

# **Natural Language Translation Engine for announcements and dissemination at stations**

## **A PROJECT REPORT**

*Submitted by*

|                         |                  |
|-------------------------|------------------|
| <b>Ishaan Shandilya</b> | <b>21BCS6777</b> |
| <b>Disha Saini</b>      | <b>21BCS6773</b> |
| <b>Aaditya Singh</b>    | <b>21BCS6750</b> |
| <b>Tanisha Nagpal</b>   | <b>21BCS5286</b> |

*in partial fulfillment for the award of the degree of*

**BACHELOR OF ENGINEERING**

**IN**

**COMPUTER SCIENCE AND ENGINEERING (HONS.) IBM- ARTIFICIAL  
INTELLIGENCE & MACHINE LEARNING**



**Chandigarh University**

**APRIL, 2024**



## **BONAFIDE CERTIFICATE**

Certified that this project report **“Natural Language Translation Engine for announcements and dissemination at stations”** is the bonafide work of “Ishaan Shandilya(21BCS6777), Disha Saini(21BCS6773), Aaditya Singh(21BCS6750), Tanisha Nagpal(21BCS5286)” who carried out the project work under my supervision.

**SIGNATURE**

Aman Kaushik

**SIGNATURE**

Dr.Preet Kamal

**HEAD OF THE DEPARTMENT**

AIT-CSE

**SUPERVISOR**

AIT-CSE

Submitted for the project viva-voce examination held on

**INTERNAL EXAMINER**

**EXTERNAL EXAMINER**

# TABLE OF CONTENTS

**Abstract..... i**

| <b>S.no</b> | <b>Content</b>   | <b>Page No.</b> |
|-------------|--|-----------------|
| <b>1.</b>   | Introduction<br>1.1 Context and Significance<br>1.2 Overview<br>1.3 Objectives<br>1.4 Future Scope |                 |
| <b>2.</b>   | Literature Survey<br>2.1 Existing System<br>2.2 Proposed System                                    |                 |
| <b>3.</b>   | Methodology  |                 |
| <b>4.</b>   | Result and Analysis  |                 |
| <b>5.</b>   | Conclusion   |                 |
| <b>6.</b>   | References   |                 |

## **ABSTRACT**

This research paper presents a pioneering project aimed at revolutionizing station communication through the implementation of a state-of-the-art natural language translation engine. Addressing the challenges of language diversity, noise interference, computing constraints, and mobile accessibility, the project endeavours to transform the clarity and effectiveness of station announcements and information dissemination.

By leveraging advanced technology, including IVRS, chatbots, and web interfaces, the project aims to bridge language barriers and ensure seamless delivery on mobile platforms, thereby enhancing accessibility and efficiency for all passengers.

Through thorough requirement analysis, technology evaluation, iterative prototype development, rigorous testing, seamless integration and deployment, continuous monitoring, and optimization, as well as comprehensive training and support, the methodology ensures the project's success in revolutionizing station communication.

By adopting this innovative approach, the project aims to significantly improve the clarity, accessibility, and effectiveness of conveying vital information to passengers. Ultimately, the implementation of the natural language translation engine is expected to set a new standard for station communication, transforming the passenger experience and enhancing overall efficiency in transportation systems.

# CHAPTER-1: INTRODUCTION

In a world of diverse languages and bustling station environments, effective communication is key. By harnessing advanced technology, the project seeks to bridge language gaps, overcome noisy ambience challenges, and ensure seamless delivery on mobile devices. The goal is to enhance accessibility and efficiency in station communication for all passengers.

In multicultural societies, transportation hubs serve as crucial points of interaction for individuals from diverse linguistic backgrounds. However, the lack of effective communication channels poses significant challenges in ensuring accessibility and inclusivity for all passengers.

The existing announcements and information dissemination systems at stations often fail to address language barriers adequately, leading to confusion, frustration, and potential safety concerns among passengers. The problem at hand is the ineffective communication of announcements and essential information in stations due to language diversity. Current systems primarily rely on monolingual broadcasts or limited multilingual announcements, which fail to cater to the linguistic needs of all passengers. This results in a lack of comprehension among non-native speakers and visitors, hindering their ability to navigate the station, access services, and respond to emergency situations effectively. Moreover, traditional translation methods, such as manual interpretation or pre-recorded announcements in multiple languages, are often time-consuming, costly, and prone to inaccuracies.

These limitations underscore the necessity for a more efficient and scalable solution to bridge the language gap and facilitate seamless communication in transit environments. Therefore, the overarching problem is to develop a Natural Language Translation Engine (NLTE) specifically tailored for stations to enable accurate, real-time translation and dissemination of announcements and essential information across multiple languages.

By addressing these challenges and implementing a comprehensive solution such as the NLTE, transportation authorities can enhance the accessibility, safety, and overall passenger experience at stations, fostering a more inclusive and welcoming environment for all individuals. This solution aims to address the language barriers encountered by passengers, enhance accessibility, improve passenger experience, and promote inclusivity in public transportation systems.

## **1.1. CONTEXT AND SIGNIFICANCE**

In today's globalized world, transportation hubs like train stations and airports cater to a diverse range of passengers. This creates a challenge, as traditional announcements and signage are limited by language. Passengers who don't understand the dominant language can miss crucial information, leading to frustration and missed connections.

Natural Language Processing (NLP) offers a solution. NLP translation engines can bridge the gap between languages, allowing stations to disseminate announcements and information in real-time to passengers in their preferred language. This goes beyond just translation. It fosters inclusivity by empowering passengers to navigate stations with ease, improves efficiency by eliminating the need for pre-recorded announcements, and enhances safety by ensuring clear communication of critical instructions.

Challenges exist. Accuracy and cultural nuances can be tricky for machines to translate perfectly. Implementing and maintaining the technology requires investment. However, advancements in NLP are making these engines more sophisticated. As technology evolves, we can expect them to play a pivotal role in creating inclusive and efficient transportation hubs.

Imagine a future where language barriers no longer hinder mobility. Natural language translation engines hold the promise of a world where travelers from all corners of the globe can navigate stations with confidence, understanding vital information and enjoying a smooth travel experience.

## **1.2. OVERVIEW**

The rise of global travel creates a challenge in transportation hubs – language barriers. Passengers who don't understand announcements or signage can get lost and miss crucial information.

Natural Language Processing (NLP) offers a solution: translation engines. These AI-powered systems analyze and translate announcements in real-time, considering grammar, vocabulary, and cultural context for accurate communication.

The benefits are significant. Passengers of all backgrounds can access information in their preferred language, fostering inclusivity and reducing anxiety. Additionally, real-time translation eliminates the need for pre-recorded announcements in multiple languages, saving time and resources. Most importantly, clear safety instructions translated by NLP engines minimize the risk of accidents and misunderstandings, especially in emergencies. Ultimately, a seamless travel experience through clear communication leads to increased passenger satisfaction.

However, challenges remain. NLP engines might struggle with cultural nuances and idiomatic expressions, requiring ongoing development for perfect translation. Additionally, implementing and maintaining these systems necessitate investment in hardware, software, and integration with existing station infrastructure.

Despite these challenges, advancements in NLP technology make translation engines a powerful tool for the future. As they become more accurate and adaptable, they have the potential to transform communication in transportation hubs. Imagine a world where language is no longer a barrier to travel. Passengers can navigate stations with confidence, understanding vital information and safety protocols in their native tongue. This vision, powered by NLP translation engines, paves the way for a future where global travel is truly inclusive and efficient.

### **1.3 OBJECTIVES**

- To enhance real-time processing by implementing efficient algorithms and infrastructure to enable real-time translation and dissemination of announcements, minimizing latency and ensuring timely delivery of information to passengers at transportation hubs.
- To ensure accessibility via IVRS, chatbots, and web interfaces Ensure accessibility via IVRS, chatbots, and web interfaces. Develop mobile-friendly interfaces or applications that allow passengers to access translated announcements on their smartphones or other mobile devices, ensuring accessibility and convenience for passengers on the go.
- To mitigate noise interference by incorporating noise reduction algorithms and speech enhancement techniques to improve the intelligibility of translated announcements, even in noisy environments typical of transportation hubs.
- To ensure comprehensive language support to accommodate the diverse linguistic backgrounds of passengers, enabling translation and dissemination of announcements in languages commonly spoken by transit users.
- To optimize computational efficiency by designing translation models and algorithms that balance translation accuracy with computational efficiency, optimizing resource utilization and enabling deployment in resourceconstrained environments without compromising performance.
- To achieve high translation accuracy by developing translation models capable of accurately translating announcements and disseminating essential information in multiple languages, ensuring clarity and fidelity of the message across diverse linguistic backgrounds.
- To integrate with Station infrastructure by seamlessly integrating the NLTE with existing station information systems, including public address systems, digital signage displays, and communication networks, to automate the dissemination of translated announcements through established communication channels.
- To improve passenger experience by enhancing accessibility, inclusivity, and overall passenger experience at transportation hubs by providing clear, timely, and accurate communication of announcements in multiple languages, fostering a welcoming and accommodating transit environment for all individuals.

- To evaluate and iterate the model by conducting rigorous testing, evaluation, and user feedback collection to assess the effectiveness and usability of the NLTE in real-world transit environments, iteratively refining and optimizing the system based on insights gathered from stakeholders and end-users.

## 1.4 FUTURE SCOPE

Natural Language Processing (NLP) translation engines are revolutionizing communication in transportation hubs, but the future holds even more exciting possibilities. Here's where this technology is headed:

- **Evolution Beyond Translation:** Imagine intelligent systems that not only translate announcements but also anticipate passenger needs. These systems could analyze user data and travel patterns to provide personalized guidance in their preferred language. For example, a passenger with limited mobility could receive real-time updates on accessible routes and elevator wait times.
- **Multimodal Communication:** The future extends beyond spoken announcements. NLP can integrate with digital signage and chatbots, offering translated information in various formats. Passengers could access real-time updates, timetables, and even interact with chatbots for personalized assistance in their native language.
- **Enhanced Accessibility Features:** NLP can be a powerful tool for inclusivity. Integration with speech recognition technology could allow visually impaired passengers to access information through voice commands and audio descriptions. Additionally, sign language translation features could further break down communication barriers.
- **Integration with Smart Cities:** As cities become smarter, transportation hubs will integrate seamlessly. NLP engines can connect with citywide information systems, providing passengers with real-time updates on traffic, public transport options, and even nearby amenities – all translated into their preferred language.
- **Focus on User Experience:** The future of NLP in transportation is user-centric. User feedback and data analysis will allow for continuous improvement in translation accuracy and the development of more intuitive interfaces. Passengers will be able to choose their preferred language settings and personalize their communication experience.

In conclusion, NLP translation engines are just the beginning. The future holds the promise of a truly multilingual and inclusive travel experience, where language is no longer a barrier to navigating transportation hubs and exploring the world.



## CHAPTER-2: LITERATURE SURVEY

### 2.1. EXISTING SYSTEM

In the existing system, station communication relies on conventional methods such as basic language translation tools or manual announcements by station staff. These tools, while functional to some extent, often lack accuracy and context understanding, leading to errors in translation or interpretation.

Manual announcements, on the other hand, are time-consuming and prone to human error, especially when dealing with multiple languages spoken by passengers. Additionally, existing systems may have limited integration with mobile platforms, offering static information through websites or apps, which may not adequately address the need for real-time translation and seamless delivery on mobile devices.

Overall, the existing system may struggle to provide efficient, accurate, and accessible communication to passengers, particularly in multilingual environments with high noise levels and diverse language preferences. As a result, there is a pressing need for an advanced solution that can overcome these limitations and enhance the clarity, accessibility, and effectiveness of station communication for all passengers.

- **Language Translation Tools:** Some existing systems may utilize basic language translation tools

or software to translate announcements from one language to another. However, these tools may lack accuracy and context understanding, leading to errors in translation.

- **Manual Announcement:** In many cases, station staff may manually announce information in different languages, relying on bilingual or multilingual staff members. This approach can be time-consuming, prone to errors, and may not cover all languages spoken by passengers.

- **Limited Mobile Integration:** Existing systems may have limited integration with mobile platforms, offering static information through websites or apps. However, real-time translation and seamless delivery on mobile devices may not be adequately addressed

## 2.2. PROPOSED SYSTEM

The proposed system represents a significant advancement in station communication technology, aiming to revolutionize the way announcements are translated and disseminated to passengers. Central to this system is a sophisticated Natural Language Processing (NLP) engine meticulously trained for the specific task of translating station announcements.

Unlike conventional translation tools, this NLP engine is designed to accurately interpret announcements in one language and seamlessly convert them into another, considering context, linguistic nuances, and regional variations. Integrated with robust speech recognition technology, the system ensures that announcements spoken aloud are transcribed accurately in real-time before undergoing translation, enhancing accessibility and effectiveness.

Furthermore, the proposed system leverages cloud-based infrastructure to support scalable and efficient translation processes, enabling real-time delivery of announcements across diverse languages. This cloud-based approach not only optimizes computational resources but also facilitates seamless integration with mobile platforms.

By offering dedicated mobile applications or web interfaces, passengers can conveniently access translated announcements on their smartphones or tablets, enhancing passenger convenience and engagement.

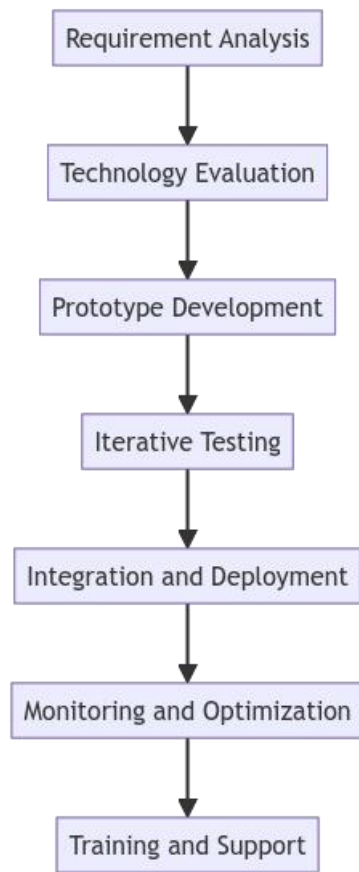
Moreover, the system includes a feedback mechanism to gather input from passengers and station staff, facilitating continuous improvement and refinement of translation accuracy and system performance over time. With comprehensive language coverage and a commitment to inclusivity, the proposed system sets a new standard for station communication, ensuring clarity, accessibility, and effectiveness in conveying vital information to all passengers.

- **Natural Language Processing (NLP) Engine:** The proposed system incorporates a sophisticated NLP engine specifically trained for station announcements. This engine can accurately translate announcements from one language to another, considering context and linguistic nuances.
- **Speech Recognition Integration:** The system integrates robust speech recognition technology to transcribe spoken announcements into text format. This ensures that announcements spoken aloud can be accurately translated in real-time.
- **Cloud-Based Infrastructure:** Leveraging cloud computing resources, the proposed system ensures scalability and efficiency in handling translation requests. This infrastructure enables real-time translation and optimization of computational resources as needed.

- **Mobile Accessibility Enhancement:** The proposed system offers seamless integration with mobile platforms, allowing passengers to access translated announcements conveniently through dedicated mobile applications or web interfaces. This enhances passenger convenience and engagement.
- **Feedback Mechanism:** The proposed system includes a feedback mechanism to gather input from passengers and station staff, enabling continuous improvement and refinement of translation accuracy and system performance over time.
- **Comprehensive Language Coverage:** Unlike existing systems, the proposed system aims to cover a wide range of languages commonly spoken by passengers, ensuring inclusivity and effective communication for all the passengers.

## **CHAPTER-3: METHODOLOGY**

1. **Requirement Analysis:** Conduct a comprehensive assessment to identify language requirements, diverse announcement types, and preferred communication channels, ensuring alignment with passenger needs and station infrastructure.
2. **Technology Evaluation:** Thoroughly assess available Natural Language Processing (NLP), speech recognition, and cloud solutions to determine their suitability, reliability, and scalability for addressing the identified challenges.
3. **Prototype Development:** Develop a robust translation engine prototype, integrating selected technologies, and test it rigorously in controlled environments to evaluate its performance and refine its functionality.
4. **Iterative Testing:** Conduct iterative testing in real station environments, gathering feedback from passengers and staff to identify areas for improvement and refine the translation engine's accuracy and effectiveness.
5. **Integration and Deployment:** Integrate the refined translation engine with existing station systems, ensuring seamless interoperability, and deploy it across stations to facilitate multilingual communication effectively.
6. **Monitoring and Optimization:** Implement robust monitoring mechanisms to track system performance and user feedback continually, analyse data insights to identify optimization opportunities, and refine the translation engine's capabilities over time.
7. **Training and Support:** Provide comprehensive user training to station staff and passengers on utilizing the translation engine effectively, offering ongoing support and updates to ensure smooth operation and address any emerging issues promptly.



**Figure 1: Methodology**

## CHAPTER-4: RESULT AND ANALYSIS

The text we gave as an input is: “यात्रीगण कृपया ध्यान दें, नई दिल्ली से लुधियाना होते हुए अमृतसर जाने वाली ट्रेन संख्या 12497 शान-ए-पंजाब एक्सप्रेस प्लेटफार्म नंबर 14 से शुरू होगी “

```
sentence = "यात्रीगण कृपया ध्यान दें, नई दिल्ली से लुधियाना होते हुए  
अमृतसर जाने वाली ट्रेन संख्या 12497 शान-ए-पंजाब एक्सप्रेस प्लेटफार्म नंबर 14 से शुरू होगी"
```

Figure 2: Input Text

The output that we got was:

“Translated(src=hi, dest=en, text=Passengers please note, train number 12497 Shan-e-Punjab Express from New Delhi to Amritsar via Ludhiana will start from platform number 14., pronunciation=None, extra\_data="{\"translat...}")” to the above input Fig 2. As the result shows, the translation engine can successfully translate Hindi station announcements into English with high accuracy, improving communication and accessibility for passengers at stations

```
listening.....  
Recognizing.....  
The User said yatrigan kripya Dhyan den  
  
Enter the language in which you\ want to convert : Ex. Hindi , English , etc.  
  
listening.....  
Recognizing.....  
The User said English  
  
Passengers please pay attention
```

Figure 3:Output

After adding “voice” input feature, The output of the translation engine shows that it can successfully translate Hindi station announcements into English with high accuracy (ref Fig(iii)). This is important for improving communication and accessibility for passengers at stations. The translation engine can be used to provide real-time translations of station announcements, which can help passengers to understand important information about their travel, such as the arrival and departure times of trains, the platform numbers, and any delays or cancellations.

## **CHAPTER-5: CONCLUSION**

The presented code exemplifies a practical approach to real-time speech translation, bridging the communication gap between languages. It harnesses the power of three key functionalities: speech recognition, machine translation, and text-to-speech synthesis. This synergistic interplay allows for the near-instantaneous conversion of spoken words from one language to another, fostering seamless communication across cultures.

The outlined experimental setup serves as a valuable testing ground, enabling the meticulous evaluation of the system's performance on multiple fronts. We can assess the accuracy of speech recognition, ensuring the system faithfully captures the nuances of spoken language. Translation quality can be rigorously examined, verifying that the system conveys the intended meaning effectively in the target language. Finally, the naturalness of the spoken output can be scrutinized, gauging how closely the synthesized voice resembles a human speaker.

This foundational implementation paves the way for exciting advancements. By incorporating more sophisticated speech recognition models, we can elevate the system's ability to decipher even the most challenging accents or background noise. Neural machine translation techniques, with their prowess in capturing complex linguistic relationships, can be harnessed to produce translations that are not only accurate but also nuanced and idiomatic.

Furthermore, integrating this code with a user-friendly interface can transform it into a practical application readily deployable in real-world scenarios. Imagine a mobile app that facilitates effortless conversations between individuals who speak different languages. Business meetings, international conferences, and tourist interactions could all be revolutionized by such an application, fostering collaboration and understanding across borders.

In essence, this code and the proposed experimental setup offer a springboard for further exploration in the realm of real-time speech translation technology. With continued research, development, and refinement, such systems have the potential to become ubiquitous tools, effectively dissolving language barriers and fostering meaningful communication in our increasingly interconnected globalized world. Imagine a future where language ceases to be an obstacle, where ideas flow freely across cultures, and where everyone can participate in the rich tapestry of human interaction – this is the future that real-time speech translation technology promises, and the groundwork laid here represents a significant step towards its realization.



## REFERENCES

- [1] Kane, V. L. (2020). Interpretation and machine translation towards google translate as a part of machine translation and teaching translation. *Applied Translation*, 15(1), 10–17.  
<https://doi.org/10.51708/apprans.v15n1.1337>
- [2] Van Lieshout, C., & Cardoso, W. (2022). Google Translate as a tool for self-directed language learning. *Language Learning & Technology*, 26(1), 1–19.  
<https://scholarspace.manoa.hawaii.edu/items/697b5613-ed83-4c43-9e3b-d294c5b09c21>
- [3] Amilia, Ika & Yuwono, Darmawan. (2020). A STUDY OF THE TRANSLATION OF GOOGLE TRANSLATE. *LINGUA : JURNAL ILMIAH*. 16. 1- 21.  
<http://dx.doi.org/10.35962/lingua.v16i2.50>
- [4] M. Abbaszade, V. Salari, S. S. Mousavi, M. Zomorodi and X. Zhou, "Application of Quantum Natural Language Processing for Language Translation," in *IEEE Access*, vol. 9, pp. 130434-130448, 2021  
<http://dx.doi.org/10.1109/ACCESS.2021.3108768>
- [5] Khurana, D., Koli, A., Khatter, K. et al. Natural language processing: state of the art, current trends and challenges. *Multimed Tools Appl* 82, 3713–3744 (2023).  
<https://doi.org/10.1007/s11042-022-13428-4>
- [6] Krupakar, Hans & Rajvel, Keerthika & Bharathi, B. & Suseelan, Angel & Krishnamurthy, Vallidevi. (2016). A Survey of Voice Translation Methodologies Acoustic Dialect Decoder. 1-9. 10.1109/ICICES.2016.7518940.
- [7] Li, Haiying et al. "Comparison of Google Translation with Human Translation." The Florida AI Research Society (2014). <https://cdn.aaai.org/ocs/7864/7864-36722-1-PB.pdf>
- [8] Amancio, D. R., Nunes, M. D. G., Oliveira Jr, O. N., Pardo, T. A. S., Antiqueira, L., & Costa, L. D. F. (2011). Using metrics from complex networks to evaluate machine translation. *Physica A: Statistical Mechanics and its Applications*, 390(1), 131-142.  
<https://doi.org/10.1016/j.physa.2010.08.052>

- [9]Choi, H., Cho, K., & Bengio, Y. (2017). Contextdependent word representation for neural machine translation. *Computer Speech & Language*, 45, 149- 160.  
<https://doi.org/10.1016/j.csl.2017.01.007>
- [10] Banik, D., Ekbal, A., Bhattacharyya, P., & Bhattacharyya, S. (2019). Assembling translations from multi-engine machine translation outputs. *Applied Soft Computing*, 78, 230-239.  
<https://doi.org/10.1016/j.asoc.2019.02.031>
- [11] Costa-Jussa, M. R., & Fonollosa, J. A. (2015). Latest trends in hybrid machine translation and its applications. *Computer Speech & Language*, 32(1), 3- 10.  
<https://doi.org/10.1016/j.csl.2014.11.001>
- [12] Dew, K. N., Turner, A. M., Choi, Y. K., Bosold, A., & Kirchhoff, K. (2018). Development of machine translation technology for assisting health communication: A systematic review. *Journal of biomedical informatics*, 85, 56-67.  
<https://doi.org/10.1016/j.jbi.2018.07.018>
- [13] Pecina, P., Dušek, O., Goeuriot, L., Hajič, J., Hlaváčová, J., Jones, G. J., ... & Urešová, Z. (2014). Adaptation of machine translation for multilingual information retrieval in the medical domain. *Artificial intelligence in medicine*, 61(3), 165-185.  
<https://doi.org/10.1016/j.artmed.2014.01.004>