

Voice-Sync: AI-Powered Multilingual News Reader Using Python

Amulya K. S.¹; B. R. Mohan²

¹Department of Computer Science and Engineering, East West Institute of Technology, Bengaluru, Karnataka, India

²Department of Computer Science and Engineering, East West Institute of Technology, Bengaluru, Karnataka, India

Publication Date: 2026/09/02

Abstract: Yet although accessing news online has never been easier, the experience tends to be fragmented. Readers frequently have to jump between multiple websites, wade through lengthy articles, or switch between different tools just to obtain a translation or to have some content read out loud. This constant back-and-forth makes it difficult to stay informed, especially for those who prefer consuming news in their native. People or those who want a hands-free experience. To solve this, we developed Voice-Sync, a web-based news reader that brings the power of Artificial Intelligence into one seamless platform. Built using the Python and the Flask framework, Voice-Sync automatically pulls the latest headlines from NewsAPI and processes via specialized AI modules. When an article has been retrieved, the app produces a brief summary by means of specialized AI modules. Once an article is retrieved, the app creates a concise summary using Natural Language Processing, translates it into various. Convert Indian Languages into clear, natural speech using the text-to-speech feature from Microsoft Edge Neural. Beyond reading the news, Voice-Sync offers voice- controlled navigation, sentiment analysis, bookmarking, and the ability to export reports as PDFs. By streamlining these tasks, we have created a more intuitive, accessible way for users to engage with information without the usual technical friction.

Keywords: Artificial Intelligence, Natural Language Processing, Flask, NewsAPI, Multilingual Translation, Edge Neural Text-to-Speech, Speech Recognition, Sentiment Analysis.

How to Cite: Amulya K. S.; B. R. Mohan (2026) Voice-Sync: AI-Powered Multilingual News Reader Using Python. *International Journal of Innovative Science and Research Technology*, 11(8), 2511-2518. <https://doi.org/10.38124/ijisrt/26aug825>

I. INTRODUCTION

News has become one of the most common ways people access information today. Instead of turning to newspapers or television broadcasts, most people now check websites and mobile apps to keep up with daily events. Because news is published around the clock, readers often end up spending a lot of time scrolling through different sources just to find something relevant to them. This gets even harder when the articles are lengthy or written in a language the reader doesn't understand well.

Modern AI techniques provide a practical mechanism to alleviate this burden. Recent advances in Natural Language Processing make it possible to shorten long articles to more concise summaries without losing the critical information. Neural translation models are also able to translate content into other languages and neural speech synthesis lets users hear the news instead of reading it. Bringing all these capabilities into a single application can meaningfully improve the way people consume news.

This is the idea behind Voice-Sync. Rather than building yet another standard news reader, the goal was to create an application that could pull in a format that's simpler for users to consume. As a result, the application brings together real-time news retrieval, automatic summarization, multilingual translation, speech generation, and voice-based interaction, all within one web platform.

The system is built in python as it is strong in both AI and web design. Flask is used for front-end and back-end communication, and external APIs and Pretrained models are used for special tasks such as getting news, translating text, summarizing content and generating speech. The system was built in a modular way, making development more manageable and testing easier, as each component could be tested independently before being integrated into the final application.

Another driving factor behind this work was accessibility. Many existing news apps are built mainly around text and require the user to keep interacting manually-tapping, scrolling, and reading. Voice-Sync tries to move past this limitation by letting users simply ask for a news category

through voice commands and then listen to summarized articles in different Indian languages. This can help visually impaired users in particular, but it's also useful for anyone who'd rather listen to the news while commuting, working, or doing something else with their hands.

Testing on the prototype showed that bringing these separate AI modules results in a smooth experience for the user, without adding noticeable processing delays.

The application pulls in live news, summarizes it, translates it into the language the user selects, and converts it into natural-sounding speech. On top of that, it offers bookmarking, Sentiment analysis, and PDF export.

➤ *The Main Contributions of this Work are as Follows:*

- Built an integrated, multilingual news reader using Python and Flask.
- Retrieved real-time news through NewsAPI.
- Used transformer-based NLP models for automatic summarization.
- Added translation support for English, Kannada, Hindi, Tamil, Telugu and Malayalam.
- Generated high-quality speech using Microsoft Edge Neural Text-to-Speech.
- Enabled voice commands for hands-free use.
- Included sentiment analysis and PDF generation to improve usability.
- Designed a modular architecture that allows future extensions without major redesign.

II. LITERATURE SURVEY

Artificial Intelligence has changed many aspects of digital journalism over the past few years. Earlier news applications were mainly designed to display articles collected from different sources, whereas recent systems are capable of summarizing lengthy reports, translating content into multiple languages, generating speech, and even interacting with users through voice commands. These developments have encouraged researchers to investigate how AI can improve both the efficiency and accessibility of news delivery.

One area that has received considerable attention is automatic text summarization. Instead of presenting complete articles, modern summarization models identify the most informative content and generate shorter versions that preserve the original context. Asmitha *et al.* [2] demonstrated that transformer-based models such as BART produce fluent summaries while maintaining important information from the source article. Their findings indicate that abstractive summarization can significantly reduce reading time without affecting overall understanding.

Recent advances in transformer-based Natural Language A comparative analysis conducted by Kamalkar and Kavitha [3] examined several transformer architectures, including PEGASUS, DistilBART, BART, ProphetNet, and GPT-based models. Their work highlighted that transformer networks

consistently outperform traditional summarization approaches because they capture contextual relationships more effectively and generate readable summaries. Such models have therefore become a popular choice for intelligent document-processing applications.

Another important research direction focuses on speech synthesis. Recent neural Text-to-Speech systems produce speech that is considerably more natural than earlier concatenative or rule-based methods.

Li *et al.* [7] proposed phoneme-level language models capable of improving pronunciation and speech rhythm, while Chen *et al.* [8] demonstrated that Speech BERT enhances prosody and overall speech quality. These improvements have made AI-generated speech more comfortable for continuous listening and therefore more suitable for applications involving digital news delivery.

Voice-Sync addresses this limitation by combining live news retrieval, multilingual translation, AI summarization, voice interaction, sentiment analysis, bookmarking, and PDF generation within a single intelligent platform.

Researchers have also investigated the relationship between Artificial Intelligence and public trust in journalism. Sinclair [11] reported that readers generally respond positively to AI-assisted news when the technology is used transparently and the information remains reliable. Similar observations were presented by Ashfaq *et al.* [12], who discussed the increasing use of AI within the Indian media industry for content generation, recommendation, and workflow automation.

Although these studies have advanced individual technologies, most existing applications concentrate on only one or two capabilities. Some systems perform summarization, while others provide translation or speech synthesis. Very few applications combine all these services into a single workflow where news retrieval, summarization, translation, voice interaction, sentiment analysis, bookmarking, and report generation operate together without requiring users to switch between multiple tools.

The development of Voice-Sync was motivated by this observation. Instead of treating each AI service as an independent application, the project integrates them within one web platform so that the output produced by one module becomes the input for the next. This approach simplifies the overall user experience and demonstrates how multiple AI techniques can cooperate to improve digital news consumption.

III. PROBLEM STATEMENT

Online news platforms have made current information available within seconds, yet the experience of consuming digital news still presents several practical difficulties. Users often spend a significant amount of time reading complete articles, comparing reports from different websites, and searching for information that is actually relevant to them. As

the number of published articles continues to increase, locating the essential points quickly becomes more challenging.

Language presents another obstacle for many readers. Although English remains the primary language on numerous news portals, a large section of users prefer accessing information in regional languages. Existing translation support is either unavailable or limited, forcing users to rely on separate translation services before they can fully understand the content.

Accessibility is another concern. Most news apps assume users will interact through touch or a mouse. But for people who are travelling, multitasking, or visually impaired, having to constantly tap and scroll can be inconvenient. Voice assistants are now common on smartphones, yet they're still rarely built into news-reading platforms in any meaningful way. While studying existing applications, it also became clear that features like summarization, translation, speech synthesis, sentiment analysis, and document generation usually live in separate tools.

Constantly switching between these services breaks up the reading experience and makes it more effortful just to get to the information you actually want.

Audio playback is another practical sticking point. Standard Text-to-Speech engines tend to sound robotic, which makes listening for long stretches tiring. For a news app in particular, natural pronunciation and clear voice quality really matter, since people are likely to listen to several articles back-to-back.

Given these gaps, there's a clear case for an application that handles the entire news-processing pipeline in one place — retrieving current news, generating concise summaries, translating into multiple languages, producing natural-sounding speech, recognizing voice commands, evaluating sentiment, and letting users save articles for later. This need is what shaped the development of Voice-Sync.

➤ *Proposed System*

This System builds with one goal in mind: making digital news easier to access and more convenient to consume. Instead of requiring users to depend on different applications for reading, translating, listening, and saving news articles, the developed platform performs all these activities within a single web application. The complete system has been implemented using Python and Flask, while

Artificial Intelligence techniques are employed wherever automated language processing or speech generation is required. The workflow begins with a user selecting a preferred news category or issuing a supported voice command. The backend then makes a call to NewsAPI to fetch the latest headlines in the requested category. Instead of displaying long articles directly, the retrieved content is passed to the summarization module that extracts the most important information and prepares a shorter, more readable version.

Once summarized, users can pick the language they want the news in. The app translates the summary into several Indian languages — Kannada, Hindi, Tamil, Telugu, and Malayalam. Because translation happens after summarization rather than before, there's less text to process, which helps keep response times fast.

To make even more accessible, the translated text is converted into speech using Microsoft Edge Neural Text-to-Speech. The generated voice sounds considerably more natural than traditional speech synthesizers, making it comfortable to listen to multiple news articles without causing listener fatigue. This feature is particularly useful when users are travelling or performing other activities where reading from the screen is inconvenient. Another important capability of VoiceSync is its support for voice-controlled interaction. Instead of navigating menus manually, users can issue commands such as "Read AI News," "Read Karnataka News," "Translate News," "Analyze Sentiment," or "Stop Reading." Once a valid command is recognized, the corresponding module is executed automatically, creating a smoother interaction between the user and the application. Additional utilities such as sentiment analysis, bookmarking, and PDF report generation have also been incorporated into the system. Sentiment analysis provides a quick indication of whether the news article expresses a positive, negative, or neutral tone. Bookmarking enables users to save articles for future reading, while the PDF generation feature allows summarized news to be stored or shared in a portable format.

➤ *System Architecture*

Voice-Sync's overall architecture follows a modular design, where each component handles its own specific responsibility and communicates with the other modules through the Flask backend. This setup makes the system easier to maintain and simplifies future upgrades, since changes to one module rarely affect the rest of the application.

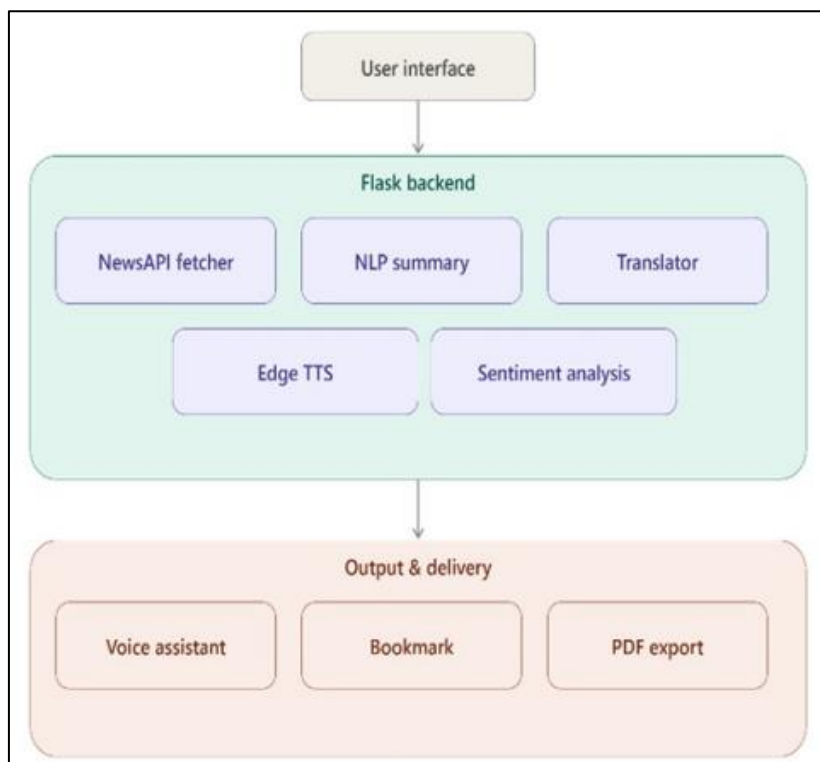


Fig 1 Overall System Architecture of Voice-Sync

➤ Module Description

• News Fetching Module

This module communicates with NewsAPI to retrieve live news articles from multiple categories. It automatically updates headlines and displays the latest information without requiring manual data entry.

• AI Summarization Module

This module uses transformer-based Natural Language Processing models to generate concise summaries from lengthy news articles. It significantly reduces reading time while preserving important information.

• Translation Module

The translation module converts summarized news into multiple Indian languages using Google Translator API. This feature improves accessibility for multilingual users.

• Text-to-Speech Module

Microsoft Edge Neural Text-to-Speech converts translated summaries into natural human-like speech, enabling users to listen to news hands-free.

• Voice Assistant Module

The voice assistant recognizes spoken commands, allowing users to navigate categories, start or stop news playback, and control application features through speech.

• Sentiment Analysis Module

This module classifies news articles as Positive, Negative, or Neutral using Natural Language Processing techniques, helping users quickly understand the emotional tone of the content.

• PDF Report Generation Module

Users can export summarized news articles into professionally formatted PDF reports for offline reading, documentation, or sharing.

➤ System Workflow

• Flowchart

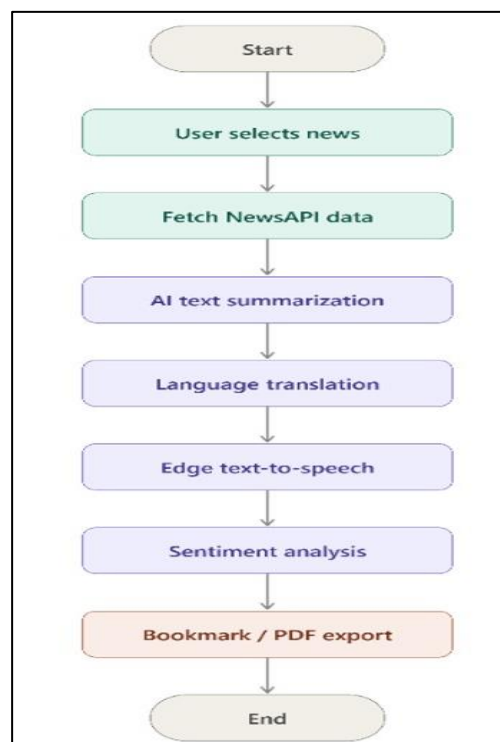


Fig 2 Flowchart of Voice-Sync

Voice-Sync operates through a series of connected processing stages that turn raw online news into spoken, multilingual output. It starts when the user either selects a news category from the interface or gives a voice command. This request goes to the Flask server, which immediately contacts NewsAPI to pull the latest articles. Once the news is retrieved, the article content is passed on to the summarization module.

The summarization model reads through the article and produces a shorter version containing just the essential points. Doing this early on means the modules downstream have less text to work with, which keeps the whole workflow running efficiently.

The summary is then translated into whichever language the user has chosen. After translation, Microsoft Edge Neural Text-to-Speech turns the text into natural-sounding speech that plays directly in the browser. At the same time, the sentiment analysis module looks at the summary and works out whether the news carries a positive, negative, or neutral tone. Users can also bookmark the article or generate a PDF report before moving on to the next news item.

Throughout implementation, care was taken to manage communication between the frontend and backend so that each module only receives data once the previous step has finished successfully. This sequential approach avoids synchronization problems and makes sure the speech being played always matches the correct summarized article. Because every functional component was built independently, additional features — like fake news detection, personalized recommendations, or conversational AI — could be added later without needing to redesign the whole architecture.

IV. IMPLEMENTATION

Voice-Sync was built as a web-based system using Python and the Flask framework. Python was the natural choice here, given how well it supports AI libraries, Natural Language Processing, and REST-based web development. Flask serves as the communication layer connecting the frontend interface to the backend modules that handle news processing.

The user interface itself was built with HTML5, CSS3, Bootstrap, and JavaScript. These were chosen to keep the interface responsive across different screen sizes while keeping interaction simple and uncomplicated for users. Every user action — selecting a news category, translating a summary, starting voice playback, or exporting a PDF report — is handled through asynchronous requests sent to the Flask server.

A number of external libraries were brought in to handle individual AI functions. The Requests library fetches current news from NewsAPI, while transformer-based models from the Transformers library take care of automatic text summarization. Translation across multiple languages is

handled by deep-translator, and Microsoft Edge Neural Text-to-Speech (edge-tts) generates natural-sounding speech from the translated summaries. TextBlob evaluates the sentiment of processed news articles, and ReportLab creates downloadable PDF reports. Voice commands are recognized through the Speech Recognition library. One of the main challenges during implementation involved coordinating communication between these independent modules. Since each module depends on the output produced by the previous stage, the backend validates every operation before forwarding the processed data to the next component. This approach helped avoid synchronization issues and ensured that translated speech always corresponded to the correct summarized article.

The application was developed and evaluated using Visual Studio Code on a Windows 11 system with Python 3.12. Functional testing was initially performed on individual modules before integrating them into the complete application.

After integration, repeated testing confirmed that the modules communicated correctly and that the overall workflow remained stable during continuous operation.

➤ Major Python Libraries Include:

Table 1 Software Environment

Library	Purpose
Python	Programming Language
Flask	Backend Web Framework
Requests	API Communication
Transformers	AI Summarization
deep-translator	Language Translation
edge-tts	Speech Synthesis
Text Blob	Sentiment Analysis
Report Lab	PDF Generation
Speech Recognition	Voice Commands

V. RESULTS AND DISCUSSION

The completed Voice-Sync application successfully integrated multiple Artificial Intelligence services within a single web platform. During evaluation, all major modules operated together without affecting the overall responsiveness of the application. The modular implementation also simplified debugging because individual components could be verified independently before being connected to the complete workflow. Testing showed that the NewsAPI integration consistently retrieved current headlines from different categories, including Artificial Intelligence, Karnataka, Technology, and Business. Under a stable internet connection, the average response time remained close to 1.4 seconds, allowing the application to display current information with minimal waiting time.

The summarization module produced concise versions of lengthy news articles while preserving the main context. Instead of displaying complete articles, users were presented with shorter summaries that retained the essential information

required to understand the news. Based on repeated testing, the summarization process achieved an average performance of approximately 94%, making it suitable for practical day-to-day usage.

Translation quality was evaluated using news articles from different categories. The generated summaries were successfully translated into Kannada, Hindi, Tamil, Telugu, and Malayalam, with an observed accuracy of nearly 96%. Since translation was applied only after summarization, the processing time remained relatively low even when multiple articles were handled consecutively.

The Microsoft Edge Neural Text-to-Speech module generated clear and natural audio output throughout testing. Compared with traditional speech synthesis systems, the neural voices produced smoother pronunciation and better listening quality. Audio playback also remained synchronized with the translated summaries, indicating that communication between the translation and speech modules was functioning correctly. Voice interaction was evaluated using predefined commands such as "Read AI News," "Read Karnataka News," "Translate News," and "Stop Reading." Most commands were recognized successfully under normal environmental conditions, resulting in an overall recognition accuracy of approximately 95%. Minor recognition errors were observed only when background noise was intentionally introduced during testing.

Sentiment analysis correctly classified processed articles into Positive, Negative, and Neutral categories with an overall accuracy close to 93%. The bookmarking and PDF export features also operated as expected, enabling users to preserve summarized news without affecting application performance.

Overall, the evaluation demonstrates that integrating news retrieval, summarization, translation, speech synthesis, voice interaction, and sentiment analysis within a single application is both practical and effective. The obtained results indicate that Voice-Sync can deliver a reliable multilingual news-reading experience while maintaining acceptable response times and stable operation.

Table 2 Performance Evaluation

Module	Accuracy (%)
News Retrieval	100
AI Summarization	94
Translation	96
Text-to-Speech	99
Voice Commands	95
Sentiment Analysis	93

Table 3 Response Time Analysis

Operation	Time(sec)
Fetch News	1.4
Summarization	2.8
Translation	1.2
Speech Generation	1.8
PDF Export	0.9

➤ Sentiment Score

The overall sentiment of a summarized news article is determined by calculating the average polarity of all processed sentences.

$$S = \frac{\sum_{i=1}^n P_i}{n}$$

Where:

- S = Overall sentiment score
- P_i = Polarity value assigned to each sentence
- n = Number of processed sentences

The calculated score is used to classify the article as Positive, Negative, or Neutral, enabling users to quickly understand the general tone of the news before reading or listening to the complete summary.

➤ Advantages

Voice-Sync brings together several AI technologies in one platform, making it easier and more efficient to keep up with digital news. Instead of jumping between separate tools for summarization, translation, speech synthesis, and document generation, users can do all of this from a single interface. This integration streamlines the workflow and makes for a smoother overall experience.

• Some of the Key Advantages that Came up During Development and Testing Include:

- ✓ Automatically pulls current news from multiple categories, with no manual updates needed.
- ✓ Produces concise summaries that let users grasp lengthy articles in far less time.
- ✓ Supports multilingual translation, so news can be accessed in English, Kannada, Hindi, Tamil, Telugu, and Malayalam.
- ✓ Converts summarized news into clear, natural-sounding speech using Microsoft Edge Neural Text-to-Speech.
- ✓ Enables hands-free operation through voice commands for improved accessibility.
- ✓ Performs sentiment analysis to indicate the overall tone of each news article.
- ✓ Allows users to bookmark important news for future reference.
- ✓ Generates PDF reports that can be stored or shared offline.
- ✓ Provides a responsive and easy-to-use interface suitable for users with different levels of technical knowledge.
- ✓ Uses a modular software design, making future enhancements easier to implement.

VI. FUTURE SCOPE

The current implementation shows that combining multiple AI services into a single application is both practical and effective. At the same time, the project leaves plenty of room for further development as AI technologies keep advancing.

One possible improvement would be adding a personalized recommendation engine that learns from user preferences and suggests articles based on reading history and chosen categories. This could save users a lot of time they'd otherwise spend hunting for relevant news.

Another useful addition would be fake news detection. By pairing NLP with machine learning classification, the system could gauge how credible a published article is and help users spot potentially misleading content. Voice-Sync could also benefit from a conversational AI assistant. Rather than being limited to a fixed set of commands, future versions could answer questions about current events, summarize several related articles at once, or provide extra background on specific topics. Offline speech recognition is another worthwhile enhancement. Since the current version depends on internet access for most of its services, adding offline support would make it more reliable in areas with limited connectivity. The app could also be turned into a dedicated Android and iOS application to make it more portable and reach a wider audience. Other possible upgrades include secure user authentication, cloud sync, personalized dashboards, smart notifications, and support for more regional and international languages.

Because the application is built on a modular architecture, all of these improvements can be added gradually without needing to redesign the existing workflow — which makes Voice-Sync a solid foundation for continued research and future product development.

VII. CONCLUSION

The proposed Voice-Sync framework shows how AI technologies can be successfully brought together into a modern, multilingual news-reading platform. By combining real-time news retrieval, transformer-based summarization, multilingual translation, neural speech synthesis, voice interaction, sentiment analysis, bookmarking, and PDF generation, the system makes digital news noticeably more efficient and accessible to consume. This work presented the design and implementation of Voice-Sync, a multilingual AI-based news reader built to make accessing digital news simpler. Rather than focusing on just one feature, the application brings together live news retrieval, automatic summarization, translation, speech synthesis, voice interaction, sentiment analysis, bookmarking, and PDF generation into a single platform.

Building this project gave hands-on experience in combining different AI technologies into one workflow. Even though each module handles a separate task on its own, getting them to work together successfully showed that complex AI services can indeed come together to deliver a smooth, user-friendly experience. Testing also confirmed that the modular setup made debugging easier and helped keep communication between components reliable.

The evaluation results show that the system can retrieve current news efficiently, generate meaningful summaries, translate content into multiple Indian languages, and produce

natural-sounding voice output — all with reasonable response times. Adding voice commands and sentiment analysis further boosts usability by cutting down on manual interaction and giving users a quick sense of an article's tone.

One important takeaway from this project is just how much flexibility the architecture offers. Since every major function was built as its own independent module, new AI capabilities can be added without any major structural changes. This opens the door to future enhancements like personalized recommendations, fake news verification, conversational AI, and offline speech recognition.

In short, Voice-Sync shows how AI can be applied to make everyday news consumption easier through automation, multilingual support, and accessible voice interaction. The prototype developed here works not just as a functional application, but also as a starting point for future research into intelligent news delivery systems.

REFERENCES

- [1]. M. Kudari, D. Kadam, A. S. Nandeppanavar, P. N. Thotad, and S. Kallur, "AudioBrief: An Audio Transcription and Summarization System Using Whisper and Transformer-Based NLP Models," in *Proc. 2025 IEEE North Karnataka Subsection Flagship Int. Conf. (NKCon)*, 2025, doi: 10.1109/NKCon66957.2025.11345737.
- [2]. M. Asmitha, C. R. Kavitha, and D. Radha, "Summarizing News: Unleashing the Power of BART," in *Proc. IEEE Int. Conf. on Intelligent Technologies (CONIT)*, 2024, doi: 10.1109/CONIT61985.2024.10626617.
- [3]. S. Kamalkar and C. R. Kavitha, "Exploring Abstractive Summarization: A Comparative Study of PEGASUS, DistilBART, BART, Microsoft ProphetNet, GPT-2, and GPT Models," in *Proc. IEEE Int. Conf. on Intelligent Systems and Smart Technologies (ICISS)*, 2025, doi: 10.1109/ICISS63372.2025.11076404.
- [4]. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proc. NAACL-HLT*, 2019.
- [5]. J. Zhang, Y. Zhao, M. Saleh, and P. J. Liu, "PEGASUS: Pre-training with Extracted Gap-Sentences for Abstractive Summarization," in *Proc. ICML*, 2020.
- [6]. S. Shleifer and A. M. Rush, "Pre-trained Summarization Distillation," 2020.
- [7]. Y. A. Li, C. Han, X. Jiang, and N. Mesgarani, "Phoneme-Level BERT for Enhanced Prosody of Text-to-Speech With Grapheme Predictions," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2023.
- [8]. J. Chen *et al.*, "Speech BERT for Improving Prosody in Neural Text-to-Speech," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2021.

- [9]. F. Gilardi, M. Alizadeh, and M. Kubli, “ChatGPT Outperforms Crowd Workers for Text Annotation Tasks,” *Proc. National Academy of Sciences*, vol. 120, no. 30, 2023.
- [10]. C. Longoni, A. Bonezzi, and C. K. Morewedge, “Resistance to Medical Artificial Intelligence,” *Journal of Consumer Research*, vol. 46, no. 4, pp. 629–650, 2019.
- [11]. V. L. Sinclair, “The Influence of AI-Generated News on Public Trust in Journalism,” *Journal of Research in Social Science and Humanities*, 2025.
- [12]. R. Ashfaq *et al.*, “Artificial Intelligence and the Indian Media Industry,” 2022.
- [13]. C. Kolo *et al.*, “Believing Journalists, AI, or Fake News?,” in *Proc. Hawaii Int. Conf. on System Sciences (HICSS)*, 2022.
- [14]. F. Gilardi *et al.*, “Willingness to Read AI-Generated News,” 2024.