Τετράδια Ανάλυσης Δεδομένων

Τόμ. 20, Αρ. 1 (2024)

Τετράδια Ανάλυσης Δεδομένων - 20

Multiple Correspondence Analysis in Social Sciences and Humanities Research: A Longitudinal Mapping

Nikos Koutsoupias

Βιβλιογραφική αναφορά:

Koutsoupias, N. (2024). Multiple Correspondence Analysis in Social Sciences and Humanities Research: A Longitudinal Mapping. *Τετράδια Ανάλυσης Δεδομένων*, *20*(1), 59–77. ανακτήθηκε από https://ejournals.epublishing.ekt.gr/index.php/dab/article/view/32902

ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ | REVIEW PAPER

Multiple Correspondence Analysis in Social Sciences and Humanities Research: A Longitudinal Mapping

Nikos Koutsoupias

Department of International and European and Studies/ School of Social Sciences, Humanities and Arts, University of Macedonia

KEYWORDS

Multiple Correspondence Analysis
Bibliometric Data Analysis
Data Analysis
Scientometrics
Bibliometrics

CORRESPONDENCE

Nikos Koutsoupias
University of Macedonia
156 Egnatia str, 54636 GR
nk@uom.gr

ABSTRACT

Multiple Correspondence Analysis (MCA) is a multivariate nominal categorical data analysis method widely utilized for decades. We use bibliometric methods to investigate MCA related research applied in various fields from 1983 to 2021. The findings include the most important publication channels, authors, papers, and topics. In total, 1208 articles from the Social Sciences, Humanities, and Arts published in 662 journals were examined. Aside from other findings, we discovered 2962 distinct authors who contributed to the development of this research topic, with an annual growth rate of 15,07%. This article provides a broad overview of different perspectives on MCA-based research over time and will be of great assistance to those interested in this area of study.

1. Introduction

Multiple Correspondence Analysis (MCA) is a multivariate exploratory technique used to assess associations between multivariate categorical variables (Rozylowicz et al., 2017) and has a wide range of applications in fields such as archaeology, environmental, health and medical sciences, social and behavioral sciences, and so on. The use of MCA in humanities and social (MCAHS) research to explore data sets of categorical variables is well established as a popular method to explore and visualize the relationships among categorical data (Beh & Lombardo, 2019).

This paper examines the application of MCA to the literature in the humanities and social sciences as a tool for detecting latent patterns in related categorical data. MCA is a dimensionality reduction factor technique similar to the more popular principal component analysis (PCA), but it is specifically designed to deal with nominal or ordinal categorical data, and it is well suited to the types of data we face when describing subject (row) characteristics. MCA works specifically with categorical data, allocating scale values to the categories of discrete variables and maximizes the variance of those scores to discover: (1) associations between variables, and (2) individual proximity.

Although this will not be a thorough discussion of the numerous aspects of MCA, we will concentrate on the important literature for the study of nominal categorical variables. For reviews of the technical, practical, and historical issues related to MCA and related methods, the reader is encouraged to review the works of Tenenhaus

and Young (1985), Gower (1990), Carlier and Kroonenberg (1996), Le Roux and Rouanet (2004), Greenacre and Blasius (2006), Choulakian (2008), Hwang, Tomiuk and Takane (2009), Markos, Menexes, and Papadimitriou (2009), Beh and Lombardo (2014, 2019), Kroonenberg (2014), Husson and Josse (2014) and Mori, Kuroda and Makino (2016).

Furthermore, various researchers published literature reviews of MCA applications in general (Kamalja and Khangar, 2017; Hjellbrekke, 2018) or within their scientific fields, including Archaeology (Macheridis, 2017), Education (Van der Heijden et al., 1997), Epistemology (Alhuzali et al., 2022), International Relations (Martin-Maze, 2018), Finance (Dungey, et al., 2018), Medicine (Meyer et al., 2004; Han et al., 2018), and Operational Research (Pierreval, 1994).

To our knowledge, MCA has not yet been reviewed comprehensively as a Social Sciences, Humanities and Arts research tool using bibliometric analysis.

A review of published research that is expanding exponentially (Linnenluecke et al., 2020) can provide scientists and policymakers with a detailed picture of the research done up to a particular point in time, as well as future perspectives (Mair et al., 2018). Nonetheless, systematic reviews of recent literature typically necessitate a significant amount of time and effort, depending on the number of publications to be investigated (Radhakrishnan et al., 2017). The process of synthesizing research evidence and discussing the state-of-the-art relies heavily on literature reviews (Aria and Cuccurullo, 2017). The bibliometric approach introduces a systematic and objective procedure that is transparent, trustworthy, and simple to replicate (Aria et al., 2020; Diodato and Gellatly, 2013), usually on massive article sets (Donthu et al., 2021). Bibliometric analysis is a systematic statistical analysis of scientific research that is most used to determine the top-cited articles in journals, to assess trends and patterns in specific research fields, to identify the most commonly used research topics or interests, and to map the structure of co-authorship, affiliations, co-citation networks (Shiau et al., 2017), and cross-border collaboration (Xu et al., 2018).

We utilize bibliometrix (Aria and Cuccurullo, 2017) for this bibliometric review on about seven hundred articles. Numerous researchers in Social Sciences, Humanities and the Arts sub-fields recently took advantage of the capabilities of bibliometrix to review interdisciplinary literature (Ho et al, 2021; Tontodimamma et al., 2021) and among others, in Culture (Youngblood and Lahti, 2018), Education (Agbo et al., 2021; Bond and Buntins, 2018; Bond et al., 2019; Guraya et al., 2021; Liu et al., 2020), Geography (Gibadullina, 2021; Latocha, 2020; Rosas-Chavoya, 2022), Information Science (Bapte, 2021; Patel and Bhatt 2021, Pham-Duc et al., 2022), Political Science and International Studies (Carammia, 2022; Jensen and Moses, 2021; Medie and Kang, 2018; Peez, 2022), Psychology (Adler and Sarstedt, 2021; Altarturi et al., 2020), Sociology (Aria et al., 2020; Maretti et al., 2019) and Transportation (Chen and Silva, 2021; Jing, 2021; Schneider and Bengler, 2020).

The purpose of this paper is to review and summarize the scientific production of MCA applications in Social Sciences, Humanities and the Arts, using bibliometric techniques implemented with bibliometrix to analyze the evolution and trend research [Objective 1, Obj1], the origin and evolution of scientific production (by countries [Obj2], authors [Obj3], papers [Obj4], keywords[Obj5], and trends [Obj5]), and the classification and analysis of the content of articles based on the keywords and themes [Obj6].

The remainder of the paper is organized as follows. We begin in Section 2 with a discussion of current areas of concern around the bibliographic review methodological framework and proceed in Section 3 to present the basic statistical results of the bibliometric analysis. Sections 4 and 5 provide results of trends analysis and multivariate mapping of themes of MCA in Social Sciences, Humanities, and the Arts. Conclusions, Limitations and a reference to Future Research are covered in Section 6.

2. Data and Methodology

We searched the Scopus database for the "Multiple Correspondence Analysis" within Title, Author Keyword and Abstract fields from 1983 to 2021. The goal was to find English-language journal research articles in Social Sciences, Humanities and Arts related disciplines. The following is the search query, which was conducted on October 16, 2022:

"multiple correspondence analysis" AND
(LIMIT-TO (SUBJAREA, "SOCI") OR LIMIT-TO (SUBJAREA, "ARTS")) AND
(LIMIT-TO (DOCTYPE, "ar")) AND
(LIMIT-TO (SRCTYPE, "j")) AND
(EXCLUDE (PUBYEAR, 2023) OR EXCLUDE (PUBYEAR, 2022))

The findings cover a total of 1208 SCOPUS-indexed journal articles since early access materials formally published in 2022 are excluded from the list. The open-source bibliometrix package in R, (Aria and Cuccurullo, 2017) was used for bibliometric analysis. Biblioshiny, the package's web interface app, allows users to perform analysis at four levels of analysis, namely sources, countries, authors, affiliations, and documents, as well as map knowledge structures in conceptual, intellectual, and social terms.

3. Descriptive Results

Answering Obj1 of the present review, the average annual growth of scientific article production is 26.13% overall. Figure 1 depicts the annual scientific MCA output.

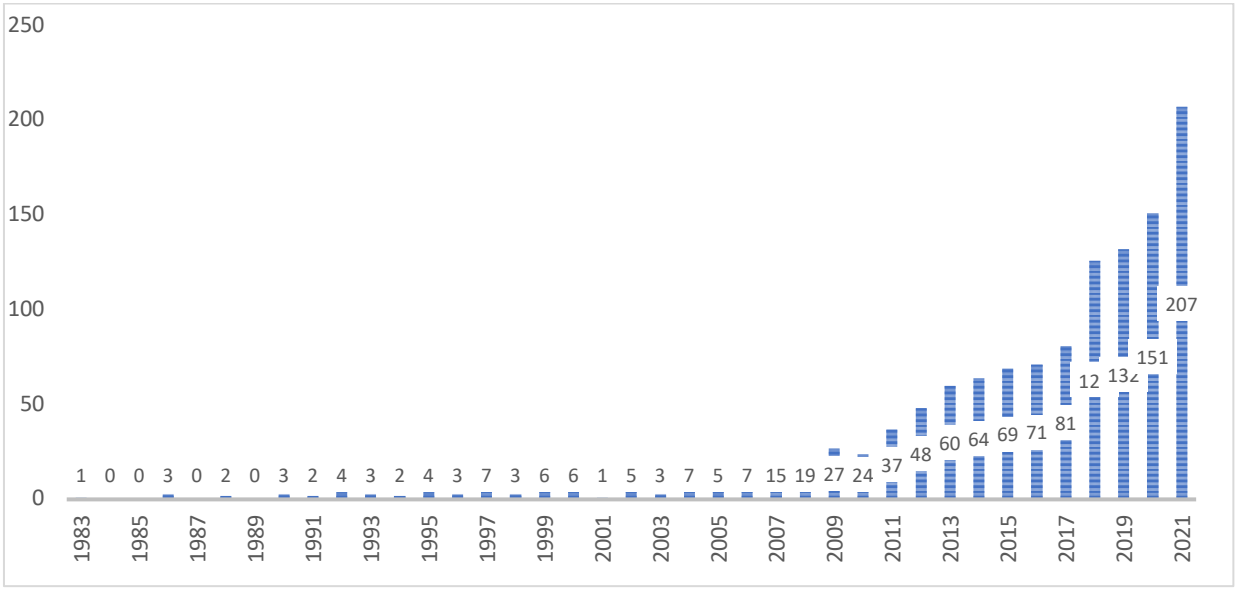

Figure 1: Annual scientific production

Table 1 summarizes the key statistics of publications on MCA within the specified period. The raw statistics indicate that the documents were published by a total of 2962 distinct authors. Approximately 5.2% of these documents were published by a single author, with the remainder written by multiple scholars.

Table 1

Descriptive statistics of research on 'MCA' in 1983–2021

Description	Results

MAIN INFORMATION ABOUT DATA	
Timespan	1983:2021
Sources (Journals, Books, etc)	662
Documents	1208
Annual Growth Rate %	15,07
Document Average Age	6,32
Average citations per doc	15,74
DOCUMENT CONTENTS	
Keywords Plus (ID)	2970
Author's Keywords (DE)	4070
AUTHORS	
Authors	2962
Authors of single-authored docs	204
AUTHORS COLLABORATION	
Single-authored docs	232
Co-Authors per Doc	2,91
International co-authorships %	24,01

(Source: Scopus)

Provided that articles with a large number of authors can skew the metrics, the statistic for co-authors per document, 2.91 in this case, takes into consideration the number of times an author appears in the entire set of documents.

Both measures indicate that, on average, about three authors were engaged in one publication on MCA between 1983 and 2021, while the cross-boundary collaboration reaches a relatively high value of 24%.

4. Analysis by Country, Author, Paper, and Keywords

This section presents a bibliometric study based on several performance indicators. This allows us to meet the Objectives 2 to 5 of the present review ([Obj2-Obj5]) .

Contributing countries in terms of productivity and impact

The publication of articles is an indicator that suggests a country's development in a certain research area (Xie et al., 2020). In this study, 91 countries were identified as having undertaken MCAHS research between 1983 and 2021. Based on author appearances by country affiliations, the USA was the most prolific country with 377 publications, followed by Italy (342), France (274), Spain (264), UK (209), Portugal (159), Germany (142), Canada (122), Brazil (111) and the Netherlands (101).

Table 2

Highest impact in total and average article citations by country

Country	TC	AC (Avg)
ITALY	3174	36,91
UNITED KINGDOM	2520	40,00
USA	1909	18,90
FRANCE	1345	17,24
SPAIN	845	12,43

GERMANY	568	11,14
BELGIUM	552	39,43
CANADA	552	17,81
SWITZERLAND	552	21,23
NETHERLANDS	532	15,20

The analysis of the top ten countries on MCAHS publications is based on three indices: total times cited, average article citations and number of articles (Table 2). Italy ranked first, with the greatest number of times cited articles (3174), followed by the UK, USA, France, and Spain (2520, 1909, 1345 and 845 respectively). Most productive scientists on MCAHS come from USA (377), Italy (342) and France (274) while the highest number of citations on average originate from UK based researchers.

Contributing authors in terms of productivity and impact

Lotka's Law was used to calculate the authors' productivity distribution based on the frequency of articles published on MCAHS. According to the results of Lotka's Law, 2637 authors out of a total of 2962 authors (89%) published one article in the field of MCAHS; 221 authors (7.5%) published two papers; 51 authors (1.7%) published three papers; 25 authors (0.8%) published four papers, 13 authors (0.4%) published five papers, and 21 authors published six or more articles.

Table 3

Top 10 authors by h-index, impact and production

Author	NP	h_index	TC
DAS S	16	8	275
BLASIUS J	14	9	230
SAVAGE M	8	8	1184
ROSENLUND L	8	7	219
BÜHLMANN	8	6	157
LE ROUX B	7	7	906
DUTTA A	7	6	119
FLEMMEN M	6	5	174
HJELLBREKKE J	5	5	814
JARNESS V	5	5	137

Note: TC represents total citations. NP represents the number of publications.

As shown in Table 3, the top authors of publication number were Das, S. (16), Blasius, J. (14), Savage, M. (8), Rosenlund, L. (8), Buhlman, F. (8), Le Roux, B. (7), Dutta, A. (7), Flemmen, M. (6), Hjellbrekke, J. (5), and Jarness, V. (5).

However, the authors' productivity in terms of publication number was not proportional to the impact they gained regarding the citation number. Citation analysis of authors was performed to identify the most influential or impactful authors in the study field (Table 3). The top ten authors out of 2962 received well over 100 citations for their publications, according to the findings.

Savage, M. was named the most influential author in MCAHS sciences fields, with a total of 1184 citations. Le Roux, B., followed Savage, M., with 906 citations and Hjellbrekke, V., who came in third with 814 citations. It's worth noting that many authors collected more citations despite having fewer publications. For instance, the

co-authors Aria M., and Cuccurullo, C., received 1939 citations from two pieces of work, indicating that their papers had a significant impact in the MCAHS field of study.

Contributing papers in terms of impact

According to research, in addition to scientific productivity counts, the number of citations for a publication serves as an indicator of its contribution and scholarly impact (Grant et al., 2000; Waheed et al., 2018). Citation analysis was performed to investigate the most important documents in the field of MCAHS. It was seen that 1049 out of 1208 articles obtained more than one citation, implying that approximately 86,8% of documents were cited more than once. The top ten most cited papers were identified as having more than 150 citations each (Table 4).

Table 4

Top papers by citation count

Author (1st)	Year	Title	Total Citations	TC per Year
Aria, M.	2017	*Bibliometrix: An R-tool for comprehensive science mapping analysis. J. Inf* 11.4: 959-975	1839	306,50
Savage, M.	2013	A new model of social class? Findings from the BBC's Great British Class Survey experiment	681	68,10
Cohen, J.	1988	Set correlation and contingency tables	392	11,20
Chauvin, N.	2013	A new model of social class? Findings from the BBC's Great British Class Survey experiment	318	31,80
Philippot, P.	1993	Inducing and assessing differentiated emotion-feeling states in the laboratory	283	9,43
Booysen, F.	2008	Using an asset index to assess trends in poverty in seven Sub-Saharan African countries	203	13,53
Prieur, A.	2013	Updating cultural capital theory: A discussion based on studies in Denmark and in Britain	195	19,50
Lieven, T.	2011	Who will buy electric cars? An empirical study in Germany	184	15,33
Beauregard, E.	2007	A descriptive model of the hunting process of serial sex offenders: A rational choice perspective	179	11,19
Dubois, G.	2019	It starts at home? Climate policies targeting household consumption and behavioral decisions are key to low-carbon futures	167	41,75

Apart from the paper that introduces the Bibliometrix package (Aria & Cuccurullo, 2017) which is the most impactful publication in the examined dataset with 1839 citations, Savage's et al. (2013) article entitled "A new model of social class? Findings from the BBC's Great British Class Survey experiment." shows 681 citations and Cohen's (1988) research comes third with 392 citations.

Keyword frequencies and co-occurrence network

We examined both keyword frequency as well as keyword co-occurrences since keyword analysis, according to Strozzi et al. (2017), and Song (2019), is important for determining latest research developments, trending topics and scholars' interests in the field. Keywords plus was the approach of choice for this part of the study, since research showed that it enhances the depth and breadth of a literature search (Garfield, 1990). This analysis is necessary because publication keywords aid in quickly identifying the theme and focus of research. The

wordcloud in Figure 8 depicts consistently used keywords+ in publications on MCAHS, and Table 5 contains the top ten Keywords Plus in the examined data set.

Figure 2: K+ Wordcloud

Table 5

Top keywords (K+)

Words	Articles
female	140
male	139
human	133
adult	123
correspondence analysis	98
humans	90
article	83
middle aged	80
aged	61
young adult	45

Apart from the term "correspondence analysis," the rest of the frequently used words emerge from general Social Studies and Humanities research fields strongly focusing on gender and age issues.

Callon et al. (1984) proposed co-word analysis as a content analysis technique. It is used to map the strength of associations between textual information items (Wang et al., 2012). It entails a keyword co-occurrence analysis in a specific body of literature. Co-occurrence analysis, a key component of association analysis in data mining, is applied to cluster keywords with high relevance (Cheng et al., 2014). Each cluster can be thought of as a research topic. Co-occurrence analysis is frequently employed to identify emerging and established research themes, as well as to trace patterns (Deng et al., 2021; Leung et al., 2017; Ravikumar et al., 2015; De la Hoz-Correa et al., 2018].

Keyword co-occurrence networks aid in the identification of important terms used in papers within a field of study and provide insights into the domain's main research themes (Van Eck and Waltman, 2010). For the purposes of network analysis, keywords or key terms and their co-occurrences in publications, were extracted using the Keywords Plus approach (Garfield, 1990).

We visualized keyword co-occurrence network as shown Figure 3. A total of 2970 keywords were extracted from our list of 1208 MCAHS related publications. Only keywords that appeared more than 15 times were considered for keyword analysis. This criterion was met by 50 keywords. The distance between circles in the illustrated keyword network (Figure 3) provides information on the relatedness of keywords, and the size of the nodes reflects the occurrence frequency of a keyword (Van Nunen et al., 2018). The frequency with which those keywords appear together within the keyword network is indicated by node overlap (Rodrigues et al., 2014). Furthermore, the keyword map shows how the keywords of MCAHS publications are grouped together.

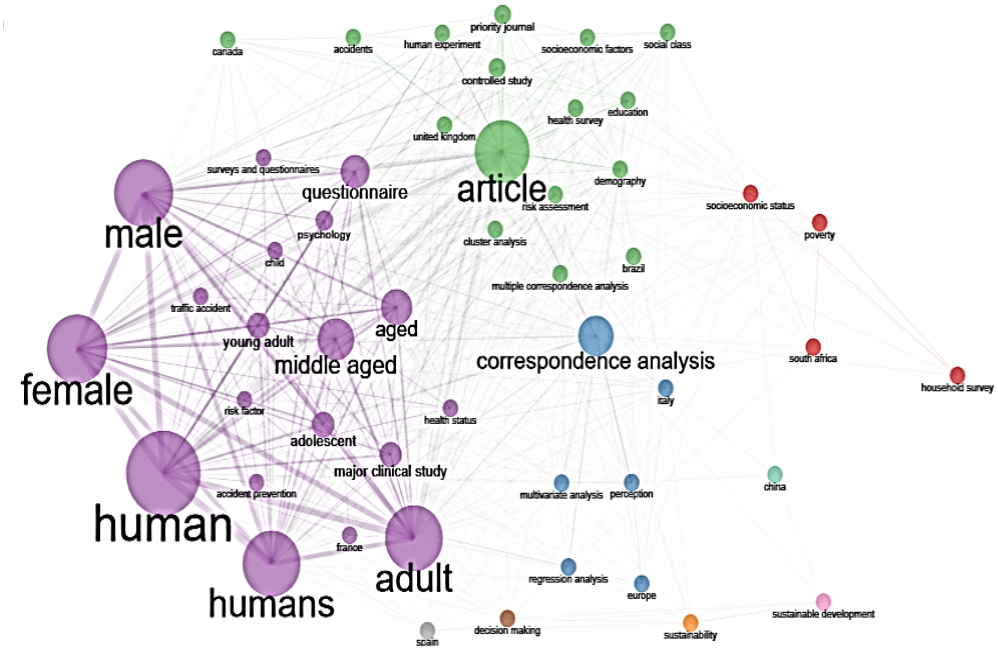

Figure 3: Keyword co-occurrence (n >15)

The keyword network appears to be divided into three distinct clusters, according to our analysis: cluster one (purple color, left side of the network), cluster two (green color, top side), cluster three (blue color, right side).

The purple cluster (cluster 1) is the largest with 38 keywords. The most frequently used keywords here are female, male, human, adult, humans, middle aged, aged, young adult, adolescent, questionnaire, child, major clinical study, psychology, France, health status, risk factor, surveys and questionnaires, traffic accident and accident prevention. According to this analysis, this cluster focuses on French publications based on gender and age-related questionnaires in the context of MCAHS. Another significant finding from our analysis is that this cluster has keywords of papers that conducted traffic accident research.

Cluster two (green) is the second largest cluster, with 16 keywords including: article, cluster analysis, multiple correspondence analysis, risk assessment, education, Brazil, priority journal, social class, United Kingdom, accidents, demography, controlled study, socioeconomic factors, Canada, health survey and human experiment. This cluster relates mainly to research papers from UK, Canada and Brazil referring to risk assessment, education, social class. This cluster has a considerable overlap with the network's first cluster and a moderate degree of coherence.

The third cluster (blue color) includes six keywords and is situated close to the middle of the network. This cluster has a low degree of coherence, and it overlaps to a large degree with cluster one. The keywords in this cluster are correspondence analysis, Italy, perception, Europe, regression analysis and multivariate analysis. In addition to Correspondence Analysis, papers in this word sub-graph mainly employed regression analysis ME that refer to data from Italy and Europe.

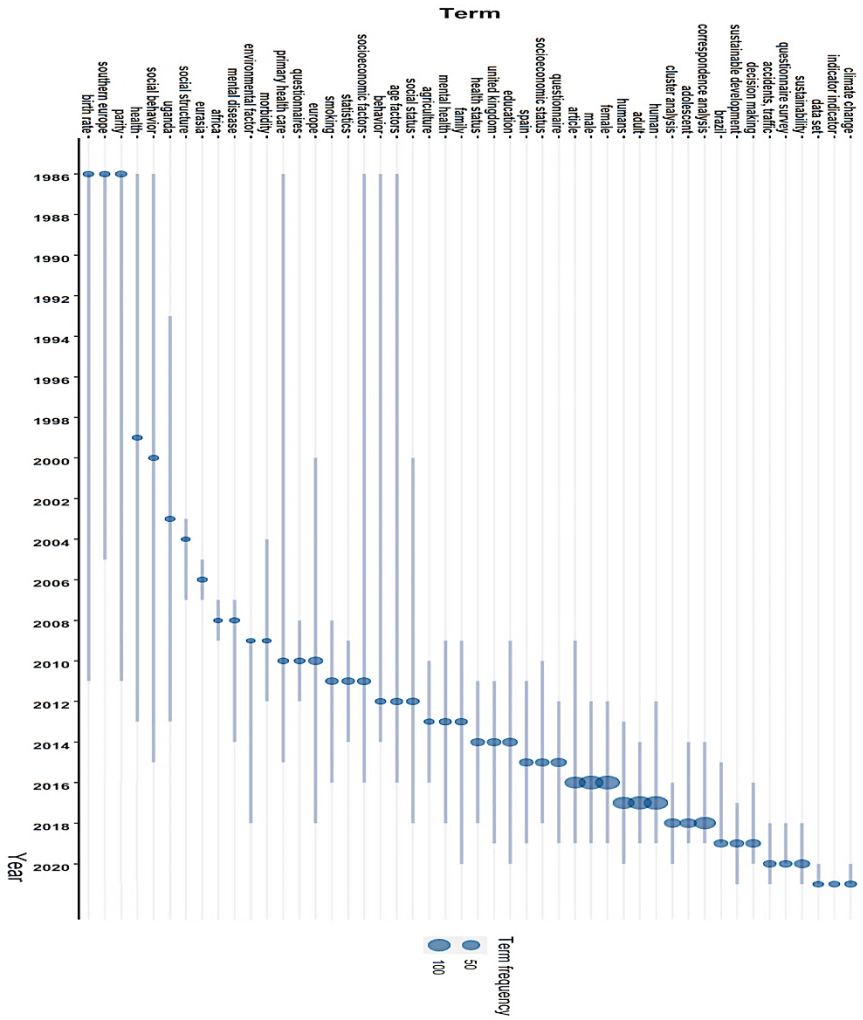

Figure 4: Trend topics

5. Trends and Thematic Analysis

This section provides trend and thematic analysis using co-word analysis, multiple correspondence analysis a k-means clustering, thus meeting Obj5 and Obj6 of the present review.

Trends based on Keywords+

Following co-word analysis, we add to the present academic exploration by highlighting quantitative trends on MCAHS. Figure 4 shows a graphical representation of the top 50 research trends over time. It has been discovered that the most popular topics (as determined by their positions on the figure) relate to climate change and sustainability. Initially, since 1986, MCAHS research seems to focus on parity and birth rate issues mainly in southern Europe with a shift to health, social behavior and mental disease topics in Eurasia and Africa from 1999 to 2008. Next, for the period between 2009 and 2015, subjects like morbidity, smoking, behavior, agriculture, family, education, health, and socioeconomic status in the European region seem to be main trends. The period from 2016 to 2018 is marked greatly based on gender and age characteristics by studies that combine Correspondence Analysis and Cluster Analysis. The most persistent terms over time (especially from 1986 to 2014) seem to relate to research on age and socioeconomic factors, as well as behavior and social behavior and

primary health care, while the shortest research period was spent on African and Eurasia issues in 2008 and 2006 correspondingly.

Themes Analysis

A thematic map was created to further investigate article themes and examine how networks evolve (Fig. 5). Topic clusters were created and plotted using centrality and density indices (Callon et al., 1991; Ding, 2001).

The map (Fig. 5) shows the strength of their internal (density) and external (centrallity) relationships. It is divided into four quadrants: topics with high density and centrality (q1), topics with low density and high centrality (q2), topics with high density and low centrality (q3), and topics with smaller values on both axes (q4) (Grivel et al., 1995; López-Fernández, et al., 2016).

A thematic or strategic map with different clusters or themes is represented on it, where centrality denotes the significance of the theme in the overall research field and density denotes the measure of the theme's development (Cobo et al., 2011). The field's thematic map parameters are the keywords plus terms with a word count of 169. Each colored circle represents a network cluster, and the name is the word with the highest occurrence value in that cluster. The circle size is proportional to the number of cluster word occurrences. The position of these circles in the Bibliometrix package is determined by their centrality and density (Callon et al. 1983); the four quadrants are as follows: The upper-left quadrant contains greatly developed but disconnected themes with strong internal interconnections but no outward connectivity; and hence, their significance is limited. The lower-left quadrant contains themes with low density and centrality that are either evolving or declining and thus underdeveloped. The motor themes with the highest density and centrality are represented in the upper-right quadrant, and they are linked to topics associated with other strongly connected concepts. These themes are well-developed and important for the advancement of a field of research. The primary and transversal themes in the lower-right quadrant are crucial for study but have not been fully developed.

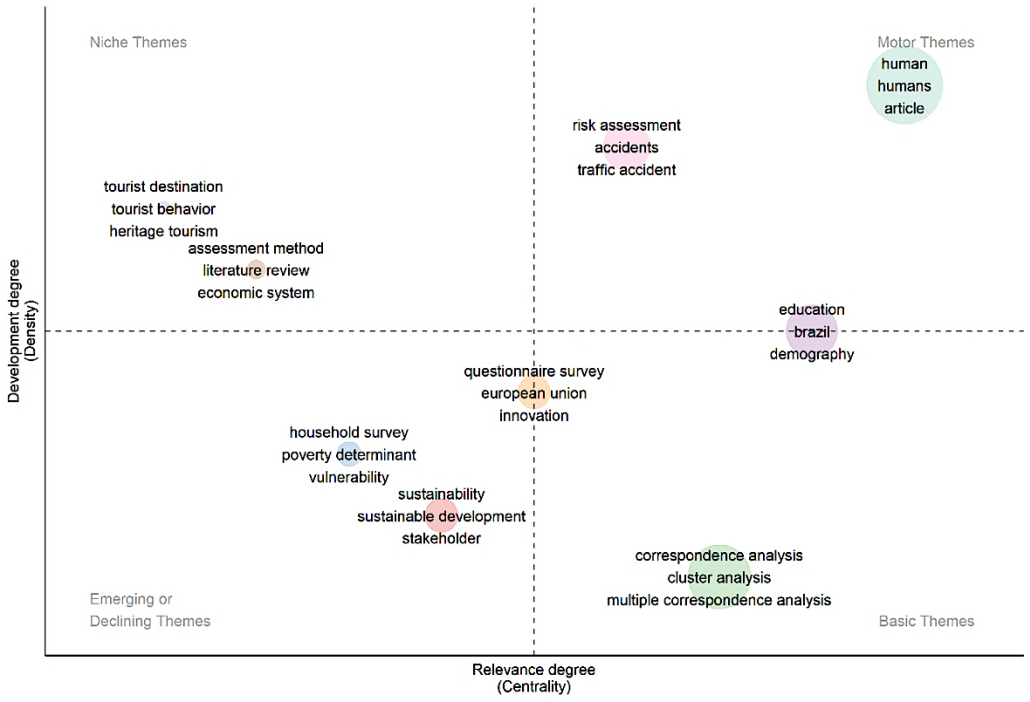

Figure 5: Theme mapping

The thematic terms in the first quadrant (upper right), with high internal and external associations, are considered mainstream topics in the analyzed set of articles on MCAHS. Themes related to humans, risk assessment and accidents are located in this quadrant.

In contrast, research topics in the fourth quadrant (lower right), which have low density and centrality values, are underdeveloped or emerging themes that need to be developed further. The main topics in this group is multiple correspondence analysis and cluster analysis. We find basic and transversal topics in the second quadrant (upper left), which has a low density and a high centrality. Consequently, this group covers general tourism and methodology topics. Finally, there are articles about specialized topics such as sustainability and poverty in quadrant three, which has a high density but a low centrality. Important issues in MCAHS research that need more development appear between third and fourth quadrants (questionnaire surveys, European Union, and innovation).

The conceptual structure map produced by multiple correspondence analysis and k-means clustering analysis puts the keywords (+) together while taking their homogeneity within the network into account. Figure 6 depicts the distribution of the most popular terms in two clusters, along with the corresponding most cited papers. The blue cluster on the left brings together keywords such as sustainability, education, demography, decision making, socioeconomics, and accidents. Highly cited articles in this cluster mainly refer to general health issues (Veenstra, 2007) and issues related to AIDS (Cavalcante, et al., 2013) and gerontology (Liotta, et al., 2017; Lucchetti, 2008; Moreno, 2016) in China, Brazil and Europe.

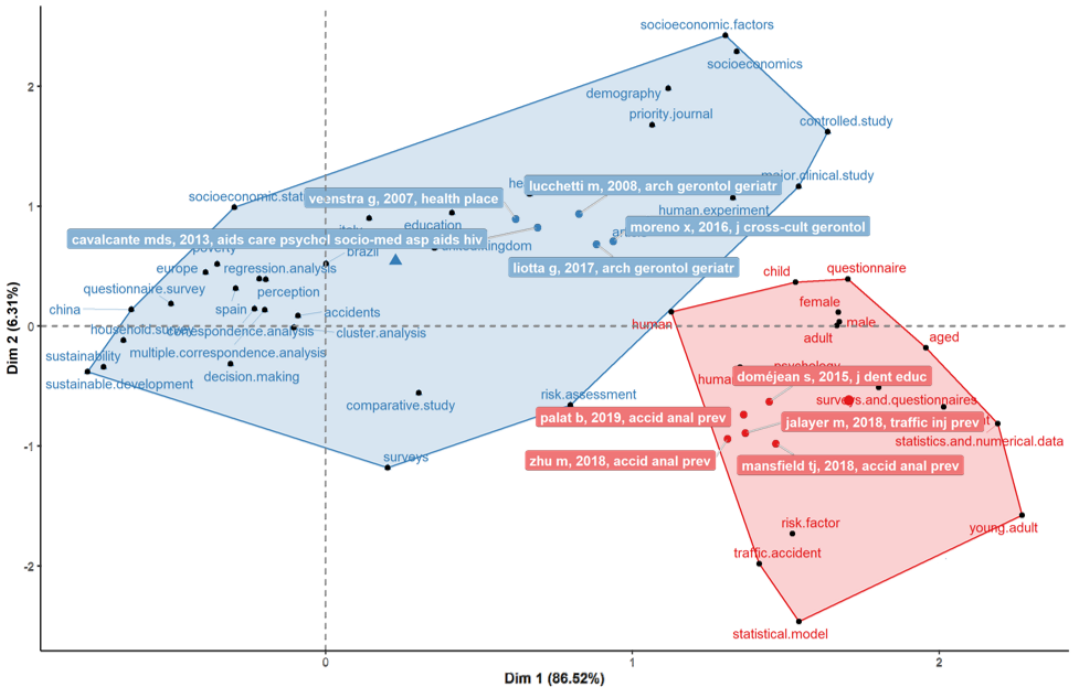

Figure 6: Conceptual map and most cited papers (tot. inertia 93%)

The red cluster (right side) on the other hand, contains themes such as age, male, female, child, psychology, traffic accident and risk factor, while the most cited articles relate to accident analysis (Jalayer et al., 2018; Mansfield et al., 2018; Palat et al., 2019; Zhu et al., 2018) and education (Doméjean et al., 2015).

6. Conclusion, Limitations and Future Research

This study has provided a review of the MCAHS literature spanning nearly four decades. The productivity ratio (e.g., total publication) and relevance (e.g., total citations, h-index) of authors, journals, and nations were assessed using several bibliometric indicators. The co-occurrence of keywords and thematic mapping of the scopus papers within the field under examination supplemented the results of the performance analysis. These analyses were carried out using R's Bibliometrix package (Aria and Cuccurullo, 2017), and allowed for the presentation of the conceptual and intellectual structures of the study field. It was demonstrated that while experts' interest in the issue has grown, it remains divided. Furthermore, despite having a small number of published publications, some experts are quite important in the field. Future investigations and systematization efforts should be encouraged in this regard. Thematically, mainstream topics seem to include health and accident analysis, while issues that need more development include questionnaire survey, European Union, and innovation. In terms of conceptual structure, and even though highly cited papers in both clusters come from health studies, two distinct topic clusters were found by keyword co-occurrence analysis: (i) sustainability, demography and education; and (ii) gender and age studies.

The purpose of this study is to reinforce the epistemology of bibliometric and structured network studies on MCAHS. This study's findings supplement the previous subjective, critical and reflective assessments of the literature on MCAHS This review article, like all research reviews, has limitations. First, the dataset was collected using Scopus search services to obtain high quality results. This, however, reduced the number of papers that could be analyzed. Second, when comparing various publications or authors, some indications might lead to discrepancies. As a result, each bibliometric parameter, such as the h-index, should be interpreted in conjunction with the others. Each of these constraints, in turn, creates opportunity for future study for researchers and practitioners interested in the growth of such promising future areas of research.

Bibliography

Adler, S., & Sarstedt, M. (2021). Mapping the jungle: A bibliometric analysis of research into construal level theory. *Psychology & Marketing*, *38*(9), 1367-1383. DOI: 10.1002/mar.21537

Agbo, F. J., Sanusi, I. T., Oyelere, S. S., & Suhonen, J. (2021). Application of virtual reality in computer science education: A systemic review based on bibliometric and content analysis methods. *Education Sciences*, *11*(3), 142. DOI: 10.3390/educsci11030142

Alhuzali, T., Beh, E. J., & Stojanovski, E. (2022). Multiple correspondence analysis as a tool for examining Nobel Prize data from 1901 to 2018. *Plos one*, *17*(4), e0265929. DOI: 10.1371/journal.pone.0265929

Altarturi, H. H., Saadoon, M., & Anuar, N. B. (2020). Cyber parental control: A bibliometric study. *Children and Youth Services Review*, *116*, 105134. DOI: 10.1016/j.childyouth.2020.105134

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of informetrics*, *11*(4), 959-975. DOI: 10.1016/j.joi.2017.08.007

Aria, M., Misuraca, M., & Spano, M. (2020). Mapping the evolution of social research and data science on 30 years of Social Indicators Research. *Social indicators research*, *149*(3), 803-831. DOI: 10.1007/s11205-020-02281-3

Bapte, V. D. (2021). Media Literacy: A Scientometric Study Based on Web of Science during 1989-2020. *DESIDOC Journal of Library & Information Technology*, *41*(4). DOI: 10.14429/djlit.41.4.16301

Beh, E. J., & Lombardo, R. (2014). *Correspondence analysis: Theory, practice and new strategies*. Chichester, England: Wiley.

Beh, E.J., & Lombardo, R. (2019), "Multiple and multiway correspondence analysis", Wiley Interdisciplinary Reviews: *Computational Statistics*, e1464. DOI: 10.1002/wics.1464

Bond, M., & Buntins, K. (2018). An analysis of the Australasian journal of educational technology 2013-2017. *Australasian journal of educational technology*, *34*(4). DOI: 10.14742/ajet.4359

Bond, M., Zawacki-Richter, O., & Nichols, M. (2019). Revisiting five decades of educational technology research: A content and authorship analysis of the British Journal of Educational Technology. *British journal of educational technology*, *50*(1), 12-63. DOI: 10.1111/bjet.12730

Callon, M., Courtial, J. P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social science information*, *22*(2), 191-235. DOI: 10.1177/053901883022002003

Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. *Scientometrics*, *22*(1), 155-205. DOI: 10.1007/bf02019280

Carammia, M. (2022). A bibliometric analysis of the internationalisation of political science in Europe. *European Political Science*, 1-32. DOI:10.1057/s41304-022-00367-9

Carlier, A., & Kroonenberg, P. M. (1996). Decomposition and biplots in three-way correspondence analysis. *Psychometrika*, 61, 355–373. DOI: 10.1007/BF02294344

Cavalcante, M. D. S., Kerr, L. R. F. S., Brignol, S. M. S., Silva, D. D. O., Dourado, I., Galvão, M. T. G., & Kendall, C. (2013). Sociodemographic factors and health in a population of children living in families infected with HIV in Fortaleza and Salvador, Brazil. *AIDS care*, *25*(5), 550-558. DOI: 10.1080/09540121.2012.726343

Chen, Y., & Silva, E. A. (2021). Smart transport: A comparative analysis using the most used indicators in the literature juxtaposed with interventions in English metropolitan areas. *Transportation research interdisciplinary perspectives*, *10*, 100371. DOI: 10.1016/j.trip.2021.100371

Cheng, B., Wang, M., Mørch, A. I., Chen, N. S., & Spector, J. M. (2014). Research on e-learning in the workplace 2000–2012: a bibliometric analysis of the literature. *Educational research review*, *11*, 56-72. DOI: 10.1016/j.edurev.2014.01.001

Choulakian, V. (2008). Multiple taxicab correspondence analysis. *Advances in Data Analysis and Classification,* 2, 177–206. DOI: 10.1007/s11634-008-0023-6

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for information Science and Technology*, *62*(7), 1382-1402. DOI: 10.1002/asi.21525

Cohen, J. (1988). Set correlation and contingency tables. *Applied psychological measurement*, *12*(4), 425-434. DOI: 10.1177/014662168801200410

De la Hoz-Correa, A., Muñoz-Leiva, F., & Bakucz, M. (2018). Past themes and future trends in medical tourism research: A co-word analysis. *Tourism Management*, *65*, 200-211. DOI: 10.1016/j.tourman.2017.10.001

Deng, S., Xia, S., Hu, J., Li, H., & Liu, Y. (2021). Exploring the topic structure and evolution of associations in information behavior research through co-word analysis. *Journal of Librarianship and Information Science*, *53*(2), 280-297. DOI: 10.1177/0961000620938120

Ding, Y., Chowdhury, G. G., & Foo, S. (2001). Bibliometric cartography of information retrieval research by using co-word analysis. *Information processing & management*, *37*(6), 817-842. DOI: 10.1016/S0306-4573(00)00051-0

Diodato, V. P., & Gellatly, P. (2013). Dictionary of bibliometrics. Routledge.

Doméjean, S., Léger, S., Rechmann, P., White, J. M., & Featherstone, J. D. (2015). How do dental students determine patients' caries risk level using the caries management by risk assessment (CAMBRA) system?. *Journal of dental education*, *79*(3), 278-285. DOI: 10.1002/j.0022-0337.2015.79.3.tb05882.x

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, *133*, 285-296. DOI: 10.1016/j.jbusres.2021.04.070

Dungey, M., Tchatoka, F. D., & Yanotti, M. B. (2018). Using multiple correspondence analysis for finance: A tool for assessing financial inclusion. *International Review of Financial Analysis*, *59*, 212-222. DOI: 10.1016/j.irfa.2018.08.007

Garfield, E. (1990). KeyWords Plus-ISI's breakthrough retrieval method. 1. Expanding your searching power on current-contents on diskette. *Current contents*, *32*, 5-9. DOI: 10.1016/j.irfa.2018.08.007

Gibadullina, A. (2021). The birth and development of Anglophone financial geography: A historical analysis of geographical studies of money and finance. *Geoforum*, *125*, 150-167. DOI: 10.1016/j.geoforum.2021.06.009

Gower, J. C. (1990). Fisher's optimal scores and multiple correspondence analysis. *Biometrics*, 46, 947–961. DOI: 10.2307/2532440

Greenacre, M. J., & Blasius, J. (2006). *Multiple correspondence analysis and related methods*. London, England: Chapman & Hall/CRC.

Grant, J., Cottrell, R., Cluzeau, F., & Fawcett, G. (2000). Evaluating "payback" on biomedical research from papers cited in clinical guidelines: applied bibliometric study. Bmj, 320(7242), 1107-1111. DOI: *10.1136/bmj.320.7242.1107*

Grivel, L., Mutschke, P., & Polanco, X. (1995). Thematic mapping on bibliographic databases by cluster analysis: A description of the SDOC environment with SOLIS. *Journal of Knowledge Organization*, *22*(2), 70-77.

Guraya, S. S., Guraya, S. Y., & Yusoff, M. S. B. (2021). Preserving professional identities, behaviors, and values in digital professionalism using social networking sites; a systematic review. *BMC medical education*, *21*(1), 1-12. DOI: 10.1186/s12909-021-02802-9

Han, L., Benseler, S. M., & Tyrrell, P. N. (2018). Cluster and multiple correspondence analyses in rheumatology: paths to uncovering relationships in a sea of data. *Rheumatic Disease Clinics*, *44*(2), 349-360. DOI: 10.1016/j.rdc.2018.01.013

Hjellbrekke, J. (2018). Multiple correspondence analysis for the social sciences. Routledge.

Ho, M. T., Mantello, P., Nguyen, H. K. T., & Vuong, Q. H. (2021). Affective computing scholarship and the rise of China: a view from 25 years of bibliometric data. *Humanities and Social Sciences Communications*, *8*(1), 1-14. DOI: 10.1057/s41599-021-00959-8

Husson, F., & Josse, J. (2014). Multiple correspondence analysis. In J. Blasius & M. Greenacre (Eds.), *Visualization and verbalization of data* (pp. 165–183). Boca Raton, FL: CRC Press. DOI: 10.1201/b16741-17

Hwang, H., Tomiuk, M. A., & Takane, Y. (2009). Correspondence analysis, multiple correspondence analysis and recent developments. *Handbook of quantitative methods in psychology*, 243-263.

Jalayer, M., Pour-Rouholamin, M., & Zhou, H. (2018). Wrong-way driving crashes: A multiple correspondence approach to identify contributing factors. *Traffic injury prevention*, *19*(1), 35-41. DOI: 10.1080/15389588.2017.1347260

Jensen, M. R., & Moses, J. W. (2021). The state of political science, 2020. *European Political Science*, *20*(1), 14-33. DOI: 10.1057/s41304-020-00297-4

Jing, P., Pan, K., Yuan, D., Jiang, C., Wang, W., Chen, Y., ... & Xie, J. (2021). Using bibliometric analysis techniques to understand the recent progress in school travel research, 2001–2021. *Journal of Transport & Health*, *23*, 101265. DOI: 10.1016/j.jth.2021.101265

Kamalja, K. K., & Khangar, N. V. (2017). Multiple Correspondence Analysis and its applications. *Electronic Journal of Applied Statistical Analysis*, *10*(2), 432-462.

Kroonenberg, P. M. (2014). History of multiway component analysis and three-way correspondence analysis. In J. Blasius & M. Greenacre (Eds.), *Visualization and verbalization of data* (pp. 77–93). London, England: CRC Press.

Latocha, A. (2020). Modern transformation of deserted settlements in the Sudetes Mountains, SW Poland. *GeoScape*, *14*(2). DOI: 10.2478/geosc-2020-0008

Le Roux, B., & Rouanet, H. (2004). Geometric data analysis: from correspondence analysis to structured data analysis. Springer Science & Business Media.

Leung, X. Y., Sun, J., & Bai, B. (2017). Bibliometrics of social media research: A co-citation and co-word analysis. *International Journal of Hospitality Management*, *66*, 35-45. DOI: 10.1016/j.ijhm.2017.06.012

Linnenluecke, M. K., Marrone, M., & Singh, A. K. (2020). Conducting systematic literature reviews and bibliometric analyses. *Australian Journal of Management*, *45*(2), 175-194. DOI: 10.1177/0312896219877678

Liotta, G., O'Caoimh, R., Gilardi, F., Proietti, M. G., Rocco, G., Alvaro, R., ... & Marazzi, M. C. (2017). Assessment of frailty in community-dwelling older adults residents in the Lazio region (Italy): A model to plan regional community-based services. *Archives of Gerontology and Geriatrics*, *68*, 1-7. DOI: 10.1016/j.archger.2016.08.004

Liu, Z., Moon, J., Kim, B., & Dai, C. P. (2020). Integrating adaptivity in educational games: A combined bibliometric analysis and meta-analysis review. *Educational technology research and development*, *68*(4), 1931-1959. DOI: 10.1007/s11423-020-09791-4

López-Fernández, M. C., Serrano-Bedia, A. M., & Pérez-Pérez, M. (2016). Entrepreneurship and family firm research: A bibliometric analysis of an emerging field. *Journal of Small Business Management*, *54*(2), 622-639. DOI: 10.1111/jsbm.12161

Lucchetti, M., Corsonello, A., & Gattaceca, R. (2008). Environmental and social determinants of aging perception in metropolitan and rural areas of Southern Italy. *Archives of Gerontology and Geriatrics*, *46*(3), 349-357. DOI: 10.1016/j.archger.2007.05.009

Macheridis, S. (2017). The use of multiple correspondence analysis (MCA) in taphonomy: The case of Middle Helladic Asine, Greece. *International Journal of Osteoarchaeology*, *27*(3), 477-487. DOI; 10.1002/oa.2571

Mair, L., Mill, A. C., Robertson, P. A., Rushton, S. P., Shirley, M. D., Rodriguez, J. P., & McGowan, P. J. (2018). The contribution of scientific research to conservation planning. *Biological Conservation*, *223*, 82-96. DOI: 10.1016/j.biocon.2018.04.037

Mansfield, T. J., Peck, D., Morgan, D., McCann, B., & Teicher, P. (2018). The effects of roadway and built environment characteristics on pedestrian fatality risk: A national assessment at the neighborhood scale. *Accident Analysis & Prevention*, *121*, 166-176. DOI: 10.1016/j.aap.2018.06.018

Maretti, M., Tontodimamma, A., & Biermann, P. (2019). Environmental and climate migrations: an overview of scientific literature using a bibliometric analysis. *International Review of Sociology*, *29*(2), 142-158. DOI: 10.1080/03906701.2019.1641270

Markos, A., Menexes, G., & Papadimitriou, T. (2009). Multiple correspondence analysis for "tall" data sets. *Intelligent Data Analysis*, 13, 873–885. DOI: 10.3233/IDA-2009-0398

Martin-Mazé, M. (2018). Multiple Correspondence Analysis in International Relations. In *Resources and Applied Methods in International Relations* (pp. 139-149). Palgrave Macmillan, Cham. DOI: 10.1007/978-3-319-61979-8_10

Medie, P. A., & Kang, A. J. (2018). Power, knowledge and the politics of gender in the Global South. *European Journal of Politics and Gender*, *1*(1-2), 37-54. DOI: 10.1332/251510818X15272520831157

Meyer, N., Ferlicot, S., Vieillefond, A., Peyromaure, M., & Vielh, P. (2004). Contribution of multiple correspondence analysis in histopathology. In *Annales de pathologie* (Vol. 24, No. 2, pp. 149-160). DOI: 10.1016/s0242-6498(04)93938-7

Moreno, X., Sánchez, H., Huerta, M., Albala, C., & Márquez, C. (2016). Social representations of older adults among Chilean elders of three cities with different historical and sociodemographic background. *Journal of cross-cultural gerontology*, *31*(2), 115-128. DOI: 10.1007/s10823-016-9288-y

Mori, Y., Kuroda, M., & Makino, N. (2016). Multiple correspondence analysis. In *Nonlinear principal component analysis and its applications* (pp. 21-28). Springer, Singapore.

Palat, B., Saint Pierre, G., & Delhomme, P. (2019). Evaluating individual risk proneness with vehicle dynamics and self-report data- toward the efficient detection of At-risk drivers. *Accident Analysis & Prevention*, *123*, 140-149. DOI: 10.1016/j.aap.2018.11.016

Patel, P., & Bhatt, A. (2021). Growth and impact of scholarly contributions for SP University: A Bibliometric Analysis. *Library Philosophy and Practice (e-journal)*, *20*.

Peez, A. (2022). Contributions and blind spots of constructivist norms research in international relations, 1980–2018: A systematic evidence and gap analysis. *International Studies Review*, *24*(1), viab055. DOI: 10.1093/isr/viab055

Pham-Duc, B., Tran, T., Trinh, T. P. T., Nguyen, T. T., Nguyen, N. T., & Le, H. T. T. (2022). A spike in the scientific output on social sciences in Vietnam for recent three years: Evidence from bibliometric analysis in Scopus database (2000–2019). *Journal of Information Science*, *48*(5), 623-639. DOI: 10.1177/0165551520977447

Pierreval, H. (1994). Using multiple correspondence analysis in the analysis of simulation experiments: a study of dynamic scheduling strategies. *International Transactions in Operational Research*, *1*(2), 147-157. DOI: 10.1016/0969-6016(94)90016-7

Radhakrishnan, S., Erbis, S., Isaacs, J. A., & Kamarthi, S. (2017). Novel keyword co-occurrence network-based methods to foster systematic reviews of scientific literature. *PloS one*, *12*(3), e0172778. DOI: 10.1371/journal.pone.0172778

Ravikumar, S., Agrahari, A., & Singh, S. N. (2015). Mapping the intellectual structure of scientometrics: A co-word analysis of the journal Scientometrics (2005–2010). *Scientometrics*, *102*(1), 929-955. DOI: 10.1007/s11192-014-1402-8

Rodrigues, S. P., Van Eck, N. J., Waltman, L., & Jansen, F. W. (2014). Mapping patient safety: a large-scale literature review using bibliometric visualisation techniques. *BMJ open*, *4*(3), e004468. DOI: 10.1007/s11192-014-1402-8

Rosas-Chavoya, M., Gallardo-Salazar, J. L., López-Serrano, P. M., Alcántara-Concepción, P. C., & León-Miranda, A. K. (2022). QGIS a constantly growing free and open-source geospatial software contributing to scientific development. *Cuadernos de Investigación Geográfica*, *48*(1), 197-213. DOI: 10.18172/cig.5143

Rozylowicz, L., Nita, A., Manolache, S., Ciocanea, C. M., & Popescu, V. D. (2017). Recipe for success: A network perspective of partnership in nature conservation. *Journal for Nature Conservation*, *38*, 21-29. DOI: 10.1016/j.jnc.2017.05.005

Savage, M., Devine, F., Cunningham, N., Taylor, M., Li, Y., Hjellbrekke, J., ... & Miles, A. (2013). A new model of social class? Findings from the BBC's Great British Class Survey experiment. Sociology, 47(2), 219-250. DOI: 10.1177/0038038513481128

Shiau, W. L., Dwivedi, Y. K., & Yang, H. S. (2017). Co-citation and cluster analyses of extant literature on social networks. *International Journal of Information Management*, *37*(5), 390-399. DOI: 10.1016/j.ijinfomgt.2017.04.007

Schneider, S., & Bengler, K. (2020). Virtually the same? Analysing pedestrian behaviour by means of virtual reality. *Transportation research part F: traffic psychology and behaviour*, *68*, 231-256. DOI: 10.1016/j.trf.2019.11.005

Song, Y., et al. (2019). Exploring two decades of research on classroom dialogue by using bibliometric analysis. Computers in Education, 137, 12–31. DOI: 10.1016/j.compedu.2019.04.002

Tenenhaus, M., & Young, F. W. (1985). An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis and other methods for quantifying categorical multivariate data. *Psychometrika*, 50, 91–119. DOI: 10.1007/BF02294151

Tontodimamma, A., Nissi, E., Sarra, A., & Fontanella, L. (2021). Thirty years of research into hate speech: topics of interest and their evolution. *Scientometrics*, *126*(1), 157-179. DOI: 10.1007/s11192-020-03737-6

Van der Heijden, P. G., Teunissen, J., & van Orle, C. (1997). Multiple correspondence analysis as a tool for quantification or classification of career data. *Journal of Educational and Behavioral Statistics*, *22*(4), 447-477. DOI: 10.3102/10769986022004447

Van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *scientometrics*, *84*(2), 523-538. DOI: 10.1007/s11192-009-0146-3

Van Nunen, K., Li, J., Reniers, G., & Ponnet, K. (2018). Bibliometric analysis of safety culture research. *Safety Science*, *108*, 248-258. DOI: 10.1016/j.ssci.2017.08.011

Veenstra, G. (2007). Social space, social class and Bourdieu: Health inequalities in British Columbia, Canada. *Health & place*, *13*(1), 14-31. DOI: 10.1016/j.healthplace.2005.09.011

Waheed, H., Hassan, S. U., Aljohani, N. R., & Wasif, M. (2018). A bibliometric perspective of learning analytics research landscape. *Behaviour & Information Technology*, *37*(10-11), 941-957. DOI: 10.1016/j.healthplace.2005.09.011

Wang, Z. Y., Li, G., Li, C. Y., & Li, A. (2012). Research on the semantic-based co-word analysis. *Scientometrics*, *90*(3), 855-875. DOI: 10.1007/s11192-011-0563-y

Xie, H., Zhang, Y., Wu, Z., & Lv, T. (2020). A bibliometric analysis on land degradation: Current status, development, and future directions. Land, 9(1), 28. DOI: 10.3390/land9010028

Xu, J., Lu, W., Xue, F., Chen, K., Ye, M., Wang, J., & Chen, X. (2018). Cross-boundary collaboration in waste management research: A network analysis. *Environmental Impact Assessment Review*, *73*, 128-141. DOI: 10.1016/j.eiar.2018.08.005

Youngblood, M., & Lahti, D. (2018). A bibliometric analysis of the interdisciplinary field of cultural evolution. *Palgrave Communications*, *4*(1), 1-9. DOI: 10.1057/s41599-018-0175-8

Zhu, M., Li, Y., & Wang, Y. (2018). Design and experiment verification of a novel analysis framework for recognition of driver injury patterns: From a multi-class classification perspective. *Accident Analysis & Prevention*, *120*, 152-164. DOI: 10.1016/j.aap.2018.08.011

ΤΕΤΡΑΔΙΑ ΑΝΑΛΥΣΗΣ ΔΕΔΟΜΕΝΩΝ | DATA ANALYSIS BULLETIN, 20(1), 59-77

ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ | REVIEW PAPER

Η Πολλαπλή Παραγοντική Ανάλυση στις Κοινωνικές και Ανθρωπιστικές Επιστήμες: Μια Διαχρονική Χαρτογράφηση

Νικόλαος Κουτσουπιάς

Τμήμα Διεθνών & Ευρωπαϊκών Σπουδών / Σχολή Κοινωνικών, Ανθρωπιστικών Επιστημών και Τεχνών, Πανεπιστήμιο / 'Πανεπιστήμιο Μακεδονίας

KEYWORDS IN GREEK

Πολλαπλή Παραγοντική Ανάλυση
Βιβλιομετρία
Επιστημομετρία
Αναλυση Δεδομένων

CORRESPONDENCE

Νικόλαος Κουτσουπιάς
Πανεπιστήμιο Μακεδονίας
Εγνατία 156, Θεσσαλονίκη
nk@uom.gr

ABSTRACT IN GREEK

Η συγκεκριμένη διερευνητική εργασία χρησιμοποιεί βιβλιομετρικές μεθόδους για την επισκόπηση της περιοχής έρευνας που αξιοποιεί την Πολλαπλή Παραγοντική Ανάλυση Αντιστοιχιών (MCA) σε διάφορους τομείς από το 1983 έως το 2021. Τα ευρήματα περιλαμβάνουν τα πιο σημαντικά κανάλια δημοσίευσης, συγγραφείς, εργασίες και θεματικές. Συνολικά, εξετάστηκαν 1208 άρθρα από τις Κοινωνικές, Ανθρωπιστικές Επιστήμες και τις Τέχνες που δημοσιεύτηκαν σε 662 περιοδικά. Εκτός από άλλα ευρήματα, εντοπίσαμε 2962 διακριτούς συγγραφείς που συνέβαλαν στην ανάπτυξη αυτού του ερευνητικού πεδίου, με ετήσιο ρυθμό αύξησης 15,07%. Στο παρόν άρθρο παρέχεται μια ευρεία επισκόπηση των διαφορετικών προοπτικών για την έρευνα που βασίζεται στην MCA διαχρονικά και μπορεί να αποτελέσει αφετηρία έρευνας όσων ενδιαφέρονται για αυτόν τον τομέα μελέτης.