ANADOLU ÜNİVERSİTESİ
Cilt: 25, Sayı: 3, Yıl: 2025
e-ISSN: 2667-8683
AÜSBD
Anadolu Üniversitesi Sosyal Bilimler Dergisi
Anadolu University Journal of Social Sciences

Bibliometric Analysis of Gender Inequality Research Using Data Mining Techniques: Trends, Key Insights, and Future Directions

Nesibe MANAV MUTLU[1] Polathan KÜSBECİ[2]

Başvuru Tarihi/Submitted: 02.12.2024 **Kabul Tarihi/Accepted:** 15.04.2025
Makale Türü/Article Type: Araştırma Makalesi/Research Article

Abstract

This bibliometric analysis study aims to explore the intersection of gender inequality and artificial intelligence research by uncovering key contemporary trends, influential works, and emerging themes in this interdisciplinary field. To achieve this, the study employed content and bibliometric analysis methods to examine 5,074 academic publications indexed in the Scopus database. These publications focus on gender inequality and AI-related areas such as data mining, machine learning, and predictive modeling. By using tools such as VOSviewer and Python-based analytical techniques, the study identified thematic trends, methodological approaches, and interdisciplinary patterns across various sectors including education, healthcare, and the workplace. The analysis also revealed significant gaps and evolving directions in the literature, offering a comprehensive view of how data-driven methods have been applied to understand and address gender inequality. By focusing on the most cited publications, prominent authors, and international collaborations, the analysis provides a comprehensive assessment of the current state of the field. Furthermore, by identifying thematic clusters and research gaps, the study sheds light on the evolving approaches to addressing gender inequality using modern data-driven methods. This research contributes to the growing body of literature that seeks to harness data science for social good and to promote a deeper understanding of gender-related challenges in contemporary societies. In addition, it addresses the relationship between gender theories and computational methodologies, particularly the intersection of gender perspectives with data mining and artificial intelligence.

Keywords: *Bibliometric analysis, gender inequality, artificial intelligence, data mining*

JEL Code: *J16, J71, C55, C88*

1 İstanbul Nişantaşı Üniversitesi, İktisadi, İdari ve Sosyal Bilimler Fakültesi Yönetim Bilişim Sistemleri, Veri Bilimi, Dr., nesibe.manavm@gmail.com, ORCID: 0000-0002-7853-6337
2 İstanbul Nişantaşı Üniversitesi, İktisadi, İdari ve Sosyal Bilimler Fakültesi Yönetim Bilişim Sistemleri, Dr., scholar@polathan.com, ORCID: 0000-0002-4858-3853

ANADOLU ÜNİVERSİTESİ
Cilt: 25, Sayı: 3, Yıl: 2025
e-ISSN: 2667-8683
AÜSBD
Anadolu Üniversitesi Sosyal Bilimler Dergisi
Anadolu University Journal of Social Sciences

Veri Madenciliği Tekniklerini Kullanarak Cinsiyet Eşitsizliği Araştırmalarının Bibliyometrik Analizi: Trendler, Temel Görüşler ve Gelecekteki Yönler

Öz

Bu bibliyometrik analiz çalışması, toplumsal cinsiyet eşitsizliği ile yapay zekâ araştırmalarının kesişim noktasını inceleyerek, bu disiplinler arası alandaki güncel eğilimleri, etkili çalışmaları ve ortaya çıkan temaları ortaya koymayı amaçlamaktadır. Bu doğrultuda, Scopus veri tabanında indekslenen, toplumsal cinsiyet eşitsizliği ve yapay zekâyla ilişkili alanlara (veri madenciliği, makine öğrenmesi, kestirimsel modelleme vb.) odaklanan 5.074 akademik yayın, içerik ve bibliyometrik analiz yöntemleriyle incelenmiştir. VOSviewer yazılımı ve Python tabanlı analiz araçları kullanılarak, eğitim, sağlık ve iş dünyası gibi çeşitli sektörlerde tematik eğilimler, yöntemsel yaklaşımlar ve disiplinlerarası örüntüler belirlenmiştir. Yapılan analizler, literatürdeki temel boşlukları ve gelişen araştırma yönelimlerini ortaya koyarak, veri odaklı yöntemlerin toplumsal cinsiyet eşitsizliğini anlama ve çözme sürecindeki rolüne dair kapsamlı bir çerçeve sunmaktadır. Analiz, en çok atıf alan makalelere, etkili yazarlara ve ülkesel iş birliklerine odaklanarak, alanın mevcut durumunun kapsamlı bir değerlendirmesini sunmaktadır. Dahası, tema kümelerini ve araştırma boşluklarını belirleyerek, modern veri odaklı yöntemleri kullanarak cinsiyet eşitsizliğiyle mücadele için geçmişten günümüze değişen yaklaşımlara ışık tutmaktadır. Bu çalışma, veri bilimini toplumsal fayda için kullanmayı amaçlayan ve modern toplumlarda cinsiyetle ilgili zorlukların daha derin bir şekilde anlaşılmasını teşvik eden büyüyen literatüre katkıda bulunmaktadır. Ayrıca, toplumsal cinsiyet kuramları ve veri madenciliği, yapay zeka arasındaki ilişkiyi ele almaktadır.

Anahtar Kelimeler: *Bibliyometrik analiz, cinsiyet eşitsizliği, yapay zeka, veri madenciliği*

JEL Kodu: *J16, J71, C55, C88*

Introduction

Gender gap

Measurable disparities between men and women in a range of social, economic, political, and cultural areas, including political representation, labor force participation, income levels, educational achievement, and healthcare access, are referred to as the "gender gap" (World Economic Forum, 2023: 5). It is a quantitative idea that draws attention to gender differences in results and frequently acts as a sign of more pervasive structural injustices. The gender gap is conceptually and causally related to gender inequality, which indicates the systematic and institutionalized discrimination that causes these differences, even while it offers concrete evidence of differing experiences and opportunities depending on gender (Ridgeway, 2011: 102). To put it another way, gender gaps are the visible signs or effects of gender inequality.

Gender in/equality

The term “gender inequality” describes how people are treated differently depending just on their gender, which leads to differences in opportunities and resources in the social, political, economic, and cultural spheres (Demirgöz Bal, 2014). This issue is especially noticeable in areas like pay discrepancies, academic visibility differences, institutional impediments to women's job advancement, and restricted access to leadership roles (UNDP, 2020; Holman, Stuart-Fox, & Hauser, 2018). One example of a persistent gender-based imbalance in knowledge production processes is the fact that women are still much underrepresented in scientific publishing compared to men (Cameron et al., 2023). Moreover, interdisciplinary comparative analyses show that these disparities have systemic roots rather than being purely individual (Huang et al., 2020).

Conversely, gender equality is the absence of any difference between women and men regarding social and cultural rights. This issue, which has been discussed for a long time, has led to many changes in the social position of women with the development of human civilization and productive forces. In modern societies, with the rapid advancement of productive forces, gender equality has become an important trend in the development of social civilization. In today's knowledge economy, mental labor stands out as one of the main factors that increase productivity by replacing manual work. Women can compensate for the lack of physical labor by developing their mental abilities (Peng et al., 2024: 1). Gender equality is a key component of sustainable development (Singh, 2023: 6210). In recent years, gender equality in management has received increasing attention as organizations recognize the importance of diversity and inclusion to achieve sustainable success. Despite progress in some areas, women are often underrepresented and face unique challenges in the workplace. Despite advances in gender equality initiatives, women continue to face systemic barriers, such as limited access to management networks, lack of mentoring opportunities, and entrenched gender norms that constrain their career progression (Irfan and Sharif, 2024: 64).

Disciplines such as history, anthropology, and sociology have been used more frequently to highlight and demonstrate the vast differences between the sexes. Gender discrimination is a social problem that has recently become notorious. It is a situation that affects women in areas such as access to employment, health, education, politics, and causes a significant difference in life choices and chances between women and men. It is fair to consider the elimination of this difference not only from a social justice perspective but also to consider that the elimination of these differences contributes to a more developed world (Dominguez et al., 2019: 242). Gender equality actions are most effective when they are implemented synergistically. Various fields of literature—such as business, economics, sociology, and psychology—offer different perspectives on the problem of achieving gender equality, often from specific and limited perspectives (Belingheri et al., 2021: 2). Inequality is considered one of the world's most serious social problems in many ways. Recently, society's attention has increasingly turned

to the issue of gender inequality. Indeed, women have difficulty asserting themselves in the workplace and are perceived as less productive than their male counterparts. This issue has been highlighted by various initiatives at national and international levels and is supported by many organizations such as the United Nations (UN) and the Global Reporting Initiative (GRI) (García-Sánchez et al., 2022: 385).

Gender equality is one of the fundamental elements of innovation. The participation of women in scientific and technological fields is increasing. Advancing gender equality is an important agenda item in terms of technological innovations and socio-economic outcomes. In addition, work-life balance is an effective way to support women's participation and performance in science and technology. This balance provides individuals with opportunities to realize their potential. Work-life balance also motivates employees, encouraging self-learning and innovation, thus increasing productivity. In particular, R&D (research and development) organizations must maintain a high level of work-life balance to improve their performance (Ko et al., 2021: 148). Achieving gender equality in the workplace in a truly sustainable way is a matter of utmost importance, given the current resistance and concerns that prevent women from advancing in the workplace structure. The challenges of promoting gender equality have become a priority in research in the field of social and corporate sustainability and the strategic plans of international institutions (Gartzia, 2021: 1). In developed countries, gender diversity is generally seen to have positive effects on innovation. In particular, gender diversity facilitates innovation by providing a wider range of knowledge and more effective decision-making processes (Ritter-Hayashi et al., 2019: 1). Knowledge plays a critical role in advancing gender equality in the workplace. Therefore, training and various experiences should be provided to increase employees' awareness of gender-related issues. Each needs to be aware that they can achieve equal opportunities in all areas of development, regardless of biological differences (Farsia, 2024: 25).

Gender equality is recognized as one of the fundamental components of sustainable development. However, current data and research reveal that significant progress is still needed on a global scale. According to the World Bank (2024), global inequalities persist in women's participation in the labor force, access to educational opportunities, and economic empowerment. Ensuring the full and effective participation of women in economic development has the potential to significantly enhance countries' overall development levels (World Bank, 2024).

The OECD's report The Pursuit of Gender Equality (2017) highlights ongoing inequalities in women's participation in the workforce and gender pay gaps. The report offers recommendations for policymakers and emphasizes that gender inequality in the labor market negatively impacts economic growth (OECD, 2017).

An IMF working paper titled Trends in Gender Equality and Women’s Advancement (Stotsky et al., 2016) underlines the positive macroeconomic effects of increasing gender equality. The study states that when gender equality is achieved, women's economic participation and productivity increase, contributing to countries' sustainable development goals.

Research specific to Turkey also provides valuable insights. Toksöz and Memiş (2018) conducted a comprehensive mapping and monitoring study on gender equality in employment in Turkey. Their research found that gender inequalities vary by region and that female employment rates remain low. Among their policy recommendations, incentive mechanisms to increase women’s participation in the labor market are particularly emphasized.

Data mining

Although data mining is a relatively new field, it has evolved rapidly since the 1990s and has made great progress (Xin, 2021: 2). Data mining, which plays a critical role in the process of knowledge discovery in databases, is used to uncover useful and previously unnoticed patterns from large data sets. This field includes various functions, techniques, and algorithms to identify and analyze interesting and valuable patterns from large data pools. Data

mining is a process that detects, analyzes, and applies discovery algorithms to data for specific patterns within acceptable computational efficiency limits (Gupta and Chandra, 2020: 1243). Data mining can be approached from both a technology and business perspective. From a technical perspective, data mining refers to the use of various technologies to extract interesting and meaningful information from large databases and data warehouses. This process presents the extracted information in the form of concepts, rules, laws, and patterns. In the business world, data mining is seen as a new generation of business analysis and processing technologies. This technology, which is used to discover hidden information and relationships in large data sets, helps decision-makers find possible connections between data and detect overlooked factors. Such information and factors are of great importance for predicting trends and guiding decision-making processes (Li and Zhang, 2022: 7). Data mining is the process of discovering or extracting data patterns that will help identify common errors (Acharya et al., 2015: 430).

Recently, the term data mining has been gaining more and more attention in both academic and popular circles, especially on social media and online platforms. Data mining involves the use of electronic methods in the process of extracting information from data to improve organizational performance. Unlike the physical tools used in traditional mining (e.g. pickaxes and shovels), data mining relies on statistical algorithms and analytical tools that automatically analyze data (Bitzenis et al., 2023: 1). Data mining is the process of automatic data extraction, processing, and modeling using a set of methods and techniques. This process is usually based on a structured methodology that specifies inputs, outputs, tasks, and how to execute these tasks to achieve the project's goals (Plotnikova et al., 2023: 141). Data mining is a widely used analysis method to obtain necessary or useful information from large data pools (Križanić, 2020: 1). The two main features of data mining are that it can deal effectively with large data sets and that it can reveal hidden problems with unknown relationships from data sources (Davari et al., 2019: 49).

Data mining technology offers exciting research opportunities due to its potential to serve a wide range of commercial applications. For example, it provides great benefits in areas such as supporting business decision-making processes and determining market strategies. Companies dealing with large data streams are in great demand for data mining applications that will allow this technology to develop and improve rapidly (Yan and Liu, 2019: 77). Data mining involves important steps such as data transformation, preprocessing, mining, and then data analysis in the process of knowledge discovery in databases. This process attempts to understand an event using data manipulation, statistical analysis, synthesis, and reasoning methods and evaluates current data to guide and predict future events, thus helping to solve practical problems (Bowen, 2021: 441). Many businesses and institutions attach great importance to data mining technology to effectively uncover valuable information in data. This technology aims to save manpower and material resources by using it in business processes. The idea of integrating data mining technology into businesses has been on the agenda for a long time, but today the huge amount of data supporting this process makes a difference. In short, data mining aims to help users optimize their experience-based decisions by providing automatic support (Zhen, 2019: 136).

Artificial intelligence

The design and execution of systems that can replicate human-like cognitive capabilities is the focus of the interdisciplinary discipline of artificial intelligence (AI), which sits at the nexus of computer science, statistics, and cognitive science (Russell & Norvig, 2021: 28-33). Complex operations like learning, reasoning, problem-solving, and natural language processing may be automated by these systems. Nilsson (1998: 4) defines artificial intelligence (AI) as the creation of "intelligent agents" that are able to sense their surroundings and act appropriately in response to those perceptions. As a potent decision-support tool, artificial intelligence (AI) has been used more and more in recent years in a variety of social sectors, such as public policy, healthcare, education, and workforce planning.

While data mining usually focuses on examining and analyzing vast amounts of data to facilitate knowledge discovery, artificial intelligence (AI) seeks to build systems that are capable of making choices on their own. In reality, AI systems commonly employ data mining techniques like network analysis, association rule learning, and clustering to organize inputs, find pertinent variables, and reveal latent trends.

However, the kind and caliber of input data are crucial for both data mining and artificial intelligence. The use of historical datasets, many of which have social biases ingrained in them, has sparked grave worries about how algorithmic systems may reinforce current disparities. The incorporation of historical and systematic prejudice into algorithmic decision-making presents serious ethical issues, especially when considering gender disparity. For example, it has been demonstrated that algorithms used in hiring disadvantage women, and AI-based healthcare systems frequently perform poorly for women because they are underrepresented in training data (Dastin, 2018; Buolamwini & Gebru, 2018: 8-9). In many situations, data mining techniques—while effective in identifying patterns—may unintentionally strengthen prejudiced systems if used carelessly.

Furthermore, the results of AI systems are strongly impacted when women and gender minorities are underrepresented or stereotyped in datasets. This undercuts attempts to create gender-sensitive policies and fair technical design in addition to perpetuating digital inequality (Crawford, 2021: 113-118). The growing use of AI and data mining in high-stakes social applications necessitates the incorporation of ethical governance and gender-sensitive principles across all phases of system development, from data collection and preprocessing to model training and deployment. In addition to being a technological need, doing so is a crucial commitment to promoting social justice and fairness in the digital era.

This study highlights the significance of applying advanced data mining techniques to address the persistent issue of gender inequality, offering a novel perspective at the intersection of computational methods and social sciences. Its originality lies in conducting a comprehensive bibliometric analysis that maps the evolving research landscape, uncovering hidden patterns, trends, and collaborative networks often overlooked in traditional studies. By systematically identifying key themes, influential contributors, and underexplored areas, the study contributes both theoretically—by integrating data-driven approaches with gender studies—and practically—by providing evidence-based insights for policymakers and researchers. Addressing a notable research gap, it bridges the lack of interdisciplinary analyses combining data mining with gender inequality research, offering a foundation for future studies and advancing scholarly discourse in this vital area.

Research Method

In this bibliographic analysis study, content analysis method was used in line with the purpose of the research. In order to determine the basic universe of the study, academic publications in the Scopus database were scanned by selecting "all fields" and a sample was created based on the keyword search "("Gender Inequality" OR "Gender Gap") AND ("Data Mining" OR "Machine Learning" OR "Artificial Intelligence")". In the sample selection process, the type of the selected publications (article, thesis, conference proceedings, etc.), the years they were published, and their interdisciplinary scope were taken into consideration. Scopus was selected as the primary database due to its comprehensive coverage, inclusion of peer-reviewed literature, and high-quality standards.

In the data collection phase, a systematic scan was performed using keywords related to the research topic. A total of 5,074 works were downloaded and examined with their information in Excel format. The scan results were examined in terms of fields such as keywords, publication fields, country associations, and publications suitable for the study were selected. Due to the very limited number of publications available before 2017, data from earlier years were not included in the scope of this study, as they were not sufficient to support a comprehensive and reliable analysis.

In the analysis of the data, parameters such as thematic distribution of publications, methods used, theoretical approaches, frequency analysis, and change over the years were considered. In the analysis process, data visualization and network analysis were performed using the bibliometric software VOSviewer and open-source Python codes for statistical analysis results provided by Scopus. These methods have been effective in revealing trends, gaps and relationships in the literature. Due to the nature of the study, informed consent or ethics committee approval was not required.

Findings and Discussion

The results of the analysis of the dataset obtained from Scopus according to years are shared in Table 1 below. Accordingly, from 2017 to 2020, there is a consistent increase in the number of publications each year, starting at 129 in 2017 and rising to 451 in 2020. This suggests a steady growth in research output during these years.

Table 1

Document Count by Year

Year	Number of Documents
2017	129
2018	205
2019	312
2020	451
2021	795
2022	1002
2023	1405
2024	1673
2025	14

2021 to 2024, a notable surge occurs during this period. The number of publications increases from 795 in 2021 to 1,673 in 2024, showing rapid growth. This could be due to a variety of factors such as increased research funding, more researchers entering the field, or greater interest in the topic.

2025 projection shows a decrease to 14 documents. This could be due to the dataset not yet being fully updated, as we are currently in 2024.

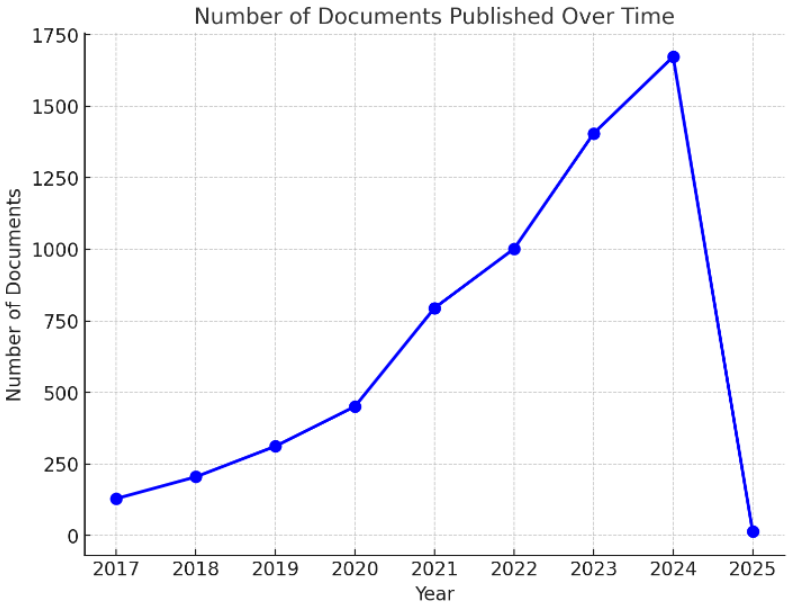

Figure 1. Number of documents published over time

From 2017 to 2024, the overall trend is a steady and significant increase in research output, particularly in 2021-2024. This reflects growing interest or advancements in the field. The year-on-year increase, especially between 2021 and 2024, suggests a sharp rise in the research being published, which could indicate emerging trends, new discoveries, or increased institutional support for research. The sudden drop for 2025 is likely due to incomplete data.

The period from 2021 to 2024 shows a strong increase in academic activity. This pattern reflects a period of significant research growth in the years leading up to the current period. The provided Figure 1 displays the number of documents published in various years, from 2017 to 2025.

The data suggests a strong emphasis on technological, sustainability, and interdisciplinary frameworks in gender inequality research. Journals and conferences in these domains are key contributors, underscoring the role of computational methods and bibliometric analysis in advancing the field can be seen in Table 2.

Table 2

Distribution of Documents

Source	**Number of Documents**
Lecture Notes in Computer Science (LNCS)	158
Sustainability Switzerland 119	119
PLOS One 82	82
ACM International Conference Proceeding Series	80
Conference on Human Factors in Computing Systems	52
Proceedings of the ACM on Human-Computer Interaction	48

According to Table 2, Lecture Notes in Computer Science (LNCS), with 158 documents, LNCS emerges as the most significant source, indicating a strong contribution from computer science and related fields. This aligns with the study's emphasis on bibliometric and data mining techniques, which are core areas in computer science.

Interdisciplinary Contributions from Sustainability: Sustainability Switzerland ranks second with 119 documents, reflecting a growing interest in applying data-driven techniques to societal challenges, such as gender inequality, through a sustainability lens.

Impact of Multidisciplinary Journals (PLOS One): PLOS One's 82 documents suggest the relevance of the topic across a broad range of disciplines, highlighting its appeal in health, social sciences, and computational studies.

Focus on Human-Computer Interaction (HCI): The ACM International Conference Proceedings Series, Conference on Human Factors in Computing Systems, and Proceedings of the ACM on Human-Computer Interaction collectively contribute a substantial number of documents (180 in total). This indicates the critical role of HCI and computational approaches in analyzing and addressing gender-related issues.

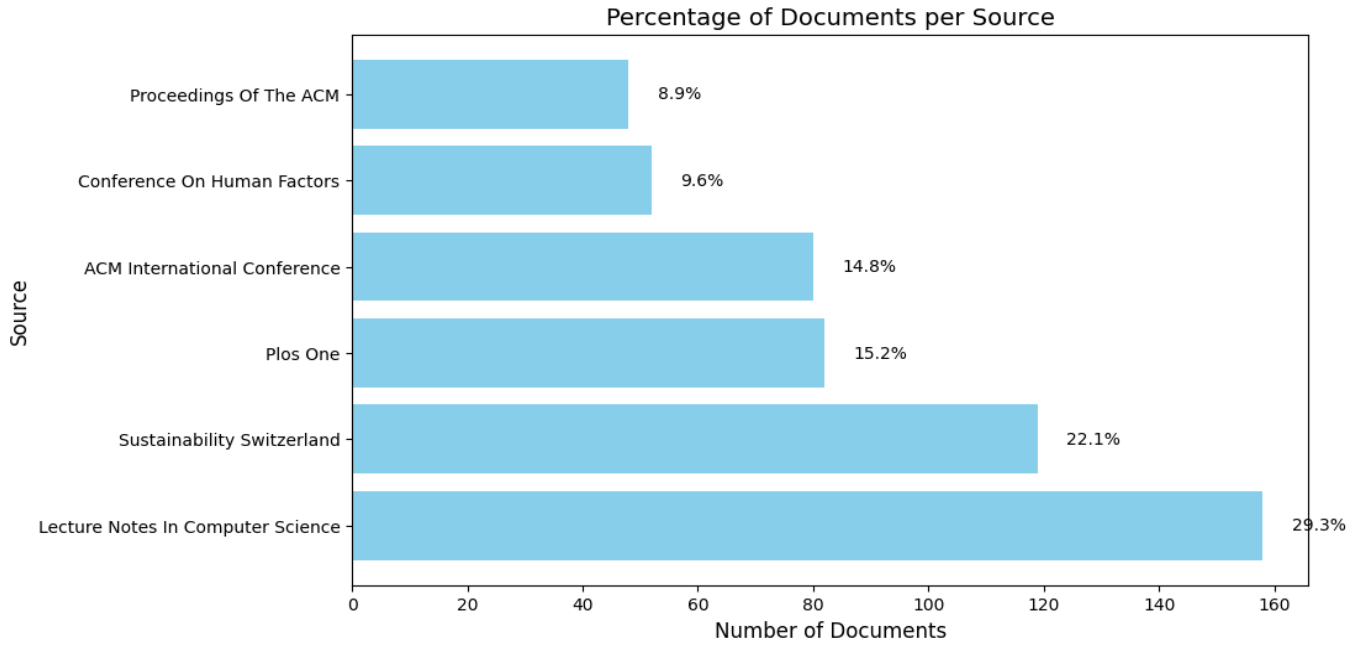

Figure 2. Distribution of documents per source (in percentage)

Figure 2 shows the distribution of documents according to their sources. Accordingly, the first journal was Lecture Notes in Computer Science with 29,3 percent. Sustainability Switzerland comes second with 22,1 percent. The third place belongs to the source called Plos One with 15,2 percent. These insights highlight the integration of gender studies with technology and sustainability, suggesting an evolving academic landscape that bridges social and computational sciences.

Table 3 is an effective way to represent the distribution of documents across different countries. This table allows readers to easily explore global trends, identify leading countries in research or publication, and interpret geographical patterns in the data.

Table 3

Distribution of Publications by Country

Country/Territory	**Documents**
United States	1987
United Kingdom	707
China	602
Germany	503
Spain	406
Italy	355
Canada	331
India	331
Australia	297

Accordingly, the United States leads with 1,987 documents, reflecting its dominant position in global research output. This is consistent with its significant investment in research and development (R&D), world-class academic institutions, and extensive funding for innovation.

United Kingdom follows with 707 documents, also showing a strong academic presence. The UK has a rich history of scientific discovery and continues to be a leader in academic research, although its output is lower than the U.S.

China with 602 documents indicates the rapid growth of research output in the country, driven by heavy investment in science and technology, a growing number of universities, and a focus on becoming a global leader in R&D.

Germany (503 documents) is a major European research hub, reflecting its strength in engineering, natural sciences, and technology.

Spain (406 documents) and Italy (355 documents) represent a solid research output in Southern Europe, though at a smaller scale compared to the leading countries like the U.S. or the UK.

Canada and India, each with 331 documents, show a comparable research output, reflecting the global spread of research. Canada is known for strong research in health, technology, and environmental sciences, while India has a growing academic community, especially in technology and healthcare.

Australia (297 documents) rounds out the list, still contributing a notable amount of research despite its smaller size compared to the other countries mentioned. When we want to show these results on a map, Figure 3 offers us a nice visualization on this subject.

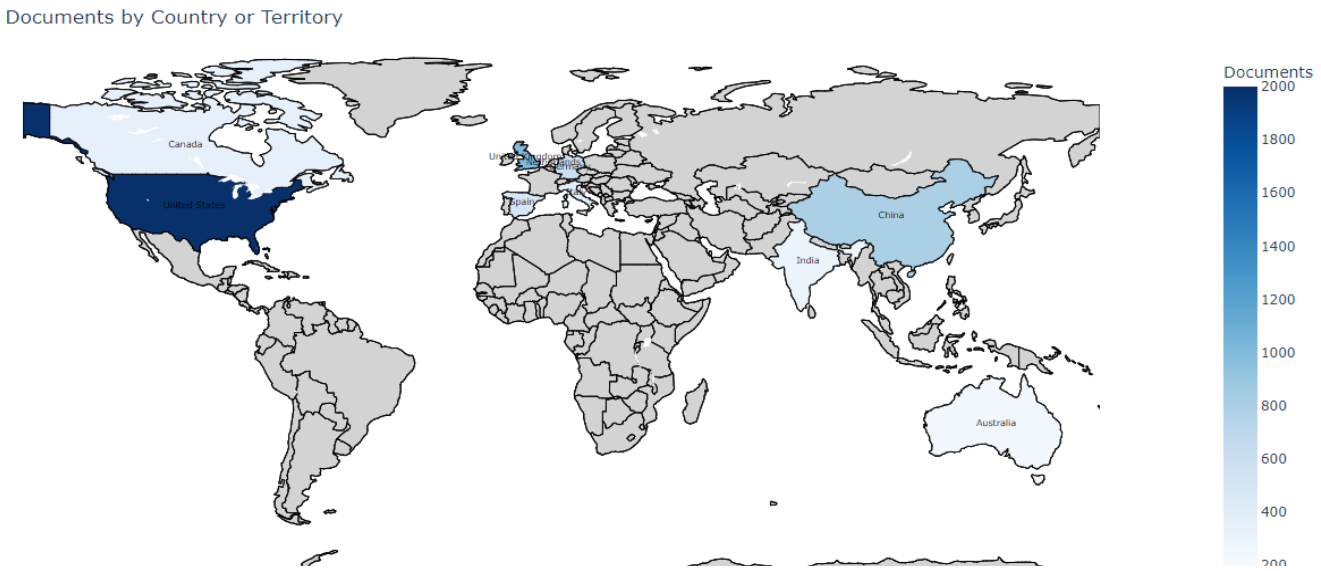

Figure 3. Distribution of documents per country

The United States is the clear leader in global research, followed by the United Kingdom and China, highlighting the concentration of research output in these countries.

China's rapid growth in research reflects its expanding academic and industrial capabilities.

European countries like Germany, Spain, and Italy maintain solid research outputs but are not at the scale of the top two.

India and Canada are on par in research volume, underscoring their importance in global academic production, particularly in fields like technology and healthcare.

Australia's research output is robust, though it lags behind the leading countries.

This distribution suggests that the largest contributors to academic research are concentrated in North America, Europe, and Asia, with these regions continuing to dominate global scholarly output.

Article (4,115) is the most common document type, which is typical for academic research based on Table 4. Articles are usually peer-reviewed and represent original research or reviews of existing work. Their high count indicates that most research publications are in this format.

Table 4

Distribution of Document by Type

Document Type	**Count**
Article	4115
Conference Paper	1069
Book Chapter	420
Book	401
Review	361
Editorial	43
Note	21
Data Paper	8
Letter	7

Conference Paper (1,069) is the second most common type. Conference papers are often preliminary reports or discussions of ongoing research presented at academic or professional conferences. This suggests a healthy conference-driven research culture.

Book Chapter (420) and Book (401) are significant but represent fewer documents. Books and book chapters tend to focus on more comprehensive, detailed explorations of topics, often targeted at academic or professional audiences. Their relatively lower count reflects the longer publication process and more focused scope compared to articles.

Review (361) papers are important for summarizing and synthesizing existing research in a particular area. A substantial number of reviews indicates active efforts to consolidate knowledge in various fields.

Editorial (43) and Note (21) are less common, typically serving specific purposes like introducing special issues or providing short updates or corrections to existing work.

Data Paper (8) and Letter (7) are the least common types. Data papers present datasets that are available for reuse by other researchers, and letters are typically shorter, concise communications, often presenting quick insights or preliminary results.

Most of the research is presented as articles, highlighting the standard format for publishing original research. Conference papers also constitute a significant portion, which suggests an active conference-driven research culture. Other types, like book chapters and reviews, indicate that the field also values comprehensive overviews and scholarly contributions in larger works. Editorials, notes, and letters are much rarer, which is expected as they serve more specialized functions. Overall, this distribution shows that the dominant modes of scholarly communication are articles and conference papers, with other formats like books and reviews playing supportive, yet still valuable, roles in academic publishing.

Table 5

Distribution of Document by Area

Subject Area	**Document Count**
Social Sciences	2930
Computer Science	2205
Business, Management and Accounting	1046
Economics, Econometrics and Finance	769
Engineering	754
Medicine	692
Psychology	614
Mathematics	485
Environmental Science	456

The distribution by document type is shared in detail in Table 5. According to the table, Social Sciences has the highest number of documents (2,930), indicating a strong focus on research in this field. It might reflect the broad nature of social sciences and its interdisciplinary reach across sociology, psychology, economics, and more.

Computer Science follows with 2,205 documents, showing its prominence in academic research, likely due to rapid technological advancements and widespread interest in areas like AI, machine learning, and data science.

Business, Management and Accounting (1,046) also sees significant research activity, highlighting the importance of understanding market dynamics, organizational behavior, and financial systems in a globalized economy.

Economics, Econometrics, and Finance (769) show a strong research presence, though slightly smaller than business-related areas, but still crucial for understanding economic systems, financial markets, and quantitative modeling.

Engineering (754) and Medicine (692) indicate that technical and healthcare fields are also well-represented, with growing interest in innovations, sustainability, and health challenges.

Psychology (614) also has a significant number of publications, reflecting ongoing research into human behavior, mental health, and cognitive processes.

Mathematics (485) is crucial but comparatively lower in publication count, suggesting that while mathematics is foundational for many other fields, it might have a narrower scope in terms of pure research publications.

Environmental Science (456) indicates growing interest in sustainability, climate change, and ecological issues, but it's still a smaller field compared to others.

Overall, the data reveals a significant emphasis on social sciences and computer science, while fields like environmental science and mathematics are slightly less represented in terms of publication volume. The trend underscores the broad nature of interdisciplinary research and its relevance to global issues.

In the analysis conducted with Vosviewer on the data downloaded from Scopus, 18,124 keywords determined and 597 words that were used at least 10 times in publications were taken from among them. The screenshot showing the analysis of these 597 words by the program is shared in Figure 4.

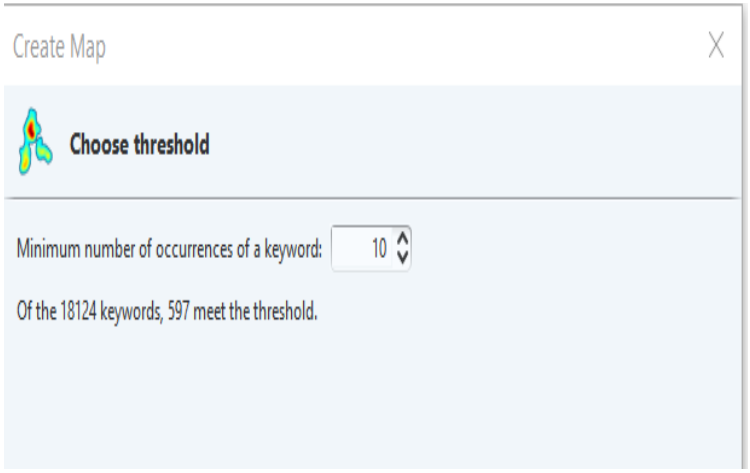

Figure 4. Choosing keywords from dataset

During this analysis, the strength of co-occurrence links is calculated for each keyword (Figure 5). The results from here are shared in the table below (Table 6).

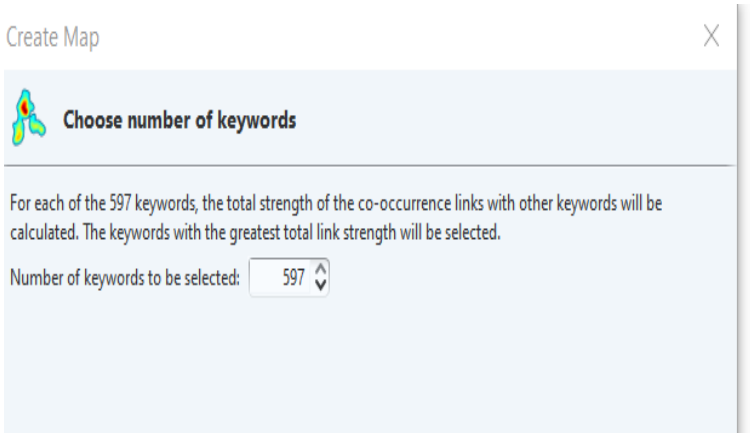

Figure 5. Choosing number of keywords

The strength of the connection indicates the intensity or importance of the relationship between two elements. Co-occurrence, for example, shows the number of times two terms appear in the same publication

Table 6

Distribution of Keywords by Occurrences and Total Link Strength

Term	Occurrences	Total Link Strength
Human	387	4358
Machine Learning	499	3259
Humans	264	3201
Article	257	3127
Female	231	2953
Male	192	2581
Adult	148	2017
Fairness	325	1983
Decision Making	280	1919
Machine-Learning	255	1836
Gender	296	1726
Artificial Intelligence	264	1363
Algorithmics	185	1202
Learning Systems	160	1127
Learning Algorithms	131	1041
Sustainable Development	158	998
Economic and Social Effects	135	966
Controlled Study	68	953
Classification (of Information)	132	907

Networks consist of items and the connections between these items. In this analysis, keywords are taken as items. Items are grouped into clusters on a map. The results can be seen as follows in Figure 6. In Vosviewer network visualization, elements are represented by labels and circles, with label and circle sizes changing depending on the element's weight. As the weight grows, the label and circle become larger, however labels are not displayed for some items to prevent overlapping. The colors of the elements signify which cluster they belong to, allowing clusters to be visually distinguished. This form makes it easy to understand the relationships and importance of network elements.

The keyword data suggests trends in the subjects and themes of the research documents.

Human and Gender-Related Terms: The terms human (387 occurrences), humans (264 occurrences), female (231 occurrences), male (192 occurrences), and gender (296 occurrences) appear frequently. This indicates a significant focus on human-centered studies and gender-related topics, which may align with research into gender equality, human factors in decision-making, and the impact of gender on various fields like business or technology.

Machine Learning and AI: Machine learning (499 occurrences) and machine-learning (255 occurrences) are prevalent, indicating a heavy emphasis on computational models and algorithms, particularly in research related to artificial intelligence (AI) and learning systems. Artificial intelligence (264 occurrences) further supports this, highlighting a strong intersection of AI and human-related studies.

Decision-Making and Fairness: Fairness (325 occurrences) and decision making (280 occurrences) suggest a focus on ethical considerations, such as fairness in algorithms, which is crucial in AI and machine learning applications. This is particularly relevant in contexts like gender equality, where fairness in decision-making processes, whether in hiring or promotions, may be analysed.

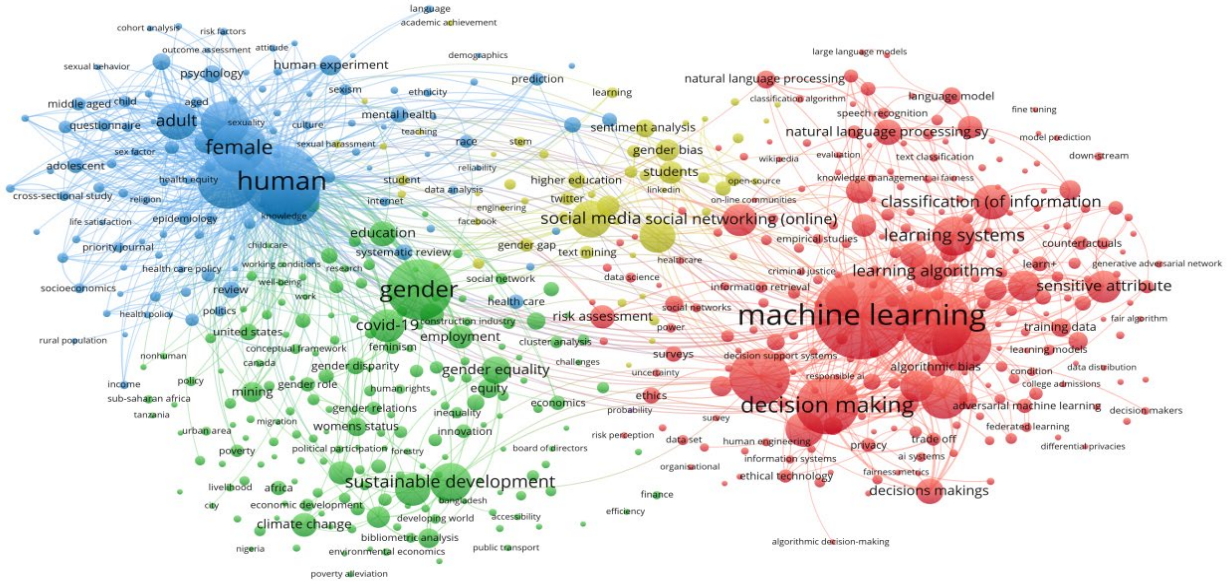

Figure 6. Analysis of keywords

Sustainable Development: Sustainable development (158 occurrences) is another key area, which ties into broader societal concerns such as the environmental impact of technologies, corporate responsibility, and sustainable practices in business and management.

Economic and Social Effects: Economic and social effects (135 occurrences) reflect the importance of understanding the broader implications of various technological and social trends, including gender equality and the adoption of AI and machine learning in different sectors.

Controlled Studies and Classification: The relatively lower mentions of controlled study (68 occurrences) and classification (of information) (132 occurrences) may indicate a methodological focus on structured, experimental research or data categorization in the field, possibly within the realms of data science or educational research.

In summary, the data shows a strong focus on gender equality, machine learning, artificial intelligence, and fairness in decision-making, with a notable connection to sustainable development and economic/social effects.

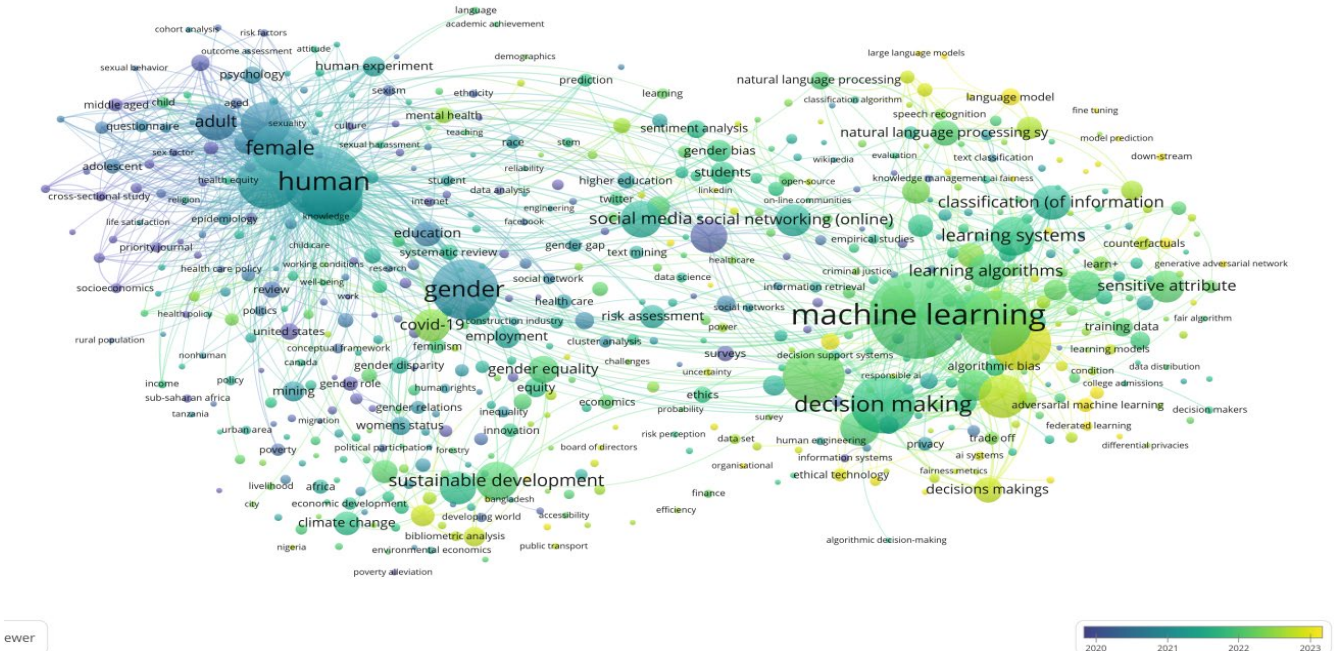

Figure 7. Changes in keywords over time

The change in keywords over time shows how popular a topic has become over time, or whether interest has increased or decreased. In Figure 7, the colour scale (bottom right bar) colors indicate the time in which a specific keyword was used more. Keywords in yellow or green represent terms that have gained popularity or have been studied more in recent years. For example, new technologies such as "large language models, decisions makings, bibliographic analysis" are often represented by yellow tones. Keywords, as "human, female, gender," in blue represent topics that have been studied extensively in previous years but have lost interest today. Gradient colours (green to blue) show the trend of a topic's development and change over time such as "machine learning, natural language processing" are often represented by green tones.

Next, the co-citation analysis of the countries is given. It can be seen in Figure 8, below that a minimum of 5 publications was selected as the threshold value. 88 countries that achieved this are shared in Figure 9.

Figure 8. Analysis of co-citation countries

The concept of bibliographic coupling means that researchers from two countries produce studies based on common references in the context of countries. In other words, when researchers from two countries work on the same academic literature, this indicates that there is bibliographic coupling between these countries. It indicates that there is a fundamental connection between the academic research of two countries. This connection shows that countries focus on similar research topics or benefit from the same sources. The academic communities of two countries may focus on the same disciplines or subfields. For example, both countries may refer to similar literature in the fields of artificial intelligence, sustainability, or gender

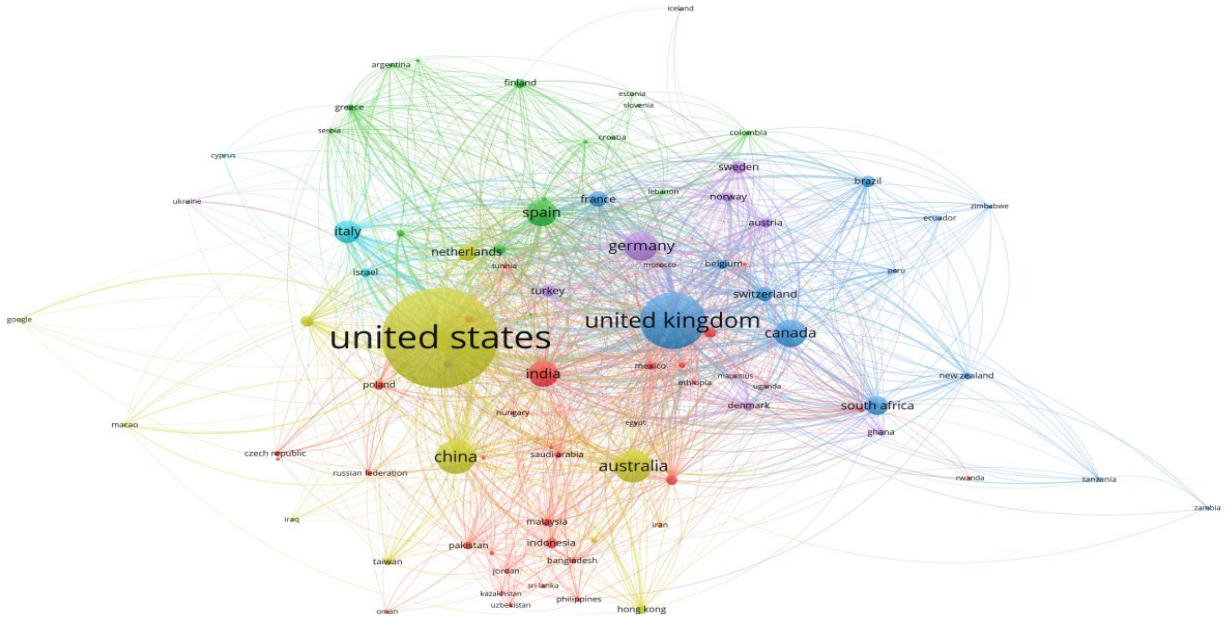

Figure 9. Bibliographic coupling of countries

Bibliographic coupling can reveal existing or potential collaborations between countries. Researchers using similar sources can collaborate on joint projects or studies. It helps to understand which countries are working intensively in which areas and on which literature they are progressing. This is a critical tool in mapping global scientific trends.

It shows which countries are closer to each other in knowledge production in the global academic system. For example, the USA, China, Australia, and the Netherlands are together in the yellow cluster in the Figure 8 above, which indicates a strong bibliographic coupling relationship and that the countries are strong together in this field. In addition, England, Canada and Switzerland also form another group together. Türkiye is in the group with Germany, Norway, Sweden, Denmark and Austria.

Conclusion

The keywords analysis suggests that this research is exploring the intersections between technology (AI and machine learning) and social issues (gender equality, fairness), while also addressing broader sustainable development goals.

Visualization (Figure 7) provides researchers with the opportunity to track developments in their field, understand which topics are new and popular, and evaluate which topics to focus on in the future.

In bibliometric analyses, bibliographic coupling relationships between countries can guide countries when determining their science policies. Countries can collaborate with countries that focus on similar literature to strengthen their partnerships. For example, from Figure 9, Türkiye, Germany, Norway, Switzerland, Denmark and Austria are in the same purple group. If researchers from these countries refer to similar sources on data mining and gender equality, it can be said that there is a strong bibliographic coupling in this area between these countries. This connection indicates potential collaboration opportunities between these countries and can form the basis for global collaborations in their research.

Recent studies have increasingly emphasized the need for data-informed and interdisciplinary approaches to address gender inequality. For instance, Demiray and Ünüvar Ünlüoğlu (2023) conducted a bibliometric analysis of research trends on gender and sustainability, revealing a growing interest in gender equality within the framework of the Sustainable Development Goals. However, they also highlighted that theoretical depth and interdisciplinary integration in the literature remain limited. Similarly, a large-scale bibliometric study by Boekhout, van der Weijden, and Waltman (2021) examined gender disparities in academic careers, showing that female researchers still lag behind their male counterparts in terms of publication and citation metrics, despite some recent improvements. Additionally, the latest UN Women (2023) report underscores that existing policy frameworks are insufficient to meet the 2030 gender equality targets and calls for urgent, comprehensive action.

In this context, the present study provides a novel contribution to the theoretical literature by examining how modern analytical techniques such as data mining and machine learning can be effectively utilized to combat gender inequality. By revealing the potential of data-driven solution strategies, the study offers practical insights for both policy makers and practitioners. The findings raise awareness of how data science, when applied critically and ethically, can serve not only as a diagnostic tool to expose gender-based disparities, but also as a transformative mechanism for designing equitable interventions and policies.

Recommendations

Future studies can be expanded to cover more areas and geographical regions. In addition, a more in-depth comparison of data mining and machine learning methods can be focused on how to make them more effective in combating gender inequality. It is also important for future studies to investigate the applicability of these methods on larger data sets with fewer resources.

Limitations

Among the limitations of this study is that due to the analysis conducted only on the Scopus database, potentially important studies in other databases were excluded. In addition, the analysis conducted in a certain time may not fully reflect the evolutionary development of studies on gender inequality and data mining. The effects of other social and cultural factors on these studies can be examined in more depth.

References

Acharya, A. A., Mahali, P., & Mohapatra, D. P. (2015). Model-based test case prioritization using association rule mining. In *Computational Intelligence in Data Mining-Volume 3: Proceedings of the International Conference on CIDM, 20-21 December 2014* (pp. 429-440). Springer India.

Belingheri, P., Chiarello, F., Fronzetti Colladon, A., & Rovelli, P. (2021). Twenty years of gender equality research: A scoping review based on a new semantic indicator. *Plos One, 16*(9), 1-27.

Bitzenis, A., Koutsoupias, N., & Boutsiouki, S. (2023, July). Business research and data mining: A bibliometric analysis. In *2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)* (pp. 1-6). IEEE.

Boekhout, H., van der Weijden, I., & Waltman, L. (2021). Gender differences in scientific careers: A large-scale bibliometric analysis. *arXiv preprint*, arXiv:2106.12624.

Bowen, D. (2021, September). Construction of business English subject system based on data mining algorithm. In *2021 IEEE 4th International Conference on Information Systems and Computer Aided Education (ICISCAE)* (pp. 441-445). IEEE.

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research, 81,* (pp. 1–15).

Cameron, C., Pierson, H., Aragon, C. M., & West, J. D. (2023). *Gender disparities in the dissemination and acquisition of scientific knowledge*. arXiv. https://arxiv.org/abs/2407.17441

Crawford, K. (2021). Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. *Yale University Press.*

Dastin, J. (2022). Amazon scraps secret AI recruiting tool that showed bias against women. In Ethics of data and analytics, *Auerbach Publications* (pp. 296-299).

Davari, M., Noursalehi, P., & Keramati, A. (2019). Data mining approach to professional education market segmentation: A case study. *Journal of Marketing for Higher Education, 29*(1), 45-66.

Demiray, B., & Ünüvar Ünlüoğlu, D. (2023). Bibliometric analysis of research trends on gender and sustainability: Future research perspectives. *İmgelem*, 7(13), 545-560.

Demirgöz Bal, M. (2014). Toplumsal Cinsiyet Eşitsizliğine Genel Bakış. *Kadın Sağlığı*

Hemşireliği Dergisi, 1(1), 15-28.

Dominguez, D., Pantoja, O., De Los Nietos, J., García, H., Sánchez, Á., & González, M. (2019, April). Soft-computing modeling and prediction of gender equality. In *2019 Sixth International Conference on eDemocracy & eGovernment (ICEDEG)* (pp. 242-248). IEEE.

Farsia, L. (2024). Ensuring equal opportunity: Gender equality in workplace. *Accentia: Journal of English Language and Education, 4*(1), 21-28.

García-Sánchez, I. M., Minutiello, V., & Tettamanzi, P. (2022). Gender disclosure: The impact of peer behaviour and the firm's equality policies. *Corporate Social Responsibility and Environmental Management, 29*(2), 385-405.

Gartzia, L. (2021). Gender equality in business action: A multi-agent change management approach. *Sustainability, 13*(11), 1-29.

Gupta, M. K., & Chandra, P. (2020). A comprehensive survey of data mining. *International Journal of Information Technology, 12*(4), 1243-1257.

Holman, L., Stuart-Fox, D., & Hauser, C. E. (2018). *The gender gap in science: How long until women are equally represented?* PLoS Biology, 16(4), e2004956. https://doi.org/10.1371/journal.pbio.2004956

Huang, J., Gates, A. J., Sinatra, R., & Barabási, A.-L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences, 117(9),* 4609–4616. https://doi.org/10.1073/pnas.1914221117

Irfan, S., & Sharif, K. (2024). Gender equality in management: Exploring social science perspectives. *Global Journal of Econometrics and Finance, 1*(02), 64-75.

Ko, Y., Ko, H., Chung, Y., & Woo, C. (2021). Do gender equality and work–life balance matter for innovation performance? *Technology Analysis & Strategic Management, 33*(2), 148-161.

Križanić, S. (2020). Educational data mining using cluster analysis and decision tree technique: A case study. *International Journal of Engineering Business Management, 12*, 1847979020908675.

Li, T., & Zhang, C. (2022). Research on the application of multimedia entropy method in data mining of retail business. *Scientific Programming, 2022*(1), 1-13.

Nilsson, N. J. (1998). Artificial Intelligence: A New Synthesis. *Morgan Kaufmann.*

OECD. (2017). The pursuit of gender equality: An uphill battle. *Retrieved from* https://www.oecd.org/en/publications/the-pursuit-of-gender-equality_9789264281318-en.html

Peng, X. Y., Fu, Y. H., & Zou, X. Y. (2024). Gender equality and green development: A qualitative survey. *Innovation and Green Development, 3*(1), 100089.

Plotnikova, V., Dumas, M., Nolte, A., & Milani, F. (2023). Designing a data mining process for the financial services domain. *Journal of Business Analytics, 6*(2), 140-166.

Ridgeway, C. L. (2011). Framed by Gender: How Gender Inequality Persists in the Modern World. *Oxford University Press.*

Ritter-Hayashi, D., Vermeulen, P., & Knoben, J. (2019). Is this a man's world? The effect of gender diversity and gender equality on firm innovativeness. *Plos One, 14*(9), 1-19.

Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach (4th ed.). *Pearson*.

Singh, G. V., Ghosh, S., & Ekbal, A. (2023). Promoting gender equality through gender-biased language analysis in social media. In *IJCAI*, 6210-6218.

Stotsky, J. G., Shibuya, S., Kolovich, L., & Kebhaj, S. (2016). Trends in gender equality and women's advancement. *IMF Working Paper Research Department and Strategy, Policy, and Review Department*.

Toksöz, G., & Memiş, E. (2018). Mapping and monitoring gender equality in employment. *CEİD Publications,* Ankara.

UN Women. (2020). The World's Women 2020: Trends and Statistics. *United Nations.*

UN Women. (2023). Global gender equality in 2023: Urgent efforts needed to reach 2030 goals. *Retrieved from* https://www.unwomen.org/en/news-stories/feature-story/2023/09/global-gender-equality-in-2023-urgent-efforts-needed-to-reach-2030-goals

UNDP. (2020). Human development report 2020: The next frontier – Human development and the Anthropocene. United Nations Development Programme. *Retrieved from* https://hdr.undp.org

World Bank. (2024). Gender overview. *Retrieved from* https://www.worldbank.org/en/topic/gender/overview

World Economic Forum. (2023). Global Gender Gap Report 2023. *Retrieved from* https://www.weforum.org/reports/global-gender-gap-report-2023/

Xin, Y. (2021). Analyzing the quality of business English teaching using multimedia data mining. *Mobile Information Systems, 2021*(1), 1-8.

Yan, Y., & Liu, Q. (2019). An overview of the application of data mining technology in e-commerce. *Academic Journal of Engineering and Technology Science, 2*(2), 76-88.

Zhen, C., & Jiang, C. (2019). Overview of data mining in the era of big data. *International Core Journal of Engineering, 5*(10), 136-139.

Genişletilmiş Özet

Amaç

Toplumsal cinsiyet eşitsizliği, kadınlar ve erkekler arasındaki sosyal ve kültürel hak farklılıklarını vurgular. Toplumsal gelişmelerle birlikte evrimleşen bu kavram, sürdürülebilir kalkınma için kritik bir rol oynamaktadır. Modern ekonomilerde zihinsel emeğin ön plana çıkması, kadınların fiziksel emek farklılıklarına rağmen başarılı olmasını sağlamıştır. Ancak, ağlara erişim eksikliği, mentorluk ve toplumsal cinsiyet yanlılığı gibi engeller, kadınların kariyer ilerlemesini sınırlamaktadır. Sosyoloji ve ekonomi gibi disiplinler, istihdam, eğitim ve siyaset alanlarında cinsiyet eşitsizliğini incelemektedir. Cinsiyet eşitliği, yenilikçiliği teşvik eder, iş-yaşam dengesini iyileştirir ve özellikle Ar-Ge sektörlerinde verimliliği artırır. Birleşmiş Milletler gibi uluslararası kuruluşlar, sistematik engelleri aşmak için toplumsal cinsiyet eşitliği girişimlerini aktif olarak desteklemektedir.

Bu bibliyometrik analiz çalışması, cinsiyet eşitsizliği ve veri madenciliği araştırmalarının kesişim noktasını incelemeyi amaçlamaktadır. En çok atıf alan makaleler, etkili yazarlar ve ülkeler arası iş birlikleri incelenerek, cinsiyet eşitsizliği ile mücadelede değişen yaklaşımlar ve araştırma boşlukları ortaya konulacaktır. Çalışma, veri

biliminin sosyal fayda yaratma potansiyeline katkıda bulunacaktır. Ayrıca, feminist teori, cam tavan sendromu ve kesişimsellik perspektifleri gibi toplumsal cinsiyet kuramlarıyla veri madenciliği ve bilgi yönetimi arasındaki ilişkiyi ele almaktadır.

Yöntem

Veri madenciliği, büyük veri kümeleri içindeki örüntüleri ve ilişkileri keşfetmek için güçlü bir araçtır. 1990'lardan itibaren geliştirilen bu teknoloji, anlamlı bilgiler elde etmek ve tahmine dayalı modeller oluşturmak için algoritmalar ve analitik teknikler kullanır. İşletmeler, veri madenciliğini eğilimleri analiz etmek, performansı iyileştirmek ve stratejilerini güçlendirmek için kullanmaktadır. Veri ön işleme, dönüştürme ve analiz gibi yapılandırılmış süreçleriyle bu teknoloji, pratik sorunları çözmede ve gizli örüntüleri ortaya çıkarmada önemli bir rol oynar.

Bu çalışma, Scopus veritabanından elde edilen verilerle gerçekleştirilen bibliyometrik analiz türündedir. Keşifsel bir yaklaşım benimsenmiş ve veriler betimleyici olarak analiz edilmiştir. Farklı alanlardaki çalışmalar için Scopus veri tabanındaki "tüm alanlar" seçilerek tarama yapılmış ve "("Cinsiyet Eşitsizliği" VEYA "Cinsiyet Uçurumu") VE ("Veri Madenciliği" VEYA "Makine Öğrenmesi" VEYA "Yapay Zeka")" anahtar kelime aramasına göre bir örneklem oluşturulmuştur. Bu tarama sonucunda toplamda 5,074 çalışma Excel formatında yazar/yazarlar, yayınlanan kitap, dergi veya konferans, yayın yılı, destekleyen kuruluş gibi bilgileriyle indirilmiş ve bu bilgiler üzerinden incelenmiştir. Çalışmada önemli yazarlar, makaleler ve iş birlikleri incelemesi yapılıp tematik kümeler ve kelime analizi gibi tekniklerle araştırma boşlukları ve eğilimler belirlenmiştir. Bu işlemlerde VOSviewer bilimsel haritalama programından ve görselleştirmelerde ayrıca Python açık kaynak yazılımından faydalanılmıştır. Çalışmada yalnızca Scopus veri tabanı kullanılmış olup, Web of Science ve Google Scholar gibi diğer büyük veri tabanları karşılaştırmalı olarak analiz edilmemiştir. Ancak, Scopus'un geniş kapsamı ve yüksek atıf indeksine sahip akademik çalışmalar içermesi nedeniyle tercih edilmiştir.

Bulgular

Araştırmanın bulguları yıllara göre çalışmalara bakıldığında, veri analizi, cinsiyet eşitsizliği ve veri madenciliği ortak anahtar kelimelerinin geçtiği dolayısıyla bu alan üzerine yapılan araştırmaların son yıllarda arttığını göstermektedir. En çok atıf alan çalışmalar, makine öğrenmesi ve modelleme tekniklerinin cinsiyet eşitsizliği araştırmalarında nasıl kullanıldığını ortaya koymaktadır. Çalışmalara bakıldığında, eğitim ve istihdam gibi sektörlerde veri madenciliği yöntemlerinin yaygınlaştığı, ülkeler arası iş birliklerinin arttığı ve bazı gelişmiş ülkelerde dikkat çekici çalışmaların yapıldığı gözlemlenmiştir. Bu artışın nedenleri arasında cinsiyet eşitliği ile ilgili farkındalığın artması, politika değişiklikleri ve büyük veri analiz yöntemlerinin yaygınlaşması gösterilebilir.

Araştırma, disiplinler arası dağılıma da ayrıca dikkat çekmektedir. Sosyal bilimler (2,930 doküman) bu araştırma alanlarının içinde başı çekerken, bilgisayar bilimleri (2,205 doküman) yapay zeka ve veri bilimine olan ilgi de açıkça görülebilmektedir. İşletme, yönetim ve muhasebe (1,046 doküman) daha çok piyasa dinamiklerine odaklanırken, ekonomi ve finans (769 doküman) ekonomik modellemeyi daha çok ele almıştır. Mühendislik (754) ve tıp (692) gibi alanlarda yenilikçilik ve sağlık konularının daha ön planda olduğu görülmektedir. Psikoloji (614), matematik (485) ve çevre bilimleri (456) gibi diğer disiplinler de kendi araştırma önceliklerini yansıtmaktadır.

Vosviewer, kelime analzinde 18,124 anahtar kelimeyi belirlemiş ve ortak görülme sıklığına göre 597 terimi analiz etmiştir. Öne çıkan anahtar kelimeler insan (387), cinsiyet (296), makine öğrenmesi (499) ve yapay zeka (264) şeklinde görülmektedir. Ayrıca kelime analizi çalışmasının sonucu, insan odaklı çalışmalar, cinsiyet eşitliği, yapay zeka uygulamaları yanı sıra karar verme süreçleri ve adalet gibi konulara da odaklanıldığını göstermektedir.

88 ülkenin eş-atıf analizi, yapay zeka, sürdürülebilirlik ve cinsiyet çalışmaları gibi ortak araştırma ilgi alanlarına dayalı küresel iş birliklerini vurgulamaktadır. ABD küresel araştırma çıktılarında başı çekerken, onu İngiltere ve Çin takip etmektedir. Yayınların çoğu makale formatında olup, bunu konferans bildirileri ve kitap bölümleri izlemektedir.

Anahtar kelime verileri, insan ve cinsiyet odaklı çalışmaların yanı sıra makine öğrenmesi ve yapay zeka gibi hesaplamalı modeller ve algoritmaların etik boyutlarına dikkat çekmektedir. Ayrıca yıl bazlı yapılan analizde anahtar kelimelerin insan ve cinsiyet konulu çalışmalardan karar verme, makine öğrenmesi, yapay zeka, dil modeli gibi güncel konulara kaydığını ve gelişen teknolojinin bu araştırmalarda kullanılabildiğini de göstermektedir.

Bibliyografik eşleşme, ülkeler arasında ortak kaynaklara dayalı mevcut veya potansiyel iş birliklerini belirleyen önemli bir yöntemdir. Bu analizden elde edilen buldular, ABD, Çin, Avustralya ve Hollanda gibi ülkelerin çalışmanın ele aldığı konu üzerine bilgi üretiminde yakın ilişkiler kurduğunu göstermektedir. Bunun yanında İngiltere, Kanada, İsviçre de birliktelikleri ile başka bir grup oluşturmaktadırlar. Türkiye ise Almanya, Norveç, İsveç, Danimarka ve Avusturya'nın olduğu grupta yer almaktadır.

Analiz, cinsiyet eşitsizliği ile mücadelede veri odaklı yeni yöntemlerin öne çıktığını ancak önemli araştırma boşluklarının hala bulunduğunu ortaya koymaktadır. Yapay zeka, makine öğrenmesi, cinsiyet eşitliği üzerine yapılan bu vurgu, disiplinlerin birbirine bağlılığını ve araştırma eğilimlerinin evrimini vurgulamaktadır.

Araştırma sınırlılıkları

Bu çalışmanın sınırlılıkları arasında yalnızca Scopus veritabanının kullanılması ve diğer veritabanlarındaki çalışmaların dışarıda bırakılması yer almaktadır. İngilizce yayınların değerlendirilmesi, diğer dillerdeki çalışmaların gözden kaçmasına neden olmuş olabilir. Anahtar kelimeler de, arama kapsamını sınırlayarak bazı ilgili çalışmaları hariç tutmuş olabilir. Paylaştığımız bu sınırlılıklar gelecekte daha kapsamlı çalışmalar yapmak isteyen araştırmacılara değerlendirmeler ve araştırmalar konusunda yardımcı olabilir ve yön verebilir.

Öneriler

Çalışma, cinsiyet eşitsizliğiyle mücadelede veri madenciliği ve makine öğrenmesi yöntemlerinin daha etkili kullanımı için mevcut durumun literatür özetini yapmakta, ayrıca tarihsel açıdan araştırmaların yönelimlerini belirtmekte, ülke işbirliklerine dikkat çekmekte, gelecek çalışmalar için teorik ve pratik öneriler sunmaktadır. Teorik olarak, araştırmacılara daha fazla veri odaklı, gelişen yeni makine öğrenmesi, yapay zeka gibi yöntemlerin kullanıldığı ve başka alanların da dahil edildiği uygulamalı çalışmalar yapabilmeleri konusunda yol göstericidir. Uygulamada ise bu çalışma, veri bilimi kullanılarak yapılan çalışmalar sayesinde cinsiyet eşitsizliğini çözmede sosyal ve başka alanlardaki pek çok konuya dikkat çekerek buradan elde edilen bilginin nasıl kullanılabileceğine dair politika geliştirmeye dair farkındalık gelişmesini sağlar. Sosyal açıdan, bu çalışmalar, cinsiyet eşitsizliğini daha iyi anlamak ve çözmek için sosyal farkındalık gelişmesine katkı sağlar. Gelecek araştırmalar için daha fazla konu ve sektörü kapsayacak şekilde, özellikle gelişmekte olan ülkelerde farklı alanlardaki cinsiyet eşitsizliğine odaklanması önerilmektedir.

Özgün değer

Bu çalışma, cinsiyet eşitsizliği ve veri madenciliği arasındaki ilişkiyi, bu alanda yapılan bilimsel çalışmaları bilimsel bir veri tabanı üzerinden kapsamlı bir şekilde inceleyen ilk çalışmalardan biridir. Bulgular, bu alandaki araştırma boşluklarını, gelecek çalışmaların yönelimlerini ve iş birliği fırsatlarını ortaya koymaktadır. Geleneksel çalışmaların göz ardı ettiği örüntüleri, eğilimleri ve iş birliği ağlarını kapsamlı bir bibliyometrik analizle ortaya çıkararak, veri odaklı yöntemlerle toplumsal cinsiyet çalışmalarına teorik ve pratik katkılar sağlamaktadır. Ayrıca, veri madenciliği ve toplumsal cinsiyet eşitsizliği araştırmalarını birleştiren disiplinlerarası analizlerdeki boşluğu doldurarak, gelecekteki çalışmalara yön verecek bir temel oluşturmaktadır.