Information Retrieval Research 1 (2025) 167-196 Submitted 12/24; Revised 05/25; Published 06/25

A survey of Inclusive Information Access

Yue Zheng YZ15U22@SOTON.AC.UK
University of Southampton
Southampton, United Kingdom

Haiming Liu H.LIU@SOTON.AC.UK
University of Southampton
Southampton, United Kingdom

Mike Wald M.WALD@SOTON.AC.UK
University of Southampton
Southampton, United Kingdom

Editor: Debarshi Kumar Sanyal

Abstract

Inclusive Information Access (IIA) is an essential field within information science and technology aimed at providing equitable information access for individuals from diverse backgrounds. Despite its significance, research on IIA remains dispersed across various domains, making it difficult to gain a comprehensive understanding of the field. This survey systematically examines the existing literature to map key themes, identify research trends, and highlight influential contributions within IIA. Using Latent Dirichlet Allocation (LDA) as a topic modeling method, this survey provides insights into the evolution of major topics, including publication trends, institutional engagement, and geographical reach. By analyzing authorship networks and citation patterns, this survey identifies key contributors, highly cited works, and influential research institutions, providing a structured overview of scholarly impact within IIA. This survey aims to support researchers in navigating the complex landscape of IIA and identifying avenues for future research.

Keywords: Inclusive Information Access, Information Retrieval, Digital divide

1 Introduction

In the digital era, access to information has evolved beyond mere convenience, becoming a fundamental human right (Bovens et al., 2000). Despite widespread access, systemic barriers related to age, gender, economic status, and disabilities continue to impede equitable information access, underscoring the importance of Inclusive Information Access (IIA) (Olphert et al., 2005; Hayes and Bulat, 2019). IIA seeks to ensure that all individuals have equitable access to the digital information landscape, irrespective of their challenges. This endeavor is not only about providing access but also about ensuring digital literacy, enabling individuals to effectively utilize and benefit from digital resources.

The scholarly exploration of IIA, while crucial, is still in its early stages. There is a growing consensus on the need for inclusive digital platforms and tools (Van Dijk and Hacker, 2003; Jaeger and Xie, 2009; Kiruki and Mutula, 2023), but comprehensive research

©2025 held by the author(s) License: CC-BY 4.0 WWW: https://irrj.org DOI: 10.54195/irrj.21092

that clearly outlines the scope of IIA and proposes actionable solutions is limited. As the world becomes increasingly interconnected digitally, a nuanced understanding of IIA is vital.

Understanding the landscape of IIA requires a comprehensive approach that captures its complexity and breadth. A survey is essential for systematically mapping the diverse facets of IIA, providing a structured overview that highlights foundational contributions, major research trends, and emerging areas (Munn et al., 2018). By consolidating fragmented knowledge across topics like digital inclusion, accessibility technology, and information retrieval, this survey offers a detailed synthesis to guide researchers, policymakers, and practitioners striving to address the information divide.

Additionally, this survey integrates machine learning techniques to enhance the traditional survey process. By employing algorithms such as Latent Dirichlet Allocation (LDA) for topic modeling (Jelodar et al., 2019), the survey uncovers thematic patterns and provides insights into the field's evolution over time. This combination of scoping methodology with advanced computational tools ensures a more robust and efficient analysis, delivering a nuanced understanding of IIA's historical development, current trends, and future directions.

Despite the breadth of research on related areas like "digital inclusion" and the "digital divide" there remains a lack of comprehensive surveys specifically addressing the unique scope of IIA. Prior studies primarily focus on policy implications or sociocultural factors, often overlooking the technological dimensions that shape IIA.

In this study, we define Inclusive Information Access (IIA) as the design, implementation, and evaluation of information systems that enable equitable and effective access to digital content for individuals across the spectrum of physical, sensory, cognitive, socioeconomic, and linguistic diversity.

By combining bibliometric analysis and topic modeling, this review provides a structured synthesis of the IIA research landscape. It examines scholarly influence, thematic evolution, and cross-disciplinary engagement within the field. The findings reveal key trends, research gaps, and emerging directions that can inform future development in IIA. To achieve this, the following research questions are addressed:

RQ1: What are the dominant research patterns including influential authors, institutions, and major publication sources in IIA?

RQ2: How have research topic in IIA evolved over time?

RQ3: What are the existing gaps in IIA research, and what directions should future studies explore?

By addressing these questions, this study provides a comprehensive overview of IIA research, offering insights into its development, key challenges, and potential future directions.

2 Related work

The digital age has transformed the framework of information rights, shifting focus from mere acquisition to equitable access for all. This change has brought Inclusive Information Access (IIA) to the center of scholarly and policy discussions, emphasizing its critical role in the digital world. The 2003 World Summit on the Information Society highlighted this by advocating for an inclusive Information Society (Bryne, 2005).

Despite growing interest, a dedicated survey on IIA is lacking. This study aims to fill this gap by examining current reviews on the "digital divide" and "digital inclusion" drawing on methodologies and findings from these studies.

The concept of the digital divide, initially about access to technology, has evolved to focus on skills and usage. Studies have shifted from physical access to digital inclusion, considering social, cultural, and economic dimensions (Van Dijk, 2006; Parsons and Hick, 2008; Nemer, 2015). However, these primarily qualitative reviews lack empirical validation and overlook potential biases, highlighting the need for more rigorous research.

The discourse on "digital inclusion" has explored various barriers, particularly for older adults, and emphasized the need for holistic approaches that address root causes of digital exclusion (Olphert et al., 2005; Weerakkody et al., 2012). Pinder (2004) and Jaeger (2006) had linked digital inclusion to broader social inclusion, noting the gap between policy and practice in ICT accessibility for people with disabilities.

Methodologically, studies have employed various approaches to surveying quantitative literature, ranging from bibliometric analysis to topic modeling, each contributing to a structured understanding of the field (Kelly and Sugimoto, 2013; Istenic Starcic and Bagon, 2014; Chen et al., 2020; Vassilakopoulou and Hustad, 2023; Perez-Escolar and Canet, 2023). However, these approaches have limitations, highlighting the need for more comprehensive, interdisciplinary, and data-driven surveys to capture the evolving landscape of digital inclusion research.

The roots of IIA lie in special education, where initial efforts focused on individualized education for students with disabilities. However, these efforts fell short in bridging systemic gaps in equitable access to educational resources. Libraries, key to public knowledge, have faced criticism for inadequate infrastructure for universal accessibility (Subramaniam et al., 2013). Additionally, the engagement of students with disabilities in modern technology remains low (Ezeani et al., 2017).

The lack of a universally accepted methodology for information sharing among disadvantaged groups further complicates this issue (Hayes and Bulat, 2019). Addressing this requires increased funding for assistive technology infrastructure (Kiruki and Mutula, 2023). Information access challenges extend beyond physical access, encompassing comprehension and interpretation difficulties (Publishing, 2013). Despite efforts to simplify content for disadvantaged groups, there's no consensus on evaluating these interventions (Jaeger and Xie, 2009). A proposed multi-dimensional approach for IIA evaluation includes usability, accessibility, findability, comprehensibility, and reusability (Grenon et al., 2023).

Accessible design benefits not just those with impairments but society as a whole (LaCheen, 2000). Considering that most people will experience disability-like conditions at some point, creating an electronic society based on IIA principles is both an ethical and practical imperative.

Bibliometric analysis is a powerful tool for examining research trends across various academic fields. It involves analyzing bibliographic data like publication volume, key sources, prominent authors, citation rates, and relevant keywords (Snyder, 2019). This method reveals the development of research themes, collaboration patterns, and the overall intellectual landscape of a field. It identifies emerging topics and trends, guiding future research and encouraging interdisciplinary work. However, for rigor, a bibliometric analysis typically requires a dataset of at least 500 publications (Donthu et al., 2021). In newer research areas,

meta-analysis (Aguinis et al., 2011) and citation analysis (De Groote, 2015) might be more appropriate.

Citation analysis, a key aspect of bibliometrics, examines the intertextual connections between scholarly works (Cronin, 2001). It identifies influential publications by analyzing citation patterns, thus assessing the impact of specific research contributions (Alzahrani et al., 2011). This analysis is useful for evaluating the citation metrics of articles, authors, institutions, and other indicators of academic productivity, aiding in reflective assessment and strategic planning (Nightingale and Marshall, 2012). However, citation analysis has limitations and should be applied carefully (Chikate and Patil, 2008). It quantitatively assesses citations from databases like Web of Science, Scopus, or Google Scholar, often visualized through citation networks or graphs to study publication trends and co-authorship dynamics.

3 Methods

For this survey, Web of Science (WoS) was selected as the principal data source. Although Web of Science (WoS) was chosen as the primary database due to its high-quality, peer-reviewed academic sources, we acknowledge its limitations (S. Adriaanse and Rensleigh, 2013). Future studies could supplement this dataset with Scopus, IEEE Xplore, or Google Scholar to enhance coverage and mitigate potential biases (Harzing and Alakangas, 2016; van Eck and Waltman, 2019). However, for this study, WoS was preferred for its structured citation data, facilitating bibliometric and topic modeling analyses.

Query statements were conducted in the Web of Science (WoS) Core Collection on June 11, 2023, using a Topic search(TS), which inherently includes Title, Abstract, Author Keywords, and Keywords Plus. The query applied the terms "Inclusive Information Access" OR "Inclusive Access Retrieval", ensuring a focused retrieval of studies addressing IIA. The search was restricted to academic journal articles, conference proceedings, and review papers. The initial search retrieved 1,349 articles, which were refined by applying subject category filters to ensure relevance to Information Science, Computer Science, and Information and Communication Technology (ICT) topics, resulting in a final dataset of 299 articles for bibliometric and topic modeling analysis.

To ensure full reproducibility, the exact search query executed in WoS was as follows: `TS=("Inclusive Information Access" OR "Inclusive Access Retrieval") AND DT=(Article OR Proceedings Paper OR Review) AND LA=English AND PY=<=2023`

Also a Bibliometric analysis software has been introduced to improve graphical representation for more organized and insightful visualizations. The goal of VOSviewer (Van Eck and Waltman, 2010) is to produce maps that display bibliometric data. It can map phrases, authors, or journals, offering a simple yet effective way to comprehend a scientific field's structure or the relationships between distinct concepts or writers. For data input, the software uses a user-friendly drag-and-drop interface. In the form of network maps, VOSviewer creates representations that are easy to see and comprehend. In contrast to CiteSpace (Chen, 2014), where many basic features require a paid upgrade, VOSviewer offers these functionalities for free, making it a more cost-effective choice

To identify journals that are most actively contributing to the field of IIA, we adopted the **h-index** and **g-index** as bibliometric indicators (Hirsch, 2005; Egghe, 2006). The **h-**

index of a journal indicates that it has published h articles that have each been cited at least h times, reflecting both productivity and consistent citation impact. The **g-index**, on the other hand, gives more weight to highly cited articles by identifying the largest number g such that the top g articles received together at least g^2 citations. These measures help surface journals that are not only frequently publishing in IIA-related topics but also whose publications are being consistently cited within this field.

To accomplish the domain topic modeling and evolution task, this study introduces the Latent Dirichlet Allocation (LDA) topic model, supplements word vectors using TF-IDF, calculates similarity using cosine distance, and visualizes topic evolution through Sankey diagrams (De Felice and Polimeni, 2020; Liu et al., 2022; Raza et al., 2019; Amjad and Ihsan, 2020; Zimmerman et al., 2021; Dhrangadhariya et al., 2020; Xiong et al., 2018; Owa et al., 2021). The workflow is presented as Figure 1, and the specific processes will be elaborated in detail in this section.

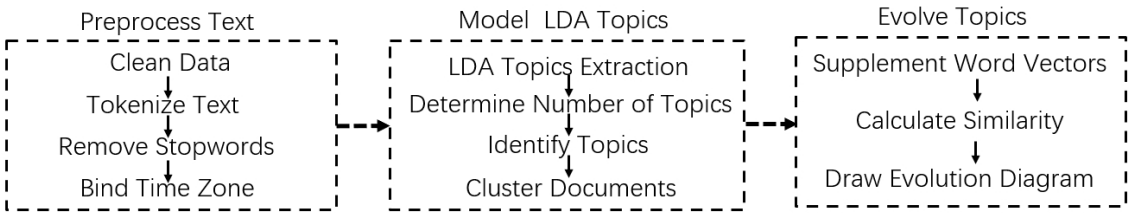

Figure 1: Workflow Diagram for Topic Modeling and Evolution

LDA is a generative probabilistic model crucial in topic modeling and Natural Language Processing (NLP). It assumes that each text in a corpus is composed of various topics, and each word in the text is associated with one of these topics. LDA uses a three-level hierarchical Bayesian model with Dirichlet distributions to uncover these latent topics (Blei et al., 2003).

In this study, LDA was employed to model the distribution of topics across documents and the distribution of words within topics. The articles' titles, contents, and publication years were retained for analysis. Text tokenization was performed using the Natural Language Toolkit (NLTK)'s word_tokenize function (Bird et al., 2009), and stopwords were filtered out using a unique list.

The evolution of the IIA field was divided into three periods: 2000–2010, 2011–2016, and 2017–2023. The rationale for selecting these periods as the basis for sorting publications is further explained in Section 4.1. LDA's perplexity metric, which measures how well the model's predicted probability distribution matches the actual distribution of words in texts, was used to determine the ideal number of topics. Lower perplexity scores indicate better model generalization (Wallach et al., 2009). The inflection point in the perplexity curve was chosen as the criterion for the optimal number of topics, adhering to the principle of parsimony.

After setting the topic parameters, the LDA model was used to identify topic terms for each period, followed by manual labeling to enhance interpretability by analyzing representative keywords and verifying them through a review of relevant articles. In this context,

topic parameters refer to the predefined values that influence the model's structure and output, including the number of topics (K), which determines the granularity of thematic categorization, and the Dirichlet priors (α and β), which regulate the distribution of topics across documents and words across topics, respectively (Blei et al., 2003). Despite the effectiveness of LDA in identifying optimal topics and word distributions, manual intervention was necessary for summarizing and generalizing topic words. To ensure the reliability of topic interpretation and minimize potential bias, a second reviewer independently examined the LDA-generated keywords, validated the manually assigned topic labels, and cross-checked the phase boundaries used in the topic evolution analysis. Any discrepancies were discussed and resolved collaboratively.

The study also faced challenges in calculating cosine distances for topic evolution analysis due to low probability distribution weights of topic keywords. To overcome this, the Term Frequency-Inverse Document Frequency (TF-IDF) technique was used to identify high-frequency words within each topic (Zimmerman et al., 2021; Sparck Jones, 1972). Documents under each topic for each year were combined into a single document for TF-IDF vector computation.

Cosine similarity, measuring the similarity between two texts based on the cosine of the angle between their vectors in vector space, was used to establish connections between similar text sets (Wang et al., 2010; Ahad et al., 2016; Park et al., 2020). After normalizing the modeling to obtain vector representations, cosine similarity was calculated between texts. The study used topic vectors from TF-IDF computations to calculate cosine similarity between topics across different phases. An evolutionary relationship was established between topics when their similarity exceeded a predefined threshold. The final theme evolution diagram was visualized using the Sankey method in the pyecharts library (Chaudhuri, 2019).

To determine the optimal number of latent topics for LDA modeling, we trained a series of models with topic numbers K ranging from 2 to 10. The selection was guided by the perplexity score, a standard metric assessing how well the model predicts a sample. The inflection point in the perplexity curve indicated that $K = 5$ offered a balance between model complexity and generalizability, aligning with the principle of parsimony.

The LDA model was implemented using the Gensim library with symmetric Dirichlet priors. We adopted the default hyperparameters, setting $\alpha = 1/K$ and $\beta = 0.01$, which are widely recognized for providing robust results across diverse corpora. Given the exploratory nature of this study and the corpus size, these defaults were considered sufficient for capturing meaningful topics.

Text preprocessing was conducted using the Natural Language Toolkit (NLTK). This process involved lowercasing all texts, tokenization, stopword removal using an extended English stopword list, and filtering out punctuation and non-alphabetic characters. Stemming or lemmatization was intentionally omitted to preserve domain-specific semantics and enhance interpretability.

Topic quality was assessed both quantitatively and qualitatively. Perplexity scores were used for model comparison, and manual topic labeling was performed based on the top keywords per topic. Two independent reviewers collaboratively validated the labels to mitigate bias and enhance semantic coherence. This semi-automated evaluation approach

is common in topic modeling applications where full automation of thematic interpretation remains limited.

4 Analysis and results

4.1 Publication Trend

Per the bibliometric framework delineated by Donthu et al. (2021), the temporal distribution of publications functions as a multifaceted indicator, invaluable for assessing the maturity and dynamism of a given research domain. Additionally, this temporal metric is an essential gauge for evaluating the field's evolutionary trajectory over a designated period. Key features such as the publication trend curve's growth rate and inflexion points yield incisive insights into the field's dynamic fluctuations. Such data prove indispensable for macro-level analyses to appraise the field's current scholarly prominence and for predictive models forecasting its developmental trajectory and emergent trends.

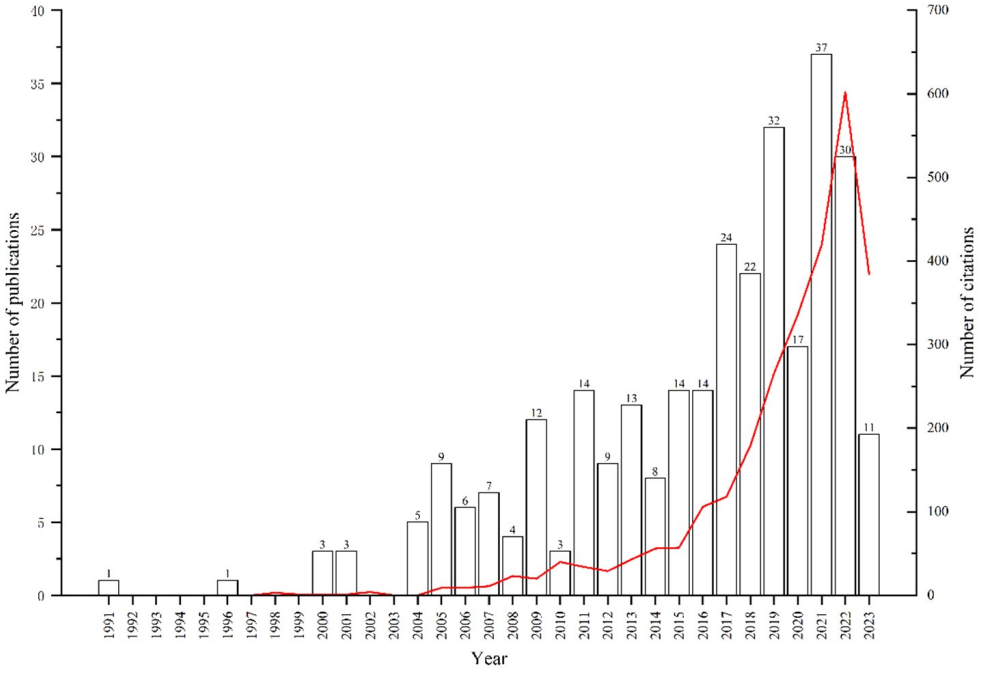

Figure 2: The number of annual research publications and citations

Upon statistical analysis of the temporal distribution of 299 publications, it became evident that IIA research originated as early as 1991. In this seminal work, Gerich (1991) explored the role of the National Science Foundation Network (NSFNET) in providing expansive network access to academic and research communities. The paper accentuated the dramatic escalation in network traffic and underscored the imperative for international collaboration and governance as the Internet began transcending national and global boundaries. This work marked the inaugural introduction of the term "Internet Inclusive" which,

although not identical to the current focus on IIA, resonated with the burgeoning influx of information in the digital age.

It is worth noting that, following the convening of the 9th International Conference on Computers Helping People with Special Needs, the concept of Design for All (DfA) gained recognition as a significant paradigm within the realm of Information Society Technologies (IST) (Bühler and Stephanidis, 2004). This concept aimed to guarantee universal access to technology products and services by eliminating existing barriers. After 2004, the volume of publications pertinent to IIA consistently escalated, signalling an intensifying scholarly engagement with the subject matter.

Specific data reveals that from 1991 to 2023, an average of 9.34 articles were published yearly. These articles have been cited 2751 times, with an average annual citation rate of 85.97. Analyzing Figure 2, it can be observed that the field has undergone three distinct development phases.

From 1991 to 2010, the first stage served as an exploration phase. The field was still young, with only three papers published a year on average during this time. Due to the limited availability of documented outputs prior to the year 2000, with only two instances, the initial phase of the study was adjusted to focus on the period from 2000 to 2010. This modification was implemented to better align the study's timeframe with the broader temporal context, thereby enhancing the validity of subsequent analyses. The second phase, which lasted from 2011 to 2016, included fast expansion. At this time, the discipline produced 12 publications annually on average, and papers were getting many more citations.

The third and current phase started in 2017 and is still in progress. Academic output during this period was solid and consistent. The culmination was a record-breaking 37 publications in 2021. These papers have rec lasted from 2011 to 2016 andations as well. It is essential to keep in mind that the data for 2023 is insufficient because it only contains items released up until August. However, according to the most recent data, the number of articles published has reached the Phase II average, and the number of citations is higher than the previous year.

In conclusion, there has been a steady increase in IIA research in recent years. Three main developmental phases can be identified for this expansion: the original stage, which lasted from 2000 to 2010, the middle phase, which lasted from 2011 to 2016, and the most recent stage, which lasted from 2017 to 2023. Given the course it is taking, it is logical to assume that the field will continue to grow in the future.

4.2 Publication Source

In the dataset under investigation, 231 academic journals are represented. Among these, a minority subset of 13 journals is remarkably prolific, each contributing more than three articles to the corpus. In addition, 25 journals have contributed exactly two papers, while the vast majority, with 193 journals, have each contributed a single piece to the field. This distribution suggests a highly skewed landscape where a few journals are the primary conduits for scholarly output in this area. The top 10 prolific publication Sources in the field of IIA are shown in Table 4.2. "Universal Access in The Information Society" despite having the highest number of articles, has an h-index of 5 and a g-index of 8, which are not significantly higher than those of other journals. This suggests that while it may

Rank	Sources	h-index	g-index	TC	Articles
1	Universal Access in The Information Society	5	8	72	10
2	Journal of Documentation	3	8	80	8
3	Telecommunications Policy	5	6	109	6
4	Telematics and Informatics	4	4	182	4
5	IEEE Access	3	4	50	4
6	Ethics and Information Technology	3	3	26	3
7	Information Society	3	3	91	3
8	Journal Of The Australian Library and Information Association	3	3	16	3
9	Interacting With Computers	2	3	14	3
10	International Journal Of Electronic Government Research	2	3	11	3

Table 1: Top 10 Publication Sources

be a cornerstone in terms of volume, its overall impact, as measured by these indices, is comparable to other journals like "Telecommunications Policy" which has an h-index of 5 but a higher total citation count of 109.

This concentration of publications within a top-tier subset of journals suggests that the field is beginning to coalesce around a central academic nexus. Yet, it's worth noting that the area also enjoys contributions from a broad spectrum of journals, indicating its inherently multidisciplinary nature. Journals like "Ethics and Information Technology" and "IEEE Access" exemplify this trend, highlighting the intersection of ethical, technological, and policy dimensions within IIA research.

4.3 Publication Institution

For the purpose of locating potential partners in the field, it is essential to comprehend the institutional landscape of IIA research. In order to do this, the linkages between various research institutes based on their contributions to IIA were mapped using a network visualization Figure 3.

The network visualization offers a number of significant insights. First, each node in the network has a size that reflects the number of publications the institution has produced, acting as a measure of research output. As an illustration, the "University System of Maryland" node is clearly larger, demonstrating its considerable contribution to IIA research. Second, the strength and temporal characteristics of institutional collaborations are indicated by the thickness and colour of the lines joining the nodes. However, the lines connecting the nodes are neither numerous nor particularly thick or even almost obscured by the institutional nodes, suggesting that institutional cooperation is not common overall. This suggests the industry is still in an exploratory phase where institutions independently look into different aspects of IIA.

The absence of significant inter-institutional collaborations in the IIA research landscape is evident. However, Table 4.3 offers a different perspective. Specifically, the "University System of Ohio" stands out with a high total citation count (TC) of 161, indicating its substantial impact on the field.

According to Table 4.3, it is clear that the institutions with the highest publication output regularly work with other institutions, indicating that inter-institutional collabora-

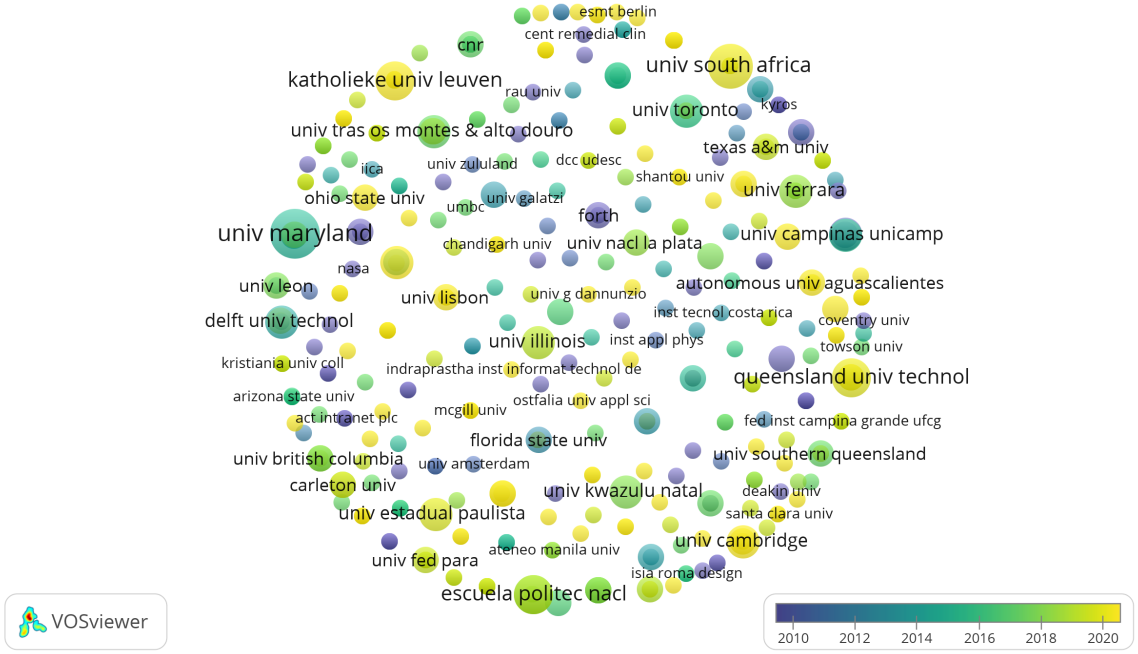

Figure 3: Institutional Cooperations

Institution	Country	NP	First Institution	Corresponding Institution	TC	PC
University System of Maryland	America	8	7	4	62	7.75
Escuela Politecnica Nacional Ecuador	Spain	6	3	4	71	11.83
Universidade Estadual De Campinas	Brazil	6	3	3	10	1.67
University of South Africa	South Africa	5	5	4	64	12.8
Consiglio Nazionale Delle Ricerche	Italy	4	3	1	7	1.75
Ku Leuven	Belgium	4	4	3	12	3
Pennsylvania Commonwealth System of Higher Education Pcshe	America	4	4	0	15	3.75
Queensland University of Technology	Australia	4	4	4	41	10.25
Universitat D Alacant	Spain	4	2	0	60	15
University of Maryland College Park	America	4	4	0	23	5.75
University System of Ohio	America	4	4	0	161	40.25

Table 2: Institutional Contributions

tion greatly boosts field developments. Notably, the Queensland University of Technology stands out for having a high preference for external cooperation; each of its four papers was co-authored with other institutions. This highlights the university's proactive strategy for creating relationships, which may work as a catalyst for encouraging innovation and research excellence in the field.

In conclusion, the state of collaboration is generally lacklustre. Nevertheless, institutions with the highest publication output exhibit a stronger propensity for cooperation. So, the

Country	Articles	SCP	MCP	Percent	MCP Ratio	TC	Avg. Article Citations
USA	57	51	6	0.191	0.105	578	10.1
UK	22	15	7	0.074	0.318	180	8.2
Brazil	21	20	1	0.07	0.048	101	4.8
Australia	19	14	5	0.064	0.263	154	8.1
Canada	16	14	2	0.054	0.125	122	7.6
South Africa	12	9	3	0.04	0.25	129	10.8
India	12	10	2	0.04	0.167	78	6.5
Italy	9	8	1	0.03	0.111	15	1.7
Portugal	8	8	0	0.027	0	9	1.1
China	7	5	2	0.023	0.286	167	23.9

Table 3: Top 10 Publication Countries

lack of connections between these academic institutions raises the possibility of collaborative research and knowledge sharing being underutilized.

Therefore, it is essential that future studies concentrate on establishing collaborative frameworks while appreciating individual institutions' contributions. Such a strategy could enhance the IIA research environment by ensuring a more comprehensive and broad spectrum of perspectives while accelerating field progress.

4.4 Publication Country

Understanding the worldwide geography of this study topic requires understanding the contributing nations in the field of IIA. With the use of citation rates, this analysis will be able to pinpoint the nations that produce the most research, contribute to it in the most meaningful ways, and have the most overall influence. The technique comprises a thorough analysis of the literature with a focus on the number of papers published in each nation, the average number of citations per article, and the proportion of multi-country to single-country publications.

According to the data, 64 nations or areas have contributed to the worldwide IIA research effort. Table 4.4 further categorizes the contributions by Single Country Publications (SCP), Multi-Country Publications (MCP), Percent of MCP, MCP Ratio, Total Citations, and Average Article Citations. For instance, the United States leads with 57 articles, of which 51 are SCPs, and has an average citation rate of 10.1. This indicates high research output and significantly impacts the academic community. With just seven articles, China has an average article citation rate of 239, highlighting the impact and quality of its research in this field.

In the realm of scientific collaboration, the network visualization Figure 4 offers a compelling snapshot of the intricate relationships between various countries in terms of their research output. The graph reveals several key trends and patterns that merit further scrutiny.

For the purpose of this analysis, data for England and Scotland are aggregated under the United Kingdom (UK). The United States and the United Kingdom prominently lead in both the volume of publications and the extent of their academic collaboration in the field

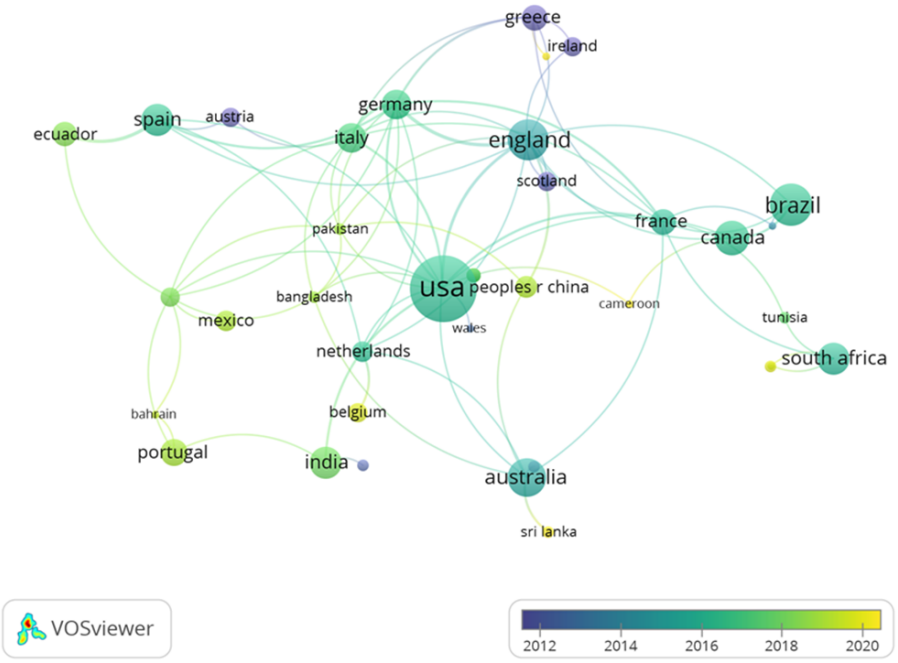

Figure 4: Collaborations of Countries

of IIA research. This suggests a significant strategic partnership, reflecting a high degree of synergy and shared research interests.

Additionally, other strong international collaborations are evident, such as the partnership between the United Kingdom and Germany, and the cooperation between Ecuador and Spain. These relationships, likely driven by shared research objectives or complementary expertise, highlight the global nature of scientific research.

Factors like common language and educational systems might facilitate the U.S.-U.K. collaboration, while cultural and historical connections could influence the Ecuador-Spain partnership.

Overall, the data underscores the global and interconnected nature of IIA research, emphasizing the importance of international collaboration in advancing knowledge. This suggests that the future of IIA research relies on a collective, globally-inclusive effort.

4.5 Authorship

Prolific authors serve as the cornerstone of any research field (Brito et al., 2023), and this is no less true for the domain of IIA. Between 1991 and 2023, a total of 840 authors have contributed to the body of literature on IIA. This study employs Price's Law, a well-established bibliometric principle, to identify core contributors within the expansive

authorship network (Price, 1963). According to this law, authors with more than two publications can be considered as core authors.

However, the application of this criterion revealed that as many as 52 authors had published more than two articles. This introduced complexity into the analysis due to a high number of tied rankings among authors with fewer publications. This is most likely due to the small total number of articles studied, as well as the fact that the field is still in a developmental stage. To provide a more focused view, the criterion was adjusted to spotlight authors with three or more publications, as detailed in Table 4.5.

Author	h-index	g-index	TC	NP	PY start
Baranauskas MCC	2	3	10	6	2009
Lujan-mora S	3	5	68	5	2015
Sitbon L	3	4	40	4	2017
Stephanidis C	3	4	37	4	2004
Bonacin R	1	1	3	4	2010
Acosta T	3	3	57	3	2018
Brereton M	2	3	33	3	2018
Neris VPD	2	2	7	3	2009
Maietti F	1	2	5	3	2018
Dos reis JC	1	1	3	3	2010

Table 4: Top 10 core authors

By adopting this more stringent criterion, the study aims to offer a more focused view of the most impactful authors, thereby elucidating the intellectual landscape of IIA research

A review of articles published by core authors in the field reveals different research focus. Baranauskas MCC, Bonacin R, Dos reis JC, and Neris VPD engaged in the design and construction of inclusive social networks tailored to the specific characteristics of Brazilian society. Lujan-mora S and Acosta T primarily evaluated the accessibility features of various web pages. Sitbon L and Brereton M concentrated on extending learning opportunities in information technology for individuals with intellectual disabilities, while also encouraging their participation in information retrieval tasks. Stephanidis C proposed semantic-based user modeling to facilitate the adaptation of web-based user interfaces. Maietti F, on the other hand, focused on tasks related to the preservation of inclusive cultural heritage.

This diversity in research topics underscored the multidisciplinary nature of IIA and highlighted the range of approaches that were employed to address its complex challenges.

Expanding upon the foundational research previously discussed, a visual representation of author collaboration networks, as shown in Figure 5, was generated using VOSviewer. In this graph, it is noteworthy that among authors with more than two publications, there are 12 who have published independently, while the rest have engaged in collaborations to varying degrees. However, the majority of these collaborations occur within the same institution.

The field may still be in its early stages because the frequency and density of partnerships are often low. The small number of partnerships also raises the probability that many writers are working alone, which may reduce the variety of viewpoints and research methods in the subject.

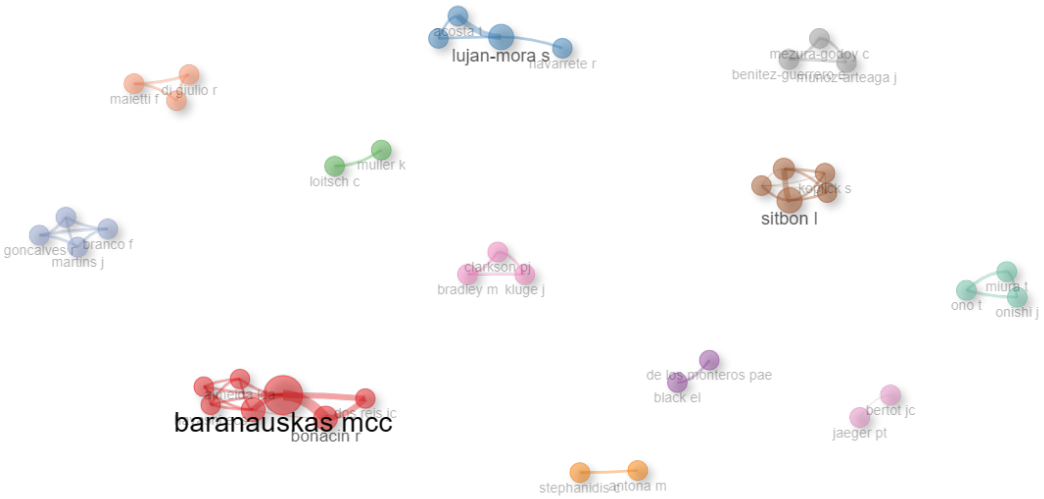

Figure 5: Collaborations of Authors

Given this context, it stands to reason that more varied and frequent partnerships ought to be promoted. The research environment in IIA may be considerably enhanced by encouraging interdisciplinary and interinstitutional cooperation, which would also provide researchers with a deeper knowledge of the complexities of the field.

4.6 Citation

To gain a deeper understanding of the influential works shaping the field of Information Inequality and Accessibility (IIA), it's useful to examine the most cited publications. Table 4.6 lists these key articles, which serve as important benchmarks in the existing body of knowledge and guide future research directions. The subsequent sections will delve into the details of each article, highlighting their contributions and significance in the IIA landscape.

Title	Global Citations	Local Citations
Web of Science (WoS) and Scopus: The Titans of Bibliographic Information in Today's Academic WorldPranckutė (2021)	269	0
Digital inclusiveness - Longitudinal study of Internet adoption by older adultsLam and Lee (2006)	138	2
Consumers' acceptance of information and communications technology in tourism: A reviewUkpabi and Karjaluoto (2017)	137	0
5G mobile technology: A surveyMitra and Agrawal (2015)	122	0
Mobile health technology adoption across generations: Narrowing the digital divideFox and Connolly (2018)	109	0

Table 5: Top 5 Cited Articles

A comprehensive overview of two major bibliographic databases, Web of Science (WoS) and Scopus, is provided by integrating information from the database owners with the latest research findings (Pranckutė, 2021). The investigation's primary area of interest is information accessibility. Similar to another study which used the review method to examine consumer adoption of web-based services, social media, and mobile information systems in the tourist industry, did not specifically address how inaccessible e-tourism is for populations with limited access to information (Ukpabi and Karjaluoto, 2017).

Lam's study explored the use of the Internet by the elderly, showing that training improves their online confidence and willingness to use the Internet (Lam and Lee, 2006). Fox's 2018 research (Fox and Connolly, 2018) found that older adults are hesitant to use mobile health technologies due to mistrust and privacy concerns, suggesting the need for designs and education that enhance their confidence and understanding of privacy.

Mitra (Mitra and Agrawal, 2015) offered a thorough review of 5G development activities. However the terms "inclusive" and "access" pertained to the hardware's compatibility with the technology. It is a different topic than the one explored in this study. This also reflects the confusing nature of nomenclature in the IIA field.

While the above publications receive the most global citations, only one paper receives two local citations, indicating a potential gap between local and global academic significance. This could imply that although these publications are acknowledged and mentioned in the larger academic world, the IIA research environment may not find them to be as significant or pertinent.

The significance and methodological rigor of the research they embody are both reflected in highly cited articles, which act as significant landmarks in the academic landscape. Table 4.6 displays the top ten cited articles in research in the field of IIA, reflecting the multifaceted nature of the field.

Title	Local Citations
Transforming our world: the 2030 Agenda for Sustainable DevelopmentResolution et al. (2015)	7
Development as freedomSen (2000)	7
Digital divide: Civic engagement, information poverty, and the Internet worldwideNorris (2001)	6
Cache template attacks: Automating attacks on inclusive {Last-Level} cachesGruss et al. (2015)	5
User-Sensitive Inclusive Design in Universal Access in the Information Society August 2011Newell et al. (2011)	5
Social network sites: Definition, history, and scholarshipBoyd and Ellison (2007)	5
Digital divide research, achievements and shortcomingsVan Dijk and Hacker (2003)	5
User acceptance of information technology: Toward a unified viewVenkatesh et al. (2003)	5
The impoverished life-world of outsidersChatman (1996)	5
Basic formal education quality, information technology, and inclusive human development in sub-Saharan AfricaAsongu and Odhiambo (2019)	4

Table 6: Top 10 Cited Reference

The most highly cited reference in the dataset is "The 2030 Agenda for Sustainable Development (Resolution et al., 2015)". Within this seminal work, the objective to "Make urban areas and habitations inclusive, secure, robust, and sustainable" emerges as the most frequently cited content, followed by discussions on inclusivity in learning and work environments. Collectively, citations to this article underscore the foundational premise in the field of IIA that access to the Internet, emergency services, and information and communication

technologies constitutes a fundamental human right. Sen (2000) emphasized the importance of IIA, advocating for equipping those without information access with the necessary tools for self-reliance. Van Dijk and Hacker (2003) identified four types of digital divide access: motivation, physical, skills, and usage, underscoring the need for comprehensive theoretical frameworks, clear definitions, multidisciplinary approaches, qualitative analysis, and long-term studies in this area. This has influenced later research to fill these gaps.

A survey of the literature revealed a large worldwide digital divide between high-income and low-income countries. This finding forms the basis for future study into "e-government"-related issues in the area of IIA (Norris, 2001). It is important to remember, though, that the growth of "e-government" should not exclude people who are already informationally underprivileged, as doing so risked widening the digital divide already present. The research on sub-Saharan Africa has garnered considerable internal citations (Asongu and Odhiambo, 2019). The study posits that the low prevalence of mobile phone usage in the region is constrained by inadequate primary education, thereby highlighting a significant digital divide in comparison to global information development levels.

To strengthen ties between designers and users, Newell et al. (2011) encouraged the adoption of a moderately inclusive and user-sensitive design philosophy. According to the research, the use of inclusive design principles successfully reduced a variety of obstacles that people with impairments encountered when using technology. In the field of "inclusive design" this work served as a foundational study, and subsequent research used this methodology as a point of reference (de Almeida Neris et al., 2021). Additionally, Newell conceded that it was impractical to create goods that were accessible to all potential users. As a result, the article substituted the terms "inclusivity" and "sensitivity" for "universality" and "centrality" respectively, defining a more achievable and, in many cases, more appropriate goal. Venkatesh et al. (2003) introduced the Unified Theory of Acceptance and Use of Technology (UTAUT) model, which synthesizes eight prominent models from the user acceptance literature. Performance expectancy, effort expectancy, social influence, and enabling conditions are the four main factors identified by the UTAUT model. This study offers important insights for future research on universal information access since it focuses largely on how information is accepted in traditional cultures.

Many studies drawing on the characteristics of social networks, propose the use of user-specific data during retrieval to offer customized responses, a strategy that holds promise in the realm of inclusive design (Boyd and Ellison, 2007). However, the practical implementation of this approach can encounter ethical challenges, particularly concerning data privacy. Additionally, the re-collection of specific data could compromise the method's broad applicability.

To summarize, the highly referenced literature is divided into three categories, including: 1. The importance of study in the field of IIA 2. The global digital divide in its context 3. The underlying theories and procedures for inclusive design.

4.7 Topic Evolution

Figure 6 shows a thorough theme evolution map for the field of IIA spanning the years 2000 to 2023, building on the LDA topic modeling methodology described before. Each phase's individual color blocks correspond to different topics. These blocks' sizes show how

many articles fall under each topic. The connections between blocks in neighboring phases represent the relationship between them and the course of progression. A relationship between two themes from different phases implies that a theme has evolved or changed through time.

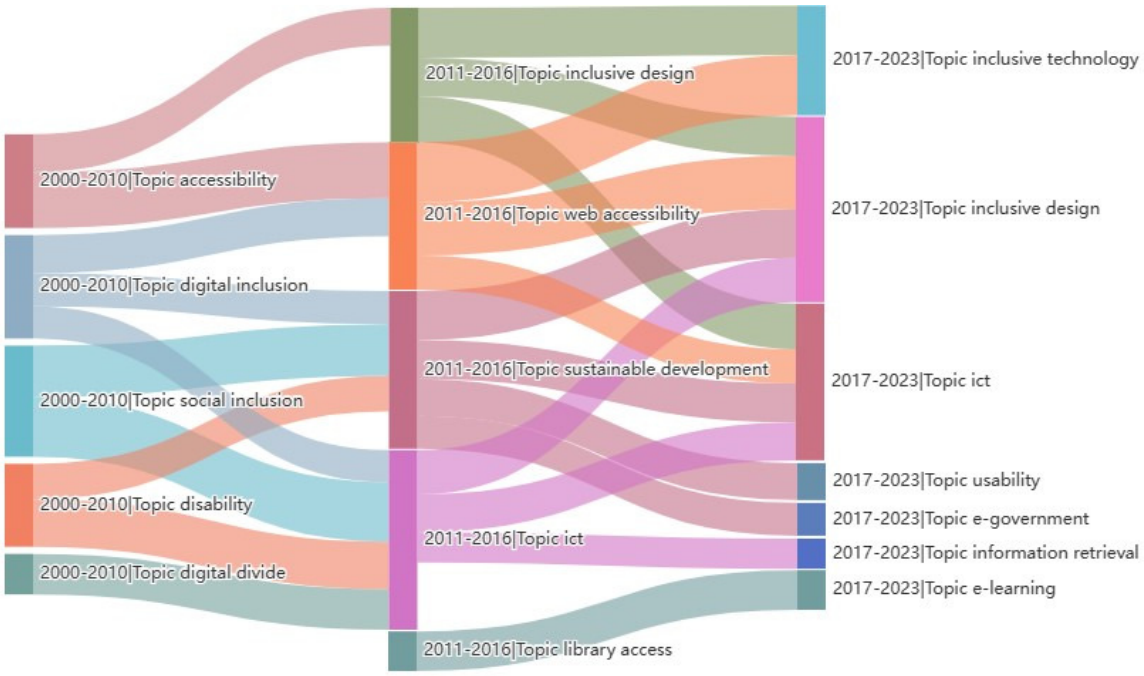

Figure 6: Theme evolution

Based on the thematic evolution diagram, the field of IIA has evolved from its initial focus on broad concepts such as accessibility, digital inclusion, social inclusion, disability, and the digital divide. It has shifted towards more specialized topics like inclusive design, web accessibility, sustainable development, ICT, and library access. Currently, the research landscape is increasingly concentrating on practical scientific and technological methods, particularly in areas such as usability, e-government, information retrieval, and e-learning, which exhibit significant potential for further development.

5 Discussion

In the evolving landscape of IIA, this section focuses on inclusive technology and information retrieval, laying the groundwork for future advancements in inclusive information retrieval study. A comprehensive literature review on the topic of inclusive design has already been conducted by Li et al. (2023), serving as a valuable resource in this domain. The goal of inclusive technology is to make information and communication technologies available to everyone, regardless of age, physical condition, or socio-economic status. This field has gained a lot of attention in recent years as researchers focus on different target populations, such as the elderly and those with visual or hearing impairments. They use

Key Contributions	Target Population	Methodology
TV-based service to enhance well-being and self-esteem Epelde et al. (2013)	Elderly	User-centered design
iLiBra platform to streamline communication using sign language icons da Costa et al. (2019)	Hearing impairments	Assistive technology platform with sign language
Drawxi, an audio-haptic tool for collaborative diagram creation Chiplunkar et al. (2019)	Visual impairments and limited sight	Sensor-based solutions
Real-time tracking model for social interactions Grayson et al. (2020)	Visual impairments	Head-mounted HoloLens device
Force-FeedbackTablet (F2T) for interacting with 2D data Gay et al. (2021)	Visual impairments	Haptic architecture

Table 7: Summary of Inclusive Technology

various methods and technologies to tackle the specific challenges faced by these groups, ultimately contributing to the larger objective of providing universal access to information.

Table 5 provides a comprehensive summary of research articles with a clear focus on technology in the field of inclusive technology, outlining their contributions, target populations, and methodologies.

Current technology in this field primarily aids those with visual impairments, using audio conversion to overcome visual challenges. However, incorporating tactile feedback could enhance understanding of complex information. For individuals with hearing loss, the focus is on converting content into sign language. Technologies for the elderly prioritize usability and clarity.

Yet, there's a notable lack of comprehensive research on IIA for people with cognitive impairments, including learning disabilities, attention deficit disorders, and dementia. Current research inadequately addresses the unique challenges faced by these individuals in understanding and retaining information.

In a sequence of investigations focused on the development of inclusive social network systems in Brazil, a novel methodology for personalized system design is introduced(Neris et al., 2009; Almeida et al., 2009; Reis et al., 2010; dos Reis et al., 2010; Dos Reis et al., 2014; de Almeida Neris et al., 2021). This methodology diverges from traditional search solutions that rely on lexical-syntactic information processing. Instead, it advances an innovative approach to web ontology design by amalgamating semantic modeling techniques with a content-based strategy for ontology creation. This inclusive search mechanism yields semantic search results that are universally applicable, while simultaneously honoring the unique preferences and needs of individual users.

Complementing this, the inclusive design literature (summarized in Table 5) provides a methodological foundation for creating digital environments that are not only accessible, but also participatory and personalized. These studies emphasize the integration of user-centered and participatory design frameworks, ergonomic data visualization, and interface design patterns such as progressive disclosure, ensuring that digital systems are flexible and responsive to diverse user needs. Importantly, they also demonstrate how co-design workshops and iterative evaluation processes help translate user insights into usable interface prototypes, reinforcing inclusivity at the design level.

Title	Author (Year)	Methodology
Attuning Speech-Enabled Interfaces To User And Context For Inclusive Design: Technology, Methodology And Practice	Neerincx et al. (2008)	Accessibility analysis and iterative refinement of speech-based multimodal interface.
Accessing User Information For Use In Design	McGinley and Dong (2009)	Tool concept development and co-design workshop to improve ergonomic user data communication.
Providing Universally Accessible Interactive Services Through Tv Sets: Implementation And Validation With Elderly Users	Epelde et al. (2011)	User-centered design using interviews and focus groups to model older adults' needs.
User-Centered Design Journey for Pattern Development	Joshi et al. (2017)	User-centered design with iterative evaluation of UI solutions; developed the progressive disclosure pattern via case study.
Addressing Brazilian diversity in personal computing systems with a tailoring-based approach	Almeida Neris et al. (2020)	Participatory tailoring approach integrating PLuRaL and FAN frameworks; applied in inclusive system design.

Table 8: Summary of Inclusive Design

Title	Author (Year)	Methodology
Addressing Universal Access In Social Networks: An Inclusive Search Mechanism	Reis et al. (2013)	Semantic search design using Web ontology and Organizational Semiotics; evaluated via case study in a social network system.
A Reference Ontology For Digital Scientific Journals Applied To Systematic Literature Review Processes	Ghisi et al. (2012)	Used reference ontology with batch/on-the-fly processing to enhance interoperability and inclusiveness in scientific literature retrieval.

Table 9: Summary of Information Retrieval

Building upon these foundations, research on inclusive information retrieval (IIR) has begun to emerge as a critical subdomain within IIA. As summarized in Table 5, current

efforts in this space have explored the use of semantic web technologies (Reis et al., 2013), organizational semiotics, and reference ontologies to develop more accessible and adaptive search mechanisms. These approaches aim to reduce the cognitive and linguistic barriers often encountered by users with low digital literacy, enabling them to retrieve information using familiar, meaningful terms rather than system-generated query syntax. However, these early investigations remain limited in scope and scale, suggesting a pressing need for more innovation in retrieval algorithms, interface design, and user modeling to support a wider range of impairments and contexts.

As attention increasingly shifts toward inclusive approaches in Information Retrieval (IR), recent research has moved beyond traditional, text-dominant keyword-based systems toward more adaptive, context-aware, and user-centered IR frameworks. These systems are designed to accommodate a broader spectrum of user profiles. For instance, several studies (Guo et al., 2018; Zhang et al., 2020; Erbacher et al., 2022) have emphasized the integration of interactive retrieval frameworks that support cognitive offloading. Such systems allow users to refine queries through dialog-based or visually guided interactions, thereby alleviating the need for precise keyword formulation. This paradigm shift aligns with the broader goal of reducing cognitive and linguistic load, particularly for users with learning difficulties, neurodiverse conditions, or limited language proficiency.

In parallel, other research (Hsieh et al., 2022; Zheng et al., 2024) has investigated the role of multimodal information retrieval, enabling users to access content through combinations of text, speech, and images. These systems are particularly beneficial for individuals with visual impairments, low literacy, or temporary accessibility challenges—such as users navigating complex physical environments. Multimodal interfaces expand the expressive capacity of users and diversify input modalities, further contributing to the development of inclusive IR systems.

Furthermore, the integration of user modeling and personalization into IR systems has become increasingly sophisticated. Recent studies (Kladouchou et al., 2025) have explored the use of real-time behavioral data, user preferences, and even physiological signals to dynamically adjust retrieval results. Such advancements enable IR systems to more accurately anticipate users' needs and deliver content in formats aligned with their individual capabilities and preferences (Zheng et al., 2024; Ji et al., 2024).

To achieve an inclusive information retrieval system, a multidisciplinary approach is essential. This involves integrating inclusive design principles to create user interfaces for various impairments, employing inclusive technology to enhance sensory accessibility, and utilizing specialized information retrieval algorithms for relevant personalized results. The aim is to address the various challenges in providing equitable access to information.

Recognizing limitations is critical, including potential data oversight from other databases and issues with keyword selection that resulted in incomplete or extraneous data. The findings are preliminary and may lose relevance over time. Although machine learning can help with theme recognition, accurate topic labeling still requires human interpretation.

Future research areas are highlighted, focusing on the need for technology that caters to a wide range of disabilities, particularly cognitive impairments, as well as the development of inclusive information retrieval systems. Addressing these issues has the potential to significantly advance the field of Inclusive Information Access, improving efficiency and accessibility for a broader audience.

6 Limitations and Future Work

Recognizing the limitations of this study is essential. First, despite the fact that the Web of Science is a credible bibliographic source that is acknowledged around the world, it is still possible that pertinent data from other databases may have been ignored. Second, the study's keyword selection procedure failed to consider any potential ambiguities, which led to the inclusion of superfluous information in the dataset. An incomplete dataset was also produced due to several important keywords being accidentally left out due to an early lack of a thorough grasp of the area. Moreover, the study's dependability may decline with time, it should be viewed as offering preliminary advice. Lastly, it is crucial to recognize that even though machine learning approaches may extract themes with topic words from hundreds of papers, correct labelling of these topics still needs thorough literature reviews. Results from machine learning can lead to future studies, but they cannot completely replace human interpretation and comprehension of the literature.

While this study offers an exhaustive analysis of the IIA domain, it also delineates several uncharted territories warranting future exploration. A critical area of concern is the creation of technology designed to accommodate individuals with a range of disabilities. Subsequent research could particularly target populations with cognitive impairments, aiming to integrate and innovate adaptive technologies that elevate the user experience for these groups.

When data were gathered, little study was done on the use of LLM in IIA. Although research on AI and LLM has been expanding rapidly, with models such as ChatGPT and GPT-4 transforming natural language processing and information retrieval (IR), their particular uses in inclusive information retrieval and accessibility were still in their infancy. In recent years, however, advances in AI-enhanced assistive technologies, adaptive retrieval systems, and LLM-driven conversational agents demonstrate the growing role of intelligent systems in breaking traditional barriers to information access (Zhu et al., 2023). These developments align with broader efforts in inclusive technology and information retrieval (IR) to ensure that digital content is available, navigable, and comprehensible for all users, including those with disabilities and those from underrepresented linguistic or cognitive backgrounds.

In order to improve accessibility for a variety of user groups, recent research highlights the potential of LLM in augmenting assistive technologies by facilitating multimodal interactions through text, audio, and image-based inputs (Adnin and Das, 2024; Raji et al., 2025). By enabling users to engage in natural discourse instead of intricate keyword-based searches, AI-driven conversational agents are proving crucial in information retrieval and improving the accessibility and navigability of digital content (Martínez et al., 2024). For users with impairments and those from low-resource language backgrounds in particular, these technologies, when combined with real-time translation and content simplification models, present a potential path toward overcoming linguistic and cognitive limitations (Fu et al., 2025).

Even with these developments, there are still significant obstacles to overcome in order to guarantee fair, transparent, and bias-free AI-driven retrieval. Research emphasizes the necessity of fairness-aware ranking algorithms that enhance accessibility-aware content ranking and reduce biases in training data (Sitbon et al., 2023). Future research must

concentrate on explainability and trustworthiness to make sure users can understand and evaluate AI-generated responses as AI-generated material becomes more and more integrated into search and retrieval systems. Furthermore, the creation of fully inclusive AI solutions that meet practical accessibility requirements will depend on the advancement of co-design approaches, in which impaired users actively participate in system development. In order to make sure that information retrieval systems empower various user groups rather than exclude them, it will be essential to incorporate universal design principles into every phase of AI model training, evaluation, and deployment going ahead.

Furthermore, there is a specific need to focus on the design of inclusive information retrieval. This would involve developing methods and technologies that offer a more inclusive approach to information retrieval for individuals with disabilities. By doing so, we can empower them to independently access information, allowing for a more seamless integration into the information-driven world.

Future academic contributions can significantly contribute to the advancement of the IIA area by addressing these highlighted gaps and obstacles. With a broad user base catered to, it would become more productive, efficient, and inclusive.

7 Conclusion

In conclusion, this study thoroughly analyzes the IIA area, shedding light on its evolving trends in terms of publication timelines, institutional contributions, and geographical distribution. Even while the area is expanding year over year and there is a definite tendency toward globalization, it has not yet developed strong academic partnerships. The study offers an early knowledge of the present status of research in the IIA sector through a citation analysis of significant researchers and notable papers. The study demonstrates the thematic evolution within the IIA field using LDA for topic modeling in conjunction with the earlier findings. It highlights the present state of research in inclusive technology and information retrieval, pointing out certain areas that need more investigation. The paper also outlines the difficulties and knowledge gaps in the IIA area, offering a guide for the next research projects. By providing a comprehensive overview of the IIA area, this study, unlike prior studies, bridges a significant gap in the literature and acts as an invaluable resource for researchers. The highlighted possible study directions also provide insightful information that can guide future scholarly activity.

8 Disclosure of Funding

None of the authors have received any funding or possess any competing interest that could have influenced the research reported in this paper.

References

Rudaiba Adnin and Maitraye Das. I look at it as the king of knowledge: How blind people use and understand generative ai tools. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–14, 2024.

Herman Aguinis, Ryan K Gottfredson, and Thomas A Wright. Best-practice recommendations for estimating interaction effects using meta-analysis. *Journal of Organizational Behavior*, 32(8):1033–1043, 2011.

Abdul Ahad, Muhammad Fayaz, and Abdul Salam Shah. Navigation through citation network based on content similarity using cosine similarity algorithm. *International Journal of Database Theory and Application*, 9(5):9–20, 2016.

Leonelo Dell Anhol Almeida, Vânia Paula de Almeida Neris, Leonardo Cunha de Miranda, Elaine Cristina Saito Hayashi, and Maria Cecília Calani Baranauskas. Designing inclusive social networks: a participatory approach. In *Online Communities and Social Computing: Third International Conference, OCSC 2009, Held as Part of HCI International 2009, Proceedings 3*, pages 653–662. Springer, 2009.

Vânia Paula Almeida Neris, Frederico Fortuna, Rodrigo Bonacin, Tatiana Silva de Alencar, Luciano de Oliveira Neris, and M. Cecília C. Baranauskas. Addressing brazilian diversity in personal computing systems with a tailoring-based approach. *Personal and Ubiquitous Computing*, 25(2):297–319, Aug 2020. doi: 10.1007/s00779-020-01444-w.

Salha Alzahrani, Vasile Palade, Naomie Salim, and Ajith Abraham. Using structural information and citation evidence to detect significant plagiarism cases in scientific publications. *Journal of the American Society for Information Science and Technology*, 63(2): 286–312, Oct 2011. doi: 10.1002/asi.21651.

Zainab Amjad and Imran Ihsan. Verbnet based citation sentiment class assignment using machine learning. *International Journal of Advanced Computer Science and Applications*, 11(9), 2020.

Simplice A Asongu and Nicholas M Odhiambo. Basic formal education quality, information technology, and inclusive human development in sub-saharan africa. *Sustainable Development*, 27(3):419–428, 2019.

Steven Bird, Ewan Klein, and Edward Loper. *Natural language processing with Python: analyzing text with the natural language toolkit.* " O'Reilly Media, Inc.", 2009.

David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

MAP Bovens et al. Information right: Citizenship in the information society. In *International Political Science Association World Congress Quebec*, 2000.

Danah M Boyd and Nicole B Ellison. Social network sites: Definition, history, and scholarship. *Journal of computer-mediated Communication*, 13(1):210–230, 2007.

Ana CM Brito, Filipi N Silva, and Diego R Amancio. Analyzing the influence of prolific collaborations on authors productivity and visibility. *Scientometrics*, 128(4):2471–2487, 2023.

A Bryne. Promoting the global information commons: A commentary on the library and information implications of the wsis declaration of principles" building the information society: a global challenge in the new millennium"(document wsis/pc-3/dt/6). 2005.

Christian Bühler and Constantine Stephanidis. European co-operation activities promoting design for all in information society technologies: Introduction to the special thematic session. In *International Conference on Computers for Handicapped Persons*, pages 80–87. Springer, 2004.

Elfreda A Chatman. The impoverished life-world of outsiders. *Journal of the American Society for information science*, 47(3):193–206, 1996.

Abon Chaudhuri. A visual technique to analyze flow of information in a machine learning system. *arXiv preprint arXiv:1908.00754*, 2019.

Chaomei Chen. The citespace manual. *College of Computing and Informatics*, 1(1):1–84, 2014.

Xin Chen, Britt Östlund, and Susanne Frennert. Digital inclusion or digital divide for older immigrants? a scoping review. In *International conference on human-computer interaction*, pages 176–190. Springer, 2020.

RV Chikate and SK Patil. Citation analysis of theses in library and information science submitted to university of pune: A pilot study. *Library Philosophy and Practice*, 222: 31–55, 2008.

Suraj Chiplunkar, Anany Maini, Dinesh Ram, Zixuan Zheng, and Yaxin Zheng. Drawxi: an accessible drawing tool for collaboration. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–6, 2019.

Blaise Cronin. Bibliometrics and beyond: some thoughts on web-based citation analysis. *Journal of Information science*, 27(1):1–7, 2001.

Simone Erbs da Costa, Carla Diacui Medeiros Berkenbrock, Lucas Eduardo Rosa de Freitas, and Fabíola Ferreira Sucupira Sell. ilibras: using assistive and collaborative technology to support the communication of deaf people. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, 14(1):11–21, 2019.

Vânia Paula de Almeida Neris, Frederico Fortuna, Rodrigo Bonacin, Tatiana Silva de Alencar, Luciano de Oliveira Neris, and M Cecília C Baranauskas. Addressing brazilian diversity in personal computing systems with a tailoring-based approach. *Personal and Ubiquitous Computing*, 25:297–319, 2021.

Francesca De Felice and Antonella Polimeni. Coronavirus disease (covid-19): a machine learning bibliometric analysis. *in vivo*, 34(3 suppl):1613–1617, 2020.

S De Groote. Measuring your impact: Impact factor, citation analysis, and other metrics: Citation analysis. *UIC Libraries Research Guides*, 2015.

Anjani Dhrangadhariya, Roger Hilfiker, Roger Schaer, and Henning Müller. Machine learning assisted citation screening for systematic reviews. *Studies in health technology and informatics*, 270:302–306, Winter 2020. doi: 10.3233/SHTI200171.

Naveen Donthu, Satish Kumar, Debmalya Mukherjee, Nitesh Pandey, and Weng Marc Lim. How to conduct a bibliometric analysis: An overview and guidelines. *Journal of business research*, 133:285–296, 2021.

Júlio Cesar dos Reis, Rodrigo Bonacin, and M Cecilia C Baranauskas. New perspectives for search in social networks-a challenge for inclusion. In *International Conference on Enterprise Information Systems*, volume 2, pages 53–62. SCITEPRESS, 2010.

Julio Cesar Dos Reis, Rodrigo Bonacin, and M Cecília C Baranauskas. Addressing universal access in social networks: an inclusive search mechanism. *Universal access in the information society*, 13:125–145, 2014.

Leo Egghe. Theory and practise of the g-index. *Scientometrics*, 69(1):131–152, 2006.

Gorka Epelde, Xabier Valencia, Eduardo Carrasco, Jorge Posada, Julio Abascal, Unai Diaz-Orueta, Ingo Zinnikus, and Christian Husodo-Schulz. Providing universally accessible interactive services through tv sets: implementation and validation with elderly users. *Multimedia Tools and Applications*, 67(2):497–528, Dec 2011. doi: 10.1007/s11042-011-0949-0.

Gorka Epelde, Xabier Valencia, Eduardo Carrasco, Jorge Posada, Julio Abascal, Unai Diaz-Orueta, Ingo Zinnikus, and Christian Husodo-Schulz. Providing universally accessible interactive services through tv sets: implementation and validation with elderly users. *Multimedia tools and applications*, 67:497–528, 2013.

Pierre Erbacher, Ludovic Denoyer, and Laure Soulier. Interactive query clarification and refinement via user simulation, 2022. URL `https://arxiv.org/abs/2205.15918`.

Chinwe N Ezeani, Scholastica C Ukwoma, Esther Gani, Prince J Igwe, and Chidimma G Agunwamba. Towards sustainable development goals: What role for academic libraries in nigeria in assuring inclusive access to information for learners with special needs? 2017.

Grace Fox and Regina Connolly. Mobile health technology adoption across generations: Narrowing the digital divide. *Information Systems Journal*, 28(6):995–1019, 2018.

Biying Fu, Abdenour Hadid, and Naser Damer. Generative ai in the context of assistive technologies: Trends, limitations and future directions. *Image and Vision Computing*, 154:105347, 2025.

Simon L Gay, Edwige Pissaloux, Katerine Romeo, and Ngoc-Tan Truong. F2t: a novel force-feedback haptic architecture delivering 2d data to visually impaired people. *IEEE Access*, 9:94901–94911, 2021.

Elise Gerich. Expanding the internet to a global environment but... how to get connected? *Computer Networks and ISDN Systems*, 23(1-3):43–46, 1991.

Fernando Benedet Ghisi, Regina Bóries, Santos Marcos, Denilson Sell, and Jean Varvakis. A reference ontology for digital scientific journals applied to systematic literature review processes. *Transinformação*, 24:91–101, Aug 2012. URL `https://www.scielo.br/j/tinf/a/z7NkZ7Lwz4HbHFQLTMFhhCp/?lang=en`.

Martin Grayson, Anja Thieme, Rita Marques, Daniela Massiceti, Ed Cutrell, and Cecily Morrison. A dynamic ai system for extending the capabilities of blind people. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–4, 2020.

Marie Michèle Grenon, Julie Ruel, Patrick Fougeyrollas, Claude L Normand, André C Moreau, Alejandro Romero-Torres, and Sylvie Gravel. Conceptualizing access to and understanding of information. *Universal Access in the Information Society*, 22(1):83–94, 2023.

Daniel Gruss, Raphael Spreitzer, and Stefan Mangard. Cache template attacks: Automating attacks on inclusive {Last-Level} caches. In *24th USENIX Security Symposium (USENIX Security 15)*, pages 897–912, 2015.

Xiaoxiao Guo, Hui Wu, Yu Cheng, Steven Rennie, Gerald Tesauro, and Rogerio Schmidt Feris. Dialog-based interactive image retrieval, 2018. URL `https://arxiv.org/abs/1805.00145`.

Anne-Wil Harzing and Satu Alakangas. Google scholar, scopus and the web of science: a longitudinal and cross-disciplinary comparison. *Scientometrics*, 106:787–804, 2016.

Anne M. Hayes and Jennae Bulat. *Disabilities Inclusive Education Systems and Policies Guide for Low- and Middle-Income Countries*. RTI Press, Research Triangle Park (NC), 2019. URL `https://www.ncbi.nlm.nih.gov/books/NBK554622/`.

Jorge E Hirsch. An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences*, 102(46):16569–16572, 2005.

Cheng-An Hsieh, Cheng-Ping Hsieh, and Pu-Jen Cheng. Mr. right: Multimodal retrieval on representation of image with text, 2022. URL `https://arxiv.org/abs/2209.13764`.

Andreja Istenic Starcic and Spela Bagon. Ict-supported learning for inclusion of people with special needs: Review of seven educational technology journals, 1970–2011. *British Journal of Educational Technology*, 45(2):202–230, 2014.

Paul T Jaeger. Telecommunications policy and individuals with disabilities: Issues of accessibility and social inclusion in the policy and research agenda. *Telecommunications Policy*, 30(2):112–124, 2006.

Paul T Jaeger and Bo Xie. Developing online community accessibility guidelines for persons with disabilities and older adults. *Journal of Disability Policy Studies*, 20(1):55–63, 2009.

Hamed Jelodar, Yongli Wang, Chi Yuan, Xia Feng, Xiahui Jiang, Yanchao Li, and Liang Zhao. Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey. *Multimedia tools and applications*, 78:15169–15211, 2019.

Kaixin Ji, Danula Hettiachchi, Flora D Salim, Falk Scholer, and Damiano Spina. Characterizing information seeking processes with multiple physiological signals. *arXiv (Cornell University)*, Jul 2024. doi: 10.1145/3626772.3657793.

Sonali Joshi, Padmalata V. Nistala, Hetal Jani, Prachi Sakhardande, and Trevor Dsouza. User-centered design journey for pattern development. *Proceedings of the 22nd European Conference on Pattern Languages of Programs*, Jul 2017. doi: 10.1145/3147704.3147730.

Diane Kelly and Cassidy R Sugimoto. A systematic review of interactive information retrieval evaluation studies, 1967–2006. *Journal of the American Society for Information Science and Technology*, 64(4):745–770, 2013.

Beatrice Wamaitha Kiruki and Stephen Mudogo Mutula. Information communication technology (ict) use for information access by visually and physically impaired persons in public university libraries in kenya. *International Journal of Knowledge Content Development & Technology*, 13(1), 2023.

Vasiliki Kladouchou, Stephann Makri, Sylwia Frankowska-Takhari, Timothy Neate, Andrew MacFarlane, Stephanie Wilson, and Abi Roper. "the internet is hard. is words": Investigating information search difficulties experienced by people with aphasia and strategies for combatting them. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25. Association for Computing Machinery, 2025. doi: 10.1145/3706598.3713808.

Cary LaCheen. Achy breaky pelvis, lumber lung and juggler's despair: The portrayal of the americans with disabilities act on television and radio. *Berkeley J. Emp. & Lab. L.*, 21:223, 2000.

Jolie CY Lam and Matthew KO Lee. Digital inclusiveness–longitudinal study of internet adoption by older adults. *Journal of Management Information Systems*, 22(4):177–206, 2006.

Guanyu Li, Dian Li, and Tang Tang. Bibliometric review of design for digital inclusion. *Sustainability*, 15(14):10962, 2023.

Chuan Liu, Rong Yu, Jixiang Zhang, Shuchun Wei, Fumin Xue, Yingyun Guo, Pengzhan He, Lining Shang, and Weiguo Dong. Research hotspot and trend analysis in the diagnosis of inflammatory bowel disease: A machine learning bibliometric analysis from 2012 to 2021. *Frontiers in Immunology*, 13:972079, 2022.

Paloma Martínez, Alberto Ramos, and Lourdes Moreno. Exploring large language models to generate easy to read content. *Frontiers in Computer Science*, 6:1394705, 2024.

Chris McGinley and Hua Dong. Accessing user information for use in design. *Lecture notes in computer science*, page 116–125, Jan 2009. doi: 10.1007/978-3-642-02707-9_13.

Rupendra Nath Mitra and Dharma P Agrawal. 5g mobile technology: A survey. *ICT express*, 1(3):132–137, 2015.

Zachary Munn, Micah DJ Peters, Cindy Stern, Catalin Tufanaru, Alexa McArthur, and Edoardo Aromataris. Systematic review or scoping review? guidance for authors when choosing between a systematic or scoping review approach. *BMC medical research methodology*, 18:1–7, 2018.

Mark A. Neerincx, Anita H. M. Cremers, Judith M. Kessens, David A. van Leeuwen, and Khiet P. Truong. Attuning speech-enabled interfaces to user and context for inclusive design: technology, methodology and practice. *Universal Access in the Information Society*, 8(2):109–122, Aug 2008. doi: 10.1007/s10209-008-0136-x.

David Nemer. From digital divide to digital inclusion and beyond. *The Journal of Community Informatics*, 11(1), 2015.

VP de A Neris, LD Almeida, LC Miranda, E Hayashi, and MCC Baranauskas. Towards a socially-constructed meaning for inclusive social network systems. In *International Conference on Informatics and Semiotics in Organisations. Beijing*, pages 247–254, 2009.

AF Newell, G Gregor, M Morgan, G Pullin, and C Macaulay. User-sensitive inclusive design in universal access in the information society august 2011. *Volume*, 10:235–24, 2011.

Julie M. Nightingale and Gill Marshall. Citation analysis as a measure of article quality, journal influence and individual researcher performance. *Radiography*, 18(2):60–67, 2012. ISSN 1078-8174. doi: 10.1016/j.radi.2011.10.044.

Pippa Norris. *Digital divide: Civic engagement, information poverty, and the Internet worldwide*. Cambridge university press, 2001.

C Wendy Olphert, Leela Damodaran, and AJ May. Towards digital inclusion–engaging older people in the 'digital world'. In *Accessible Design in the Digital World Conference 2005*, pages 1–7, 2005.

Denis Luiz Marcello Owa et al. Identification of topics from scientific papers through topic modeling. *Open Journal of Applied Sciences*, 10(04):541, 2021.

Kwangil Park, June Seok Hong, and Wooju Kim. A methodology combining cosine similarity with classifier for text classification. *Applied Artificial Intelligence*, 34(5):396–411, 2020.

Cheryl Parsons and Steven F Hick. Moving from the digital divide to digital inclusion. *Currents: Scholarship in the Human Services*, 7(2), 2008.

Marta Perez-Escolar and Fernando Canet. Research on vulnerable people and digital inclusion: toward a consolidated taxonomical framework. *Universal Access in the Information Society*, 22(3):1059–1072, 2023.

Andrew Pinder. Report of the digital inclusion panel. *The Stationery Office*, 2004.

Raminta Pranckutė. Web of science (wos) and scopus: The titans of bibliographic information in today's academic world. *Publications*, 9(1):12, 2021.

Derek J De Solla Price. *Little science, big science*. Columbia University Press, 1963.

OECD. Publishing. *The survey of adult skills: Reader's companion*. OECD Publishing, 2013.

NR Raji, CL Biji, and V Vineetha. Multi-modal generative ai for people with disabilities. In *Multimodal Generative AI*, pages 271–296. Springer, 2025.

Hassan Raza, M Faizan, Ahsan Hamza, Mushtaq Ahmed, and Naeem Akhtar. Scientific text sentiment analysis using machine learning techniques. *International Journal of Advanced Computer Science and Applications*, 10(12), 2019.

Júlio Cesar Reis, Rodrigo Bonacin, and Maria Cecília Calani Baranauskas. Prospecting an inclusive search mechanism for social network services. In *Enterprise Information Systems*. 2010.

Julio Cesar Reis, Rodrigo Bonacin, and M. Cecília C. Baranauskas. Addressing universal access in social networks: an inclusive search mechanism. *Universal Access in the Information Society*, Feb 2013. doi: 10.1007/s10209-013-0290-7.

General Assembly Resolution et al. Transforming our world: the 2030 agenda for sustainable development. *UN Doc. A/RES/70/1 (September 25, 2015)*, 2015.

Leslie S. Adriaanse and Chris Rensleigh. Web of science, scopus and google scholar: A content comprehensiveness comparison. *The Electronic Library*, 31(6):727–744, 2013.

Amartya Sen. Development as freedom. *Development in Practice-Oxford-*, 10(2):258–258, 2000.

Laurianne Sitbon, Gerd Berget, Margot Brereton, et al. Perspectives of neurodiverse participants in interactive information retrieval. *Foundations and Trends® in Information Retrieval*, 17(2):124–243, 2023.

Hannah Snyder. Literature review as a research methodology: An overview and guidelines. *Journal of business research*, 104:333–339, 2019.

Karen Sparck Jones. A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 28(1):11–21, 1972.

Mega Subramaniam, Rebecca Oxley, and Christie Kodama. School librarians as ambassadors of inclusive information access for students with disabilities. *School library research*, 16, 2013.

Dandison C Ukpabi and Heikki Karjaluoto. Consumers' acceptance of information and communications technology in tourism: A review. *Telematics and Informatics*, 34(5): 618–644, 2017.

Jan Van Dijk and Kenneth Hacker. The digital divide as a complex and dynamic phenomenon. *The information society*, 19(4):315–326, 2003.

Jan AGM Van Dijk. Digital divide research, achievements and shortcomings. *Poetics*, 34 (4-5):221–235, 2006.

Nees Van Eck and Ludo Waltman. Software survey: Vosviewer, a computer program for bibliometric mapping. *scientometrics*, 84(2):523–538, 2010.

Nees Jan van Eck and Ludo Waltman. Accuracy of citation data in web of science and scopus. *arXiv preprint arXiv:1906.07011*, 2019.

Polyxeni Vassilakopoulou and Eli Hustad. Bridging digital divides: A literature review and research agenda for information systems research. *Information Systems Frontiers*, 25(3): 955–969, 2023.

Viswanath Venkatesh, Michael G Morris, Gordon B Davis, and Fred D Davis. User acceptance of information technology: Toward a unified view. *MIS quarterly*, pages 425–478, 2003.

Hanna M Wallach, Iain Murray, Ruslan Salakhutdinov, and David Mimno. Evaluation methods for topic models. In *Proceedings of the 26th annual international conference on machine learning*, pages 1105–1112, 2009.

Jinlong Wang, Can Wen, Shunyao Wu, and Huy Quan Vu. A visual mining system for theme development evolution analysis of scientific literature. *International Journal of Digital Content Technology and its Applications*, 4(3):215–223, 2010.

Vishanth Weerakkody, Yogesh K Dwivedi, Ramzi El-Haddadeh, Ahlam Almuwil, and Ahmad Ghoneim. Conceptualizing e-inclusion in europe: An explanatory study. *Information Systems Management*, 29(4):305–320, 2012.

Zhaohan Xiong, Tong Liu, Gary Tse, Mengqi Gong, Patrick A Gladding, Bruce H Smaill, Martin K Stiles, Anne M Gillis, and Jichao Zhao. A machine learning aided systematic review and meta-analysis of the relative risk of atrial fibrillation in patients with diabetes mellitus. *Frontiers in physiology*, 9:835, 2018.

Neng Zhang, Qiao Huang, Xin Xia, Ying Zou, David Lo, and Zhenchang Xing. Chatbot4qr: Interactive query refinement for technical question retrieval. *IEEE Transactions on Software Engineering*, page 1–1, 2020. doi: 10.1109/tse.2020.3016006.

Yue Zheng, Lei Yu, Junmian Chen, Tianyu Xia, Yuanyuan Yin, Shan Wang, and Haiming Liu. Inclusive design insights from a preliminary image-based conversational search systems evaluation, 2024. URL `https://arxiv.org/abs/2403.19899`.

Yutao Zhu, Huaying Yuan, Shuting Wang, Jiongnan Liu, Wenhan Liu, Chenlong Deng, Haonan Chen, Zheng Liu, Zhicheng Dou, and Ji-Rong Wen. Large language models for information retrieval: A survey. *arXiv preprint arXiv:2308.07107*, 2023.

John Zimmerman, Robin E Soler, James Lavinder, Sarah Murphy, Charisma Atkins, LaShonda Hulbert, Richard Lusk, and Boon Peng Ng. Iterative guided machine learning-assisted systematic literature reviews: a diabetes case study. *Systematic Reviews*, 10(1): 1–8, 2021.