Support for Multilingual Interactions for Rural, Tribal, and Semi-Urban Students for Career Guidance : A systematic review

Neerja Kasture
neerja.kasture@iitgn.ac.in
Computer Science and Engineering
Indian Institute of Technology Gandhinagar

Tapananshu Manoj Gandhi
tapananshu.gandhi@iitgn.ac.in
Mechanical Engineering
Indian Institute of Technology Gandhinagar

Ashmit Chhoker
chhokerashmit@iitgn.ac.in
Mechanical Engineering
Indian Institute of Technology Gandhinagar

Dewansh Singh Chandel
dewanshsingh.chandel@iitgn.ac.in
Computer Science and Engineering
Indian Institute of Technology Gandhinagar

Abstract

Access to career counselling in India is often limited for students from rural, tribal, and semi-urban regions due to linguistic, cultural, and infrastructural barriers. Most existing platforms primarily operate in English, restricting effective guidance for learners who are more comfortable in regional or low-resource Indian languages. This systematic review investigates the use of artificial intelligence (AI) and multilingual technologies to bridge this gap. Following PRISMA guidelines, we screened 17 studies that focused on AI-driven career guidance systems, including chatbots, apps, and machine learning models, tailored to underserved student populations. Our analysis highlights four key areas: predictive accuracy of AI systems, interactive platforms for personalized guidance, multilingual and culturally adapted educational tools, and persistent socio-economic and infrastructural challenges. While AI and multilingual technologies show promise in enhancing accessibility and personalization, significant gaps remain in inclusivity, language coverage, cultural relevance, and field validation. We conclude that effective career guidance for underserved students requires integrating technological innovation with human support, contextual adaptation, and ecosystem-level interventions to ensure equitable and meaningful educational and professional outcomes.

Keywords

Career guidance, Multilingual AI, Chatbots, Rural education, Tribal students, Semi-urban students, Systematic review, PRISMA

ACM Reference Format:
Neerja Kasture, Tapananshu Manoj Gandhi, Ashmit Chhoker, and Dewansh Singh Chandel. 2025. Support for Multilingual Interactions for Rural, Tribal, and Semi-Urban Students for Career Guidance : A systematic review. In . ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

Career counselling is vital in guiding students' educational and professional choices, yet access to such support in India is hindered by linguistic and cultural barriers. Most existing platforms operate primarily in English, limiting their effectiveness for students from rural, tribal, and semi-urban regions who may lack fluency in these languages. As a result, many learners are unable to fully articulate their interests and aspirations, leading to incomplete or misaligned career assessments.

The absence of native language support also reduces cultural relevance. For instance, a student from a tribal area in Odisha who communicates best in Odia may struggle to engage with English-based platforms, receiving advice disconnected from their context. This linguistic and cultural gap perpetuates inequities in career opportunities for underserved communities.

While AI-powered multilingual platforms and state-level portals have made progress, current solutions remain limited. Challenges include:

- Narrow language coverage, often excluding tribal and low-resource dialects.
- Poor translation quality and lack of cultural nuance.
- Minimal support for voice-based interaction.
- Insufficient career-specific training data.
- Inability to switch seamlessly between languages during counselling sessions.

To address these gaps, this study aims to explore and design AI-driven multilingual career counselling solutions tailored to the needs of rural, tribal, and semi-urban students. The objective is to enhance accessibility, inclusivity, and contextual relevance, enabling learners to engage in career guidance using the language they are most comfortable with and thereby make more informed educational and professional decisions.

2 Methodology

2.1 Eligibility Criteria

Studies were considered eligible if they:

- Addressed the use of large language models (LLMs), chatbots, or AI tools in the context of career guidance, counselling, or student support.

- Focused on multilingual interactions, particularly Indian languages, regional/low-resource languages, and rural, tribal, or semi-urban contexts.
- Provided either empirical results (e.g., prototype development, field testing, app/website usability) or methodological insights (e.g., chatbot design, dataset creation, integration with job-market data).

Exclusion criteria included:

- Studies limited to teaching pedagogy without career relevance.
- Reports without accessible metadata (e.g., missing title, abstract, or keywords).
- Non-English studies. Only peer-reviewed journal articles and conference proceedings were included; unpublished manuscripts were excluded.

2.2 Information Sources

Eight bibliographic databases and repositories were systematically searched:

- Google Scholar (467 records)
- SpringerLink (33 records)
- IEEE Xplore (87 records)
- ScienceDirect (11 records)
- Scopus (75 records)
- arXiv (1 record)
- ACM Digital Library (1 record)
- International journal portals (2 records)

2.3 Search Strategy

A Boolean query was consistently applied across all sources, with minor syntax adjustments per database:

```
("multilingual LLM" OR "career counselling AI" OR "multilingual chatbot" OR "AI in education")
AND ("career guidance" OR "counselling" OR "student support")
AND ("India" OR "Indian languages" OR "low-resource languages" OR "regional languages")
AND ("rural" OR "tribal" OR "semi-urban" OR "underserved communities" OR "vernacular")
```

2.4 Selection Process

(1) All retrieved papers were scraped and merged into a master dataset.
(2) Deduplication was performed based on DOI and title.
(3) Positive and negative filtering was applied to the merged dataset:
   - Positive filter: papers containing “career”, “multilingual”, “chatbot”, or “student” in the title or author keywords.
   - Negative filter: exclusion of papers containing “health” or “teaching” in title/keywords.
(4) A Boolean flag system was applied at the full-text level to mark whether each paper addressed:
   - **When:** Year of publication
   - **Who:** Developers, researchers, educators
   - **What:** Chatbot, app/website, ML model for career prediction
   - **Why:** Motivation for multilingual support, equity, and inclusion
   - **How:** Technical aspects, including voice I/O, job market integration, open datasets, and field studies in rural/tribal/urban settings

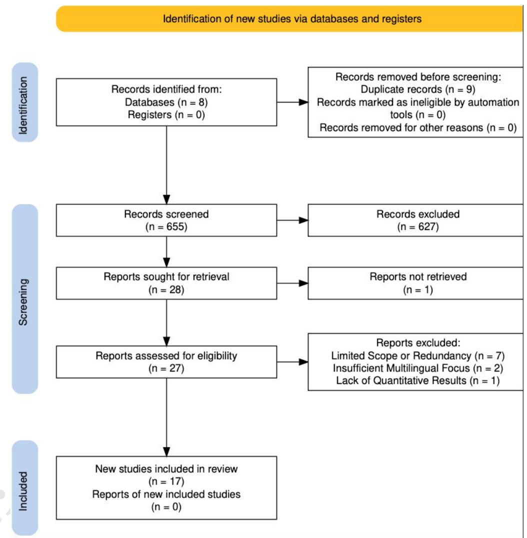

Figure 1: PRISMA flow diagram

3 Findings

This systematic review examines 17 key research papers to understand how technology can help bridge the career guidance gap for rural, tribal, and semi-urban students in India. Despite making up nearly 70% of India's population, these students often lack access to quality career counseling due to language barriers, limited resources, and geographic isolation. Recent advances in artificial intelligence and multilingual technologies offer promising solutions, but questions remain about their effectiveness and accessibility for diverse student populations. Our analysis covers four critical areas: how well AI systems can predict and personalize career advice, the development of interactive platforms like chatbots, progress in multilingual education and language technologies, and the persistent challenges around funding and infrastructure that continue to limit access to career guidance in underserved communities.

3.1 How Career Guidance Has Changed: From Simple Tests to Smart Systems

The research shows that career guidance has moved away from basic personality tests to smart AI systems that can both predict good career matches and provide personalized advice.

3.1.1 How Well Machine Learning Can Predict Careers. Many studies show that computer algorithms can be surprisingly good at suggesting careers. Some of the most successful approaches use what researchers call "ensemble methods" - basically systems that combine multiple algorithms. For example, one study got **96.31% accuracy using Random Forest** [1] and another reached nearly **90% with XGBoost** [2].

But other algorithms can work even better in some cases. One research team found that the **K-Nearest Neighbors (KNN) method achieved 97.1% accuracy** when combined with a chatbot system [3]. The "Career Navigator" platform also used KNN successfully [4].

Neural networks have also been successfully applied. The "VIBE" platform uses something called a **Recurrent Neural Network (RNN)** and gets over **96% accuracy**, while a system built for Sri Lankan students combines different types of neural networks for both course and career recommendations [5, 6]. What this tells us is that there are many different ways to build accurate career prediction systems, and the best choice depends on what kind of data you have to work with [1–3, 5].

3.1.2 Making Systems More Personal and Interactive. Beyond just predicting careers, the most advanced systems use Natural Language Processing to offer personalized help. One common feature is automated resume analysis - platforms like "Career Assist," "VIBE," and "CareerAlly" can read a student's resume, figure out what skills they have, and suggest improvements [5, 7, 8]. One study that focused just on this task reported **91% accuracy in screening resumes** using Random Forest [9].

A newer approach uses something called **Retrieval-Augmented Generation (RAG)** with Large Language Models like Google's Gemma and Mistral-7B. These systems can give much more detailed, context-aware advice by connecting a student's profile with huge databases of job information [10]. Instead of just saying "you might like engineering," they can explain why and what steps to take.

3.2 AI Chatbots and Interactive Platforms for Career Help

One of the biggest trends in recent research is building chatbots and other interactive systems to make career guidance more accessible and engaging.

Many current platforms use solid, open-source tools. One system combines its KNN prediction model with a chatbot built on the **RASA framework**, which helps it understand what users are asking and manage conversations better [3]. Another platform called the "Intelligent Career Counselling Chatbot (ICCC)" uses multiple machine learning models along with Python's Chatterbot tool and even includes an "emotional intelligence" feature to provide empathetic responses [11].

Some researchers are going beyond text-based chat. One interesting approach uses **VideoBot systems** - instead of just text responses, the system selects pre-recorded video clips of a real counselor using **BERT-based similarity matching**. This creates a more human-like interaction, though the researchers note it requires a lot of storage space and computing power as the video database grows [12].

3.3 Breaking Down Language and Economic Barriers

While these technologies are impressive, making them work for rural, tribal, and semi-urban students requires dealing with both language problems and economic challenges.

Why Mother Tongue Education Matters: The study of the **Multilingual Education (MLE) program in Odisha** provides clear evidence that students learn better when they can use their native language. The research showed that tribal students who got early education in their mother tongue had better classroom participation and were less likely to drop out [13]. This is important because it means any career guidance system needs to work in languages students actually understand, not just Hindi and English.

Technology Solutions for Language Support: Several studies show practical ways to make this happen. The "VocabuJoy" system uses real-time language processing and game-like features to help students improve their vocabulary, with users showing **35–40% improvement** [14]. For Indian languages specifically, research shows that specialized models like **IndicBART** work much better than general-purpose language models [15]. When good datasets aren't available, some teams have used **web scraping** to build their own training data [16].

For handling multiple languages, research on machine translation shows that using **English as a bridge language** and converting different scripts to a common format can help support dozens of Indian languages [17]. Some systems even tackle economic barriers directly - for example, the Sri Lankan platform includes a **chatbot that helps students find scholarships and financial aid** [6].

3.4 Key Gaps in Current Career Guidance Systems

Despite the promising technical advances shown in the literature, several important gaps remain that limit how well these systems serve rural, tribal, and semi-urban students.

The most obvious problem is the lack of trained counselors who can work with diverse student populations. The MLE program evaluation in Odisha clearly shows there aren't enough educators who speak local languages and understand the unique challenges these students face [13]. Even when good technology exists, it doesn't help much if there's no one to guide students through using it or interpreting the results.

Another major issue is the mismatch between what these systems offer and what students actually need. Most current platforms focus on matching skills to careers, but they miss bigger challenges like family expectations, financial constraints, and limited exposure to different career options. A student might have aptitude for engineering, but if their family needs immediate income or doesn't understand what engineering involves, the recommendation becomes meaningless.

The data problem also creates serious limitations. Many systems rely on information that doesn't represent rural and tribal communities well. When researchers have to scrape data from the internet just to train their models [16], it raises questions about whether these systems truly understand the local context and opportunities available to these students [15, 17].

Language barriers remain a significant challenge too. While some systems claim multilingual support, few actually handle the complexity of local dialects and cultural expressions that students use when talking about their interests and concerns. This creates a communication gap that affects both understanding the student's needs and explaining career options effectively.

There's also a timing problem. Most career guidance happens too late - when students are already in 10th or 12th grade and major educational decisions have already been made. By then, many rural and tribal students have already dropped out or been channeled into limited educational paths.

Finally, current systems largely ignore the support ecosystem these students need. Career guidance isn't just about individual advice - it requires family education, teacher training, and community engagement to be truly effective. Most existing platforms treat career guidance as an individual process, missing the collective nature of decision-making in many rural and tribal communities.

These gaps suggest that while technology can be part of the answer, the real challenges are more fundamental and require understanding the broader context of how career decisions actually happen in these communities.

4 Design Space Overview

Design space analysis is a systematic process used to explore and map out the range of possible solutions for a design problem. By making both design choices and their underlying rationale explicit, this analysis facilitates a deeper understanding of trade-offs and guides informed decision-making throughout the design process.

This design space highlights the key dimensions along which prior work can be characterized, compared, and evaluated. By organizing studies across these dimensions, we are able to identify recurring clusters of approaches, surface trade-offs, and expose gaps where future innovations are most needed.

We have followed the inductive approach towards building our design space, keeping in mind the 6Ws, taking into account for whom the system is intended, what functions it provides, when and why it was developed, how it is implemented and evaluated, and the technological models and languages that underpin it.

4.1 Dimensions of the Design Space

To synthesise findings from the systematic review we derived a set of recurring *dimensions* that characterise multilingual career-guidance systems for rural, tribal and semi-urban students. These dimensions were identified inductively from the data-extraction spreadsheet (see Appendix A).

4.1.1 Stakeholders and Target Users (Who?)

Definition. Who the system is designed for, and who is involved in development, deployment and evaluation.

Categories. The main categories in consideration are researchers who may use the system for further research purposes, developers who can implement the system as per their needs and educators/students at the senior secondary level who may use the system directly. The papers are categorised in terms of which category they serve directly

4.1.2 Functions and Features (What?)

Definition. The concrete capabilities offered by the system.

Categories.

- **Career counselling support:** skill and interest assessment, and job prediction
- **Multimodal support:** Support for multimodal interaction through voice input & output
- **Market integration:** real-time job market feeds, vacancy matching
- **Multilingual features:** native support for multiple Indian languages, translation, localized content, cultural adaptation

4.1.3 Motivations & Goals (Why?)

Definition. The driving rationale behind the system made and why studying such a system is helpful to our research goals.

Categories.

- **Equity / inclusion :** aims to reduce access gaps for underserved communities, especially rural or tribal communities
- **Reproducibility** Whether the datasets and models used in the system are open-sourced.
- **Benchmarking** Whether the models have provided accuracy estimates for their model in order to establish a baseline

4.1.4 Implementation details (How?)

Definition. The details of how the system has been designed, developed and deployed.

Categories.

- **Builds chatbot :** Whether the system has trained a chatbot to interact with users for career guidance
- **Conducts field study/testing in target group** Whether the authors of the study have done a field study/ tested their system in our target group: rural, tribal or semi urban senior secondary school students
- **Deployment status** Whether the system is a usable website/app
- **ML models for career prediction** Whether the system has trained a ML model for career prediction

Additionally, we also note which ML models have been used for the career prediction purpose and which languages are supported by the systems that support multilingual interactions.

4.2 Design Space Analysis

A analysis and synthesis of our findings helps us construct a design space. The following table summarizes the salient points covered by the research done.

Table 1: Summary of Design Space Features

Feature	Range in Dataset
Model Types	SVM, KNN, RNN, GNN, Decision Tree, RAG, Open source
Language Support	English, Hindi, Gujarati
Input Modality	Resume upload, text chat, voice chat, academic profile
Output Modality	Recommendations, career forecasts, questionnaires, voice

We can observe that there is a rising interest in this space as the number of papers tripled in the past year alone in Figure 5. It is a very recent area of research and relevant papers were not found before 2020.

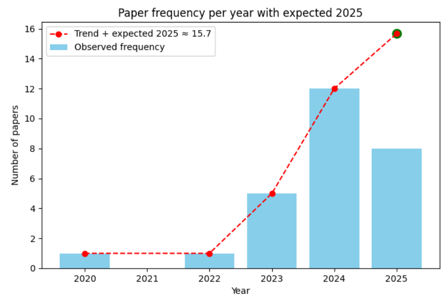

Figure 2: The number of papers has grown 3 times in one year (2022-2023) and is expected to rise this year.

As shown in Figure 3, of the several papers found with career guidance systems, few offer multilingual support in Indian languages and even fewer offer multimodal support. Real time job market integration was a feature we had not thought of before conducting the analysis but is also included in some papers. Of the papers offering multilingual support, the only languages supported are Hindi and Gujarati[15].

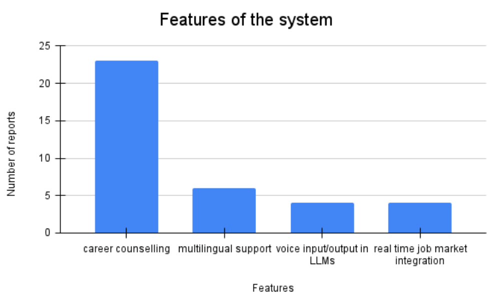

Figure 3: The frequency of papers covering a particular feature that a comprehensive career guidance system should have.

As shown in Figure 4, while the reviewed papers contribute useful benchmarking metrics for their machine learning models, they generally lack a strong emphasis on equity and inclusion during model development. Moreover, most of the models and datasets have not been made openly available, limiting reproducibility and slowing research progress in this field.

A majority of papers have built a web/app interface with machine learning models for career prediction but no studies have explicitly focused on conducting studies/tests on students from under-served communities who need it most.

We have added a chord diagram in Figure 6. The diagram has dense connections in the blue section signifying 'What?' showing that these features are complementary to each other.

5 Discussion

This systematic review identifies the potential as well as the limitations of technology-facilitated career guidance in reaching India's

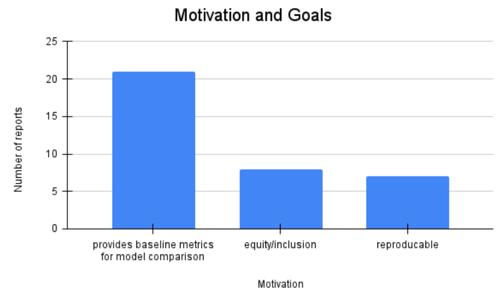

Figure 4: The frequency of papers that were motivated by a particular part of our problem statement

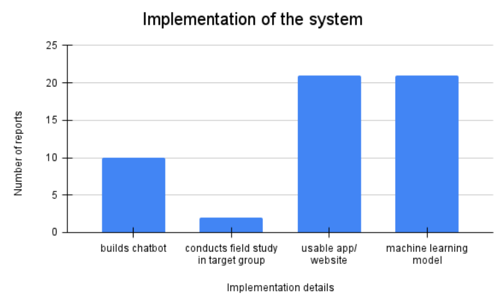

Figure 5: The frequency of papers that included a particular implementation detail.

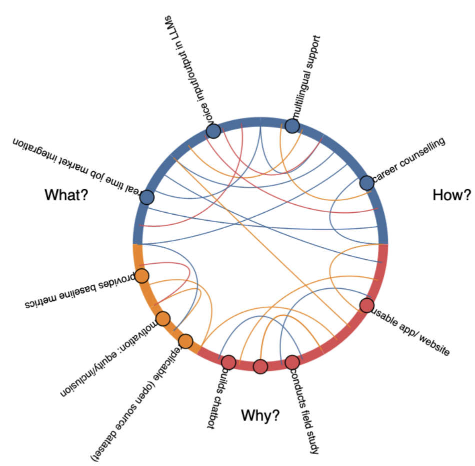

Figure 6: Chord diagram relating all factors considered for the review.

underserved student groups. The results indicate that although machine learning and natural language processing approaches attain impressive prediction accuracies, their actual impact is still bounded by socio-cultural, linguistic, and infrastructural limitations. Tools like "VIBE" or "Career Navigator" prove that AI can effectively personalize guidance, but their reliance on curated data and standardized inputs introduces uncertainty regarding contextual fit for rural and tribal groups.

The common thread throughout much of the literature is tension between technological sophistication and accessibility. Chatbot systems, for instance, provide scalability and interactive interaction but are limited in usability by linguistic diversity, low digital literacy, and the necessity for continuous internet access. Multilingual frameworks such as IndicBART and MLE programs also point toward the importance of mother-tongue support to enhance inclusivity, yet few platforms move beyond Hindi–English bilingualism to substantial coverage of regional dialects.

A second key aspect relates to the wider ecosystem of career choice. While AI technologies tend to emphasize skills-to-career mapping, they rarely take into account considerations such as economic limitations, family pressure, or cultural expectations that profoundly influence student decisions. This implies that technology, however vital, is not enough by itself. The addition of human guides, local stakeholders, and contextual content to online platforms can prove crucial for impact at scale.

Lastly, we observe deployment and evaluation gaps. Many promising prototypes are left in the experimental phase, not having been thoroughly field-tested with rural or tribal students. Low reproducibility of datasets and models also hinders further advancement. Closing these gaps does not just need technical innovation but also interdisciplinary collaborations between computer scientists, education specialists, linguists, and policy makers.

5.1 Conclusion

This review consolidates evidence from 17 studies to evaluate whether multilingual and AI-based technologies can enhance career guidance for semi-urban, tribal, and rural students. We observe significant advancements in machine learning–based career forecasting, interactive chatbot development, and multilingual education assistance. But issues of inclusivity, cultural translation, infrastructure constraints, and community-based views remain ongoing challenges.

Three directions should be addressed in future research

1.Deep multilingualism and cultural sensitivity – ensuring that systems respond to the linguistic and socio-cultural facts of Indian diversity.

2.Ecosystem integration – marrying AI-supported guidance with family, teacher, and local counselor support.

3.Scalable field validation – moving from lab outcomes to intensive testing in actual rural and tribal environments.

In the end, the success of such systems will not be gauged simply on predictive accuracy but on expanding access, equity, and worthwhile career opportunities for students who have historically been underserved. Technology can close the career guidance gap, but its implementation must be context-sensitive, inclusive, and buttressed by systemic support.

6 Citations and Bibliography

This section provides the list of all primary studies synthesized in the Findings section of this systematic review, along with direct links to the source documents for reader convenience.

References

[1] Upadhyay, V., & Gautam, P. (2025). *IT Career Skills Prediction Based on Students Interest, Abilities and Computer Language Proficiency.* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10940693

[2] Mahure, S. J., et al. (2024). *Enhancing Career Pathways: Advancing Guidance Systems with XGBoost Algorithm.* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10607597

[3] Goyal, A., et al. (2024). *Conversational AI for Career Counseling.* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10575379&tag=1

[4] (Various Authors). (2024). *Career Navigator: Your AI Career Counselor.* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10842724

[5] (Various Authors). (2024). *VIBE: A Data-Driven Approach to Career Guidance and Skill Development.* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10533596

[6] (Various Authors). (2024). *Cutting-Edge AI-Driven Higher Educational and Career Advisory Platform.* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10910770

[7] Soni, N., et al. (2023). *Leveraging Machine Learning for Career Prediction and Resume Analysis in Career Assist.* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10434306

[8] (Various Authors). (2024). *CareerAlly: An Intelligent NLP-Driven Chatbot.* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=11113962

[9] (Various Authors). (2024). *AI Assisted Career Guidance and Resume Screening System.* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10923842

[10] (Various Authors). (2024). *Enhancing Career Guidance for Students Through Retrieval Augmented Generation.* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=11005166

[11] Kamble, S., et al. (2023). *Machine Learning based Intelligent Career Counselling Chatbot (ICCC).* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10128305

[12] Patil, A., et al. (2023). *Advancement in Communication using Natural Language-based VideoBot System.* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10037380

[13] Panigrahi, S. S. (2025). An Evaluation of Multi Lingual Education Programme in Odisha. *International Journal for Multidisciplinary Research (IJFMR)*, 7(3). Available: https://www.ijfmr.com/papers/2025/3/43664.pdf

[14] Ray, R., et al. (2025). Interactive Multilingual Learning System using Real-Time Natural Language Processing. *International Journal of Computer Techniques*, 12(3). Available: https://ijctjournal.org/wp-content/uploads/2025/06/Interactive-Multilingual-Learning-System-using-Real-Time-Natural-Language-Processing.pdf

[15] Taunk, D., & Varma, V. (2023). Summarizing Indian Languages using Multilingual Transformers based Models. *arXiv preprint arXiv:2303.16657.* Available: https://arxiv.org/abs/2303.16657

[16] (Various Authors). (2024). *NLP-Powered Data-Driven Career Decision Methodology Based on Machine Learning.* Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10612636

[17] Das, S. B., et al. (2024). Multilingual Neural Machine Translation for Indic to Indic Languages. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 23(5). DOI: https://doi.org/10.1145/3652026

[18] Website: https://www.prisma-statement.org/prisma-2020

7 Acknowledgments

We thank Professor Yogesh Meena for his guidance which helped us conduct this systematic literature review.

A Appendices

A.1 Dataset Description

The dataset used in this systematic review is provided as a supplementary file in `.xlsx` format. It contains the bibliographic metadata and coded variables for all studies included in the review. The following columns are available:

- Authors
- Title
- Year
- Source Title
- DOI
- Link
- Abstract

- Author Keywords
- Index Keywords
- Publisher
- Coding dimensions: When, Who, What, Why, How

A.2 Access to Dataset

The full dataset can be accessed in the attached Excel file (Appendix_Dataset.xlsx) or through the following Google Sheet link:

https://docs.google.com/spreadsheets/d/17guVbdb6dIae53RYQ2Q0vxo-yepogbvErw3yKR4sguM/edit?usp=sharing