iJDASE International Journal of Data Analytics and Smart Education

iJDASE | ISSN: 2736-979X | Vol. 1 No. 1 (2025)
https://doi.org/10.64817/7eggbj57

PAPER

Understanding the Role of Web Intelligence in the Data Economy

Sunday Adewale Olaleye[1], **Oluwafemi Samson Balogun**[2]

[1]School of Business, JAMK University of Applied Sciences, Rajakatu 35, 40100 Jyväskylä, Finland

[2]School of Computing, University of Eastern, FI-70211. Finland

sunday.olaleye@jamk.fi
samsb@uef.fi

ABSTRACT

Based on bibliometric data, this paper presents an in-depth analysis of the evolution of Web Intelligence (WI) in the past two decades. These objectives are to trace thematic shifts and influential papers, understand prevailing topics within the WI community, and derive insights for future research directions. This paper traced the results of a citation network and co-citation, outlining the significant contributions and emerging trends of WI and its critical role in the data economy. The findings demonstrate how WI offers robust frameworks and innovative solutions for leveraging data to enhance data-driven decision-making processes. As such, the study bridges the divide between theory and practice to foster a more connected and intelligent society.

KEYWORDS

Web Intelligence, Bibliometric Analysis, Data Economy, Data-Driven Economy, Emerging Trends

1 INTRODUCTION

In recent years, there has been a growing interest in Web Intelligence (WI) and its role in the data economy. Web Intelligence refers to using artificial intelligence (AI) and data science techniques to analyze and extract valuable knowledge from the vast data available on the World

Sunday Adewale Olaleye, Oluwafemi Samson Balogun. (2025). Understanding the Role of Web Intelligence in the Data Economy. *International Journal of Data Analytics and Smart Education (iJDASE),* 1(1) pp. 25-54 https://doi.org/10.64817/7eggbj57

Article submitted: 12.02.2025 Revision uploaded: 16.06.2025 Final Acceptance: 11.07.2025

Wide Web (WWW) and social networks. Customer Relationship Management provides researchers and practitioners unprecedented opportunities to uncover hidden patterns, trends, and insights that can be leveraged for various purposes, such as decision-making, recommendation systems, and customer relationship management [1].

Web Intelligence is closely related to the broader field of artificial intelligence and its goal of achieving human-level AI. Understanding human intelligence, particularly brain intelligence, has long been the cornerstone of reaching the ultimate AI. The literature has explored this connection between AI and brain science, and researchers have looked towards the future vision of AI in the connected world [1].

One area of research that has gained significant attention is Brain Informatics (BI), which focuses on studying and applying brain-machine intelligence. BI aims to combine AI and brain science with big data to accelerate the development of a human-level AI society. By connecting AI and brain science with the vast amount of data available on the web, a new vision of brain-machine intelligence research and its application in various domains has emerged [1]. Some other studies explored the affective models for web intelligence [2], soft web intelligence as a tool for data collection, processing, and data mining [3,4], and web intelligence as an agent of world connection [5].

Integrating Web Intelligence (WI) into the data economy in the rapidly evolving digital landscape represents a transformative frontier. As an interdisciplinary field, Web Intelligence harnesses the power of artificial intelligence (AI), data mining, and social network analysis to extract meaningful insights from the vast expanse of web data [6]

This burgeoning field is pivotal in deciphering complex data patterns and driving innovations across various sectors, including e-commerce, social media, and beyond. However, despite the growing interest in Web Intelligence and its potential applications, gaps in the literature need to be addressed. One of the gaps is the need for a longitudinal analysis of the evolution of the Web Intelligence community. While there have been studies on specific aspects of Web Intelligence, a comprehensive review that examines how the field has evolved and identifies the most influential papers and emerging research topics is needed [7].

Furthermore, there is a need to explore the impact of Web Intelligence research on the data economy. With the increasing availability of big data, understanding how Web Intelligence can contribute to analyzing and utilizing these data is crucial. This development includes examining how Web

Intelligence techniques can extract knowledge and insights from web data, social networks, and online reviews and how these findings can improve customer relationship management and support decision-making in various industries [8].

Therefore, this study aims to understand the role of Web Intelligence in the data economy by addressing the gaps in the literature and exploring the potential applications of Web Intelligence techniques in various domains.

The significance of this study lies in its ability to bridge the gap between theoretical research and practical applications. By delving into how Web Intelligence can enhance data-driven decision-making processes, this research aims to highlight its potential to foster a more connected and intelligent society. Recent advancements have underscored the role of WI in improving web-based services and applications using AI and machine learning algorithms [9]. Despite these advancements, there remains a notable gap in understanding the longitudinal evolution of WI and its impact on the data economy, as well as a need for comprehensive bibliometric analyses to chart this progress.

The current study employs bibliometric data analysis to investigate the evolution of Web Intelligence over the past two decades and address the existing issues. The research questions guiding this investigation include: (a) How has the thematic focus of WI research shifted over time? (b) What are the most influential papers and prevailing topics within the WI community? (c) How can these insights inform future research directions? These questions are crucial for mapping out the trajectory of WI and identifying critical areas for future exploration.

Methodologically, this study is structured into several vital sections. The Methodology section details the bibliometric techniques used to analyze the evolution of WI, including citation network analysis and co-citation analysis, which help in identifying influential works and emerging trends [7]. The results section will present the findings from these analyses, showcasing how the field has developed and highlighting significant contributions. Finally, the conclusion synthesizes the findings, discussing their implications for the future of WI in the data economy and suggesting potential areas for further research. Exploring Web Intelligence is not merely an academic endeavour but a practical imperative in our increasingly data-driven world. As we continue to harness the vast amounts of data generated daily, the insights provided by WI can play a critical role in shaping a more innovative, more interconnected global society.

By answering these research questions, this study aims to contribute to the existing literature on Web Intelligence and provide insights into its potential applications in the data economy. By advancing our understanding of WI, this study aims to contribute to creating robust frameworks and innovative solutions that leverage the full potential of the data economy. The findings of this study can guide the direction of future research projects in the field and update the scope and areas of interest in current trends and relevant journals.

2 METHODOLOGY

2.1 Databases Accessed

The bibliometric data for this study were collected from two prominent databases: Web of Science and Scopus. These databases were selected due to their comprehensive coverage of high-quality research articles across various disciplines. The data collection was performed on September 1, 2023.

2.2 Search Strategy

A specific search strategy was employed. The keyword "web intelligence" was used as the main topic of interest to identify relevant literature. The search was refined using the following Boolean operators and filters:

I. Document Type: Article
II. Publication Years: 2002 - 2022
III. Language: English

The detailed search query included:

I. Topic: "web intelligence"
II. Document Types: Article
III. Publication Years: 2022, 2021, 2020, 2019, 2018, 2017, 2016, 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006, 2005, 2004, 2003, 2002
IV. Languages: English

2.3 Time Frame

The study examined research articles published over 20 years, from 2002 to 2022.

2.4 Document Types

Only articles were included in the analysis to maintain consistency and ensure high data quality.

2.5 Data Cleaning Procedures

Data cleaning was performed using R-Studio to ensure accuracy and remove redundancy. The process involved combining data from Web of Science and Scopus, which initially yielded 103 and 122 articles, respectively. After identifying and removing 71 duplicate records, 154 unique articles were retained for the final analysis.

2.6 Software and Tools

The data analysis was conducted using the Biblioshiny App, a specialized bibliometric analysis and visualization tool. This software facilitated the examination of various bibliometric indicators and provided a robust platform for data visualization.

2.7 Bibliometric Indicators

The study focused on three main bibliometric indicators:

I. Conceptual Structure: Analyzing the key themes and concepts within the literature.
II. Intellectual Structure: Examining the foundational theories and influential works in the field.
III. Social Structure: Investigating the network of authors and collaboration patterns.

2.8 Analytical Methods

A quantitative methodology was employed, following a systematic bibliographic workflow. The process comprised six steps, as outlined by Olaleye [10]; Agjei et al, [11]; Olaleye et al, [12]:

1. **Research Question Formulation**
   The study developed three unique research questions to guide the study.
2. **Data Source Selection**
   The study identified Web of Science and Scopus as appropriate sources for data collection.
3. **Data Analysis**
   The study utilized the Biblioshiny App to analyze bibliometric data [13].

4. **Data Visualization**
   The study used the Biblioshiny App to create visual representations of the data.
5. **Results Presentation and Interpretation**
   The study displayed the results generated by the Biblioshiny App and

provided interpretations.

6. **Conclusion and Future Directions**
   The study **s**ummarized the findings, discussed limitations, and suggested future research areas.

2.9 Validation Methods

Multiple validation techniques were employed to ensure the validity and reliability of the results. These included cross-referencing data between the two databases, consistent application of search filters, and thorough data cleaning to remove duplicates and errors.

2.10 Ethical Considerations

The study adhered to ethical standards in data handling and reporting. All data used was publicly accessible and did not involve personal or sensitive information. The research was conducted with transparency and integrity, ensuring an accurate representation of the findings

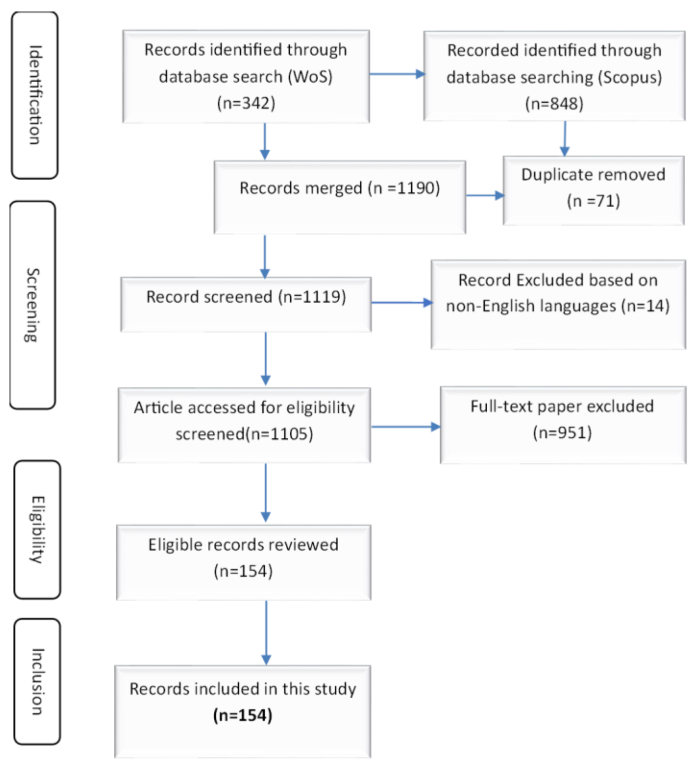

Fig. 1: Web Intelligence PRISMA

Fig. 2: Web Intelligence Literature Descriptive

Figure 1 shows a PRISMA flow diagram, which is often used in systematic reviews to clearly outline how research articles were identified, screened, and selected for inclusion. Figure 2 provides a bibliometric overview of the research dataset, offering insights into publication trends and patterns, usually generated using tools like Bibliometrix or VOSviewer to better understand the structure and impact of the scientific literature.

3 RESULTS

This section presents the findings from the quantitative bibliometric analysis. The results are structured according to the research questions addressed in the study, aiming to provide readers with a thorough understanding of the research focus, particularly concerning the role of web intelligence in the data economy. This study conducted two analyses to examine annual scientific production and citation rates, evaluating the progress in understanding the research landscape of data intelligence. These quantitative assessments provide precise data on the field's performance and prospects. This study focuses on exploring historical and projected trends in data intelligence research, specifically in the context of understanding the role of web intelligence in the data economy.
According to Figure 3, our dataset shows that research output in the field of web intelligence began modestly in 2002 with one published article. Between 2002 and 2022, a total of 154 publications from various sources have been recorded. Figure 3 indicates a steady increase in article production from 2002 to 2023. Significant interest in data intelligence emerged in 2005, with notable expansions in 2013. The most productive years for scientific articles were 2005, 2006, 2011, and 2013, yielding 14, 10, 11, and 14 documents, respectively.

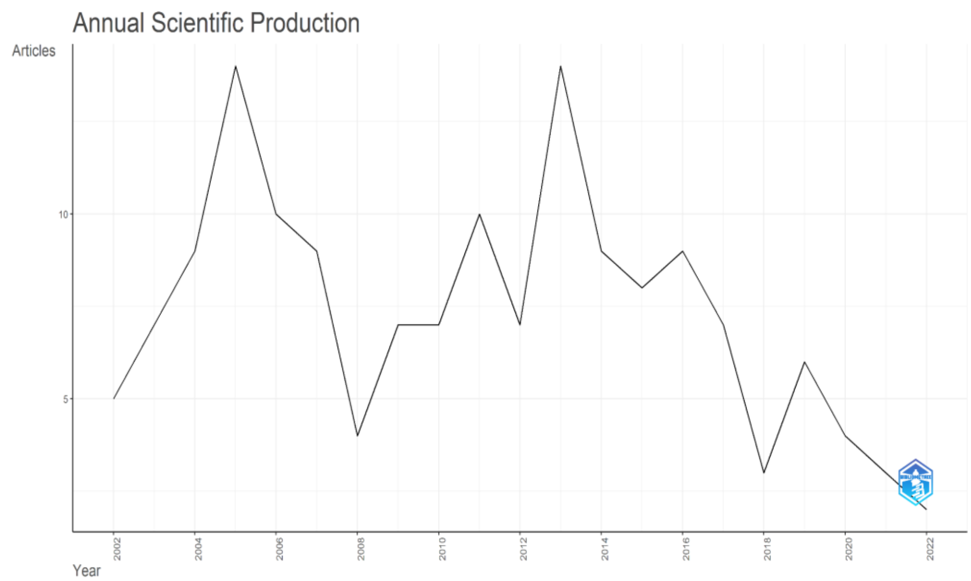

Fig. 3: Web Intelligence Annual Scientific Production

The highest average total citations per year, 2.50, were recorded in 2009 across seven publications over 15 citable years. The second highest was 2.14 in 2011, based on ten documents with 13 citable years, followed by 2.07 in 2019 across six documents over five citable years. Conversely, the lowest average total citation per year was 1, recorded for one year (2022). Interestingly, 2005 and 2013 saw 14 articles each, with average total citations per year of 6.21 and 18.5, respectively, and citable years of 19 and 11. These results suggest that the number of scientific papers does not necessarily lead to increased citations, which depend on the quality of the papers and their accessibility to researchers (see Figure 3).

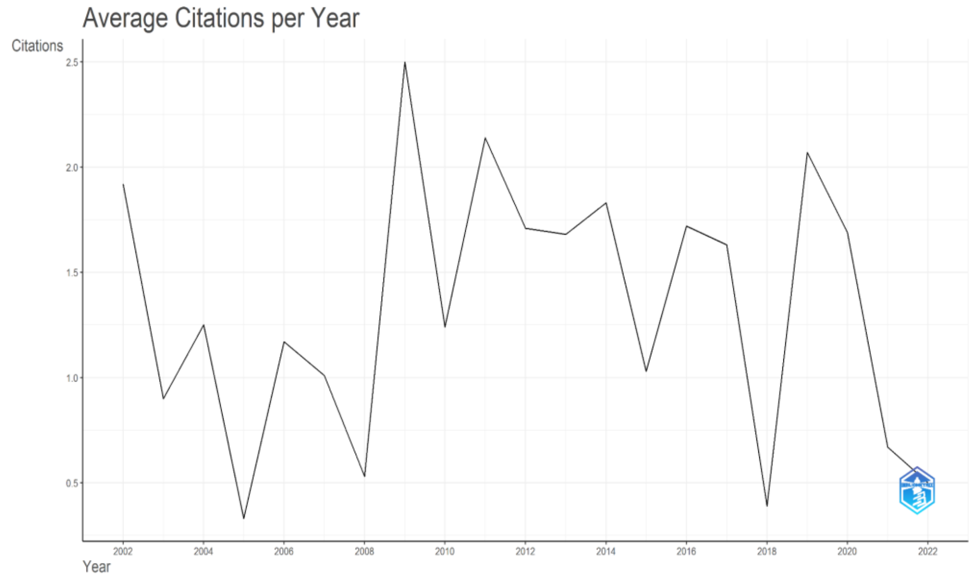

Fig. 4: Web Intelligence Average Citations per Year

Figure 4 provides additional analysis by summarizing the top ten authors in the field of web intelligence from 2002 to 2022. A larger circle diameter in Figure 4 represents more documents published annually by an author. In contrast, the intensity of the circle's shade indicates the number of citations received that year. The count of output years for an author begins with their first publication, resulting in different total production years for each writer. With seven papers each, Weichselbraun A, Yao Y, and Zhong N are the three most prolific authors according to this analysis. Li J. and Scharl A. are the third most productive, each contributing six papers to the field of web intelligence. The remaining five authors among the top ten have four and three publications each, as shown in Figure 4. This study evaluates the prolific authors in web intelligence by considering the number of documents and citations per author to assess their relevance.

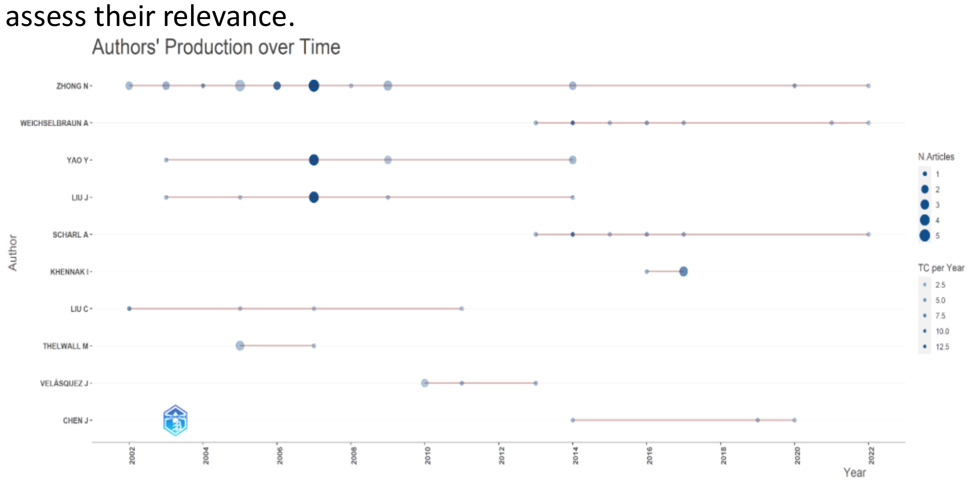

Fig. 5: Web Intelligence Authors' Production over Time

Some researchers, including Liu J., Liu C., Zhong N., and Yao Y., have investigated web intelligence since the 2000s. Among them, Zhong N. has extensively researched this topic, as illustrated in Figure 5. Thelwall M. has also significantly contributed to web intelligence research, achieving a notable local impact based on their author, H-Index. Since 2002, Zhong N. has been a prominent figure in the field, producing 22 documents and attaining a local author H-Index of 11. These findings indicate that authors like Liu J., Liu C., Zhong N., and Yao Y. have substantially influenced web intelligence research.

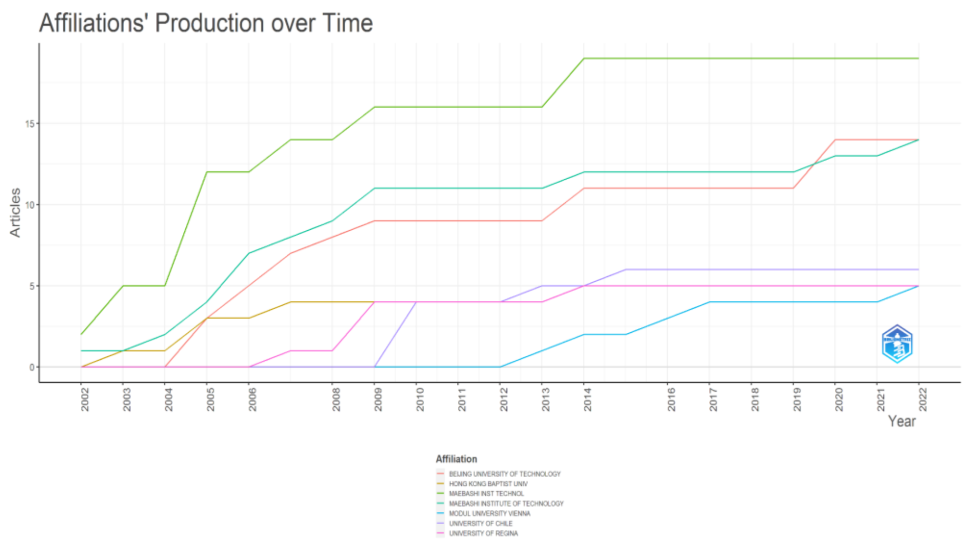

Fig. 6: Web Intelligence Affiliations' Production over Time

Figure 6 indicates that the research outputs of all leading affiliations have been consistently growing over the years, with a significant increase in recent times. However, no African country is among the top 10 in scientific output or most cited publications on web intelligence research. This limited research presence in Africa could be due to various factors. The Maebashi Institute of Technology ranks first with 19 documents from 2014-2022, followed by the same institution, and the Beijing University of Technology in second place with 14 documents each in 2022 and 2020-2022, respectively. The University of Chile holds the third position with six documents from 2015 to 2022.

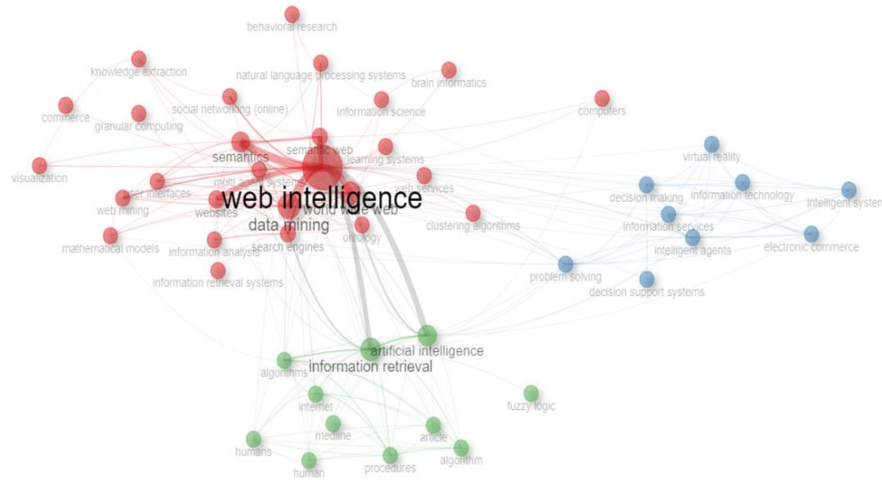

Fig. 7: **Web Intelligence Keywords Co-occurrence Network**

Factorial analysis aims to map a framework's conceptual structure using word co-occurrences in a bibliography. This analysis can use dimensionality reduction techniques like MDS, CA, or MCA. This study shows two clusters of documents that reveal some familiar concepts. Figure 7 shows the conceptual structure map of web intelligence research papers published between 2002 and 2022. Creating it involved using factorial analysis to examine numerous keyword correspondences. The following three clusters emerged: blue, green, and red. The blue cluster has nineteen topics of web intelligence that dwell on regression, neural network, prediction, artificial intelligence, decision making, mean square error, learning systems, data mining, data analytics, machine learning, algorithms, forecasting, and intelligence, while the red cluster has thirty topics of web intelligence that include cohort, article, human, male, big data, adult, aged, nasopharyngeal, nasopharynx, tumour, cancer survival, pathology, carcinoma, overall survival, neoplasms, controlled study, and female.

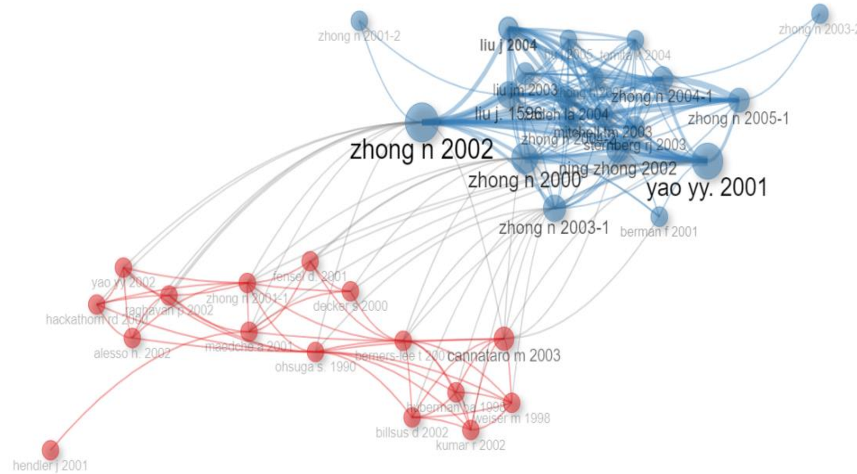

Fig. 8: **Web Intelligence Co-citation Network**

This figure 8 shows two clusters of the co-citation: blue and red. The red clusters comprise Berners-Lee T., Cannataro M., Decker S., Maedche A., Ohsuga S., Raghavan P., Zhong N, Alesso H. and the blue cluster consists of Zhong N., Liu J., Mitchell T. M., Sternberg R. J., Hu J., Tomita K., Berman F., Zadeh L. A. Co-citation occurs when two documents are cited together in one or more subsequent documents. If two published articles show the action of frequently being co-cited, then there is some relationship between them: they talk about the same thing, relate to the same topic, or are relevant to a research area. Besides, Figure 8 shows some relationships between the authors

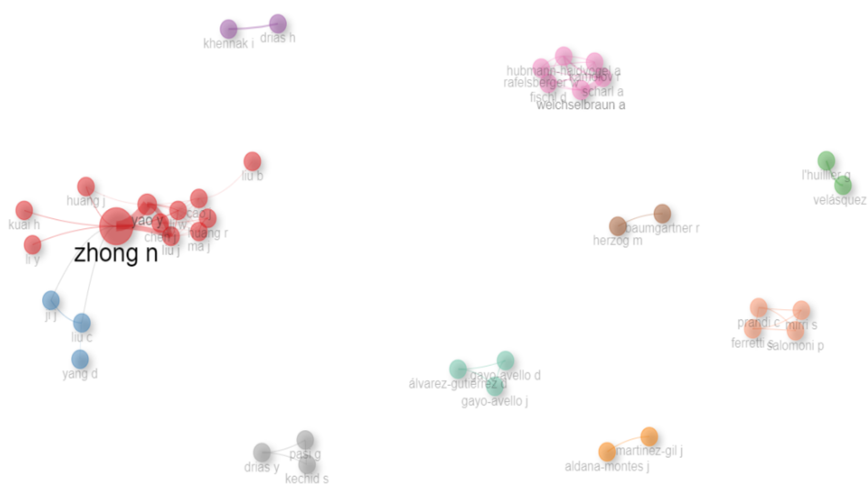

Fig. 9: **Web Intelligence Collaboration Network**

Our analysis of the authors' collaboration network depicted in Figure 9 shows that leading scholars in the field have established robust social connections with colleagues, as indicated by distinct clusters in different colours. Salomoni P., Mirri S., Prandi C., and Ferretti S. are associated with 10 clusters, followed by Gayo-Avello D. and Gayo-Avello J. with 9 clusters. Drias Y., Kechid S., and Pasi G. occupy the third position with 8 clusters. Furthermore, Figure 9 illustrates the scientific collaboration among top-ranked scholars in web intelligence from 2002 to 2022, based on our findings. Nodes represent authors, and links denote co-authorships, a widely accepted measure of scientific collaboration indicating the strength of an author's network.

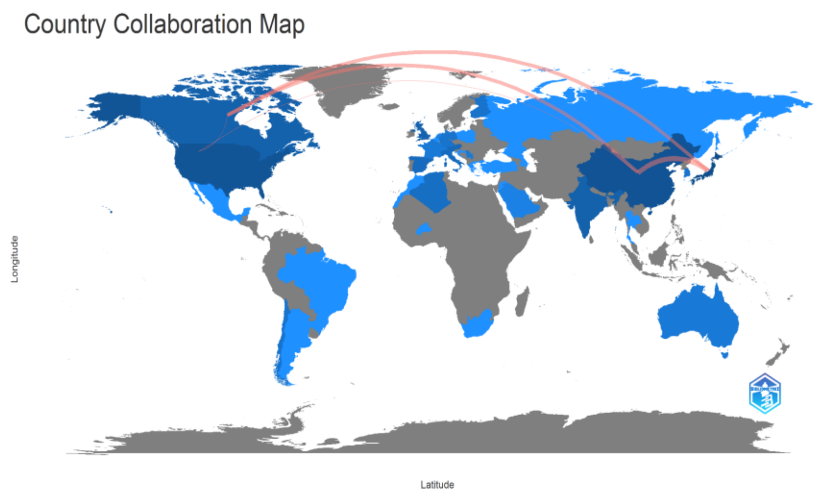

Fig. 10: Web Intelligence Country Collaboration

The partnership between countries highlights the collaborations among prolific writers in web intelligence. Japan and China have the highest collaboration, with four articles. China, Canada, and Japan each have three articles, making them prominent collaborators. Austria and Germany, China and the United States, and the United States and Canada are the third-most collaborative pairs, with two articles each. These countries are among the top producers of web intelligence research, so their collaborations are not surprising (see Figure 10).

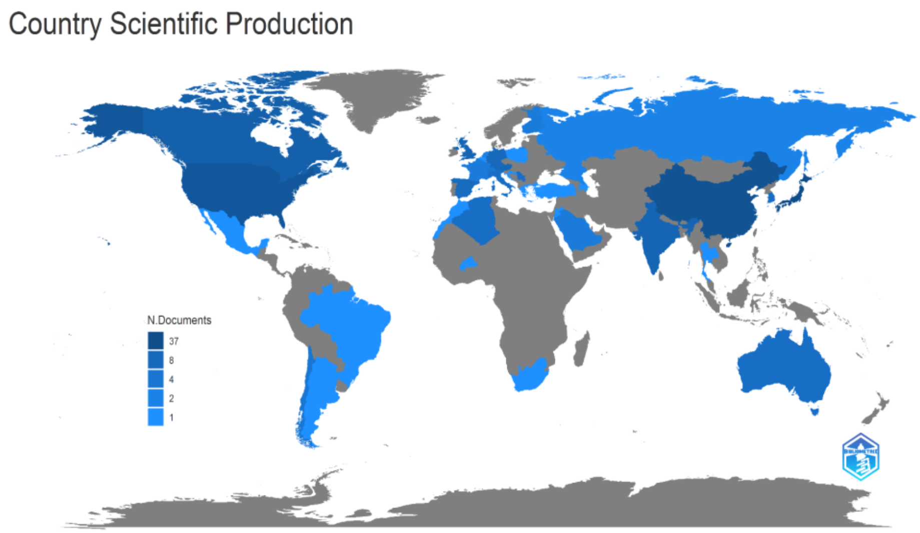

Fig. 11: Web Intelligence Country Scientific Production

The world map in Figure 11 illustrates the global impact and varying levels of engagement in web intelligence across countries and continents. Europe, North America, South America, Asia, Australia, and Africa all show varying degrees of involvement. Examining the impact of web intelligence on each continent is crucial due to its multifunctional role. Thirteen European countries have contributed to the field of web intelligence research. Among the six continents involved, Asia leads with 90 papers, Japan at the forefront with 40 papers, and China with 30 papers.

Japan is the most productive country in Asia's web intelligence research. Europe ranks second in productivity with 65 papers, with the UK leading (10 papers), followed by Austria and Germany (9 papers each), Spain (7 papers), and Italy and Serbia (6 papers each). In South America, Brazil, Chile, and Argentina are active in web intelligence research, with Chile leading (4 papers) and Argentina and Brazil contributing one paper each. North America has 37 papers, with the USA leading (22 papers), followed by Canada and Mexico. Despite the USA's strong research presence, web intelligence is still growing. Africa lags with nine papers, led by Algeria (6 papers) and contributions from Morocco, Burkina Faso, and South Africa (one paper each). Australia has only recorded six papers in total.

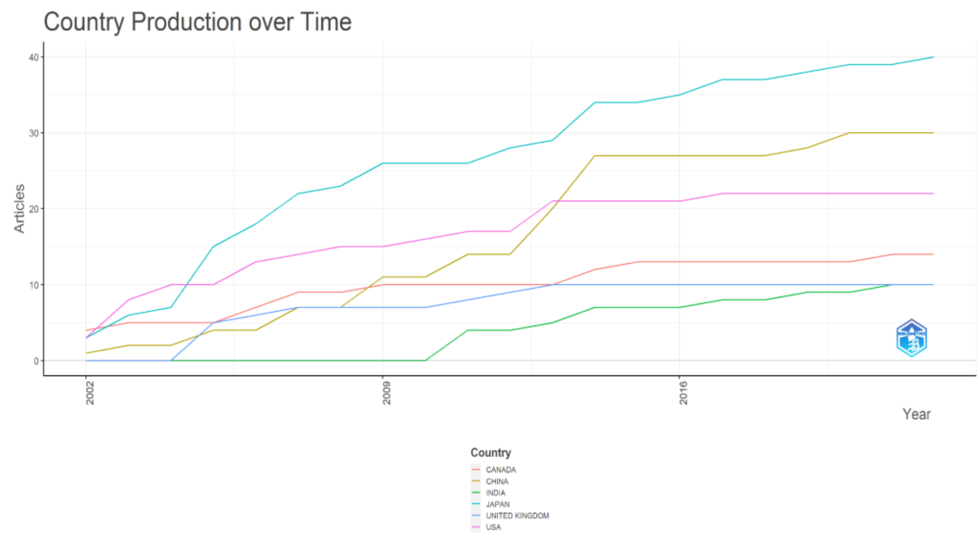

Fig. 12: Web Intelligence Country Production over time

Figure 12 illustrates the production trends of web intelligence articles and shows that for the topmost countries, their research outputs have steadily increased over the years and have sharply increased in recent years. However, African countries still need to make it to the list of the top 10 countries regarding scientific output or the most cited publications on web intelligence research. There could be several reasons why there has been limited research on the use of web intelligence in Africa from 2002 to 2022 across different countries. In 2022, China led with 40 articles, followed by China and the USA with 30 and 22 articles, respectively, from 2020 to 2022. India and the UK had the lowest production, with India from 2002 to 2010 and the UK from 2002 to 2004 (see Figure 12).

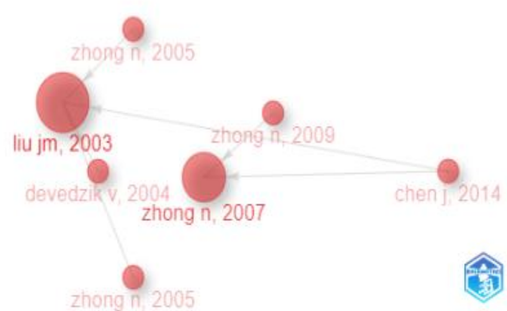

Fig. 13: Web Intelligence Historiography

A historiograph visually represents the chronological development and historical progression of a specific field of study, research topic, or scientific discipline. It illustrates the relationships between key publications, authors, and their contributions over time, showing the evolution and influence of ideas and knowledge. Several authors shared commonalities in their research titles, keywords, and keyword Plus: Liu J.M. (2003) published in J. Intell. Inf. Syst., Devedzik, V. (2004) in Edu Technol. Soc., and Zhong N. (2005) in Modeling Decisions for Artificial Intell. and Rough Set, Fuzzys, Data Mining and Granular Computing, among others. Notably, Zhong N. stands out with multiple publications, appearing more than twice (Figure 13).

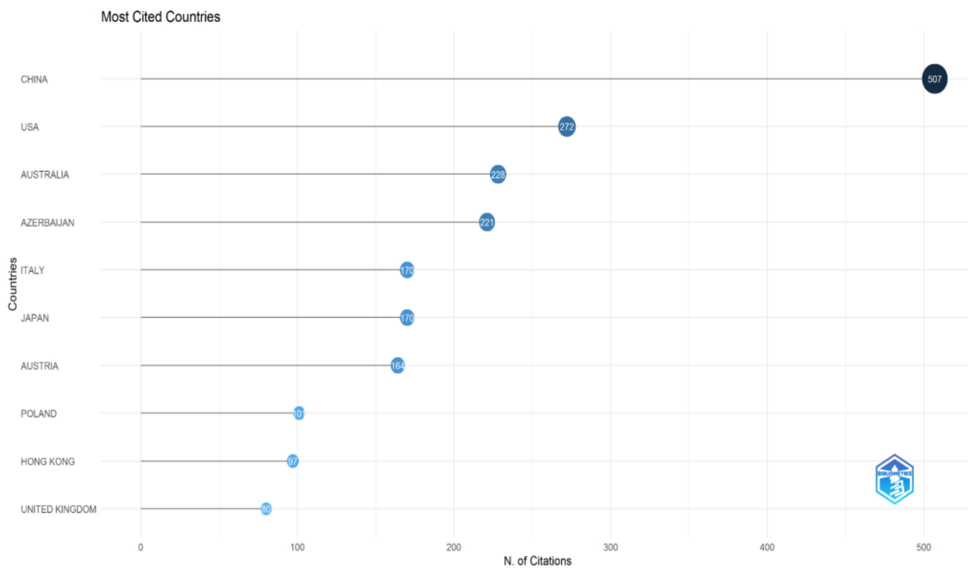

Fig. 14: Web Intelligence Most Cited Countries

Figure 14 indicates that China is the most frequently mentioned country in web intelligence from 2002 to 2022, with 507 citations. The USA follows with 272 citations, and Australia ranks third with 228 citations. China's prominence in web intelligence confirms its reputation as a leading nation in adopting emerging technologies. Additionally, China has been a pioneer in achieving the web intelligence goals set by the United Nations. Despite these accomplishments, China is actively working to improve the spread of web intelligence further [14].

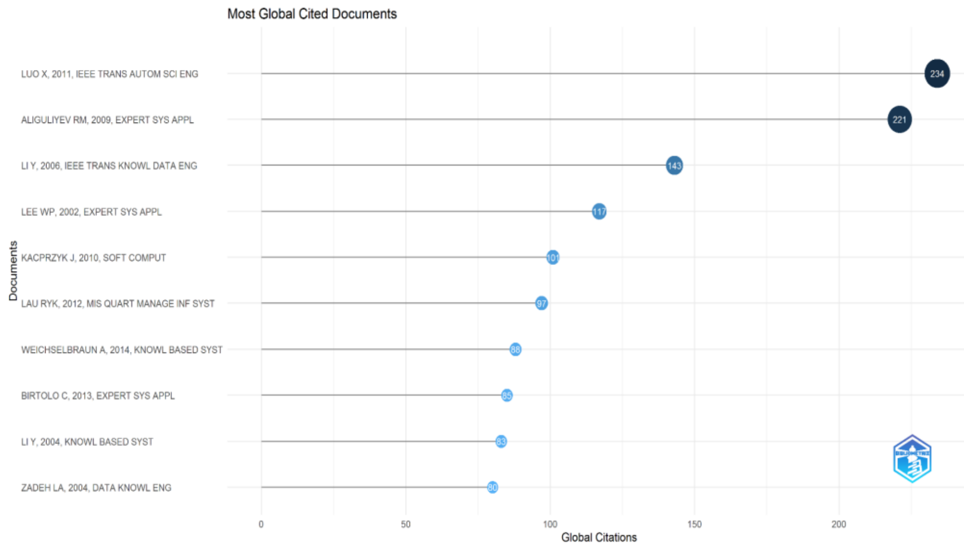

Fig. 15: Web Intelligence Most Global Cited Documents

The most globally cited document in the web intelligence research domain is by Luo X, with 234 citations (Figure 15). This leading paper was published in IEEE Trans. Autom. Sci. Eng. A journal with an impact factor of 5.9. The second most globally cited document is authored by Aliguliyev RM, which recorded 221 citations. The outlet of this paper is the Expert System Application, with an impact factor of 7.5. The third most globally cited document is by Li Y, which received 143 citations. IEEE Trans published the article. Knowl. Data Eng. with an impact factor of 8.9. Our results show that some of the leading authors have published in high-impact journals. Though the impact factor is very relevant, in this study, a lower impact factor (5.9) commands more citations than a higher impact factor (8.9); as explained earlier in this paper, it is attributed to different factors rather than the figures showcased. The results indicate that topic relevance, contest positioning, the topic trends lifespan, and accessibility of the paper count in web intelligence.

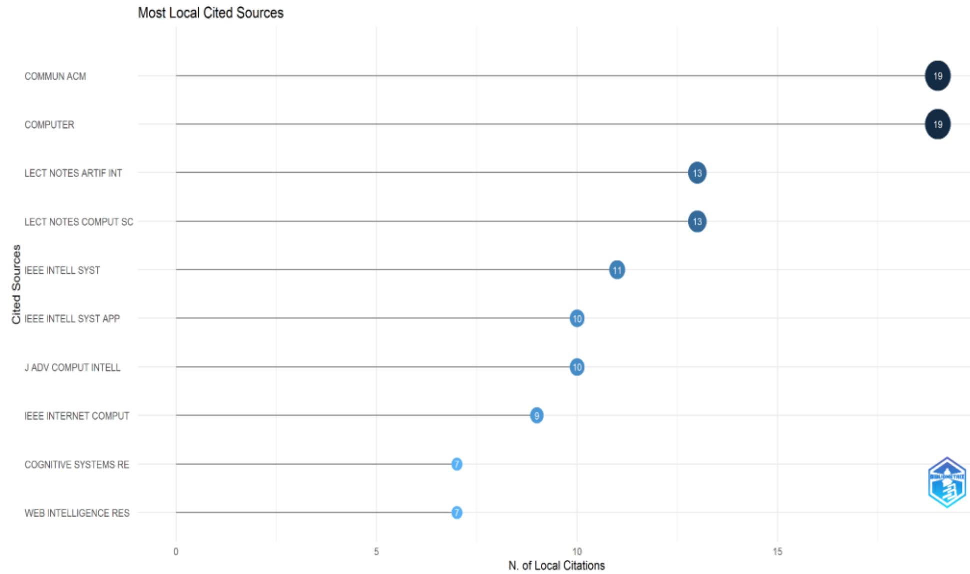

Fig. 16: Web Intelligence Most Local Cited Sources

The most locally cited source in the web intelligence research domain leading paper was published in Communication ACM and Computer with an impact factor of 22.9 and 2.2, respectively, with 19 citations each (Figure 16). The second most locally cited source recorded 13 citations each. The outlet of this paper is the Lecture Notes in Artificial Intelligence. and Lecture Notes in Computer Science. IEEE Intelligent System publishes the third most locally cited source with an impact factor of 6.744, which received 11 citations. The least locally cited source from the ten locally cited sources on web intelligence is Cognitive Systems Research, with an impact factor of 2.1, and Web Intelligence Research, with 7 citations each.

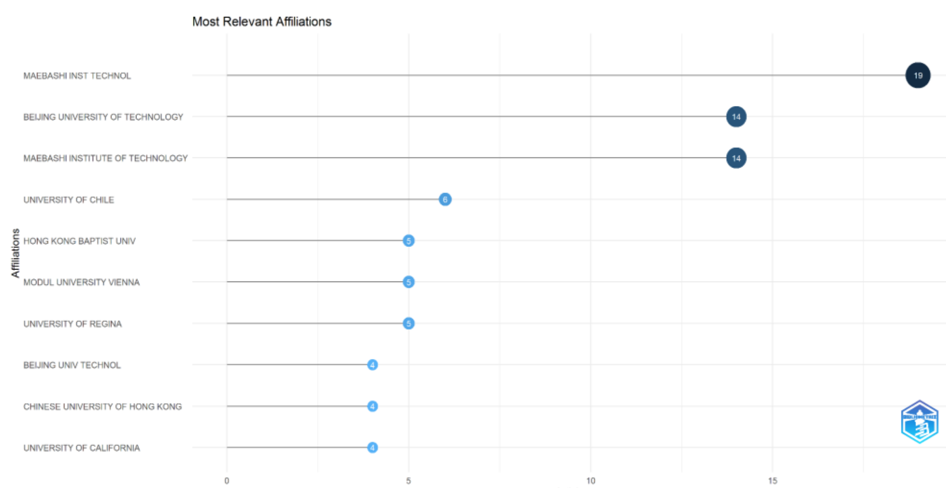

Fig. 17: Web Intelligence Most Relevant Affiliations

According to the data in Figure 17, the leading five universities in the field of web intelligence are the Maebashi Institute of Technology with 19 publications, the Beijing University of Technology and Maebashi Institute of Technology with 14 articles each, the University of Chile with six articles, and Hong Kong Baptist University with five papers. It is not surprising that the Maebashi Institute of Technology is at the top in terms of relevant affiliations. Authors from this university have been actively engaged in web intelligence projects.

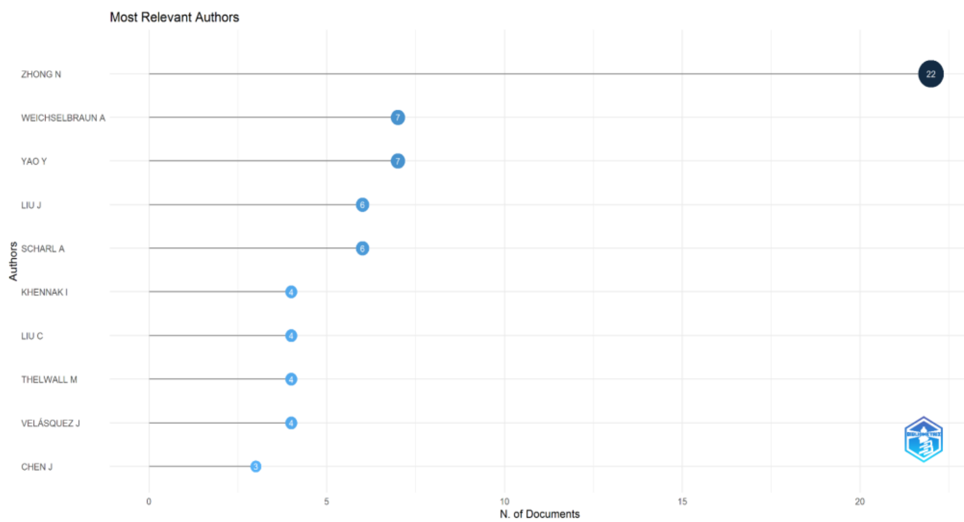

Fig. 18: **Web Intelligence Most Relevant Authors**

Figure 18 provides additional analysis by summarizing the top ten authors in the field of web intelligence from 2002 to 2022. In the figure, the size of each circle represents the number of documents an author release annually. In contrast, the darkness of the circle indicates the number of citations received that year. The count of output years for each author begins with their first publication, resulting in different total production years for each. According to this data, Yao Y and Weichselbraun A, each with seven papers, are the next most prolific authors. Additionally, Scharl A and Khennak I are identified as the third most productive authors in web intelligence, each with six publications. Among the top ten authors, the remaining five have four or three publications each, as shown in Figure 18.

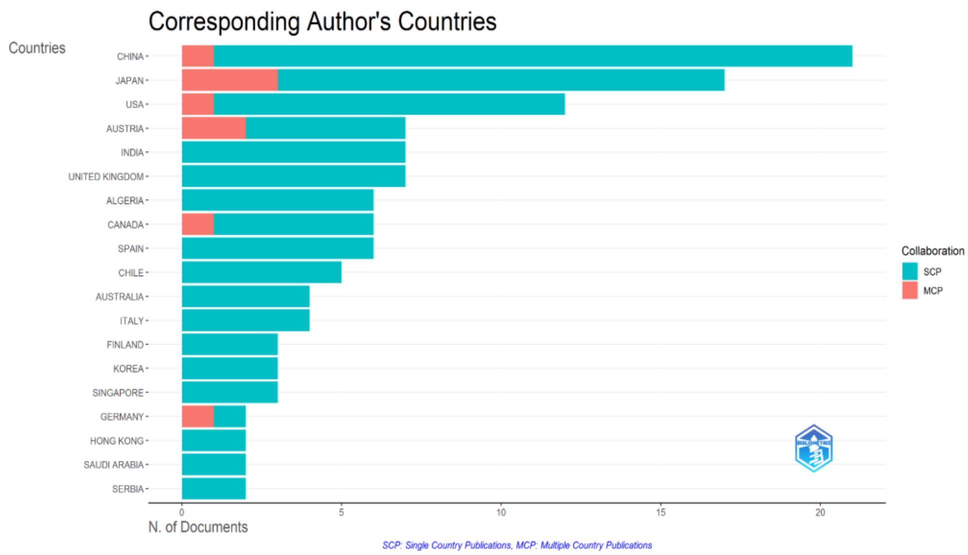

Fig. 19: **Web Intelligence Corresponding Author's Countries**

Figure 19 categorizes the critical contributors to "Global Web Intelligence Research" into single-country publications (SCP) and multiple-country publications (MCP). In SCP, all authors of a document are from the same nation. Conversely, MCP refers to documents with at least one co-author from a different country, indicating international collaboration. China leads the SCP list with 20 papers, followed by Japan with 14, the USA with 11, and Austria with 5. India and the UK have seven articles highlighting significant contributions through domestic collaborations and publications. Japan leads the MCP list with three publications, while Austria has 2, and China and the USA have one each, showcasing notable international cooperation.

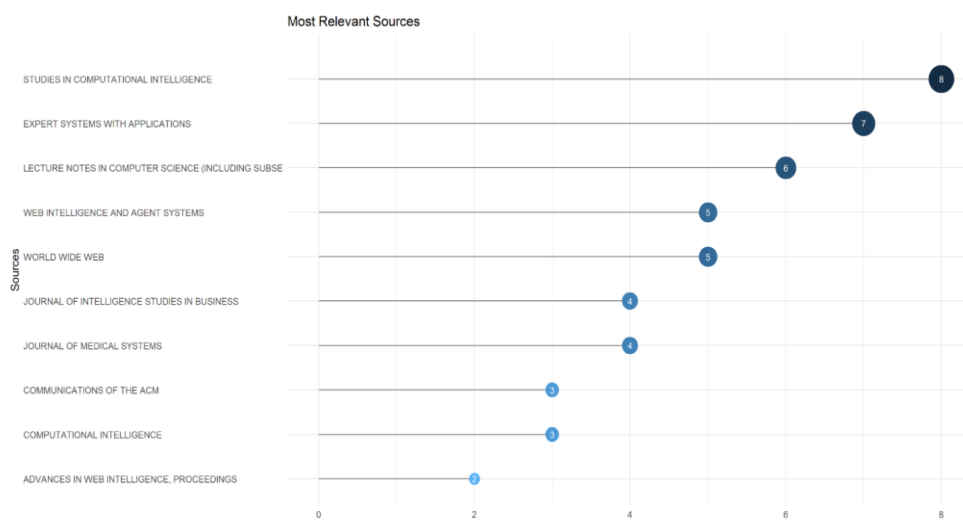

Fig. 20: **Web Intelligence Most Relevant Sources**

Figure 20 summarizes the ten most influential sources in web intelligence from 2002 to 2022. The figure shows that the size of a circle corresponds to the number of documents released annually by each source, while the darkness of the circle indicates the number of citations received that year. The production years for a source start from its first publication. Studies in Computational Intelligence emerge as the most productive source with eight papers, Expert Systems with seven papers, and Lecture Notes in Computer Science with six papers. The analysis also reveals that Advances in Web Intelligence proceedings are the least productive, with only two papers published. Among the top ten sources, the final source also contributed two publications, as depicted in Figure 20.

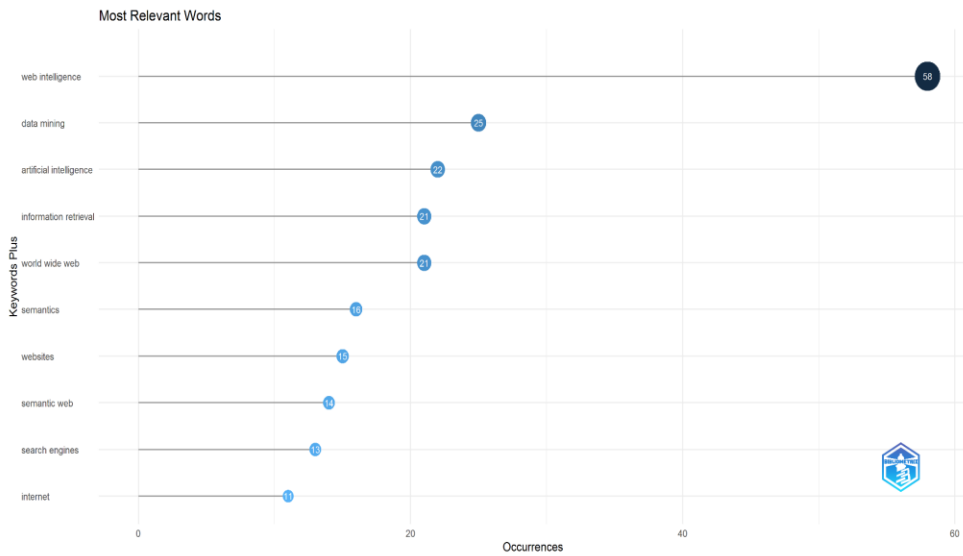

Fig. 21: Web Intelligence Most Relevant Words

From 2002 to 2022, this chart illustrates the most frequently used keywords in research on web intelligence. Among the keywords, web intelligence was the most common, appearing 58 times, followed by data mining, which appeared 25 times. Artificial intelligence comes in third, followed by information retrieval and the World Wide Web, which are mentioned as many as 22 times. The fifth keyword that appears sixteen times is semantics, which is also the fifth keyword. Internet is the keyword used the least, only 11 times (see Figure 21).

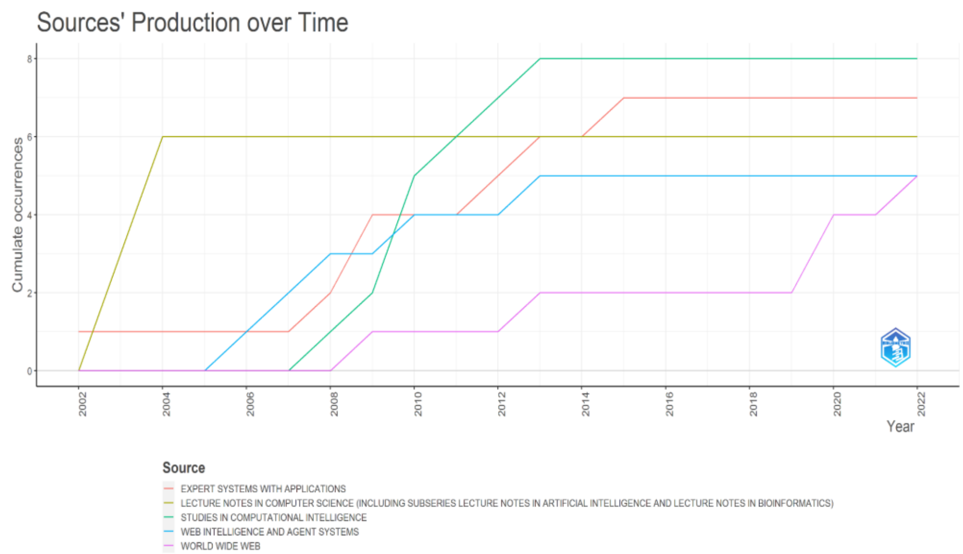

Fig. 22: **Web Intelligence Research,' Production over Time**

Figure 22 shows the sources' production of the most common publication outlets commonly used by authors of web intelligence research, and it demonstrates that the research outputs from all leading sources have been steadily rising over the years, with a notable surge in recent times. However, African countries must appear in the top 10 list for scientific output or most cited publications in web intelligence research. Various factors could explain the limited research on web intelligence in Africa. Most of these appeared in the literature from 2002 to 2022. However, from 2013 to 2022, more articles were produced in all the outlets. Based on the article production, the years coming in second are from 2009 to 2012, while the least production years were from 2002 to 2008.

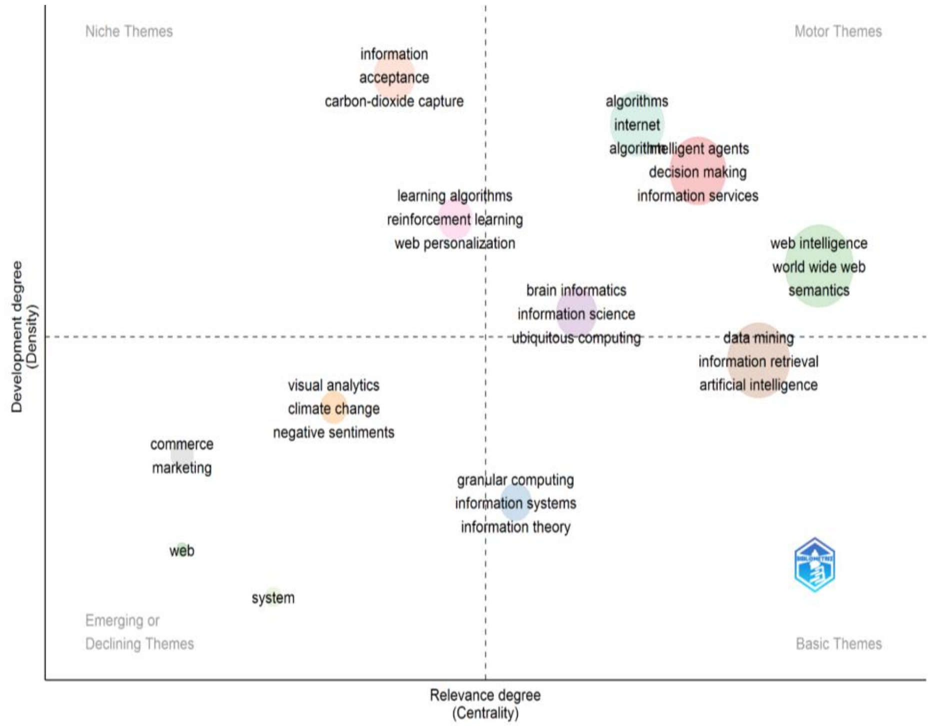

Fig. 23: **Web Intelligence Thematic Map**

Figure 23 uses the following four quadrants to represent the thematic map of the field of web intelligence: the upper right quadrant (Q1), the upper left quadrant (Q2), the lower left quadrant (Q3), and the lower right quadrant (Q4). In web intelligence production, Q1 denotes the “driving themes,” for example, algorithm, internet, algorithm, intelligent agents, information services, web intelligence, world wide web, brain informatics, information science, and semantics. Q2 represents the “much-specialized themes” such as information, acceptance, and carbon-dioxide capture. Q3 signifies the “emerging or disappearing themes,” for example, web, commerce, marketing, system, visual analytics, climate change, and negative sentiments; Q4 denotes the “underlying themes”, such as information retrieval and artificial intelligence.

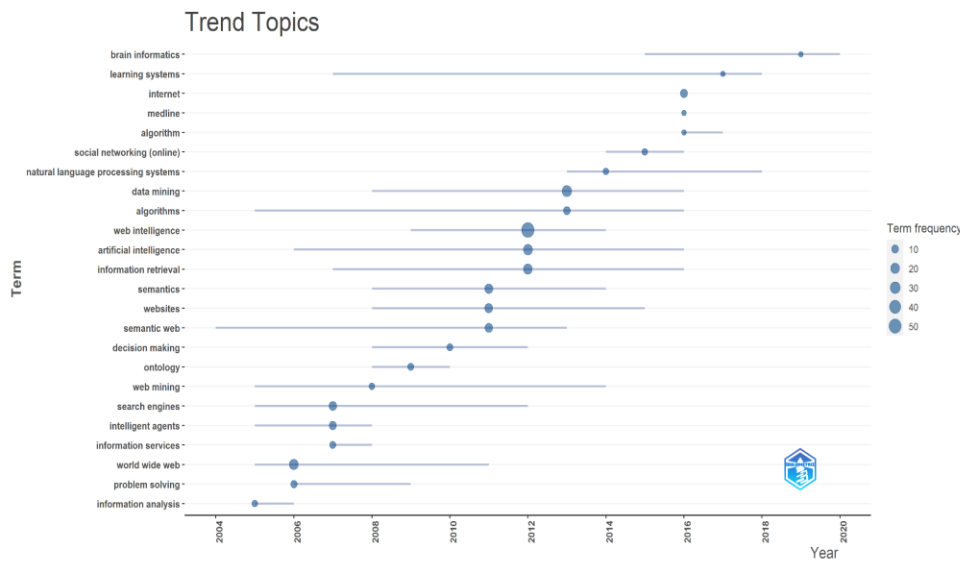

Fig. 24: **Web Intelligence Trend Topics**

Based on the author's keywords from our dataset, Figure 24 gives an overview of the trending topics in the scientific production of research publications in web intelligence from 2002–2022. Further, the topics are presented in a hierarchical arrangement each year. For instance, in 2012, the trending topics were artificial intelligence, web intelligence, and information retrieval. In 2013, data mining and data analytics were trending. It is exciting that some trending topics with higher frequency were predominant between 2005 and 2019, but not extended to 2022. An example is artificial intelligence, with 22 frequencies and trends between 2006 and 2016, and web intelligence, with a trend between 2009 and 2016, with 22 frequencies. Likewise, information retrieval trends between 2007 and 2016 had 21 frequencies.

4 FINDINGS

WI is the research and development of techniques to improve the understanding and exploitation of information resources on the Web [15,16]. WI is essential for the data-based economy, as it helps facilitate the collection, processing, and analysis of vast amounts of unstructured data residing on the Web.

The significant WI techniques are web content mining, web structure mining, and web usage mining. These techniques are useful in extracting meaningful information based on web pages, hyperlink structures, and user interactions. For example, web content mining can gather customer feedback and reviews, social media posts, and other user-generated content and analyze it to realize the data economy [17].

AI technologies, in particular machine learning, natural language processing, and computer vision, have helped gain insights from and automate decision-making in the huge troves of data in the data economy [18]. AI models can detect hidden patterns and trends in data, predict future trends, and respond much like human beings [19].

AI elaborated as used for data-driven applications and services, would critically underpin further innovation and value creation within the data economy. Examples include AI-powered chatbots, predictive analytics, and autonomous decision-making systems that shift ways of doing business in this new age of being data-driven [20].

Data mining involves identifying patterns and correlations in large data sets [21]. Based on the data, various techniques in clustering, classification, regression, and anomaly detection will be applied to discover potentially unknown relationships and business opportunities.

Data mining thus becomes critical in customer segmentation, identification of revenue-generating opportunities, and optimization of operations within the data economy [22]. It allows organizations to uncover otherwise hidden insights, helping them gain an edge over rivals by enabling innovative, data-driven products and services.

Data analytics is simply a process through which data is collected, organized, analyzed, and interpreted to provide actionable insights [23]. It combines statistical methodologies with visualization techniques and business intelligence tools to transform raw material into helpful information that could support decision-making.

Data analytics, being one of the core elements in the data economy, allows organizations to gain competitive advantages through the management and development of data-driven products and services [24]. Examples of how it transforms the data economy include informatics-driven decision-making, performance monitoring, and predictive modeling.

These four technologies are deeply interwoven and intertwined for the data economy

1. Web intelligence serves as a base for picking up and processing the huge volumes of unstructured data on the Web, thereby acting as an essential input to AI, data mining, and data analytics [15,16].

2. AI techniques, such as machine learning and natural language processing, have been utilized for extending the power of web intelligence, data mining, and data analytics toward even more sophisticated data-driven insights and decision-making [18].

3. The web intelligence gathered is fed into data mining techniques to discover patterns, relationships, and oddities that can further be analyzed and interpreted using data analytics [21].

4. Data analytics is one of the tools that enable the insights gained through web intelligence, artificial intelligence, and data mining to be integrated into concrete business strategy, new products, and services, and data-driven decisions within the data economy [23].

The interplay of these four technologies has created synergies that prove very potent, fueling the rapid growth and change in the data economy. They have helped organizations derive more value from data in most aspects of staying competitive within the digital environment [22].

5 LIMITATIONS AND FUTURE RESEARCH

The limitations of this study are that it only included articles in English and data from two databases, and hence, it might need to be more comprehensive in returning all the available literature. Further, excluding other types of documents besides articles may eliminate essential contributions from other modes of scholarly communication. Besides, sticking to 20 years could compress the search for more recent and dynamic research trends. More documents and sources should be added to give a broader view of WI research. Longitudinal studies should be conducted to better understand how the evolving impact of WI will essentially affect the data economy.

Further research is also required in the practical applications of WI techniques with broad application in almost all domains, for example, healthcare, finance, or public policy, to fully realize their potential in the provision of support for the decision-making process. A subject for further study can also cover the place of the most recent technological advances, for example, artificial intelligence and machine learning, on the frontiers of WI. By addressing these areas, future studies will achieve depth and solidify the contributions needed to develop solid frameworks for leveraging WI in the data economy.

6 CONCLUSION

The paper presents a comprehensive review of the development of WI during the past twenty years, tracing theme shifts, influential papers, and hot topics. In this regard, the bibliometric analyses prove that WI is central to the economy of data in facilitating the extraction of meaningful insights from substantial web data, driving innovations in many sectors, such as e-commerce and social media. While WI has made significant advances recently, there are gaps, especially in longitudinal analyses and the practical application of WI research. The present paper contributes to a call for further ongoing research to examine the effect of WI on the data economy and formulate robust frameworks for its application. This study enhances the frontline of present wisdom in WI and thus provides impetus toward creating innovative solutions that utilize the full potential of the data economy to guide future research directions and update current trends in this field.

7 REFERENCES

1. Kuai, H., Tao, X., & Zhong, N. (2022). Web Intelligence meets Brain Informatics: Towards the future of artificial intelligence in the connected world. *World Wide Web*, *25*(3), 1223-1241.
2. Weichselbraun, A., Steixner, J., Braşoveanu, A. M., Scharl, A., Göbel, M., & Nixon, L. J. (2022). Automatic expansion of domain-specific affective models for web intelligence applications. *Cognitive Computation*, *14*(1), 228-245.
3. Fosci, P., & Psaila, G. (2022). Towards Soft Web Intelligence by Collecting and Processing JSON Data Sets from Web Sources. In *WEBIST* (pp. 302-313).
4. Gui, Y., Gan, W., Chen, Y., & Wu, Y. (2022, April). Mining with rarity for web intelligence. In *Companion Proceedings of the Web Conference 2022* (pp. 973-981).
5. Zhong, N., Liu, J., & Yao, Y. (2024). Web intelligence: In search of a better connected world. In *Web Intelligence* (No. Preprint, pp. 1-9). IOS Press.
6. Leung, C. K., Madill, E. W., & Singh, S. P. (2021, December). A web intelligence solution to support recommendations from the web. In *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* (pp. 160-167).
7. Domingos, D. A., Moura, V. E. S., Junior, A. F. L. J., & Lobato, F. M. F. (2024). Web Intelligence Journal in perspective: an analysis of its two decades trajectory. *arXiv preprint arXiv:2405.05129*.
8. Alquhtani, S. A., & Muniasamy, M. A. (2022). Development of Effective Electronic Customer Relationship Management (ECRM) Model by the Applications of Web Intelligence Analytics. In Building a Brand Image Through Electronic Customer Relationship Management (pp. 44-63). IGI Global.
9. Mejia-Escobar, C., Cazorla, M., & Martinez-Martin, E. (2023). A Large Visual, Qualitative, and Quantitative Dataset for Web Intelligence Applications. *Computational Intelligence and Neuroscience*, *2023*(1), 1094823.
10. Olaleye, S.A. Visualizing cultural emotional intelligence literature: a bibliometric review 2001–2020. In:Laine, P., Némethová, I., Wiwczaroski, T. (eds.) Inter-cultural Competence at Work. Seinäjoki: Seinäjoen ammattikorkeakoulu. Publications of Seinäjoki University of Applied Sciences B. Reports, 2020: 160, pp. 142–156.https://urn.fi/URN:NBN:fi-fe20201215100768.
11. Agjei, R. O., Adusei-Mensah, F., Balogun, O. S., & Olaleye, S. A. (2023). The Bibliometric Global Overview of COVID-19 Vaccination.

In *International Conference on Intelligent Systems Design and Applications* (pp. 287-298). Springer, Cham.

12. Olaleye, S. A., Balogun, O. S., & Adusei-Mensah, F. (2023). Bibliometric structured review of tuberculosis in Nigeria. *African Health Sciences*, *23*(2), 139-60.
13. Olaleye, S., Olaleye, E., Balogun, M., & Balogun, O. (2023). Global spotlight of students and teachers wellbeing. A bibliometric viewpoint. 15th International Conference on Education and New Learning Technologies, Palma, Mallorca, Spain, 3658 – 3666.
14. Wei, L. (2025). A Chinese perspective on AI development. Horizons, (30). Retrieved from https://www.cirsd.org/en/horizons/horizons-spring-2025--issue-no-30/a-chinese-perspective-on-ai-development
15. Zhong, N., Liu, J., & Yao, Y. (2003). *Web intelligence*. Springer Science & Business Media.
16. Zhong, N., Liu, J., Yao, Y., Wu, J., Lu, S., Qin, Y., ... & Wah, B. (2007). Web intelligence meets brain informatics. In *Web Intelligence Meets Brain Informatics: First WICI International Workshop, WImBI 2006, Beijing, China, December 15-16, 2006, Revised Selected and Invited Papers 1* (pp. 1-31). Springer Berlin Heidelberg.
17. Xu, Z., Cheng, X., Zhou, L., & Yao, L. (2022). Web Content Mining: Techniques and Applications. Information Systems, 62, 101--112.
18. Agrawal, A., Gans, J., & Goldfarb, A. (2022). *Prediction Machines, Updated and Expanded: The Simple Economics of Artificial Intelligence*. Harvard Business Press.
19. Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business horizons*, *62*(1), 15-25.
20. Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.
21. Mining, W. I. D. (2006). Data mining: Concepts and techniques. *Morgan Kaufinann*, *10*(559-569), 4.
22. Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
23. Davenport, T., & Harris, J. (2017). *Competing on analytics: Updated, with a new introduction: The new science of winning*. Harvard Business Press.
24. Provost, F., & Fawcett, T. (2013). *Data Science for Business: What you need to know about data mining and data-analytic thinking.* " O'Reilly Media, Inc.".

8 AUTHORS

Sunday Adewale Olaleye is a Senior Lecturer at Jamk University of Applied Sciences, Finland got his Doctor of Science (D.Sc.) in Marketing from the University of Oulu, Oulu Business School (AACSB), Finland. He received his Master of Science in Information Systems from Abo Akademi University, Turku, Finland; his MBA from the Lapland University of Applied Sciences, Tornio, Finland; his NMS iICT Certificate, Innovation, and Entrepreneurship from the Nordic Master School of Innovative ICT, Turku Centre for Computer Science (TUCS), Turku, Finland; and his Certificate of Leadership and Management in Health from the University of Washington, USA. He had his post-doctoral studies at the University of Oulu in Finland, and he is a visiting professor at the Universidad de las Américas Puebla (UDLAP) in Mexico. He has presented papers at conferences and published in academic journals.

Oluwafemi Samson Balogun is a former lecturer at the Department of Statistics and Operations Research, Modibbo Adama University of Technology, Yola, Adamawa State, Nigeria. He is an International Visiting Scholar at the Department of Mathematics, Covenant University, Ota, Ogun State, Nigeria and he is currently a post-doctoral researcher at the School of Computing, University of Eastern Finland. He holds a Ph.D. in Statistics from the Department of Statistics, University of Ilorin, Kwara State, Nigeria. He is a senior research fellow at the Centre for Multidisciplinary Research and Innovation (CEMRI) and He is a member of the International Biometric Society (IBS), International Statistical Institute (ISI), Nigerian Statistical Association (NSA), Nigerian Mathematical Society (NMS). He has several conference and journal papers in reputable international bodies, and he has also served as a guest reviewer for several journals. His research interests are data science, machine learning, data mining, biostatistics, categorical data analysis, modeling, and probability distribution models.