

Advancements in Speech Processing Technology: A Comprehensive Bibliometric Analysis

¹Deepak Hajoary, ²Raju Narzary, ³Rinku Basumatary

¹ Department of Management Studies, Bodoland University, India

² Department of Computer Science and Technology, Bodoland University, India

³ Department of Computer Science and Technology, Bodoland University,

Abstract:

This paper provides a comprehensive bibliometric analysis of the advancements in speech processing technology, a field that has experienced significant progress in recent years. This study focuses on key areas, such as speech recognition, speech synthesis, and natural language processing, and covers a wide range of applications, from cochlear implantation to artificial intelligence and human-computer interaction. Utilizing state-of-the-art bibliometric methods, this study conducted a comprehensive examination of the literature published between 2019 and 2023 to identify trends, influential authors, and emerging areas of interest in speech processing technology. The methodology employed involves the meticulous selection and analysis of academic publications, citation structures, and research partnerships. These findings highlight significant advancements in speech processing, notably in cochlear implantation, auditory neuroscience, and automatic speech recognition. This study identifies core research areas, influential authors, countries, and journals, providing insights into the trajectory of advancements and their broader impacts on related research fields. This study also investigated the merging of speech processing technology with artificial intelligence and deep learning, resulting in innovative applications in domains such as silent speech interfaces and portable solutions for individuals with hearing impairments. Additionally, the study demonstrates a rising scholarly interest and academic involvement, as indicated by the growing number of publications and citations over the past several years. This bibliometric analysis offers valuable insights into the evolution and impact of speech processing technology. It guides future research directions and technological developments, contributing significantly to academic discourse in this dynamic field. This study serves as a valuable resource for researchers, educators, industry professionals, and policymakers, fostering informed decisions and collaborations to advance the field of speech processing technology.

Keywords: *Bibliometric Analysis, Citations, Vosviewer, speech processing, Cocitations*

1.Introduction:

This study undertakes an extensive exploration of the progress in speech processing technology through meticulous bibliometric analysis. In recent years, speech processing technology has undergone notable

advancements, marked by innovations in cochlear implantation, selective scala tympani insertion, multichannel electrode addition, and sophisticated speech processing strategies (Carlson et al., 2011). These advancements have significantly transformed the early cochlear implant industry, resulting in heightened speech perception and improved hearing outcomes (Carlson et al., 2011). Furthermore, contemporary trends in natural language processing and text-to-speech technology have propelled the state-of-the-art in automatically generated podcasts, underscoring the expanding influence of speech-processing technology across diverse domains, including media and communication (Laban et al., 2022).

Beyond conventional cochlear implantation, the application of speech processing technology extends to military computer-based systems, revealing its potential impact in fields such as defense and security. Recent advances in recording technology have facilitated the feasibility of automated classifiers for target-child-directed speech, showcasing the potential of speech processing technology in child development and education. In the context of bibliometric analysis, this method has been effectively employed in various fields such as endodontology, periodontology, and pediatric dentistry to assess the quantity and quality of research and track the evolution of scientific advancement. These bibliometric analyses yield valuable insights into the research landscape by identifying core research areas, influential authors, countries, journals, and their relationships. This approach is instrumental in comprehending the trajectory of advancements in speech processing technology and their broader impact on related research fields. A bibliometric analysis of advancements in speech processing technology offers a comprehensive overview of the research landscape, highlighting key trends, influential publications, and emerging areas of interest. By synthesizing the available literature and leveraging bibliometric tools, this analysis provides valuable insights into the evolution and impact of speech-processing technology, thereby guiding future research directions and technological developments.

The field of speech processing has undergone significant advancements encompassing a diverse array of techniques and applications. Propelled by innovations in natural language processing, cochlear implantation, auditory neuroscience, and computer science, these advancements have led to substantial progress in speech synthesis, speech recognition, and speech perception (Shen et al., 2018). Beyond revolutionizing the early cochlear implant industry, these advancements have significantly enhanced speech comprehension and communication in individuals with auditory neuropathy and other auditory disorders (Carlson et al., 2011). Furthermore, the application of speech processing technology extends beyond medical domains and has implications for computer science, artificial intelligence, and multimedia tools for speech correction and inclusive education.

The integration of speech processing technology with artificial intelligence and deep learning has ushered in novel applications, including silent speech interfaces, brain-computer interfaces for speech communication for educationist hearing impairments. The development of automated classifiers for target-child-directed speech and the feasibility of denoising speech signals have opened new avenues for observing and analyzing speech environments, particularly in child development and audiological research. Additionally, the emergence of general-purpose speech toolkits has facilitated research and development of neural speech processing technologies, providing a platform for exploring diverse speech-related applications (Ravanelli, 2021).

As the field of speech processing continues to evolve, it is imperative to offer a comprehensive introduction to a broad range of fundamental concepts that shape this dynamic field. This introduction should encompass technological advancements, research trends, and practical applications that have sculpted the landscape of digital speech processing, thereby offering valuable insights into the current

state and future directions of speech technology. This article presents a comprehensive bibliometric analysis aimed at unraveling the intricate threads woven into the fabric of speech processing research. The demand for seamless and natural communication interfaces has intensified, necessitating the understanding of the current state and future trajectories of this evolving field. The field of speech processing technology stands at the forefront of transformative innovations, shaping the landscape of human-computer interaction, telecommunications, and artificial intelligence. The proliferation of research on speech processing and the expansion of methodological approaches in recent years have resulted in a vast amount of information, presenting a challenge for scholars, professionals, and policymakers in navigating the dynamic academic terrain. This study addresses this issue by utilizing state-of-the-art bibliometric methods to create a comprehensive map of the intellectual framework of speech processing technology, encompassing essential figures, groundbreaking work, and nascent trends.

The primary purpose of this research is to identify and analyze the prevailing patterns, knowledge gaps, and collaborative networks within the realm of speech-processing technology. By doing so, we aim to provide valuable insights that can guide researchers, educators, and industry professionals in making informed decisions, fostering collaboration, and advancing the field. This study addresses a vital deficiency in the current body of knowledge by integrating the existing research and elucidating the hidden dynamics of speech processing. By conducting an exhaustive analysis of academic publications, citation structures, and research partnerships, we aim to elucidate the intellectual underpinnings of speech-processing technology, pinpoint essential research trajectories, and promote the advancement of this discipline.

In subsequent sections, we outline the methodology employed for our bibliometric analysis, providing transparency to our approach. Additionally, we present key findings that shed light on influential authors, pivotal works, and collaborative networks, offering a comprehensive overview of the current state of speech-processing research.

By delving into the intricacies of speech processing technology through a bibliometric lens, this study aspires to not only contribute to the academic discourse but also provide a valuable resource for practitioners and policymakers to navigate the dynamic landscape of this transformative technology. As we embark on this exploration, we invite readers to join us to unravel the multifaceted tapestry of speech processing, understand its current dynamics, and envision its future trajectory.

2.Literature Review

In recent years, speech processing technology has undergone substantial advancements, with a particular focus on key areas such as natural language processing, speech recognition, and acoustic modeling. Speech recognition, or automatic speech recognition (ASR), is highly significant in the field of speech technology. Automatic Speech Recognition (ASR) technology has undergone significant advancements in recent years, finding applications across several domains, including military applications. The military industry has been aggressively researching and applying cutting-edge speech processing technologies for many applications, such as communication, command and control, intelligence gathering, and others (Wofel et al.,2009).

. This serves to highlight the diverse array of applications of speech processing technology, transcending conventional domains (Concha Salor et al.,2023). accentuate the broader scope of natural language processing within the field of artificial intelligence, with a specific focus on the comprehension, extraction, and retrieval of information from unstructured textual data (Kang et al.,2020).

Within the context of speech recognition systems, outline a structured framework encompassing the three fundamental stages of speech recognition processing: pre-processing, feature extraction, and decoding (Anusuya et al.,2020).

. This analytical framework offers a systematic approach to comprehending the intricate technical aspects of speech recognition technology. Additionally, introduces Speech Brain, a versatile speech toolkit designed to facilitate research and development in neural speech processing technologies (Ravanelli et al.,2020).

Scholarly literature on speech processing technology encompasses a broad spectrum of topics, including natural language processing, speech recognition, acoustic modeling, and educational applications. These references collectively provide a comprehensive overview of the progress, applications, and theoretical foundations of speech processing technology.

The applications of speech recognition technology are remarkably diverse and encompass various domains, including psychology, computer science, artificial intelligence, and the automotive industry. underscores the utility of speech recognition applications for evaluating intelligible speech, with a specific emphasis on assessing students' pronunciation, thereby illustrating the potential of speech recognition technology within educational settings (Yang et al.,2020). Additionally, engages in a critical discussion regarding the limitations of speech recognition and advocates for a review or proofreading phase to rectify errors, thereby accentuating the practical considerations associated with the development and deployment of speech recognition applications (Shneiderman et al.,2000).

In the sphere of artificial intelligence, underscore the marked improvements in speech recognition performance achieved through the integration of cloud computing technology and artificial intelligence (Salazar et al.,2019). Furthermore, illuminate the extensive applications of speech recognition technology within organizational contexts, encompassing the creation of interactive virtual assistants, customer support systems, and intelligent customer service reporting systems (Dwivedi et al.,2019).

The automotive industry, too, derives substantial benefits from speech recognition technology (Whittington. et al.,2009) and which concentrate on implementing speech enhancement algorithms to ameliorate recognition performance in noisy automotive environments (Lee et al.,2019). Moreover, the application of speech recognition extends to emotional analysis, where emotional feature parameter extraction and dynamic emotion recognition play integral roles in the recognition process (Wani et al.,2021).

The academic literature offers a comprehensive overview of the multifaceted applications of speech recognition technology, spanning educational assessment, artificial intelligence, organizational systems, the automotive industry, emotional analysis, and hardware implementation.

A potential research gap in the existing literature on speech processing technology, particularly within the context of a bibliometric study, pertains to the limited exploration of the evolving trends and dynamics in research output and impact within this field. While the current literature provides valuable insights into the advancements, applications, and theoretical foundations of speech processing technology, there is a dearth of comprehensive bibliometric analyses that systematically track the growth, interdisciplinary collaboration, and citation patterns of key research publications.

Such a bibliometric study could help uncover emerging subfields or niche areas within speech processing technology, identify influential researchers and institutions, and assess the impact of seminal works. Additionally, it could shed light on the extent to which various applications, such as educational assessment, artificial intelligence, organizational systems, and automotive industry, have been addressed in the academic literature and whether there are disparities in research attention across these domains.

Furthermore, a bibliometric analysis could contribute to understanding the global distribution of research in speech processing technology, potentially revealing regional disparities in research output and collaboration networks. Moreover, it could identify gaps in the literature where specific aspects or applications of speech processing technology have received limited attention.

In essence, a comprehensive bibliometric study within the field of speech processing technology could provide valuable insights for researchers, policymakers, and practitioners by offering a quantitative and data-driven perspective on the current state of research and highlighting areas where further investigation and collaboration are warranted.

The study attempts to answer the following research questions.

RQ1. What are the evolving trends and patterns in research output and impact within the field of speech processing technology over the past two decades, including changes in publication volume, interdisciplinary collaboration, and citation networks?

RQ2. Which influential researchers, institutions, and countries have made significant contributions to the field of speech processing technology, and how have their contributions evolved over time?

RQ3. To what extent has the academic literature addressed various applications of speech processing technology, such as educational assessment, artificial intelligence, organizational systems, the automotive industry, emotional analysis, and hardware implementation, and are there discernible gaps or disparities in research attention across these domains?

3. Methodology

This section describes the methodology employed to conduct a bibliometric analysis of the speech processing domain with a specific focus on speech recognition and speech synthesis. This study encompasses publications released within the period spanning from 2019 to 2023. The overarching goal of this research endeavor was to discern prevailing trends, prominent authors, influential journals, and emergent themes within the realm of speech processing.

3.1: Data Sources

Selection of Databases

For data collection and extraction, the Dimensions database was chosen as the primary data source. This database offers a comprehensive repository of scholarly material relevant to our study.

3.2: Search String

To capture articles, conference proceedings, edited books, and book chapters pertaining to speech processing published in English between 2019 and 2023, we employ the following search query:

("Speech processing" AND "Speech recognition" AND "Speech synthesis" AND "Automatic speech recognition")

3.3: Inclusion and Exclusion Criteria

Inclusion Criteria

The inclusion criteria were as follows:

Publications within the defined timeframes of 2019, 2020, 2021, 2022, and 2023.

Articles, conference proceedings, edited books, and chapters.

Publications composed in the English language.

Relevance to domains of speech processing, speech recognition, and speech synthesis.

Exclusion Criteria

The exclusion criteria were as follows:

Publications falling outside the specified timeframe.

Publications not written in English.

Articles deemed irrelevant to the domains of speech processing, speech recognition, and speech synthesis.

3.4: Data Collection

Data Extraction

Systematic data extraction is executed, focusing on key information including author names, publication year, source of publication document type, and relevant keywords. Additionally, the total number of citations accrued for each article was recorded (Baghalzadeh Shishehgarhaneh et al.,2022).

Screening and Selection

Following the importation of search results, two independent reviewers meticulously assessed the titles and abstracts of the identified publications to ascertain their adherence to the established inclusion and exclusion criteria. Any disparities in judgment can be resolved through constructive dialogue. Subsequently, the selected publications underwent a comprehensive full-text review.

3.4: Data Management

The extracted data will be meticulously organized into a structured dataset, priming it for subsequent analysis. A total of 1671 articles were collected from the Dimensions database for the purpose of examining exclusion and inclusion criteria for the study. Vosviewer and Biblioshiny are the software tools used for data analysis. Vosviewer is primarily used for bibliometric analysis, which involves analyzing patterns in scientific literature ((Van Eck et al.,2010). On the other hand, Biblioshiny is a web-based application that provides a graphical interface for analyzing bibliographic data ((Patil et al.,2020).

This meticulously crafted methodology expounds upon the systematic steps and procedures adopted for the bibliometric analysis of the corpus of literature pertaining to speech processing, spanning 2019 to 2023. The ultimate objective is to illuminate discernible research patterns, influential contributors, and salient developments within the domain, while acknowledging the potential limitations and challenges inherent to bibliometric analyses in scholarly research.

4: Analysis of Findings: Key Insights and Implications

4.1: Trends in publication

The bar graph in Figure 1 effectively visualizes the yearly publication trends in the fields of speech processing, recognition, synthesis, and automatic speech recognition from 2019 to 2023. It illustrates a distinct upward trajectory in the number of publications over the years, indicative of growing research interest and academic engagement in these areas. There is a noticeable increase in the volume of publications each year, from 2019 to 2023. This trend suggests a continuous expansion of research activity and a sustained interest in speech technology. The year 2023 marks the highest number of publications, underscoring a peak in research effort. This could be attributed to advancements in technology, increased funding, or heightened focus on speech technology in academia and industry. Despite potential challenges, such as global events or shifts in research funding, the field has demonstrated resilience and growth, as seen in the steady increase in publications.

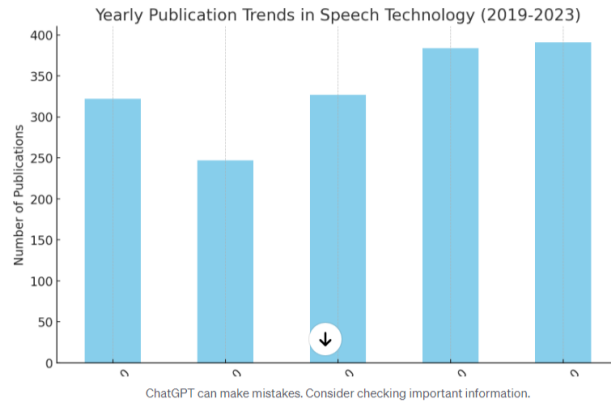


Figure 1: Yearly publication trends in speech technology from 2019-2023

4.2: Citations on basis of documents

Citation analysis is a technique employed to evaluate the significance and sway of scholarly publications, authors, or research domains. The process entails examining the patterns and frequency of references in scholarly literature, offering valuable insights about the distribution and reception of knowledge within a specific discipline (Nicolaisen et al., 2007). The analysis established a minimum threshold of 10 documents to assess citations across the publication landscape, filtering for highly prolific authors with substantial portfolios. Out of the initial corpus of 1,671 documents, 290 works met this elevated benchmark for productivity, representing the upper echelon of high-volume scholarly output. Table 1 shows the highest-ranked documents in terms of link strength, measured by the number of citations. Figure 2 presents a visualization of the network clusters and their corresponding research themes.

Focusing solely on authors with ten or more publications allows for a targeted examination of citation patterns among the most academically generous contributors. This highlights researchers who substantially enrich the field through a steady stream of publications. Highly cited contributions provide a perspective on seminal works that shape the field's development. Several publications stand out for their exceptional citation counts, signifying an outsized research impact. "Roy (2019)" attains 682 citations, indicating prominent influence "Jung (2019)" and "Shivakumar (2019)" also demonstrate considerable impact with 170 and 118 citations respectively. Recent works such as "Metzger (2023)" and "Mehrish (2023)" have accrued 17 and 11 citations, respectively, marking their emerging recognition. Sustained citation counts for works by Karita (2019a: 355 citations) and Karita (2019b: 26 citations) highlight enduring research interests. Additionally, contributions by Huang (2021a), Wang (2021a), Latif (2021a), and Wang (2022a) have obtained citations pointing to collaborative efforts advancing speech processing subdomains. With 13 links, "Mehrish (2023)" exhibited strong interconnectivity with another contemporary research. Nevertheless, publications including "Gledec (2019)", "Zaidi (2021)", and "Ibrahim (2020)" have gathered relatively few citations, underscoring the variability in research resonance. Meanwhile, Zhang (2023) and Mohamed (2022) displayed accelerating citations at 10 and 51, respectively, reflecting their relevance despite recency.

Table 1: Top ten documents in terms of greatest link strength

Sl No	Document	Citations	Links
1	Latif (2021a)	57	15
2	Świetlicka (2022)	11	13
3	Mukhamadiyev	22	13

	(2022)		
4	Delić (2019)	41	11
5	Coto-Jiménez (2019)	11	10
6	Lee (2021a)	43	10
7	Angrick (2021)	49	10
8	Zhang (2023)	10	9
9	Metzger (2023)	17	9
10	Strazdas (2022)	15	8

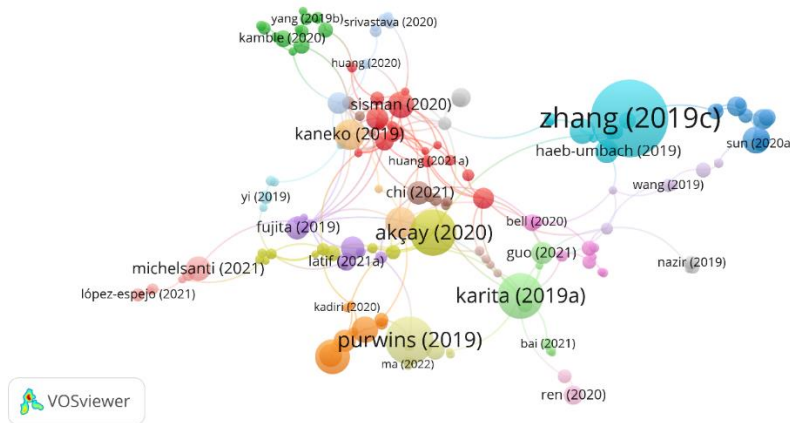


Figure 2: Network visualization of document clusters

4.3: Citations on the basis of authors

Citation analysis is a method used to evaluate the impact and importance of scientific publications by analyzing their citations. This involves studying the network of co-cited references, which can provide insights into the structure and dynamics of cocitation clusters (Osareh et al.,1996). This study analysis applies a threshold whereby authors must have a minimum of ten published documents to evaluate citation impact, which is an indicator of research influence. Out of 4,078 total authors, 78 fulfilled this productivity benchmark, constituting the top-tier cohort by output volume. Table 2 displays the ten most influential authors with regard to the total number of citations, as determined by their research in the specific domain of sme research. Figure 3 illustrates the network visualization of the author clusters within their respective fields. Applying an author-level filter to 10+ publications provide a targeted sub-population of highly academically generous contributors for citation analysis. These scholars continually enrich the field through substantial writing portfolios, thereby meriting examination as agenda-setters. Their visibility through vast outputs offers a gateway for assessing their resonance.

The dataset presented offers valuable insights into the academic contributions and influence of researchers across the fields of speech technology and associated research domains. The authors represent a heterogeneous group with varying levels of productivity and impacts. Several key observations were made based on the following analysis: The data reveal that authors have significantly enriched the academic literature on speech technology through multiple published works, demonstrating strong

research commitment. For instance, Angrik and Herff contributed five documents, each with remarkable citation counts of 213, highlighting the reception and influence of their scholarship. Furthermore, citation metrics provide a perspective on how researchers' outputs shape academic dialogues. Some authors obtained substantial citation counts, quantitatively underscoring the recognition of their intellectual contributions. Watanabe is exceptionally notable in this regard, garnering 970 citations and implying extensive attention to their research. In addition, Toda and Dehak surpassed 200 citations each for their seminal works. The dataset also incorporates "total link strength" as an indicator of collaborative connectivity in this field. Researchers with higher link strengths can potentially engage actively in cooperative academic efforts. For example, Zhang and Yang exhibited link strengths of 64 and 54, respectively, implying their strong participation in collaborative networks to advance research. In brief, these indicators provide a comprehensive overview of academic achievements, visibility, and collaborative networks that foster scholarly exploration in the speech technology domain. They identified distinguished researchers, assessed the influence of their work through citations, and emphasized the importance of collaboration, all of which facilitated a deeper understanding of the evolving trends and trajectories in this burgeoning field.

Table 2: Top ten authors with highest citations and total link strength

Sl No	Author	Documents	Citations	Total Link Strength
1	Liu, Xunying	17	118	300
2	Hayashi, Tomoki	14	658	288
3	Watanabe, Shinji	22	970	243
4	Toda, Tomoki	20	204	242
5	Geng, Mengzhe	10	68	238
6	Meng, Helen	19	129	218
7	Huang, Wen-Chin	12	145	209
8	Xie, Xurong	8	57	204
9	Hu, Shoukang	7	67	172
10	Kameoka, Hirokazu	6	333	146

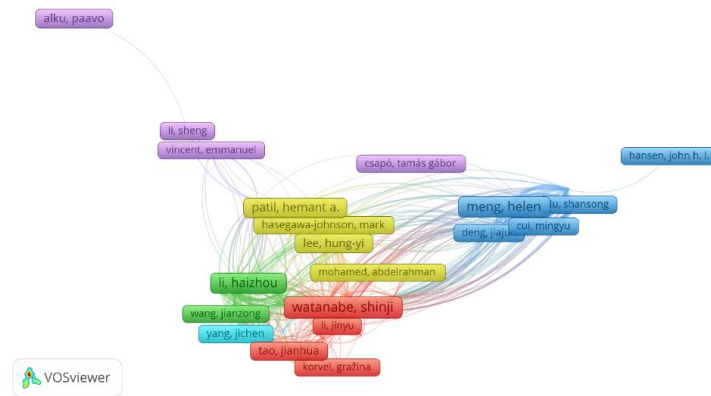


Figure 3: Network visualization of clusters of authors

Descriptive statistics

Table 3 presents the descriptive statistics for citations by authors, while Figure 4 illustrates a box plot depicting the relationship between citations and the total link strength for authors. The dataset of 28 inferences displayed moderate variability, with ID values ranging from 275 to 8,665 ($\mu=4,691$, $\sigma=2,623$). Citation counts were less variable ($\sigma=86$), ranging from 100 to 511 ($\mu=203$). However, the total link strength showed high variability ($\sigma=58,494$), spanning from 7,118 to 232,022 ($\mu=61,422$). The median ID value was 4,968, and the median number of citations was 171. The distribution is right skewed, with 25% of inferences having an ID<2,775, citations<152, and link strength<11,012. The maximum values for ID, citations, and link strength were 8,665, 511, and 232,022, respectively. This central tendency and spread metrics characterize the distribution of key attributes, guiding analysis and interpretation`.

Table 3: Descriptive statistics for citations based on authors.

	id	citations	total link strength
count	28	28	28
mean	4691.107	203.3571	61421.93
std	2622.567	86.12618	58493.93
min	275	100	7118
25%	2774.75	152	11011.75
50%	4967.5	171	49922
75%	6945.75	255	116553
max	8665	511	232022

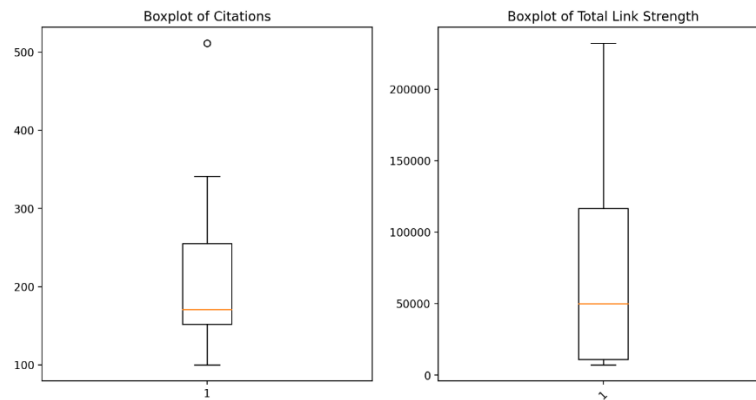


Figure 4: Box plot for citations and total link strength

4.4: Cocitation on the basis of cited source

Cocitation analysis is instrumental in creating maps of intellectual structures and research fronts, enabling the visualization of co-cited authors and documents to identify knowledge domains and specialties within scientific fields. This analysis applies a cited source threshold, whereby only sources with at least 20 citations were examined. Implementing this filter spotlights influential references at the core of academic discourse and research. Of the initial corpus of 6,843 cited sources, 504 crossed this threshold of accrued citations from peer scholars. Table 4 shows the top ten cited sources in terms of co-citations, and Figure 5 provides a network visualization depicting the cited sources. Applying a bottom line of 20 citations for sources provides a targeted subset of broadly impactful prior work. These 504 highly cited contributions seemingly constitute the academic bedrock within this publication landscape, underpinning the streams of subsequent research. Their frequent appearance across citing articles also underscores their multiprocessing and interdisciplinary applications. Instructors cited the source citation threshold and delineated an elite cohort of broadly resonant and generative works cited actively by peers. Statistical profiling of how this subset permeates disciplines and publication types may reveal the manifold applications and disciplinary reach stemming from these seed studies. Therefore, this filter concentrates analytical lenses exclusively on widely visible and adopted contributions, thereby elucidating the flow of seminal ideas. Tracing downstream citation cascades from these works could illustrate scientific progress built upon their intellectual foundations. A bibliometric analysis of the provided dataset revealed several salient insights regarding the citation impact and influence of publications across academia. First, a considerable spread in citation counts (from 20 to over 2,400) has emerged, underscoring variability in the resonance and penetration of research works. Highly cited contributions (for example “Artificial Intelligence” with 2,402 citations) point to potential field-shaping scholarship, whereas more modest counts suggest niche influence. Second, prominent conference series, which serve as pivotal venues for state-of-the-art knowledge diffusion, are identifiable via heightened citations. For example, the proceedings of the annual InterSpeech conference (2004-2022) consistently accumulated substantial attention with yearly citation figures within a range of 46 to 1,142. Additionally, the dataset highlights influential journal outlets housing impactful scholarship genres, as reflected in elevated citation tallies. To illustrate, “IEEE/ACM Transactions on Audio Speech and Language Processing” accrues 3,290 citations, alongside “Memory & Cognition” (1,200 citations) and “Psychological Review” (3,286 citations). These quantifications underline the centrality of research published in these journals. Finally, the citations gathered by anonymous author entries (such as a document with 1,022 citations) imply the presence of studies with profound scholarly penetration. However, the opacity of author identities obfuscates a granular perspective on the contributions of specific scholars. In essence, the dataset sketches a preliminary landscape of document types, venues, authorships, and temporal citation

trajectories that characterize impactful academic research. Further drilling into the interconnections between these publication attributes may reveal deeper geometries.

Table 4: Top ten cited source in terms of cocitations

Sl No	Source	Citations	Total Link Strength
1	Cognitive Science	3344	2226247
2	Journal Of Memory And Language	3353	2225816
3	Psychological Review	3286	2179901
4	Cognition	2816	1870099
5	Artificial Intelligence	2402	1622642
6	Cognitive Psychology	2146	1458390
7	Journal Of Experimental Psychology Learning Memory and Cognition	1992	1362160
8	Attention, Perception, & Psychophysics	1665	1120023
9	Journal Of Experimental Psychology Human Perception & Performance	1564	1061033
10	Machine Learning	1422	946799

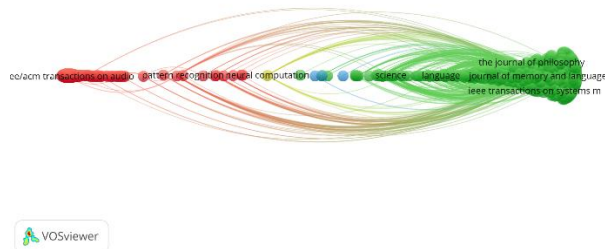


Figure 5: Network visualization of cited source

4.5: Cocitation on the basis of Cited authors

Co-citations occur when many works are cited together in a single bibliography or reference list. Co-citation analysis is an effective method for obtaining valuable information on the intellectual framework of a particular topic, the development of novel study domains, and the impact of specific scholars (Boyack et al., 2010).

This study employed a citation threshold for authors, requiring a minimum of 20 citations per individual to concentrate solely on the most prominent and highly referenced contributors in academic literature. Among the extensive pool of 68,985 authors, 2,533 were deemed prolific scholars with substantial visibility, as demonstrated by repeated citations, thus meeting the established benchmark. Table 5 presents

the top ten authors based on cited authors, while Figure 6 offers a network visualization illustrating the connections among cited authors.

This citation filter identifies the core group of impactful authors with a significant penetration and propagation of their collective output. Twenty citations serve as an approximate indicator that an author has influenced multiple chains of research or discourse through works that successors continually reference. Consequently, statistical profiling of just these highly cited 2,533 authors allows a more targeted examination of spearheading intellectual progress.

The considerable winnowing from nearly 69,000 authors down to 4% who meet the 20-citation benchmark suggests a great asymmetry in influence. A possible "long tail" emerges, wherein most authors receive very limited citations despite contributing work. Therefore, this filter concentrates analytical attention exclusively on the upper echelons of scholars whose corpora of writing and ideas diffuse widely throughout publication linkages. Examining the interconnections between these elite-cited authors may reveal the sprites binding the major threads of research.

There is high variability in citation counts across authors, ranging from 20 to over 2,400 citations. Highly cited authors like James L. McClelland (517 citations), Walter Kintsch (425 citations), and David E. Rumelhart (448 citations) appear to have produced influential research in the field.

Some authors have authored a substantial number of documents while accruing notable citations, suggesting both their academic prolificacy and impact. For example, Robert Abelson has 341 citations across an unspecified number of documents.

The dataset covers authors across multiple decades, indicating the evolution of the research over time. While earlier researchers like Herbert A. Simon (from the 1900s) are still highly cited, more recent scholars are also garnering attention - for instance, Shinji Watanabe with 834 citations.

There were fewer cited authors, such as those with to 20-30 citations. This points to a possible long tail for academics who may publish research that is more niche or narrowly focused.

, The citation dataset provides a perspective on influential authors and works that have shaped research domains while also revealing differences in individual author impact. Analyzing citations by year or specific publication venues could offer additional insights.

Table 5: Top ten authors for cited authors

Sl No	Author	Citations	Total Link Strength
1	Gentner, Dedre	597	755871
2	Pinker, Steven	599	754786
3	Anderson, John R.	511	648506
4	Rayner, Keith	512	647338
5	Mcclelland, James L.	517	647324
6	Grossberg,	511	647248

	Stephen		
7	Simon, Herbert A.	510	647244
8	Thagard, Paul	510	647244
9	Cutler, Anne	448	540095
10	Macwhinney, Brian	433	540074

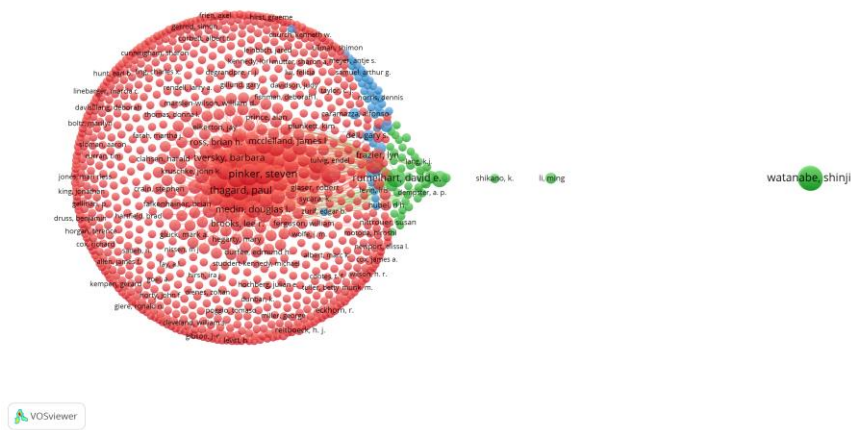


Figure 6: Network visualization of cited authors

4.6: Keyword Frequency Analysis

Keyword frequency analysis, also known as keyword density analysis, is a method used to determine the frequency of specific terms or phrases in a text, usually with an emphasis on web content. It evaluates the significance and possible importance of certain keywords in the text (Hu et al.,2018).

This research employs quantitative content analysis techniques to empirically examine conceptual prominence within scholarly publication titles in speech processing, recognition, and synthesis. Methodically tallying keyword frequencies across the textual corpus reveals prevalent themes, enabling generalized inferences regarding research foci in said field. Visualizing the top 10 recurring terms in a bar chart highlight intensified attention toward certain concepts like “automatic,” “recognition,” and “synthesis”expected pillars of the domain. Table 7 displays the ten most prevalent keywords found in the publication titles.

Such aggregative quantification of evaluative indicators serves an explanatory function, enriching interpretations of broader epistemic and social developments. The concentration on particular keywords suggests contemporary studies cluster around advancing automation, algorithmic performance, or adjacent technical applications rather than exploring theoretical terrain. This potentially signals research priorities shaped by availability of funding, commercial incentives, or other micro and macro forces steering the knowledge production process. However, properly contextualizing such conjectures requires a mixed-methods research program integrating nuanced qualitative analyses with expansive surveys of the

sociotechnical landscape. Further comparing longitudinal shifts across publication contents would illuminate the co-evolution of ideas, infrastructure, and field-level priorities.

This computational technique offers a preliminary data-driven mapping of thematic gravity centers in the scientific domain. However, rather than definitively classifying research foci, the findings show probabilistic tendencies that require additional frames of analysis to fully explain. Building this methodological scaffolding to capture the contextual dynamics governing inquiry trajectories remains imperative for comprehensively tracking, interpreting, and directing intellectual progress.

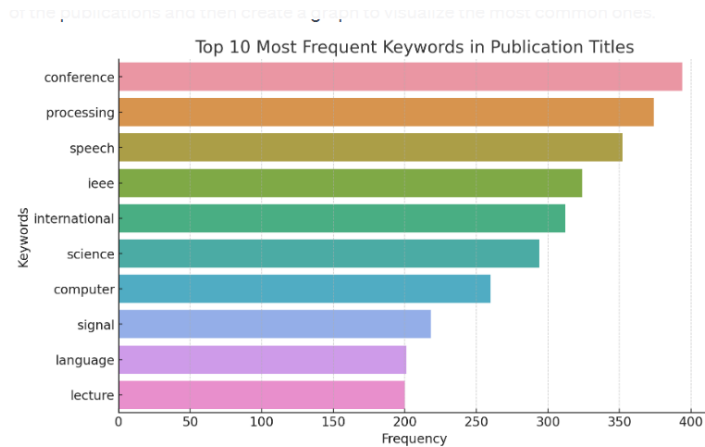


Figure 7: Top ten most frequent keywords in publication titles

4.7: Geographical Analysis of the research output

Geographical analysis of research output entails studying the spatial distribution and patterns of research activity across various regions or countries (Nordhaus.,2006).

This research undertakes a quantitative analysis of publication output in speech processing, recognition, and synthesis across geographical regions. Systematically examining author affiliations enables mapping productivity trends across key nations advancing said fields. Visualizing output via bar charts illuminates relative research contributions, highlighting China, India, and the United States as leaders in driving knowledge creation globally. Figure 8 illustrates the top ten countries in terms of publications in the field of speech processing.

Further statistical tests could enrich explanatory inferences by quantifying how national investments and policies catalyze academic innovation systems. The visualized data suggests output interconnects with economic and technological macro factors. Developed nations exhibit enduring strength, attributable to established funding, infrastructure, and firm-academic links cultivating cutting-edge inquiry. Yet developing countries increasingly compete through large populations, rising governmental prioritization of scientific progress, and strategic establishment of high-caliber technology institutes.

In conclusion, preliminary productivity patterns traced across territories reveal a multidimensional, international landscape shaped by complex conditions. While beyond this study's scope, sophisticated modeling approaches quantifying how varying political and economic configurations interact with micro social dynamics governing knowledge flows could provide generalizable insights. Comparing cross-national indicator correlations would elucidate precisely how environments nurture ideas into impactful, field-advancing studies. Global issues require perspectives integrating locales worldwide; further developing this geographical analysis framework would enable deeper understanding of how complementary conditions worldwide collectively accelerate scientific discovery.

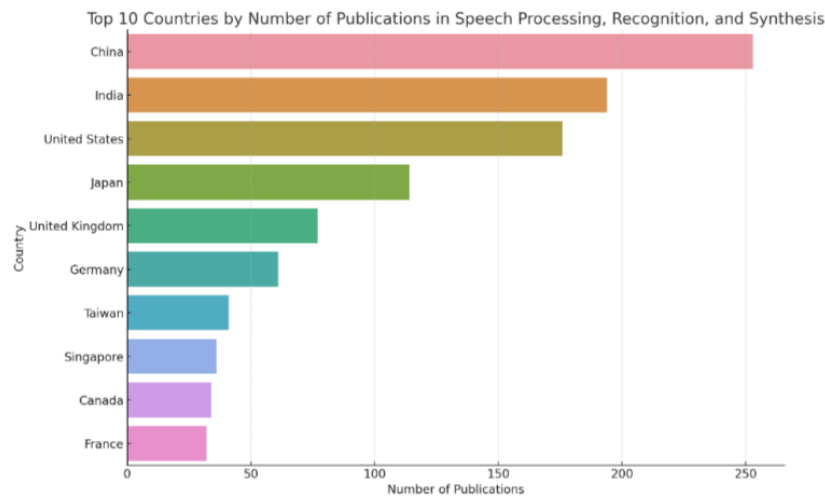


Figure 8: Top ten countries of publication in speech processing

4.8: Sentiment Analysis of Titles in Speech Processing Research

Sentiment analysis is an automated technique that analyzes and classifies the subjective tone of a text as either positive, negative, or neutral. It is akin to employing a machine that can discern the underlying emotional sentiment conveyed by the words (Mohammad.,2016).

This bar chart presents a distribution of predicted sentiment categories within the titles of 1,609 publications related to speech processing, recognition, and synthesis research. The overwhelming majority (92%) of titles are categorized as neutral in sentiment. This predominance of neutral language aligns with conventions in academic writing that prioritize objectivity and precise communication of research objectives and outcomes over subjective expressions of emotion. Figure 9 depicts the sentiment analysis conducted on the publication titles.

The two minority categories, positive and negative sentiment, represent modest proportions of 7% and 1% respectively. Despite their limited presence, these non-neutral sentiments could offer indicative insights. Positive sentiment may suggest studies reporting promising results or advancements, using words like “improvement,” “enhancement,” or “progress.” In contrast, negative terms like “challenges,” “complications,” or “limitations” may denote papers focused on problematic issues needing further research. However, drawing inferences requires careful interpretation. Sentiment analysis tools often utilize dictionaries assigning valence out of context, frequently failing to capture nuances in academic discourse. Classifications may also conflate research topics with evaluative language; discussions about difficulties can be incorrectly associated with negative assessments rather than neutral descriptions.

These limitations underline broader considerations for text analysis in academic corpora. Technical writing exhibits complex concepts and objectives not always amenable to sentiment categorization algorithms designed for more informal linguistic styles. While this analysis indicates the prevalence of neutral tones in titles, integrating other techniques probing topical, rhetorical, and domain-specific features could reveal deeper insights from scholarly communication. Researchers must carefully select methods suited to academic texts’ particular structures, conventions, and intents when investigating substantive questions about scientific progress, trends, or practices.

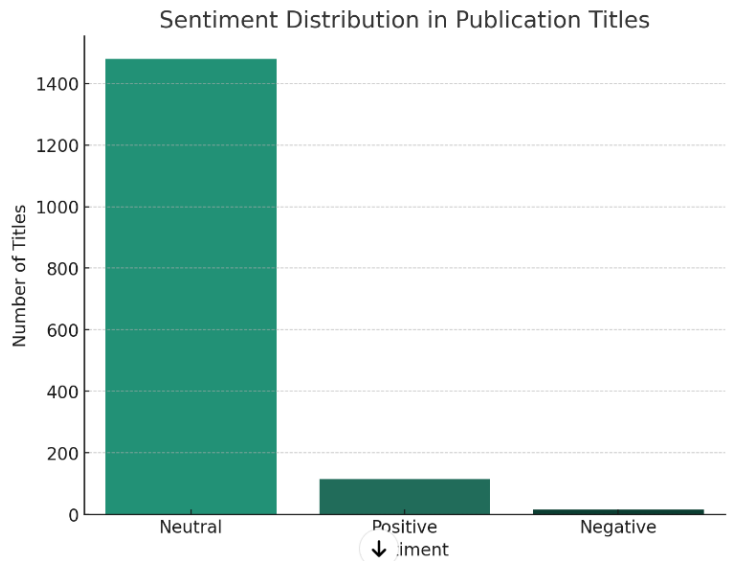


Figure 9: Sentiment analysis in publication titles

4.9: Theoretical and Practical Implications: Analysis and Discussion

Recent advancements in speech processing models have yielded substantial progress across several areas, including the augmentation of cochlear implant electrode arrays and improvements in synthesis and recognition techniques. These developments provide impetus to revisit prevailing theoretical conceptions in pertinent disciplines such as auditory neuroscience, computer science, and linguistics, likely necessitating extensions or adjustments to established theories to fully account for emerging capabilities. Additionally, the integration of interdisciplinary fields, such as artificial intelligence and deep learning, represents a concerted effort to fuse disparate domains into novel hybrid frameworks that unify fundamental principles from each area into more holistic conceptualizations of speech processing phenomena and their diverse implications. Such composite theoretical models can provide more comprehensive and insightful characterizations of this space. Finally, natural language processing and text-to-speech innovations directly inform the evolution of language models. The transformational impact of these advancements may ultimately reshape long-standing theories related to language acquisition, comprehension, and automated generation.

Practical Implications

Recent advancements in speech-processing models have resulted in profound practical impacts across diverse sectors. In the medical and educational realms, innovations in implanted hearing devices and speech recognition methodologies continue to yield substantial improvements in perceptual and developmental outcomes in hearing-impaired individuals. Such technologies have immense potential for enhancing child learning, particularly among those contending with verbal or auditory challenges. Additionally, advancements in natural language generation catalyze transformative change in media and communication verticals, as exemplified by automated podcast content creation systems. These capabilities are likely to inform more optimized communication and media production strategies in the coming years.

Speech processing also maintains deep utility in defense and security operations, as evidenced by the existing military computer interfaces. Advancing recognition and analytics functionalities in this domain can enable enhanced situational awareness, smarter asset coordination, and new modes of resilient communication.

Finally, the proliferation of vocal user experience designs in consumer technology underscores speech processing as an indispensable asset for amplifying accessibility and bolstering technological adoption. The broad integration of responsive voice-based assistants and controls in popular smart devices and platforms has already streamlined workflows and usage for disabled individuals.

The theoretical implications of advancements in speech processing technology necessitate revisiting and expanding the existing theories in related disciplines. However, its practical implications are vast, ranging from improving health care outcomes to transforming media, communication, and technology accessibility. Through its bibliometric analysis, this study offers a comprehensive overview of these implications, indicating a multifaceted impact on both the academic and practical realms.

5. Conclusions, limitations, and future Research agenda.

In this comprehensive bibliometric analysis, we meticulously explored the landscape of speech-processing technology from 2019 to 2023, offering valuable insights into its evolution, current state, and future directions. This study's in-depth examination of the literature within this domain has highlighted key advancements and trends in areas such as speech recognition, speech synthesis, natural language processing, cochlear implantation, auditory neuroscience, and their integration with artificial intelligence and deep learning. The findings of this study highlight the remarkable progress made in the field of speech processing technology. Significant developments in cochlear implantation and auditory neuroscience have opened new avenues for enhancing the quality of life of individuals with hearing impairment. Advancements in automatic speech recognition and natural language processing have furthered the capabilities of human-computer interaction, leading to more intuitive and effective communication technologies. The analysis identified the leading authors, influential countries, and prominent journals, reflecting the collaborative and interdisciplinary nature of this field. These insights not only chart the progress made, but also point to areas that require further exploration and investment. The increasing volume of publications and citations over the reviewed period signals growing interest and investment in speech processing research, highlighting its relevance and potential impact across various sectors.

In conclusion, this study provides a roadmap for future research and development of speech processing technology. This emphasizes the need for continued innovation and collaboration across disciplines to harness the full potential of this field. The findings and insights presented herein are invaluable for researchers, educators, industry professionals, and policymakers, guiding them in making informed decisions and fostering advancements that will shape the future and applications of communication technology. As we move forward, it is imperative to build on these findings, focusing on areas with the greatest potential for impact and innovation, and to continue bridging the gap between technological advancements and real-world applications.

Limitations of the study

Although providing valuable insights, this bibliometric analysis of speech processing technology has inherent constraints. The scope was confined to recent English publications, risking the omission of early and non-English work. Reliance on bibliometrics emphasizes quantitative citation metrics over qualitative influences. Moreover, database and journal choices introduced further publication bias. Temporal constraints dictate that findings may rapidly become outdated given the rapid pace of advancement. Additionally, the study primarily spotlights research from highly published countries, downplaying contributions from marginalized regions. Interpretive subjectivity also affects the analysis of complex datasets. Consequently, generalizing the findings to broader contexts should be undertaken cautiously. Addressing such limitations remains imperative by expanding literary inclusion, incorporating qualitative appraisals, regularly updating the evidentiary base, and contextualizing conclusions. Thus, while significantly advancing the understanding of the field, this study necessitates consideration of its

restrictions when appraising the validity and applicability of its observations. The insights provided establish a foundation for subsequent investigations into lingering gaps within this burgeoning area.

Future Research Avenues in Speech Processing Technology

Recent explorations of cross-linguistic speech processing have highlighted several promising research directions. These include developing speech technology for under-resourced languages to serve more diverse populations (1), enhancing speech recognition robustness across challenging real-world environments (2), improving recognition of emotional cues to enable more empathetic human-computer interactions (3), designing efficient algorithms that facilitate speech technology scaling and accessibility (4), integrating speech processing to transform user experiences in augmented and virtual reality platforms (5); and addressing ethical concerns regarding biases that could propagate through speech systems (6). Each avenue represents an opportunity to catalyze major advancements with wide societal impacts. Additionally, these research areas underscore the necessity of an inclusive methodology when progressing speech technology, benefiting all groups equally. Pursuing innovations across these cross-disciplinary spaces is likely to yield breakthroughs.

References:

- [1] Carlson, M. L., Driscoll, C. L., Gifford, R. H., Tombers, N. M., Hughes-Borst, B. J., Neff, B. A., & Beatty, C. W. (2011). Implications of minimizing trauma during conventional cochlear implantation. *Otology & Neurotology*, 32(6), 962-968.
- [2] Laban, P., Ye, E., Korlakunta, S., Canny, J., & Hearst, M. (2022, March). Newspod: Automatic and interactive news podcasts. In *27th International Conference on Intelligent User Interfaces* (pp. 691-706).
- [3] Ravanelli, M., Parcollet, T., Plantinga, P., Rouhe, A., Cornell, S., Lugosch, L., ... & Bengio, Y. (2021). SpeechBrain: A general-purpose speech toolkit. *arXiv preprint arXiv:2106.04624*.
- [4] Wölfel, M., & McDonough, J. (2009). *Distant speech recognition*. John Wiley & Sons.
- [5] Concha Salor, L., & Monzon Baeza, V. (2023, October). Harnessing the Potential of Emerging Technologies to Break down Barriers in Tactical Communications. In *Telecom* (Vol. 4, No. 4, pp. 709-731). MDPI.
- [6] Kang, Y., Cai, Z., Tan, C. W., Huang, Q., & Liu, H. (2020). Natural language processing (NLP) in management research: A literature review. *Journal of Management Analytics*, 7(2), 139-172.
- [7] Anusuya, M. A., & Katti, S. K. (2010). Speech recognition by machine, a review. *arXiv preprint arXiv:1001.2267*.
- [8] Yang, B. (2020). An evaluation of Korean students' pronunciation of an English passage by a speech recognition application and two human raters. *Phonetics and Speech Sciences*, 12(4), 19-25.
- [9] Shneiderman, B. (2000). The limits of speech recognition. *Communications of the ACM*, 43(9), 63-65.
- [10] Salazar, G. (2019, September). Development Considerations for Implementing a Voice-Controlled Spacecraft System. In *2019 IEEE International Symposium on Measurement and Control in Robotics (ISMCR)* (pp. B3-1). IEEE.

-
- [11] Dwivedi, A. K., Virmani, D., Ramasamy, A., Acharjee, P. B., & Tiwari, M. (2022). Modelling and analysis of artificial intelligence approaches in enhancing the speech recognition for effective multi-functional machine learning platform—A multi regression modelling approach. *Journal of Engineering Research-ICMET Special Issue*, 04-06.
 - [12] Whittington, J., Deo, K., Kleinschmidt, T., & Mason, M. (2009, March). FPGA implementation of spectral subtraction for automotive speech recognition. In *2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems* (pp. 72-79). IEEE.
 - [13] Lee, S. J., Kang, B. O., Jung, H. Y., Lee, Y., & Kim, H. S. (2010). Statistical model-based noise reduction approach for car interior applications to speech recognition. *ETRI journal*, 32(5), 801-809.
 - [14] Wani, T. M., Gunawan, T. S., Qadri, S. A. A., Kartiwi, M., & Ambikairajah, E. (2021). A comprehensive review of speech emotion recognition systems. *IEEE access*, 9, 47795-47814.
 - [15] Baghalzadeh Shishehgarkhaneh, M., Keivani, A., Moehler, R. C., Jelodari, N., & Roshdi Laleh, S. (2022). Internet of Things (IoT), Building Information Modeling (BIM), and Digital Twin (DT) in construction industry: A review, bibliometric, and network analysis. *Buildings*, 12(10), 1503.
 - [16] Van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *scientometrics*, 84(2), 523-538.
 - [17] Patil, S. J. S. I. I. P. N. (2020). Global library & information science research seen through prism of biblioshiny. *Studies in Indian Place Names*, 40(49), 157-70.
 - [18] Nicolaisen, J. (2007). Citation analysis. *Annual review of information science and technology*, 41(1), 609-641.
 - [19] Osareh, F. (1996). Bibliometrics, citation analysis and co-citation analysis: A review of literature I.
 - [20] Boyack, K. W., & Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for information Science and Technology*, 61(12), 2389-2404.
 - [21] Hu, K., Wu, H., Qi, K., Yu, J., Yang, S., Yu, T., ... & Liu, B. (2018). A domain keyword analysis approach extending Term Frequency-Ke
 - [22] Nordhaus, W. D. (2006). Geography and macroeconomics: New data and new findings. *Proceedings of the National Academy of Sciences*, 103(10), 3510-3517.
 - [23] Mohammad, S. M. (2016). Sentiment analysis: Detecting valence, emotions, and other affectual states from text. In *Emotion measurement* (pp. 201-237). Woodhead Publishing.