



**SINDHI LANGUAGE AUTHORITY**

**ARTIFICIAL  
INTELLIGENCE, TECHNOLOGY  
AND FUTURE OF  
SINDHI LANGUAGE**

**National Conference**

**ABSTRACT BOOK**

**September 17-2023**



**SINDHI LANGUAGE AUTHORITY**  
National highway Hyderabad, Sindh, Pakistan  
Tell: 0222-9240050-3  
Website: [www.sindhila.org](http://www.sindhila.org)  
Email: [contact@sindhila.edu.pk](mailto:contact@sindhila.edu.pk)





Artificial Intelligence, Technology  
and Future of Sindhi Language

مصنوعي ذهانت، ٽيڪنالاجي  
۽ سنڌي ٻوليءَ جو آئيندو



INTERNATIONAL SINDHI LANGUAGE CONFERENCE  
SINDHI LANGUAGE

**ARTIFICIAL INTELLIGENCE, TECHNOLOGY  
AND FUTURE OF SINDHI LANGUAGE**

Editor:

**Khalid Azad**

Editor Desk:

**Anees Kaka**

**Shoukat Chachar**

**Shahnawaz Arbab**

**Irfan Ali Mmeen**

**Faisal Memon**

*Conference Committee*

- **Prof. Dr. Ishaq Samejo**  
Chairman & Convenoer
- **Prof: Dr. Ahsan Danish**  
Secretary & Co- Convenoer
- **Ameen Muhammad Laghari**  
Deputy Director / Member
- **Amar Fayaz Buriro**  
Director Ambele / Member
- **Khalid Azad**  
Assistant Director / Member
- **Anees Kaka**  
Assistant Director / Member
- **Rabella Abro**  
Assistant Proffer / Member
- **Zahid Otho**  
Program Officer / Member
- **Ahsan Ahmed Memon**  
Assistant Director / Member
- **Saleem Jarwar**  
PRO / Member

**SINDHI LANGUAGE AUTHORITY**

National Highway Hyderabad, Sindh, Pakistan

Tell: 0222-9240050-3

Email: [contact@sindhila.edu.pk](mailto:contact@sindhila.edu.pk)

Website: [www.sindhila.org](http://www.sindhila.org)



**RESEARCH PAPER**

**مقالا**

Session_1 Sindhi Language, Script, Linguistics & Artificial Intelligence		پهرين ويھڪ: سنڌي ٻولي، رسم الخط، لسانيات ۽ مصنوعي ذهانت	
Topic	Scholar	مقالا نگار	موضوع
Decipherment of Indus Script thought AI	Dr. Syed Azhar Ali Shah	ڊاڪٽر سيد اظھر علي شاھ	مصنوعي ذهانت وسيلي سنڌو لکت جي پاڇ جو تحقيقي جائزو
Standard code registration of Sindhi characters and the proper use of UTF_8	Dr. Ahsan Ahmed Ursani	ڊاڪٽر احسان احمد عرسائي	سنڌي اکرن جي معياري رمز جي رجسٽريشن ۽ يو ٽي ايف 8 جو معياري استعمال
Sentimental Analysis of Sindhi language Tweets Using Machine Learning	Dr. Muhammad Yaqoob Kondhar	ڊاڪٽر محمد يعقوب ڪونڌر	مشيني سکيا جي استعمال سان سنڌي ٻوليءَ ۾ ڪيل نوٽيس جو احساساتي تجزيو
POS Tagging of Sindhi Noisy Social Media Text	Aqsa Khoso	اقصيٰ ڪوسو	سوشل ميڊيا جي لکتن اندر سنڌي ٻوليءَ جي منجهيل ترڪيبڪاري
Incorporating Sindhi Language Resources into the Stanza Pipeline: Enhancing NLP Capabilities	Dr. Sakeena Shah	ڊاڪٽر سڪينه شاھ	سنڌي ٻوليءَ جي وسيلن ۽ ذريعن جون فطري ٻولين جي سرشتيڪاريءَ ۾ صلاحيتون
Ontological corpus for bilingual (Sindhi_English)	Maqsoodah Khatoon	مقصوده خاتون	سنڌي ۽ انگريزي ٻولين جي ڀنڊار (ڪارپس) جو جائزو
Exploring Speech Emotion Recognition System for the Sindhi Language	Hasnain Mahboob	حسنين محبوب	سنڌي ٻوليءَ لاءِ آوازي جذبن جي سڃاڻپ جي ڳولا
The Role of machine Learning in Natural Language Processing	Zulfiqar Ahmed Mahar	ذوالفقار احمد مهر	ٻوليءَ جي قدرتي عمل ۾ مشيني سکيا جو ڪردار





**RESEARCH PAPER**

**مقالا**

Session_2 Sindhi Language, Education & Artificial Intelligene		بي ويھڪ: سنڌي ٻولي، تعليم ۽ مصنوعي ذهانت	
Topic	Scholar	مقالا نگار	موضوع
Linguistic Variations, Am- biguities & Acoustics of Sindhi Language Sounds	Dr. Ayaz Keerio	ڊاڪٽر اياز ڪيريو	ٻوليائي رنگارنگي، ابهام ۽ سنڌي ٻوليءَ جي آوازن جو سرشتو
Preserving of Folk Wisdom of the Sindhi Language in Artificial Intelligence.	Amar Fayaz Buriro	امر فياض بريرو	مصنوعي ذهانت ذريعي سنڌي ٻوليءَ جي لوڪ ڏاهپ جو بچاءُ
An Analysis and Syntactic Parsing of Sindhi Text: A Comprehensive Study based on Sindhi NLP Tools	Dr. Mazhar Dotio	ڊاڪٽر مظھر ڌوٽيو	سنڌي متن جو نحوي تجزيو: سنڌي اين ايل بي ٽولس جي بنياد تي جامع اڀياس
Challenges for Sindhi Lan- guage in AI: Resource Scar- city, Linguistic Complexity, and Community Engagement	Dr. Muteeu- ul- Rehman	ڊاڪٽر مطيع الرحمان	مصنوعي ذهانت ۽ سنڌي ٻوليءَ لاءِ درپيش خطرا: وسيلن جي گهٽتائي، لسانياتي پيچيدگيون ۽ ماڻهن جي شموليت
The Role of AI in Busi- ness Education	Dr. Muhammad Nawaz Tunio	ڊاڪٽر محمد نواز تنيو	ڪاروباري تعليم ۾ مصنوعي ذهانت جو ڪردار
Integrating AI and Chat GPT in Sindhi Lan- guage Teaching (SLT)	Dr. Abdul Waheed Kalwar	ڊاڪٽر عبدالوحيد ڪلوار	سنڌي ٻوليءَ جي تعليم ۾ چيٽ جي پي ٽي ۽ مصنوعي ذهانت کي سلهاڙڻ
Integrating Artificial In- telligence for Sindhi Lan- guage Education: Oppor- tunities and Challenges	Rafiq Ahmed Bhutto	رفيق احمد پٿو	سنڌي ٻوليءَ ۾ تعليم لاءِ مصنوعي ذهانت: موقعا ۽ درپيش مسئلا





## CONFERENCE COMMITTEE



### **Prof. Dr. Ishaq Samejo**

Chairperson/Convenor  
Sindhi Language Authority, Hyderabad

Professor Dr. Ishaq Samejo, a highly reputed scholar with a Ph.D. in Sindhi Language and Literature, has had a significant impact on both education and leadership. He taught at the University of Sindh, Jamshoro, and later became the chairman of the Department of Sindhi, where he inspired students to appreciate Sindhi language and culture. Currently, he leads as the Chairman of the Sindhi Language Authority (SLA) in Hyderabad, Pakistan, focusing on preserving and promoting Sindhi language. He also previously served as the director of Mirza Kalich Baig chair university of Sindh and Director the Institute of Sindhology at the University of Sindh, further demonstrating his dedication to Sindhi literature and heritage. Dr. Samejo has authored numerous books and articles, particularly on Sindhi poetry and literary criticism, enriching Sindhi literature and academic work. His diverse career continues to inspire scholars and enthusiasts, leaving a lasting impact on Sindhi language and culture.



### **Dr. Ahsan Danish**

Secretary/Co-Convenor  
Sindhi Language Authority, Hyderabad

Dr. Ahsan Danish, a luminary in Sindhi Literature with over three decades of dedicated engagement, earned his Ph.D. in Sindhi Literature from the esteemed University of Sindh in 2015, where he delved into the 'Social impact of Shah Latif's Poetry.' His extensive literary contributions encompass a diverse range of genres, including short stories, historical articles, poetry, and scholarly research, such as 'Besukoon Khwab jo Sach' (Short Stories, 2005), 'Larkano-Tareekhi a'een Tahqiqi Mutaliyo' (Articles on the History of Larkana, 2005), and 'Shah Latif ji Shaeri jo Samaji Karij' (Research on Shah Latif, 2016). Dr. Danish, a seasoned scholar, has presented his research in numerous conferences. Since 1998, he has been a dedicated Associate Professor and Head of the Sindhi Department at Government Degree College Larkana, leaving an indelible mark on Sindhi Literature and academic pursuits.





**Mr. Amar Fayaz Buriro**

Director MBILE/Member  
Majid Bhurgri Institute of Language Engineering

Amar Fayaz Buriro is a language engineer, linguist, author and lexicographer of Sindhi language. Currently he works as Project Director in Department of Culture, Tourism, and Antiquities for the government of Sindh. He previously served as an Information Technology specialist for the Sindhi Language Authority where he was responsible for Sindhi computing and the development of a Sindhi Codex for advanced language engineering. Amar was the creator of the first Optical Character Recognition (OCR) system for the Sindhi language, as well as the largest dictionary portal for the Sindhi Language Authority.



**Ameen Muhammad Laghari**

Deputy Director (Admin & Finance) / Member  
Sindhi Language Authority, Hyderabad

Ameen Muhammad Laghari is a Deputy Director (A/F) of Sindhi Language Authority nowadays. He has written 3 books on different displaces, ie Lexicography, Mass Media and others while his research papers have been publishing in different research Journals of Sindhi Language.



**Ms. Anees Kaka**

Assistant Director / Member  
Sindhi Language Authority, Hyderabad

Ms. Anees Kaka is a computer programmer who heads the 'Sindhi Informatics Department' of the Sindhi Language Authority. He has master's degree in computer science from Sindh University. She works on various projects including Sindhi Learning, Sindhi OCR, Sindhi Dictionary, Sindhi Language Research Journal and Encyclopaedia Sindhiana. Ms. Anees Kaka is also the convener and committee member of Sindhi language seminars.



**Miss. Rabella Abro**

Assistant Professor / Member  
University of Sindh, Jamshoro.

Rabella Abro is a graphic designer and faculty member at the University of Sindh. She completed a M.Phil. in Art and Design from the University of Sindh, Jamshoro. In addition to teaching, she has





worked as a graphic designer, utilizing skills and knowledge from both education and professional experience. Ms. Abro, has been an active member of various national & international conferences as a member and convener committee member, contributing through graphic and campaign design.



**Mr. Khalid Azad**

Assistant Director / Member  
Sindhi Language Authority, Hyderabad.

Mr. Khalid Azad, a writer and researcher of Sindhi language, holds the position of Assistant Director (Publication) at the prestigious Sindhi Language Authority in Hyderabad. He's a prolific writer and dedicated researcher, and his scholarly work has been published in respected research journals in Pakistan. He has presented his research in universities and institutions in the United States, Singapore, China, and Taiwan. Khalid Azad has written, translated, and compiled more than 15 books.





## FOREWORD



It is with great pleasure and anticipation that we introduce this abstract book for the one-day conference on "Artificial Intelligence, Technology, and the Future of Sindhi Language," organized by the Sindhi Language Authority, Government of Sindh. This conference, to be held on September 17, 2023, at the Mumtaz Mirza Auditorium of the Sindh Provincial Museum in Hyderabad, Sindh, Pakistan, marks a pivotal moment in our collective journey to safeguard and enhance the Sindhi language in the digital age.

In this age of rapid technological advancement, languages face both unique challenges and unprecedented opportunities. The primary objective of this conference is to delve into the pressing issues that Sindhi faces in the digital era and to explore the computational advancements that can be harnessed for its growth and preservation.

Within these pages, you will find a curated collection of abstracts, offering a glimpse into the wealth of knowledge, expertise, and innovation that will be shared during the conference. Esteemed experts in Sindhi computing and linguistics will present their insights, research, and strategies for addressing the evolving landscape of language preservation and development.

We extend our heartfelt gratitude to all participants, speakers, and attendees who have joined us in this noble endeavor. Your dedication and enthusiasm in the pursuit of preserving and advancing the Sindhi language are commendable.

As we embark on this intellectual journey together, we anticipate meaningful discussions, collaborations, and revelations that will shape the future of the Sindhi language. We hope that this abstract book will serve as a valuable resource and source of inspiration for all those who share our commitment to this cause.

With sincere appreciation and best wishes,

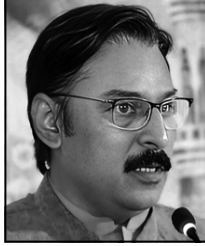
**Prof. Dr. Ishaq Samejo**  
Chairman







## SECRETARY'S NOTE



In the cultural capital of Sindh, Pakistan, an exciting event is on the horizon – the "Artificial Intelligence, Technology, and the Future of Sindhi Language" conference, organized by the Sindhi Language Authority. This conference signifies a critical juncture in the ongoing mission to safeguard the Sindhi language while ushering it into the digital age.

Sindhi, a language steeped in history and culture, faces the challenges of modernization and globalization. The conference, with its 17 carefully curated presentations by linguists, language preservation experts, and computational language development specialists, will shed light on the intricacies of Sindhi, making it more accessible to all.

Importantly, this conference seeks to blend technology with language preservation. With rapid advancements in artificial intelligence and computational linguistics, it will explore how these tools can rejuvenate Sindhi. Topics like machine translation and natural language processing will take the spotlight, promising a brighter future for the language.

Yet, this event goes beyond technology; it underscores the profound cultural significance of Sindhi. It's not just words; it's identity and heritage. The conference acknowledges and celebrates this cultural richness.

Additionally, the conference fosters networking opportunities, uniting linguists, technologists, and culture enthusiasts to collaborate in shaping Sindhi's future.

The release of the conference's abstract book is a symbolic step toward a future where tradition and technology coexist harmoniously, signifying a commitment to embrace the digital age while preserving linguistic and cultural roots..

Sindhi, like many languages, is a unique cultural thread in the global tapestry. This conference exemplifies how technology can be a powerful tool for protecting and promoting linguistic diversity.

As we eagerly await this conference, let's reflect on Sindhi language preservation's remarkable journey. Together, let's embark on this digital voyage, where Sindhi thrives and bridges tradition and technology in the 21st century and beyond.

**Prof. Dr. Ahsan Danish**  
Secretary





Dr. Azhar Ali  
Shah

## A Research Review of Indus Script Analysis through Artificial Intelligence

مصنوعي ذهانت وسيلي سنڌو لکت جي پڇا جو تحقيقي جائزو

In 1922, remnants of the Indus Valley civilization and samples of Indus inscriptions were unearthed from Mohenjo-Daro's ancient ruins. Since then, global scholars, including those from Britain, America, Finland, Russia, France, Egypt, India, and Sindh, have dedicated themselves to deciphering the Indus script. However, despite a century of effort, progress has been limited. To aid this quest, experts in various fields, such as antiquities, linguistics, computer science, and information technology, have proposed artificial intelligence-based algorithms for Indus script identification. These encompass neural networks, Markov models, blockchain approaches, and cryptography, with research findings published in various journals. Yet, much of this valuable research is primarily in English, lacking a comprehensive analysis. This paper fills that gap by providing a chronological and categorized review of these materials, detailing methodologies, analyses, successes, and limitations. Diagrams and flowcharts enhance accessibility for non-experts. The aim is to foster interdisciplinary dialogue among experts, bridging gaps between computer science and linguistic experts. Additionally, this review compares findings with research by Sindhi language experts, such as Sirajul Haque Memon and Atta Muhammad Bhanbharo, on the Indus script.

**Keywords:** Indus script, artificial intelligence, research review, Indus Valley civilization, comprehensive analysis, linguistic analysis, interdisciplinary dialogue, Sindhi language experts.

**Short Profile:** Dr. Azhar Ali Shah, a Professor of ICT at the University of Sindh's Faculty of Engineering & Technology, and Pro Vice Chancellor at Sindh University Campus Larkano, is a distinguished academic. He is a pivotal member of the Executive Committee at the Islamic World Science Citation Center (SCI) in Iran, the Board of Governors of the Abdul Majid Bhurgri Institute of Language Engineering, and Girls Cadet College Larkano. With a Ph.D. in Computer Science & IT from the University of Nottingham, UK, Dr. Shah specializes in cluster, grid, and cloud computing. His extensive awards reflect his excellence as a teacher, researcher, and administrator, while his research spans Sindhi language processing and the application of ICT in Pakistani education and business systems.





Dr. Ahsan Ahmed  
Ursani

## Standard Code Registration of Sindhi Characters and the Proper Use of UTF-8

سنڌي اکرن جي معياري رمز جي رجسٽريشن ۽ يو ٽي ايف 8 جو معياري استعمال

**Abstract:** In the realm of digital information exchange on the Internet, Sindhi, like many other languages, relies on Unicode's UTF-8 codes. Despite Sindhi's successful adaptation to the Information Age, a persistent challenge arises from the incomplete compatibility between the standard Sindhi alphabet and Unicode, mainly due to Unicode's allocation of multiple characters for the letter "h." Consequently, various keyboards and fonts used to input Sindhi characters on computers employ distinct Unicode characters for "H," resulting in visually identical words being stored with diverse Unicode characters in digital systems. Additionally, Unicode lacks two critical letters from the Sindhi alphabet, "Jha" and "Gha," impacting the standard use of Sindhi, hindering word searches in online dictionaries, and posing challenges to language processing systems like Text-to-Speech. Even the creation of modern reference materials, including thesauri, for Sindhi faces similar hurdles. This article asserts the importance of establishing a standard for precise UTF-8 implementation, necessitating the development of a standardized Sindhi character code, officially registered with the International Organization for Standardization (ISO) and the Internet Assigned Numbers Authority, with the active involvement of the Sindhi Language Authority, which plays a pivotal role in facilitating this endeavor.

**Keywords:** Standard code registration, Sindhi characters, UTF-8, Unicode, compatibility, confusion, keyboards, fonts, Internet, Sindhi Language Authority, language processing systems, online dictionaries, reference materials.

**Short Profile:** Dr. Ahsan Ahmad Ursani, part of Hyderabad's renowned Ursani family, holds an engineering degree from Mehran University of Engineering and Technology and a Ph.D. in Signal and Image Processing from France. With nearly eleven years as a professor in Biomedical Engineering at Mehran University, he also dedicates himself to promoting Sindhi language, establishing a Sindhi dictionary on Wiktionary, and sharing educational content on his "Sindhi Science society" YouTube platform, reflecting his dual commitment to Sindhi language and scientific knowledge.





Muhammad  
Yaqoob  
Koondhar

## The Role of machine Learning in Natural Language Processing

ٻوليءَ جي قدرتي عمل ۾ مشيني سکيا جو ڪردار

**Abstract:** Natural Language Processing (NLP) stands as a pivotal domain within Artificial Intelligence, bridging the gap between computers and human language to automate various language-related tasks. NLP empowers computers to comprehend written and spoken language, discern sentiment, and identify crucial elements within text. When coupled with machine learning algorithms, NLP gives rise to systems that can autonomously learn and improve. These NLP-driven tools find versatile applications, from analyzing sentiment in social media posts to extracting named entities from business emails. This paper delves into the distinct role of machine learning within the realm of Natural Language Processing, exploring its diverse applications in text analytics. Machine learning for NLP encompasses the use of algorithms and specialized AI techniques to extract meaning from textual documents, ranging from social media comments and online reviews to financial, medical, legal, and regulatory texts. Essentially, machine learning and AI in NLP and text analytics enhance, expedite, and automate text analysis functions, transforming unstructured text into valuable data and actionable insights.

**Keywords:** Natural Language Processing, NLP, machine learning, text analytics, artificial intelligence, sentiment analysis, named entity extraction, unstructured text, data insights, communication, automation.

**Short Profile:** Dr. Muhammad YaqoobKoondhar, Regional Director of ETSS Management in Malaysia and an Assistant Professor at Sindh Agriculture University, Pakistan, boasts a 16-year career in teaching, research, and web development. He holds a BCS (Hons) and an MBA from the University of Sindh, as well as a professional diploma in Science and Technology Policy from Mehran University of Engineering and Technology. His PhD in Information Technology from IIUM focused on "An integrated model of acceptance for Pervasive Learning from Students' Perspective." His research spans web-based teaching, behavioral studies, and information systems. Dr. Koondhar is celebrated for receiving awards, publishing 44 research papers, and organizing 16 IEEE International Conferences. He also plays a pivotal role as Founder Counsellor and Founder Advisor of IEEE student branches and societies at Sindh Agriculture University.





Aqsa Khoso

## POS Tagging of Noisy Sindhi Social Media Text

سوشل ميڊيا جي لکتن اندر سنڌي ٻوليءَ جي منجهيل ترڪيبڪاري

**Abstract:** Part-of-speech (POS) tagging assigns grammatical function labels to words in a sentence, facilitating various natural language processing tasks. However, social media text introduces noise through elements like emoticons, hashtags, and links, posing challenges for conventional POS tagging methods.

This study focuses on POS tagging of noisy Sindhi social media text. We leverage an existing Urdu Tagset to annotate Sindhi text with noise. By utilizing this pre-existing tagset, we address specific linguistic phenomena in noisy user-generated text.

Using an experimental approach, we extract and preprocess Sindhi social media data from Twitter, creating a small machine learning corpus that is manually tagged. We train the Stanford POS tagger on this corpus, achieving an initial accuracy of 83% f-measure. Through iterative bootstrapping, we increase the training data and improve accuracy to 94%.

This research advances the POS tagging of Sindhi language, enabling better linguistic analysis and facilitating effective natural language processing applications.

**Keywords:** POS tagging, Sindhi language, Noisy text, Social media, Urdu Tagset, Machine learning.

**Short Profile:** Aqsa Khoso is an accomplished Teaching Assistant and a dedicated student currently pursuing a Master's degree in the Department of Computer Science at Isra University, Hyderabad. Throughout her academic journey, she has developed a profound interest in the captivating field of Natural Language Processing (NLP). As a Teaching Assistant, she wholeheartedly contributes her expertise. Her passion for NLP continues to evolve, fueled by the endless possibilities and exciting advancements in this domain. Within the Department of Computer Science, she eagerly explores the intricacies of NLP, pushing the boundaries of her knowledge and skills.





Dr. Sakeena Shah  
Rashdi

## Incorporating Sindhi Language Resources into the Stanza Pipeline: Enhancing NLP Capabilities

سنڌي ٻوليءَ جي وسيلن ۽ ذريعن جون فطري ٻولين جي  
سرشتيڪاريءَ ۾ صلاحيتون

**Abstract:** In recent years, Natural Language Processing (NLP) has become a crucial domain in Artificial Intelligence (AI), focusing on how machines understand, process, and interact with human language. NLP pipelines involve various tasks like sentence segmentation, tokenization, Part-of-Speech (POS) tagging, lemmatization, dependency parsing, and named entity recognition (NER), crucial for NLP-powered applications. However, while NLP has made great strides for widely spoken languages, low-resource languages face challenges due to limited language resources. This work addresses this gap by integrating Sindhi language into Stanza, a well-known NLP library developed by Stanford NLP. The methodology for training Sindhi-specific tokenization and Named Entity Recognition modules is presented. Experimental results show high accuracy, with tokenization and sentence segmentation achieving a 99.96% f-measure accuracy rate and NER reaching 84.27%. These findings demonstrate the adaptability of existing NLP models to effectively serve Sindhi, offering promising opportunities for enhanced linguistic analysis and NLP applications in low-resource languages.

**Keywords:** Natural Language Processing (NLP), Artificial Intelligence (AI), NLP Pipeline, Sindhi language, Stanza, pertained models, tokenization, Named Entity Recognition (NER), f-measure accuracy, low-resourced languages, and linguistic analysis.

**Short Profile:** Sakeena Shah Rashdi is a Teaching Assistant and a dedicated student currently pursuing her Master's degree in the Department of Computer Science at Isra University, Hyderabad. Her academic journey has led her to immerse herself in the fascinating field of Natural Language Processing (NLP), where she has cultivated a strong passion. Her research focus centers on the incorporation of Sindhi language resources into the Stanza Pipeline, with the aim of enhancing NLP capabilities. This endeavor reflects her commitment to advancing linguistic technologies and contributing to the field of NLP.





Maqsooda  
Khatoon

## Ontological corpus for bi-lingual (Sindhi-English)

سنڌي ۽ انگريزي ٻولين جي پنڊار (ڪارپس) جو جائزو

**Abstract:** Maqsood Khatoon Rattar, also known as Rattar MK, currently serves as a Computer Science Lecturer at the prestigious Benazir Bhutto Institute of Management Sciences, Dadu. Holding an impressive academic background with a B.E. and M.E. in Software Engineering from MUET Jamshoro, she brings a wealth of educational experience to her role. Her teaching journey has spanned renowned institutions, including her previous position as a Lecturer at MUET Jamshoro in 2022 and her dedicated service as a faculty member at SZABIST Hyderabad from 2018 to 2021. Beyond her teaching commitments, Maqsood is a dedicated researcher, presenting a noteworthy paper titled "Pre-Web Search Bi-Lingual (Sindhi English) Ontological Corpora" at IMTIC'18 and co-authoring the significant journal article "Sindhi-English Bilingual Parallel Ontological Dictionary," contributing substantially to the Sindh University Research Journal in 2018. Her multifaceted contributions to academia and research exemplify her unwavering commitment to advancing knowledge and promoting the Sindhi language, establishing her as a notable figure in her field.

**Keywords:** Semantic web technologies, semantics, information retrieval, semantic multilingual information retrieval, ontological corpus, bilingual, Sindhi-English, web information retrieval, Google search engine, prototype application, SPARQL query language, ontologies, family, fruit, Friend of a Friend (FOAF).

**Short Profile:** MaqsoodKhatoonRattar, known as Rattar MK, is a Computer Science Lecturer at the Benazir Bhutto Institute of Management Sciences, Dadu. With qualifications including a B.E. and M.E. in Software Engineering from MUET Jamshoro, she brings valuable teaching experience. Maqsood previously served as a Lecturer at MUET Jamshoro in 2022 and at SZABIST Hyderabad from 2018 to 2021. Her research contributions include a short paper titled "Pre-Web Search Bi-Lingual (Sindhi English) Ontological Corpora" presented at IMTIC'18 and a co-authored journal article, "Sindhi-English Bilingual Parallel Ontological Dictionary," published in the Sindh University Research Journal in 2018.





**Hasnain  
Mahboob**

## Exploring Speech Emotion Recognition System for the Sindhi Language

### سنڌي ٻوليءَ لاءِ آوازي جذبن جي سڃاڻپ جي ڳولا

**Abstract:** Speech emotion recognition is a dynamic field within affective computing and social signal processing. However, most studies have focused on languages like English, French, and German due to limited availability of datasets in other languages. This has led to Sindhi and Urdu being referred to as low-resource

languages without publicly accessible datasets. To address this, we present a newly collected dataset for AI-based Speech Emotion Recognition in the Sindhi language. Sindhi is the native language of the Sindhi people residing in southeastern Pakistan. Our dataset consists of 1,897 scripted audio recordings obtained from 89 individuals, covering five distinct emotions: neutral, happiness, sadness, anger, and fear. Data collection involved online platforms like What Sapp and MS Teams, as well as face-to-face interactions with native Sindhi individuals aged between 18 and 25 years. The dataset was carefully labeled, annotated, and cleaned using Label Studio and Audacity. It is the world's largest dataset of Sindhi language. Additionally, we developed and trained machine learning models using audio signal processing and classification algorithms to accurately predict emotions expressed by individuals. This study has the potential to contribute significantly to the introduction of AI-based applications at national and international levels for Sindhi. We aim to make the dataset accessible to the public for non-commercial and academic research purposes.

**Keywords:** Speech emotion recognition, affective computing, social signal processing, Sindhi language, low-resource languages, dataset.

**Short Profile:** HasnainMahboob is currently pursuing a Masters of Engineering in the CIE (Computer and Information Engineering) program at Mehran University of Engineering and Technology, Jamshoro. He completed his Bachelors of Engineering in Telecommunication Engineering from the same university in 2021. His primary area of interest lies in Artificial Intelligence (AI) and machine learning (ML). Hasnain has conducted research in AI, specifically in Speech Emotion Recognition for the Sindhi language. Currently, he is engaged in research focused on Alzheimer's Dementia Recognition using deep neural embeddings. In addition to his academic pursuits, he has gained practical experience through internships at PTCL (Pakistan Telecommunication Company Limited), PBC (Pakistan Broadcasting Corporation), and NTC (National Telecommunication Corporation).







Zulfiqar Ahmed  
Mahar

## The Role of machine Learning in Natural Language Processing

ٻوليءَ جي قدرتي عمل ۾ مشيني سکيا جو ڪردار

**Abstract:** Natural Language Processing (NLP), a pivotal domain of Artificial Intelligence (AI), facilitates human-computer communication in native languages and encompasses various language-related tasks. NLP enables text comprehension, speech recognition, sentiment analysis, and key information extraction. When integrated with machine learning algorithms, NLP evolves into autonomous systems that continuously improve with experience. This paper elucidates the crucial role of machine learning within NLP and explores its diverse application areas. Machine learning (ML) in natural language processing and text analytics entails the utilization of ML algorithms and specialized artificial intelligence (AI) techniques to discern meaning from textual documents. These documents span a wide spectrum, including social media comments, online reviews, survey responses, financial reports, medical records, legal documents, and regulatory texts. Essentially, machine learning and AI's role in natural language processing and text analytics is to enhance, expedite, and automate fundamental text analysis functions, transforming unstructured text into valuable data and actionable insights.

**Keywords:** Natural Language Processing (NLP), machine learning, text analytics, Artificial Intelligence (AI), sentiment analysis, information extraction, human-computer communication, unstructured text, data insights, language-related tasks, autonomous systems.

**Short Profile:** Dr. Zulfiqar Ahmed Mahar is an Assistant Professor at Sindh Agriculture University (SAU), Tandojam, with over 18 years of versatile experience in teaching, research, and industry. He holds an MS in Software Engineering from the University Putra Malaysia and a Ph.D. in Information Technology from the International Islamic University Malaysia. Dr. Mahar has published over 40 research articles and supervised 20 MPhil students. He is an active member of various university committees and currently serves as Assistant Professor (Information Technology), Coordinator ITC, and Deputy Hostel Provost at SAU, showcasing his dedication to academia and professional growth.





Dr. Ayaz  
Keerio

## Linguistic Variations, Ambiguities & Acoustics of Sindhi Language Sounds

ٻوليائي رنگارنگي، ابهام ۽ سنڌي ٻوليءَ جي آوازن جو سرشتو

**Abstract:** Sindhi, a widely spoken language in the Sindh province of Pakistan and parts of India, consists of six language dialects and a diverse set of vowel and consonant sounds. Despite its significance, Sindhi remains one of the least studied languages in terms of linguistic analysis, particularly in the domains of articulatory, auditory, and acoustic phonetics. This study aims to explore the linguistic variations that lead to ambiguities in understanding the articulatory, auditory, and acoustic phonetics of Sindhi.

By utilizing acoustic phonetic features such as fundamental frequency, formants, and durational properties, this research investigates the linguistic variations and ambiguities that exist within the Sindhi phonemic inventory. The objective is to establish a non-deterministic association between utterances and the acoustic properties of Sindhi sounds, enabling the normalization, mapping, and classification of Sindhi language sounds. Through an analysis of speech perception and production ambiguities, as well as the acoustic-phonetic features, this study explores dialectic variations and phonemic redundancies, shedding light on the elusive and significant differences among the geographically distinct dialects of Sindhi. Additionally, the study describes the general acoustic features of the Sindhi phonemic inventory as a whole.

**Keywords:** Phonetics, Formants, Fundamental Frequency, Linguistics.

**Short Profile:** DR. AyazKeerio, a distinguished academic and researcher in computer science, holds a Ph.D. from the University of Sussex, UK, and a background in Computer Science from the University of Sindh, Jamshoro. With over two decades of teaching experience at the University of Sindh, he has excelled as the In-charge of Internet Centers. His teaching expertise spans Computer Networks, Data Communication, Operating Systems, Object-Oriented Programming Languages, Visual Programming Languages, and Database Management Systems. Keerio's research focuses on Digital Signal Processing, Speech Recognition and Synthesis systems, and Computer Networks, resulting in numerous national and international research publications and conference presentations. He actively contributes to education and research through training programs and workshops, underscoring his dedication to advancing computer science and technology.





Amar Fayaz  
Buriro

## Preserving Sindhi Folk Wisdom with Artificial Intelligence

مصنوعي ذهانت ذريعي سنڌي ٻوليءَ جي لوڪ ڏاهپ جو بچاءُ

**Abstract:** Preserving the cultural heritage and linguistic diversity embedded within indigenous languages is crucial. This abstract focuses on the Sindhi language and the challenges of preserving its rich folk wisdom in the face of modernization and globalization. We explore the potential of artificial intelligence (AI) to address this issue.

This presentation highlights the use of natural language processing (NLP) techniques to analyze Sindhi proverbs, folktales, and customs. By leveraging AI, we can document and digitize these cultural expressions, ensuring their longevity. Additionally, AI-powered language models aid in translating and interpreting Sindhi folk wisdom, making it more accessible and fostering renewed interest.

Our research-based presentation provides insights into existing initiatives at the intersection of AI and the preservation of Sindhi folk wisdom. We present successful case studies and offer recommendations for policymakers, researchers, and technology developers to effectively integrate AI, safeguarding the Sindhi language and its cultural heritage.

**Keywords:** Folk wisdom, Sindhi language, Artificial intelligence, Natural language processing, Cultural preservation, Linguistic diversity, Translation.

**Short Profile:** Amar Fayaz Buriro is a versatile language engineer, linguist, author, and lexicographer with a specialization in the Sindhi language. Currently serving as the Project Coordinator at the Abdul Majid Bhurgri Institute of Language Engineering in Hyderabad, Amar previously held a role as an Information Technology specialist at the Sindhi Language Authority. His responsibilities included pioneering Sindhi computing and crafting a Sindhi Codex for advanced language engineering. Amar's noteworthy achievements include developing the first-ever Optical Character Recognition (OCR) system for Sindhi and establishing the Sindhi Language Authority's expansive dictionary portal, the largest of its kind. He also created the pioneering web-based Urdu nastaleeq font, "Amar Nastaleeq," exclusively for online Urdu publishing.





Mazhar Ali  
Dootio

## Comprehensive Analysis and Syntactic Parsing of Sindhi Text: Leveraging Sindhi NLP Tools

سنڌي متن جو نحوي تجزيو: سنڌي اين ايل پي ٽولس جي بنياد تي جامع اڀياس

**Abstract:** This research paper delves into the linguistic analysis and syntactic parsing of Sindhi text using advanced SindhiNLP tools. Sindhi, spoken by millions in Pakistan's Sindh region and worldwide, presents challenges for natural language processing due to its intricate grammar and rich linguistic features. The study explores how SindhiNLP tools address these challenges, enhancing Sindhi text comprehension and processing.

The research surveys NLP and computational linguistics techniques for other languages, assessing their relevance to Sindhi. It emphasizes the necessity of dedicated SindhiNLP tools to accurately handle the language's nuances. The paper introduces SindhiNLP tool architecture and algorithms, highlighting their effectiveness in tasks like morphological analysis, part-of-speech tagging, syntactic parsing, and sentiment analysis. It proposes a comprehensive evaluation framework using linguistic benchmarks, corpora, and Sindhi-specific metrics.

In summary, this research underscores the significance of dedicated SindhiNLP tools for linguistic analysis and parsing. It showcases computational linguistics advancements to accommodate Sindhi's unique characteristics, offering potential applications. This study contributes to NLP and computational linguistics, particularly for less-resourced languages like Sindhi, laying the groundwork for future developments.

**Keywords:** Sindhi; SindhiNLP; Syntactic parsing; Linguistic analysis; NLP

**Short Profile:** Dr. Mazhar Ali Dootio, an accomplished academic, serves as the Chairman and Associate Professor in the Department of Computer Science and Information Technology at Benazir Bhutto Shaheed University Lyari, Karachi. With over 14 years of experience, he excels in research, teaching, and academic administration. His expertise lies in artificial intelligence and natural language processing (NLP), focusing on computational linguistics and NLP. Dr. Dootio's current project tackles linguistic challenges in Sindhi and right-hand written languages, leading to innovative NLP tools and corpora for Sindhi. He's a prolific author, an HEC-approved Ph.D. supervisor, and actively participates in global conferences on higher education and youth opportunities, underscoring his dedication to academia and language advocacy.





Dr.  
Mutee u Rahman

## Challenges in AI for the Sindhi Language: Resource Scarcity, Linguistic Complexity, and Community Engagement

انڪارپوريٽنگ سنڌي لئنگويج ريسورسز ان ٽوڊي اسٽيٽس  
پاڻپ لائين: انھاننگ اين ايل پي ڪيپيبلٽيز

**Abstract:** The Sindhi language faces significant challenges in the field of language processing, including language understanding, translation, and generation, due to its status as a low-resource language. Despite the availability of large language models like OpenAI's GPT and Google's BARD, Sindhi language processing and generation in these models lag behind languages with abundant linguistic resources, such as English and German. Furthermore, popular NLP pipelines like Stanza, Spacy, NLTK, Gensim, and CoreNLP lack support for Sindhi due to limited machine learning data availability. In addition to resource scarcity, the linguistic complexity of Sindhi, encompassing phonetics, rich morphological and syntactic constructions, and dialectal variations, poses significant challenges for language processing tasks. Overcoming these challenges requires the involvement of communities, including native speakers, linguists, computational linguists, NLP experts, and AI professionals, in collecting, annotating, and validating linguistic resources for Sindhi. This research explores the specific challenges faced by the Sindhi language in the context of AI, emphasizing the importance of resources, linguistic expertise, and community engagement in overcoming these obstacles.

**Keywords:** Sindhi, AI, Language processing, Resource scarcity, Linguistic complexity, Community engagement

**Short Profile:** Dr. Mutee U Rehman, a seasoned researcher in Natural Language Processing, Computational Linguistics, and Machine Learning, is a Professor at Isra University's Computer Science Department in Hyderabad. His Ph.D., obtained through the German Academic Exchange Service Program, delved into Computational Grammar Modeling of Sindhi Morphology and Syntax. Dr. Rehman actively contributes to diverse research projects, including DAAD-funded Text to Speech for Urdu, Erasmus+ initiatives like Active Learning in Engineering Education, and the ongoing ICT-INOV project. He plays a pivotal role in developing the NLP pipeline for Sindhi within Stanford University's Open-Source Stanza framework and co-founded the Society for Natural Language Processing Pakistan. Dr. Rehman mentors numerous students in AI and NLP projects for Pakistani languages.





Dr. Muhammad  
Nawaz Tanyu

## The Role of Artificial Intelligence in Business Education

### ڪاروباري تعليم ۾ مصنوعي ذهانت جو ڪردار

**Abstract:** Artificial intelligence (AI) is a powerful tool revolutionizing education by addressing various challenges in teaching, technology, and research. Traditional education methods often face limitations, making AI a game-changer. In general education, AI plays a pivotal role in tailoring learning experiences, selecting appropriate materials, and promoting active student engagement, essential for skill development.

This research paper extensively reviews contemporary literature to explore AI's impact on education. It highlights how AI transcends geographical and national boundaries, breaking down physical barriers. Access to educational content is facilitated by online learning and web-based platforms, making learning materials globally available. Additionally, AI tools, like language translation, further enhance accessibility, aiding students in comprehending subjects of interest.

**Keywords:** artificial intelligence, business education, teaching methods, technology development, educational research, traditional education methods, general education, learners' needs, learning materials, student empowerment, physical barriers, geographical boundaries, national boundaries, international boundaries, Internet, World Wide Web, online learning, web-based learning, accessibility, language translation tools, desired subjects.

**Short Profile:** Dr. Muhammad Nawaz Tunio, an Assistant Professor at the University of Sufism and Modern Sciences, Bhitshah, Pakistan, earned his Ph.D. in Entrepreneurship, Innovation, and Economic Development from AlpenAdria University, Austria, with a government scholarship from Pakistan's Higher Education Commission. He's a young scientist research fellow for Kent State University, Ohio. Dr. Tunio's research focuses on entrepreneurship, Corporate Social Responsibility, youth development, and self-employment. He has an extensive publication record, international conference presentations, and research workshops. As an editor for esteemed publications and special issues in entrepreneurship journals, he contributes significantly to the field. Contact him at [mntunio@gmail.com](mailto:mntunio@gmail.com) or find his ORCID profile at <https://orcid.org/0000-0003-1376-5371>.





Dr.  
Abdul Waheed  
Kalwar

## Integrating AI and Chat GPT in Sindhi Language Teaching (SLT)

سنڌي ٻوليءَ جي تعليم ۾ چيٽ جي پي ٽي ۽ مصنوعي ذھانت کي سهائڙڻ

**Abstract:** This paper explores the application of Artificial Intelligence (AI) and ChatGPT in Sindhi Language Teaching (SLT) within the context of the Sindhi Language Authority's National Conference on Artificial Intelligence, Technology & Future of the Sindhi Language. The aim of this article is to investigate the potential benefits and productivity of incorporating ChatGPT in Sindhi language education. Specifically, it examines how ChatGPT can enhance the learning process, teaching methodologies, and assessment techniques in Sindhi language education. Additionally, the paper highlights the increasing utilization of ChatGPT and its broader implications for language users and education. The research methodology employed in this study includes qualitative content review and incorporates scholastic views from the fields of computer science, transliteration, AI, language development, and education. Based on an analysis of Sindhi language usage in conjunction with ChatGPT, the paper presents recommendations for leveraging this technology effectively in Sindhi language teaching.

**Keywords:** Sindhi Language Teaching (SLT), Language Learning, Language Development, AI, ChatGPT, etc.

**Short Profile:** Dr. Abdul Waheed Kalwar, an eminent academic, is an Associate Professor at Sukkur's Government College of Education. In 2019, he earned a Ph.D. in Sindhi Language and Literature from the University of Sindh, Jamshoro, specializing in "A Sociolinguistic Study of Borrowed Words in Shaikh Ayaz's Sindhi Poetry." His expertise in Sindhi studies is showcased through two published books and numerous articles. Dr. Kalwar's commitment to advancing Sindhi language and literature has earned him high esteem in academia. Based in Sukkur, he continues to be a significant contributor to the field, enriching our understanding of Sindhi culture and heritage.





Rafique Ahmed  
Bhutto

## Integrating Artificial Intelligence for Sindhi Language Education: Opportunities and Challenges

سنڌي ٻوليءَ ۾ تعليم لاءِ مصنوعي ذهانت:  
موقعاً ۽ درپيش مسئلا

**Abstract:** Artificial intelligence (AI) is a powerful force reshaping various sectors, including education. This paper examines AI's role in Sindhi language education, emphasizing its impact. Sindhi, a significant language in Pakistan's Sindh region and parts of India, faces preservation and educational challenges like other regional languages. AI-driven language learning platforms offer personalized, interactive experiences. AI systems, utilizing natural language processing (NLP) algorithms, analyze learners' needs, tailoring content for more effective language acquisition. Virtual assistants and chatbots act as language tutors, providing immediate guidance. AI facilitates immersive learning via virtual reality (VR) and augmented reality (AR), enabling practice in real-world contexts. AI also aids Sindhi literature preservation and accessibility. Ethical collaboration among stakeholders is vital to promote linguistic diversity. AI in Sindhi education empowers learners, preserves culture, and enriches the language ecosystem.

**Keywords:** Artificial intelligence (AI), Sindhi language education, personalized learning, natural language processing (NLP), virtual reality (VR), cultural heritage, linguistic diversity.

**Short Profile:** Rafique Ahmed Bhutto, Director at IIT, is a distinguished AI researcher specializing in large-scale language models and natural language processing. Holding an MPhil in IT from the University of Sindh, he delivers keynote addresses at national and international conferences, shaping discussions on AI's future implications and socio-economic impact. As a visiting faculty member at the College of Modern Sciences and IMSA, he imparts software engineering and IT expertise to students. Rafique is renowned for his unwavering dedication and substantial contributions to AI research, inspiring breakthroughs and nurturing a comprehensive understanding of AI's vast capabilities. His influence in the field is widely acknowledged.

