



Digital Humanities: A Decade of Growth and Interdisciplinary Collaboration

Rahat Gulzar
Aasia Maqbool
Farzana Gulzar
Iqra Shafi
Sangita Gupta

Abstract

This study investigates the evolution and growth of digital humanities research from 2012 to 2022 using data from the Web of Science, analysed through Orange Data Mining, MS Excel, VOSviewer, and Microsoft Office tools. It examines key research areas, thematic trends, top contributing countries, and interdisciplinary linkages, revealing a steady increase in publications since 2002, with dominant themes including digital humanities, libraries, databases, archives, mathematics, and neurosciences. The United States leads global contributions, followed by England and Germany, with diverse journals across literature, arts, humanities, history, and computer science publishing in this field. As a novel work, the study highlights global participation and the integration of digital humanities into multiple academic disciplines, offering valuable insights for researchers and experts across domains.

Keywords: *Digital humanities, humanities, digital, research, interdisciplinary, libraries, subject domain, technologies*

Introduction

Digital humanities, also known as DH, is a field of study that has been developing for more than seven decades and is centred on humanity computing. As a result, the definitions of digital humanities are shifting in tandem with the development of the discipline (Su et al. 2023). DH is defined as an interdisciplinary field of study, research, teaching, and invention that focuses on the intersection of computing and humanities disciplines, according to this study. According to Rockwell and Passarotti (2019), people consider Roberto Busa to be a pioneer in the field of digital humanities because of his work with IBM in the 1940s. At this point, the area of digital humanities has already encompassed a wide variety of fields, including statistical linguistics, digital history, and library statistics. A specific subject, field, or scientific concept is the focus of bibliometric studies, which analyse quantitative research publications on that subject. The analysis of research trends in specific domains or themes is currently possible because of the availability of dependable statistical and bibliometric methods as well as indexing databases. New research domains are analysed by bibliometricians, who then produce tabulated results utilising widely known metrics and data that can be relied upon. As is the case with digital humanities, a significant number of bibliometric studies have previously been carried out to gain a deeper comprehension of the research trend that

is occurring within the subject. In the process of shifting from the traditional humanities to the digital humanities, one of the most important aspects is the use of bibliometric, statistical, and scientific instruments. As the twenty-first century has progressed, the digital humanities have transformed from an academic frontier into an independent field. Numerous educational institutions now provide undergraduate, master's, and doctorate programmes in digital humanities, and the amount of research conducted in academic institutions has increased. The discipline of digital humanities is defined by its openness, collaboration, experimentation, and diversification, which reflects the tremendous educational value and societal utility of the field, as well as its promotion of knowledge innovation and knowledge service (Spiro, 2016). Among the many activities and research topics that are conducted in this area are the archiving of historical materials, the study of media and art, the evaluation of infrastructure, the mining of data, and the planning of the most effective instructional methods. According to the Planning and Review Committee (2014), digital humanities is one of the seven developmental trends of future academic libraries. This trend involves the application of digital technology to the research of various disciplines, including history, geography, linguistics, literature, and others. This study has a significant impact on the research content and progress of the humanities subject area.

Objectives

Because of the interdisciplinary scope of Digital Humanities (DH), the main objective of this research is to identify the association of digital humanities with other subject fields and perceive the global participation with research trends of the same so that it can be known to which subject fields this discipline relates to and what will be the future of the Digital Humanities in the field of research. To achieve these objectives, the researcher carried out the scientometric analysis and pinpointed some objectives as:

- To see the evolution and annual growth of the digital humanities literature with time.
- To identify the major research areas where extensive research on digital humanities is carried out.
- To identify and illustrate different themes around which the whole literature related to digital humanities revolves.
- To identify the top contributing countries in digital humanities research.
- To identify and illustrate the subject domains of the web of science, among which the digital humanities are accommodated.

Review of Literature

Schreibman et al. (2015) argue that the digital humanities should be recognised as a well-established discipline due to the existence of its

scholarly journals and conferences. Smith (2002) found that the Index Thomisticus was among the earliest sources of humanities available in electronic format. Dalbello (2011) discovered that the origins of digital humanities can be traced back to the 1950s, specifically with the commencement of Index Thomisticus. This was further propelled by technical advancements in the 1980s. According to Wang & Inaba (2009), digital humanities are characterised by their methodological nature and multidisciplinary breadth. A co-word analysis and correspondence analysis were conducted to examine the development of digital humanities. That study explored the recent developments in the field of digital humanities and analysed the organisational framework and semantic aspects of digital humanities. Sula (2023) examined the correlation between digital humanities and libraries by analysing publications in library science-specific journals. Specifically, focused on publications related to digital humanities in the “Library, Information Science and Technology Abstract” database. The study revealed a significant increase in the number of publications related to digital humanities starting from 2005, with the number nearly doubling by 2012. Yang et al. (2020) provided a quantitative analysis that contributes to the continuing debate surrounding the transdisciplinary nature of research in the field of Digital Humanities. An examination of bibliometric data from published publications in the field of digital humanities is conducted to examine the structure and patterns of collaborative work across different disciplines, as well as the changing overall trends. The findings indicate that the fields involved in DH research exhibit a wide range of disciplines, yet there is an uneven distribution among these disciplines. The various disciplines associated with DH research are consolidated into communities focused on a select few essential disciplines, leading to the emergence of multiple interconnected study areas and disciplines within DH research. The expanding network of disciplines provides support for the transdisciplinary character of DH. This study utilises bibliometric analysis and visualisation methods to investigate the structure and patterns of transdisciplinary collaboration in the field of digital humanities. The study’s findings illustrate the interdisciplinary nature of digital humanities research, which is motivated by its wide-ranging relevance across several sectors (Amanullah, 2023; Ahmad et al., 2019; Bhadauria & Singh, 2022; Das et al., 2023). In addition, Wang (2018) performed a retrospective bibliometric analysis of writings related to digital humanities, which is a growing and interdisciplinary field. This study examines the dissemination of research findings and languages used, analyses the network of categories and cited references, and assesses the prevailing research topics. The data used is sourced from the Web of Science (WoS) core collection. To depict the general geographical features of this region, VOSviewer and CiteSpace, two famous visualisation tools, were employed to analyse bibliographic data including authorship,

categorisation, references, and keywords. In addition, the parameter design of the visualisation tools follows the standard methods and methodologies for bibliometric analysis. There is a noticeable increase in the study of digital humanities. English continues to be the predominant academic language in this field. The field of digital humanities emerges from the dynamic interplay between humanistic study and digital technology. This research field is closely interconnected with history, literary and cultural heritage, and information and library science (Given & Willson, 2018). A study titled “The Evolution of Digital Humanities in China” was undertaken by (Wang et al., 2020), revealing that Digital Humanities (DH) has experienced significant growth as a research field in China over the past decade. An examination of historical records and bibliometric data allows for an assessment of the origins and progress of digital humanities (DH) studies in China. During the initial phase, Digital Humanities (DH) was recently introduced to China, with a focus on exploring its meaning and experienced a diverse range of DH research and practical applications in areas such as library and information science, linguistics, history, literature, art, and cultural heritage preservation. Currently, Digital Humanities (DH) in China is experiencing a phase of swift advancement, resulting in significant transformations in Chinese academic research through the use of novel techniques, tools, platforms, and research frameworks.

Methodology

For the current study, data has been extracted from one of the popular indexing databases, “Web of Science”. Web of Science core collection has been used, and the search term “Digital Humanities” was put in the search box as the topic. Period filtration was used, and the data was extracted for 10 years, starting from 2012 till 2022. Tools like Orange Datamining Software and MS Excel were used to extract accurate and needful data in the form of tables for further comparison and analysis so that objectives could be achieved. The visualisation software VOSviewer was used to display visualisation maps, in addition to Microsoft Office tools.

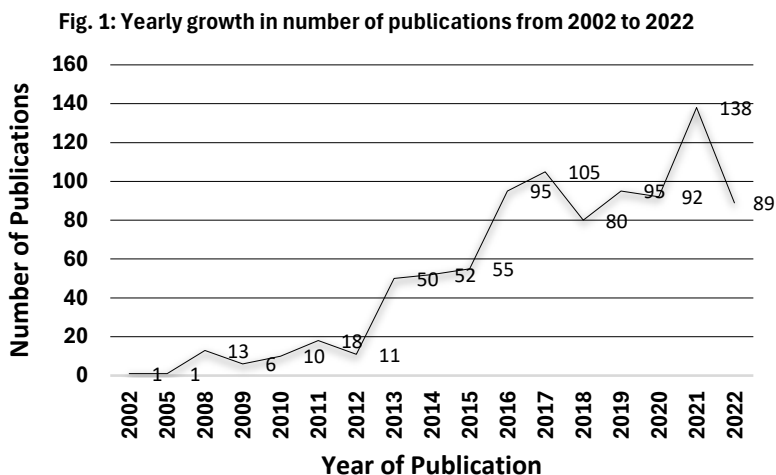
Analysis and Discussion

The search for the articles was confined to the “Arts and Humanities Citation Index” of Web of Science (WoS). The “Digital Humanities” search term was executed using the field option – Topic. The Topic field option searches the given term in the title, abstract, author keywords, and Keywords Plus of the indexed resources. The search for the given term was limited till 2022, and the data were downloaded for further execution and analysis. A total of 912 articles related to “Digital Humanities” were retrieved, and the refinement and pre-processing of this data were carried out using various software and tools, such as MS Excel and Orange data mining software. The refined data

were then migrated to VOSviewer Data Visualization Software for the visualisation of the keyword clusters and analysis of author and country co-occurrences. To ascertain the main themes covered under “Digital Humanities”, Word Clouds were generated with the aid of Orange Data Mining Software. Moreover, the data was also treated with the aid of MS Excel to create heatmaps to ascertain the subjective/thematic density revolve.

Year wise growth

The data extracted from WoS against the query and selected filters reveal that the publications started in 2002 and increased over time. There has been an increasing trend since 2005, and the year 2021 has the highest score of publications (138), followed by the year 2017. It is evident from Fig. 1 that the research on Digital Humanities is increasing yearly, which depicts that the literature on “Digital Humanities” is growing with time, accelerating its acquaintance and worth globally.



Subject clusters

Fig. 2 Subject Clusters deciphering collaborations and associations among subjects.

A total of 59 items distributed among twelve keyword clusters were created from the extracted data with the help of VOSViewer Data visualisation software. Firstly, the dataset was migrated to VOSViewer through the import column of WOS, which is specially meant for the dataset of the Web of Science. A map was created based on bibliographic data; the selected source for the data is “read data from bibliographic database files”, which support databases like WOS. After selecting the file, the analysis started. In

the analytical and counting part (Fig. 3), the type of analysis was selected as co-occurrence, and the unit for analysis was selected as all keywords(author keywords + keyWords plus). The counting method was set as full counting. The analysis filters the most dominant subject themes of the dataset, which are 59 in number and clustered among 12 big clusters (Fig. 2). The leading cluster is digital humanities and other clusters are libraries, databases, human methods, digital research, mathematics, archives, psychological methodology and neurosciences. This illustrates that Digital Humanities are used and discussed in these topics. Based on occurrences and co-occurrences, the items are positioned in the respective suitable clusters. The network map also displays the relation among these subjects by creating a network among these items. These links are created based on their co-occurrences and clusters. and the strength of the link shows the strong relationships among the subjects.

Fig. 2: Subject Clusters deciphering collaborations and associations among subjects

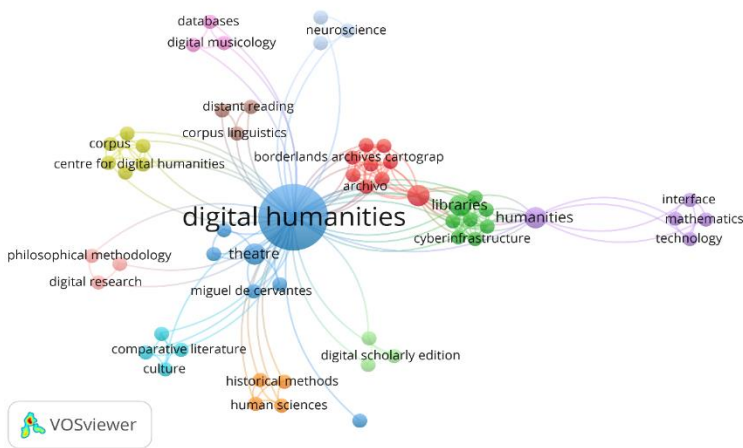
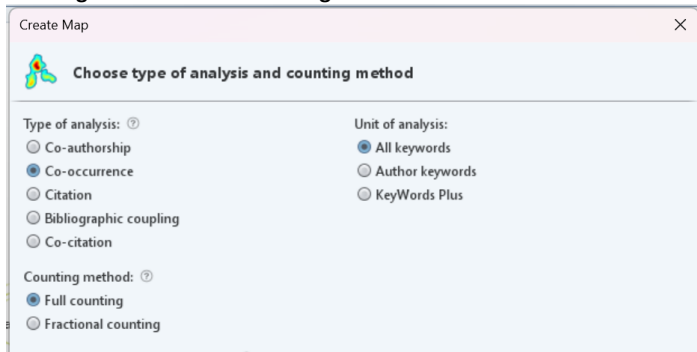


Fig. 3: Selection of counting methods in VosViewer software



Word Clouds

To reveal and illustrate the hot research topics, the dataset was migrated to the Orange data mining software. A Corpus of three columns was selected for further operations. The selected three columns were Article Title, Author Keywords and Keyword Plus. Pre-processing was done using filters, which helped remove the stopword list of commonly used terms and helping verbs. Document frequency was set to 70-10000. The frequency was set to 70, which instructed the tool to consider only those terms that had a minimum of 70 instances. This increases the soundness of that word by making it more reliable and considerable. POS tagger was used to make it more delicate. N-grams frequency was set as 2-2, commonly called bigrams. A bigram is a sequence of two adjacent words in a text. A word cloud of 60 keywords was created. It is evident from Fig. 4 that Digital Humanities has the highest centrality (the most frequently used keyword) with 722 occurrences with a strong relationship with other keywords, especially humanities digital, history digital, digital data, humanities history, digital research, and data analysis. It was also found that some terms in the word cloud belong to more than one n-gram. The word cloud illustrates the set of themes around which the whole literature related to digital humanities revolves. The authors frequently used these topics in the title, topic, and keywords (including keyword plus) related to their work.

Fig. 4 Word Cloud of themes revolving around Digital Humanities



Country wise contribution

The publications on Digital Humanities originate from 61 different countries. Fig. 5 illustrates the global distribution of the contributions of these countries. A total of 43 countries (70.49%) produce less than ten

publications, 16 countries (26.22%) produce between 11 and 50 publications, and 5 % of countries or territories have produced more than 80 publications. The USA produces the highest number of publications. Table 1 illustrates the top 20 most productive countries and publications. Countries with economic development are those that contribute and provide scientifically and academically. According to data, the USA has 34.7% of publications (n=317/912), England 12.6% (n=115/912), and Germany 9% (n=82/912).

Fig. 5 Geographical location of Contributors

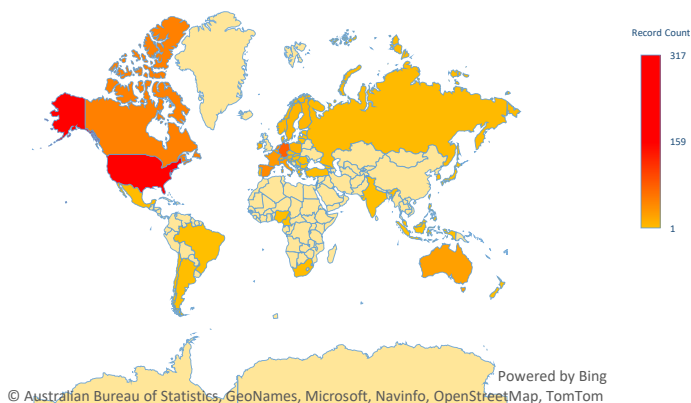


Table 1: Top twenty Contributors

Rank	Country	No. of publications	Percentage
1	United Kingdom	104	20.08
2	United States of America	102	19.69
3	Germany	32	6.18
4	Japan	20	3.86
5	Canada	17	3.28
6	Australia	16	3.09
7	Italy	14	2.7
8	Turkey	14	2.7
9	Austria	12	2.32
10	Croatia	12	2.32
11	France	12	2.32
12	Poland	12	2.32
13	Indonesia	10	1.93
14	China	9	1.74
15	Netherlands	9	1.74
16	South Africa	8	1.54

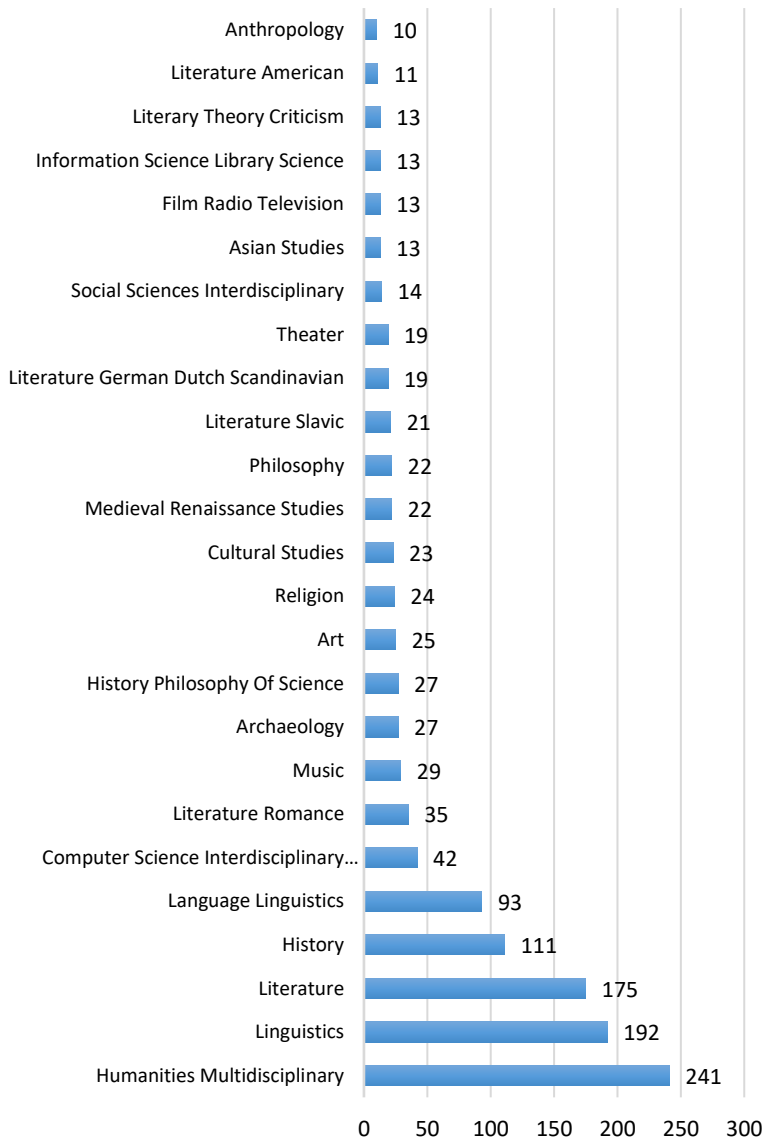
Rank	Country	No. of publications	Percentage
17	Argentina	7	1.35
18	Serbia	7	1.35
19	Brazil	6	1.16
20	Spain	6	1.16
	Others	89	17.18
Total		518	100

Subject Domains

The literature related to digital humanities is published under different subject categories of Web of Science. The data relating to our query was collected from 13 subject categories. It is evident from Fig. 6 that the humanities multidisciplinary category encompasses a huge amount of literature on digital humanities, followed by literature, literature linguistics, archaeology, and other categories. Film, Radio, Television, Literary Theory Criticism and Anthropology cover the least literature on digital humanities. Humanities Multidisciplinary is a broader subject that covers areas like art, history, language, literature, music, philosophy, and classics. These are the seven disciplines of the humanities, which make it a wide subject category that covers a large literature area.

Conclusion

This study utilised analysis techniques to investigate the current research trends in the DH research area from the years 2002 to 2022. The study demonstrates that there is a consistent increase in the amount of research that is published in DH literature. The majority of publications are typically communicated through articles that are published in the humanities interdisciplinary and literature. The United Kingdom continues to be the most productive nation, followed by the United States of America, Germany, Japan, and Canada. The primary keywords are Digital Humanities, Humanities Digital, History Digital, Digital Data, Humanities History, Digital Research, and Data Analysis, which describes the research performed in the Digital Humanities literature. The leading cluster, according to associations is digital humanities, followed by libraries, databases, human methods, and digital research. Countries with economic development are those that contribute and provide scientifically and academically, as the United Kingdom produces the highest number of publications, followed by the USA and Germany.

Fig. 6 Number of Publications under different Subject domains

References

- Ahmad, K., JianMing, Z., & Rafi, M. (2019). Assessing the literature of knowledge management (KM) in the field of library and information science. *Information Discovery and Delivery*, 47(1), 35–41. DOI: 10.1108/idd-06-2018-0021

- Amanullah, A. (2023). An investigation in the interdisciplinary nature of digital humanities. *DESIDOC Journal of Library & Information Technology*, 43(04), 234–240. DOI: 10.14429/djlit.43.04.19229
- Bhadauria, S., & Singh, V. (2022). Blending absorptive capacity with open innovation: A bibliometric review. *Benchmarking: An International Journal*, 30(4), 1110–1136. DOI: 10.1108/bij-02-2021-0111
- Dalbello, M. (2011). A genealogy of digital humanities. *Journal of Documentation*, 67(3), 480–506. DOI: 10.1108/00220411111124550
- Das, A., Kondasani, R. K. R., & Deb, R. (2023). Religious tourism: A bibliometric and network analysis. *Tourism Review*, 79(3), 622–634. DOI: 10.1108/tr-08-2022-0387
- Given, L. M., & Willson, R. (2018). Information technology and the humanities scholar: Documenting digital research practices. *Journal of the Association for Information Science and Technology*, 69(6), 807–819. DOI: 10.1002/asi.24008
- Planning, A. R. & Review Committee. (2014, June 1). Top trends in academic libraries: A review of the trends and issues affecting academic libraries in higher education. *College & Research Libraries News*. <https://crln.acrl.org/index.php/crlnews/article/view/9137/10062>
- Rockwell, G., & Passarotti, M. (2019). The index thomisticus as a big data project. *Umanistica Digitale*, 5. DOI: 10.6092/issn.2532-8816/8575
- Schreibman, S., Siemens, R., & Unsworth, J. (2015). A new companion to digital humanities (pp. 98–105). John Wiley & Sons.
- Smith, M. N. (2002). Computing: What's American literary study got to do with IT? *American Literature*, 74(4), 833–857. DOI: 10.1215/00029831-74-4-833
- Spiro, L. (2016). "This Is Why We Fight": Defining the Values of the Digital Humanities. In *Debates in the Digital Humanities 2016* (pp. 16–35). University of Minnesota Press. <https://www.jstor.org/stable/10.5749/j.ctttv8hq.6>
- Su, F., Li, S., Liu, Q., & Zhang, Y. (2023). A bibliometric study of digital humanities research in China from 2012 to 2021. *Information Research an International Electronic Journal*, 28(4), 62–82. DOI: 10.47989/ir284336
- Sula, C. A. (2013). Digital humanities and libraries: A conceptual model. *Journal of Library Administration*, 53(1), 10–26. DOI: 10.1080/01930826.2013.756680
- Wang, Q. (2018). Distribution features and intellectual structures of digital humanities. *Journal of Documentation*, 74(1), 223–246. DOI: 10.1108/jd-05-2017-0076
- Wang, X., & Inaba, M. (2009). Analysing structures and evolution of digital humanities based on correspondence analysis and co-word

- analysis. *Art research*, 9, 123-134.
<https://www.arc.ritsumei.ac.jp/lib/GCOE/info/kiyou09/11.pdf>
- Wang, X., Tan, X., & Li, H. (2020). The evolution of digital humanities in China. *Library Trends*, 69(1), 7-29. DOI: 10.1353/lib.2020.0029
- Yang, M., Wang, M., Wang, H., Yang, G., & Liu, H. (2020, August). Exploring the transdisciplinary nature of digital humanities. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020*. <http://dx.doi.org/10.1145/3383583.3398552>

Corresponding Author

Rahat Gulzar can be contacted at: rahat1320@gmail.com

Author Biographies

Rahat Gulzar is working as a Junior Professional Assistant at University of Kashmir, India and is currently enrolled as Ph.D. Scholar in the Department of Library and Information Science, University of Kashmir. He holds a Bachelor's degree in Computer Applications and has qualified UGC-NET JRF and JKSET. His research focuses on sentiment analysis, information retrieval, digital libraries, and knowledge organization, with publications in journals like *Global Knowledge, Memory and Communication*, and *Digital Library Perspectives*. He has studied online public sentiment, global dataset repositories, and scientific misconduct during COVID-19, and presented papers at national and international conferences. Technically skilled in tools such as Orange Data Mining, VOSviewer, KOHA, and Greenstone. ORCID: 0000-0001-8026-7730.

Aasia Maqbool is a Associate Professor in the Department of Education, University of Kashmir, India, teaching PG and UG courses including Psychological Foundations of Education, Guidance and Counselling, Mental Hygiene, Creativity, and Measurement and Evaluation. She holds M.Ed and M.A. degrees, earning a Gold Medal at the master's level, and qualified SET (2010) and UGC NET (2012). Her M.Phil focused on adolescents' self-concept, mental health, self-esteem, and academic achievement. She has published research papers, co-authored three books, coordinated MOOCs, served as a subject expert for MHRD e-contents, and is Principal Investigator and Co-PI on an NCERT-funded project.

Dr Farzana Gulzar is a Professor in the Department of Management Studies, University of Kashmir, India. She has many projects credited to her sponsored by reputed bodies like European Union (EU); Edventure Partners USA, Homeland Security, USA, and Facebook; Indian Council of Medical Research (ICMR), New Delhi; Indian Council of Social Science Research (ICSSR). She is also associated as an author with reputed publishing houses

like Springer, Emerald, & Sage.

Dr Iqra Shafi Bhat is an academician and researcher specializing in Marketing and Entrepreneurship. She holds an Integrated MBA from the University of Kashmir and qualified UGC-JRF in Management. Her Ph.D. on Women Technopreneurship explores the intersection of gender, technology, and entrepreneurship. She has served as Assistant Professor at the University of Kashmir and NIT Srinagar and worked as an Agricultural Branding Consultant at SKUAST, Kashmir. Dr Bhat's research, published in national and international journals, focuses on digital marketing, entrepreneurship, and technology-driven business models. Her expertise in marketing, branding, and women's entrepreneurship bridges academia and industry.

Sangita Gupta is Professor and Head of the Department of Library and Information Science at the University of Jammu, India. She also serves as Dean of Planning and Development and Head of the Department of Psychology, University of Jammu. With over 35 years of teaching and research experience, she has guided numerous doctoral and master's students and led major research projects funded by ICSSR and RUSA. A prolific author and editor, Professor Gupta has published books and over 70 research papers. She is a member and reviewer for several international journals and library associations. Prof. Gupta has been honored with Lifetime Achievement Awards from Mahatma Gandhi Central University (2025) and ASLIP (2024), as well as the IATLIS Best LIS Woman Teacher Award (2022).