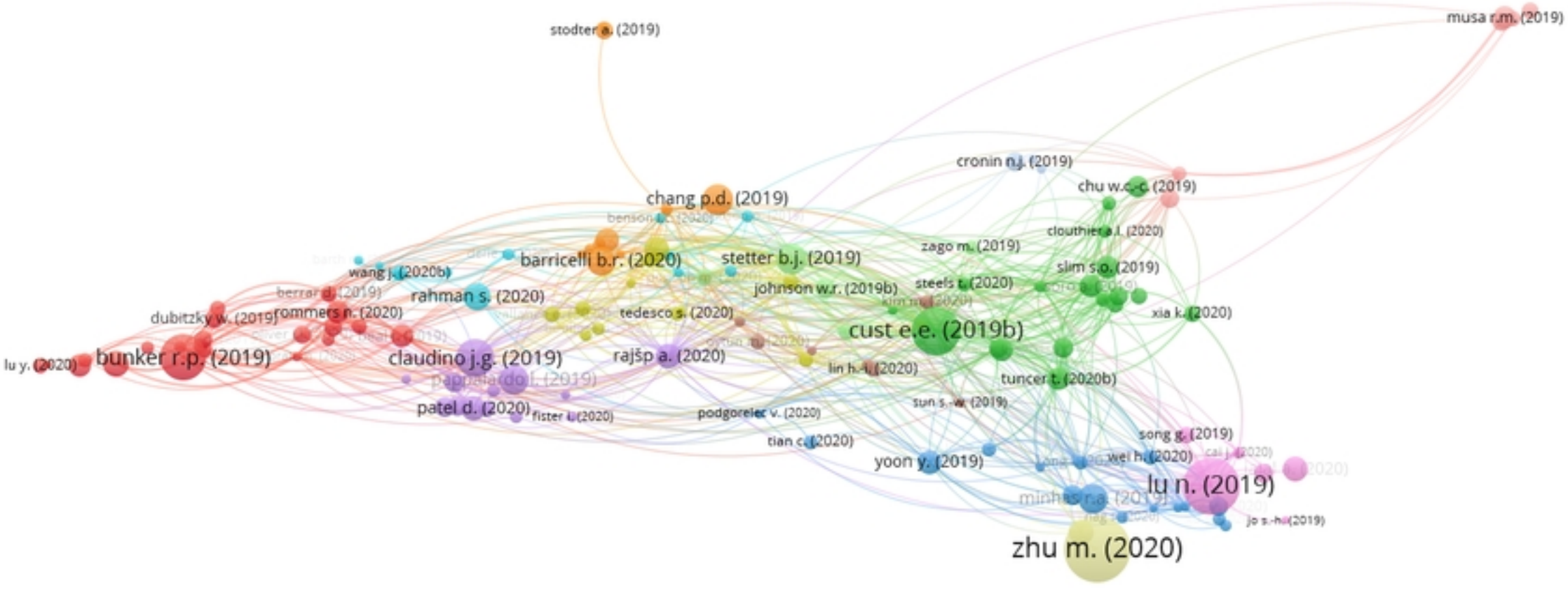
stodter a. (2019)
musa r.m. (2019)
cronin n.j. (2019)
chu w.c.-c. (2019)
chang p.d. (2019)
clouthier a.l. (2020)
zago m. (2019)
barricelli b.r. (2020)
stetter b.j. (2019)
slim s.o. (2019)
wang j. (2020b)
steefs t. (2020)
johnson w.r. (2019b)
rahman s. (2020)
xia k. (2020)
dubitzky w. (2019)
rommers n. (2020)
tedesco s. (2020)
cust e.e. (2019b)
lu y. (2020)
bunker r.p. (2019)
claudino j.g. (2019)
rajšp a. (2020)
lin h.-i. (2020)
tuncer t. (2020b)
sun s.-w. (2019)
patel d. (2020)
fister i. (2020)
podgorelec v. (2020)
song g. (2019)
tian c. (2020)
yoon y. (2019)
wei h. (2020)
lu n. (2019)
jo s.-h. (2019)
zhu m. (2020)

Figure 10

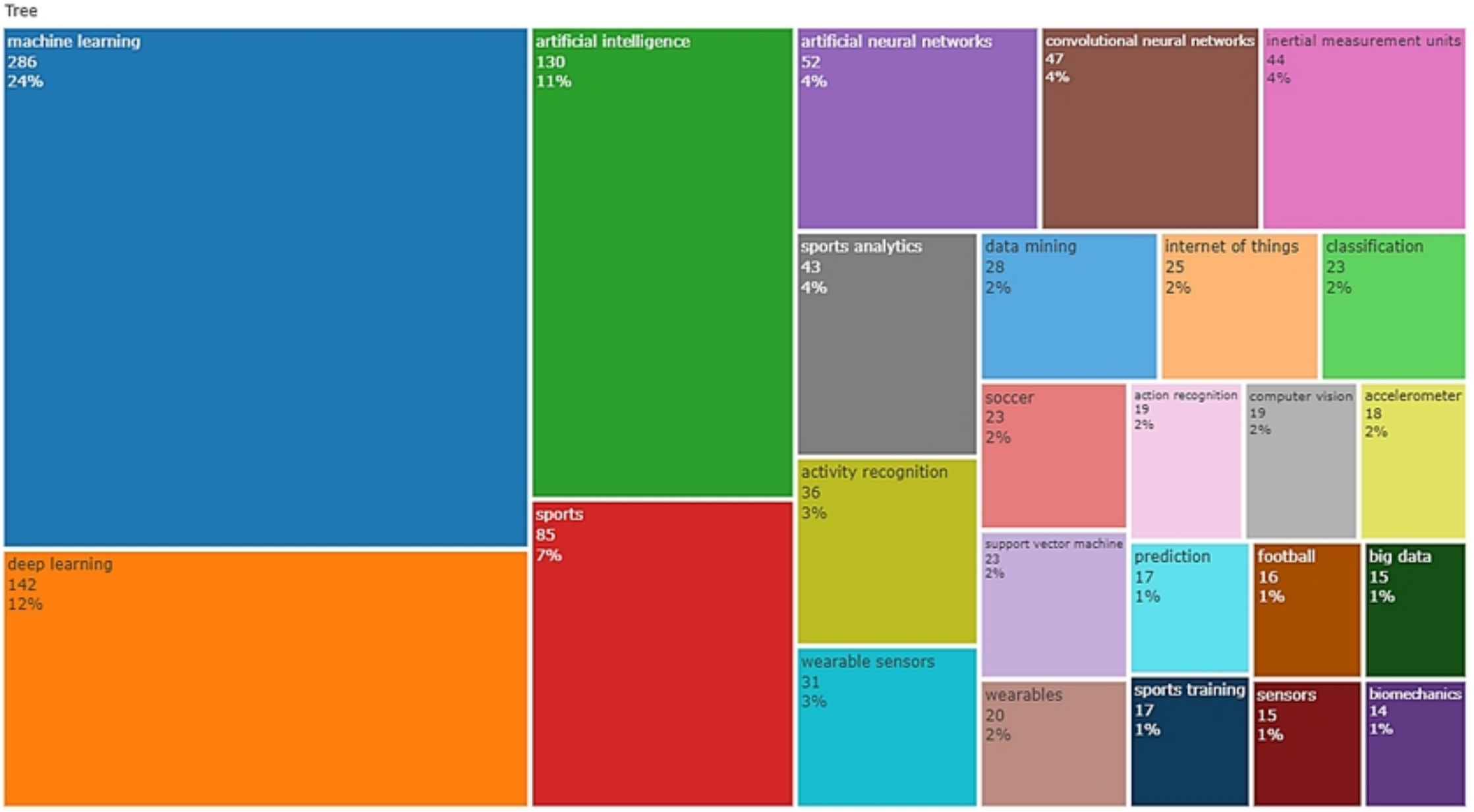
Tree
machine learning
286
24%
artificial intelligence
130
11%
artificial neural networks
52
4%
convolutional neural networks
47
4%
inertial measurement units
44
4%
sports analytics
43
4%
data mining
28
2%
internet of things
25
2%
classification
23
2%
soccer
23
2%
action recognition
19
2%
computer vision
19
2%
accelerometer
18
2%
activity recognition
36
3%
deep learning
142
12%
sports
85
7%
support vector machine
23
2%
prediction
17
1%
football
16
1%
big data
15
1%
wearable sensors
31
3%
wearables
20
2%
sports training
17
1%
sensors
15
1%
biomechanics
14
1%

Figure 11

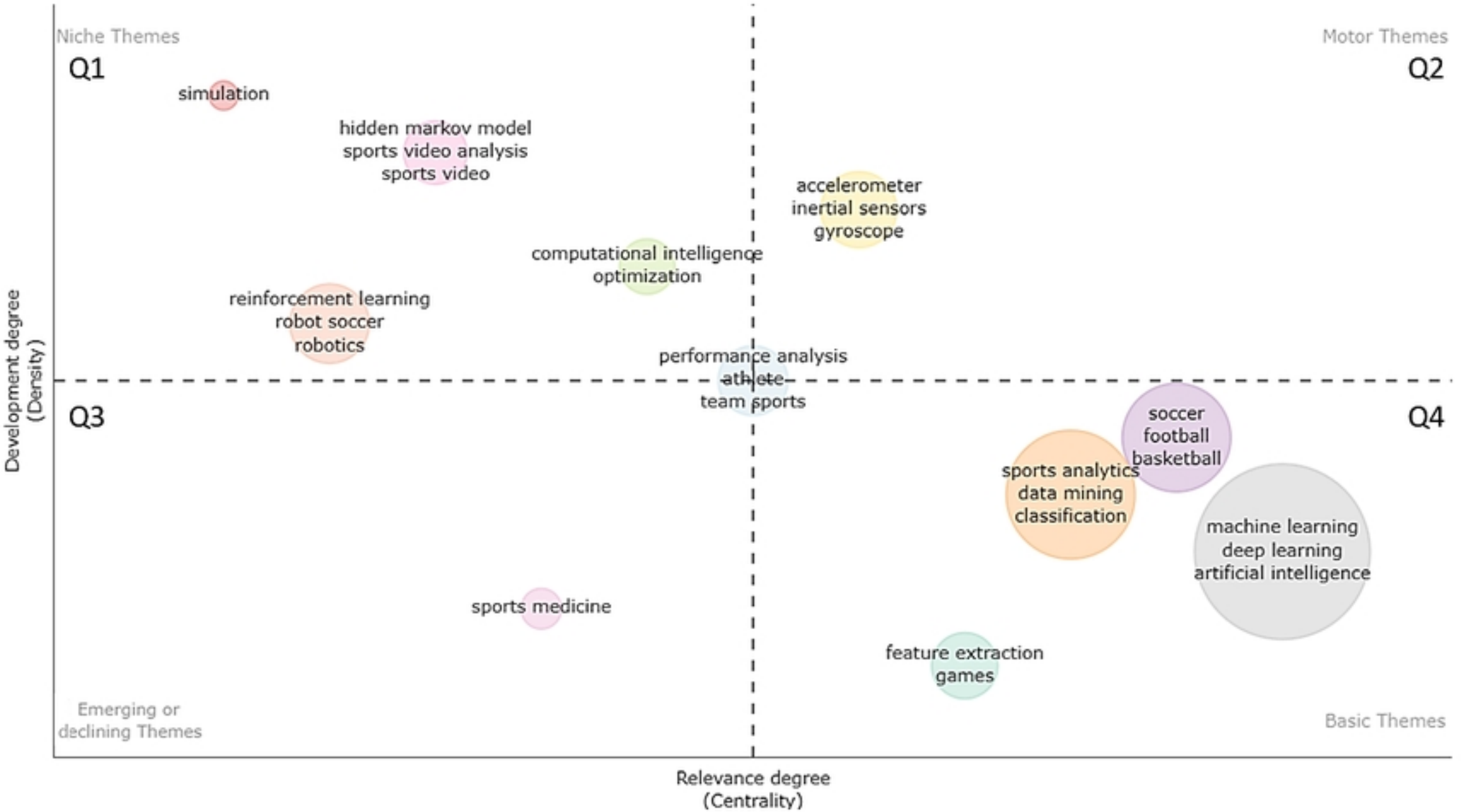
Niche Themes
Motor Themes
Q1
Q2
simulation
hidden markov model
sports video analysis
sports video
accelerometer
inertial sensors
gyroscope
computational intelligence
optimization
reinforcement learning
robot soccer
robotics
performance analysis
athlete
team sports
Development degree
(Density)
Q3
Q4
soccer
football
basketball
sports analytics
data mining
classification
machine learning
deep learning
artificial intelligence
sports medicine
feature extraction
games
Emerging or
declining Themes
Basic Themes
Relevance degree
(Centrality)

Figure 12

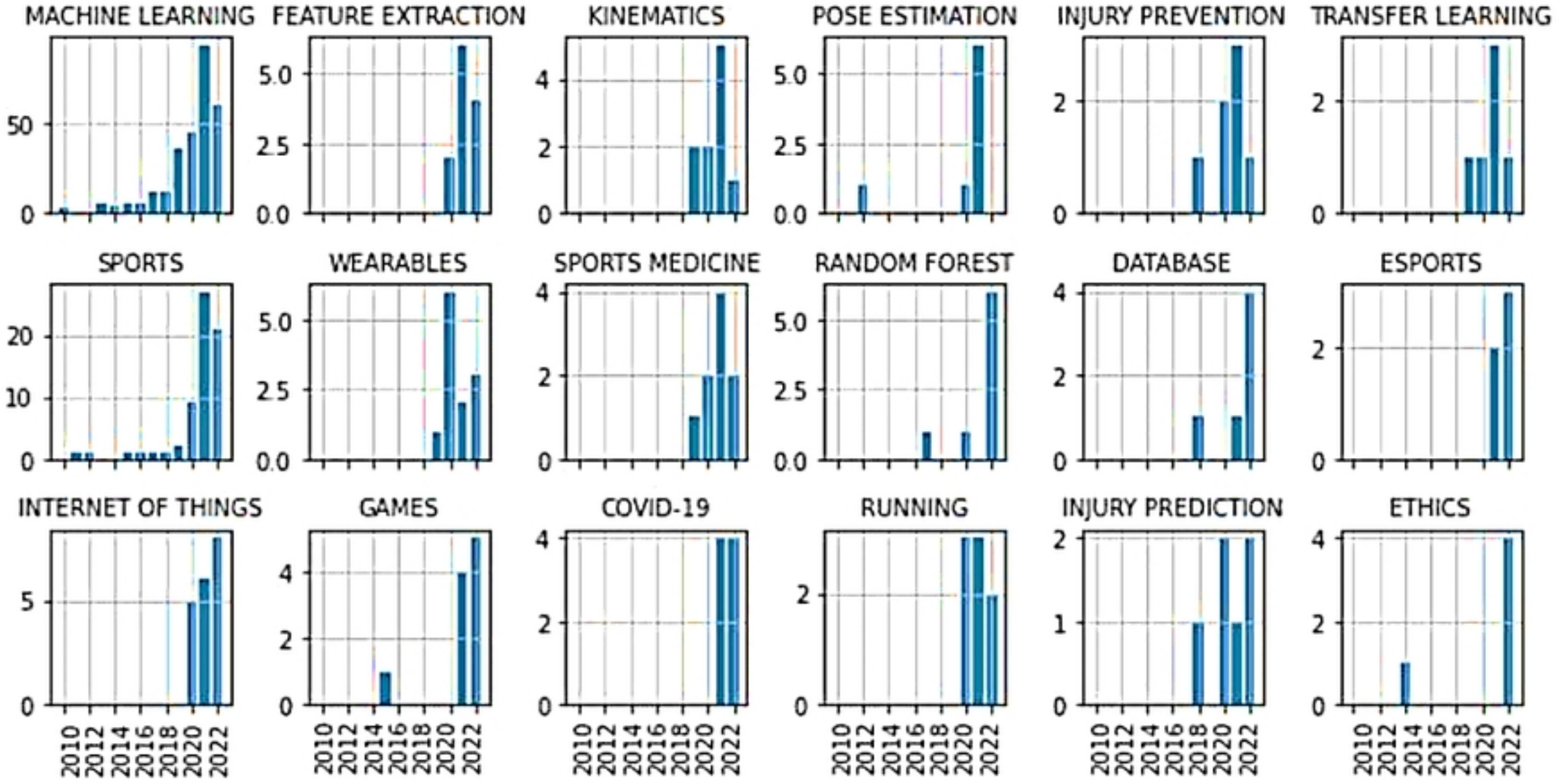

Figure 13

View publication stats