

International Journal of Engineering Research and Science & Technology



www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 3 (2026)



ijerst.editor@gmail.com
editor@ijerst.com

Research Paper

SLIDE DECK GENERATOR FROM A VOICE NOTE/TEXT

¹ A Satyanarayana, ² B Soumya, ³ M Bodanapu Noor, ⁴ C Sai Satya, ⁵ G Divya

¹Assistant Professor, ^{2,3,4,5}Students

Department of CSE(AIML)

Siddhartha Institute of Technology & Sciences, Narapally

drsatanarayanaakella_cse@siddhartha.co.in, 23TQ1A6673@siddhartha.co.in,

23TQ1A66C9@siddhartha.co.in, 23TQ1A66C4@siddhartha.co.in, 23TQ1A66A1@siddhartha.co.in

Abstract

Creating professional presentation slides from raw ideas, voice notes, or unstructured text can be a time-consuming task. Users often have useful information in the form of voice recordings, meeting notes, handwritten content, or simple text, but converting this information into a well-structured presentation requires significant manual effort. The Slide Deck Generator from a Voice Note/Text is proposed as an AI-powered system that automatically converts voice or text input into an organized and visually appealing presentation.

The proposed system accepts voice recordings or text content from users through a web or mobile interface. When a voice note is provided, a Speech-to-Text module converts the audio into textual information. Natural Language Processing techniques then analyze the transcribed or directly entered text to identify the main topic, important ideas, keywords, sections, and relationships between concepts. The system uses these extracted elements to understand the overall purpose and structure of the presentation.

A Generative AI engine transforms the analyzed information into a structured slide outline. It can determine suitable slide titles, subtitles, bullet points, descriptions, examples, summaries, and conclusions based on the provided content. The system can divide lengthy information into multiple logical slides and maintain a consistent flow between them. Users can also specify the presentation type, target audience, number of slides, language, and preferred style.

The system can automatically generate visual recommendations for each slide, including charts, icons, diagrams, illustrations, and relevant images. A presentation-generation module combines the generated content with predefined templates, layouts, fonts, and design elements to produce a complete slide deck. Users can preview the generated presentation, edit individual slides, regenerate content, rearrange slides, and export the final presentation in a suitable format.

Overall, the proposed system aims to simplify presentation creation by transforming unstructured voice or text information into a professional slide deck with minimal manual effort. By combining Speech Recognition, Natural Language Processing, Generative AI, content summarization, visual recommendation, and automated presentation design, the system can save time and improve productivity. Future enhancements can include multilingual voice processing, real-time presentation

generation, automatic speaker notes, advanced diagram generation, brand-template integration, and interactive presentation creation.

I. Introduction

Presentations are widely used in education, business, research, meetings, training, and professional communication. Preparing an effective presentation requires more than simply placing information on slides. The creator must identify important ideas, organize them into a logical sequence, summarize lengthy content, select suitable visuals, and maintain a consistent design. These activities can require considerable time, especially when the initial information is unstructured.

Many users first capture their ideas through voice notes, meeting recordings, brainstorming sessions, or simple text notes. Although these formats are convenient for collecting information, they are not directly suitable for presentation delivery. Users normally need to manually transcribe voice recordings, identify key points, create a slide outline, write slide content, select images, and format each slide.

Artificial Intelligence provides new opportunities for automating this process. Speech Recognition can convert spoken content into text, while Natural Language Processing can identify important concepts and summarize information. Generative AI can then transform the extracted content into structured narratives, slide titles, bullet points, explanations, and conclusions.

A slide-generation system can further improve this workflow by automatically selecting appropriate layouts and visual elements. Instead of simply generating text, an intelligent system can determine whether a particular idea is better represented through a bullet list, process diagram, timeline, comparison table, chart, or image. This allows the generated presentation to communicate information more effectively.

The proposed Slide Deck Generator from a Voice Note/Text integrates these technologies into a single platform. Users can upload a voice note or enter text, after which the system transcribes and analyzes the content. The AI engine identifies the main topic and creates a structured presentation outline according to the user's requirements.

The main objective of the system is to reduce the time and technical effort required to create presentations. Users can review and modify the generated slides before finalizing them. Human review remains important because AI-generated content and visual choices may require correction, especially for academic, professional, or factual presentations.

II. Literature Survey

Traditional presentation-creation tools require users to manually prepare slide content, choose layouts, insert images, and format text. Applications for presentation design provide templates and visual components, but the user remains responsible for converting raw information into an appropriate slide structure. This can be challenging for users who have large amounts of unstructured content.

Speech Recognition technology has enabled systems to convert spoken language into machine-readable text. Modern speech-to-text models can process conversations, lectures, meetings, and voice notes with increasing accuracy. Transcription provides a useful foundation for presentation generation because spoken ideas can be converted into text before further analysis.

Natural Language Processing has been widely used for text summarization, topic extraction, keyword identification, sentiment analysis, and document classification. These techniques can identify important information from lengthy documents and reduce large amounts of text into concise summaries. Such capabilities are useful for determining what information should be included in presentation slides.

Generative AI and Large Language Models have significantly expanded automated content-generation capabilities. These models can create summaries, outlines, explanations, titles, bullet points, and other forms of structured content from natural-language input. They can also adapt generated content according to audience, tone, presentation length, and subject requirements.

Research into automated presentation generation has explored methods for converting documents into slide structures and selecting appropriate visual layouts. Some systems focus mainly on extracting headings and paragraphs, while others use templates to create basic slides. However, automatically generating a coherent narrative and selecting appropriate visuals remains a challenging problem.

Recent AI-based presentation systems combine speech recognition, NLP, Generative AI, visual generation, and automated document creation. These approaches can transform raw user input into more complete presentation artifacts. However, challenges remain in maintaining factual accuracy, avoiding excessive text, selecting relevant visuals, preserving logical flow, and producing consistent designs. The proposed system addresses these challenges through structured content analysis, AI-based slide planning, visual recommendations, template-based generation, validation, and human review.

III. System Analysis

The proposed system provides an intelligent platform for converting voice notes and text into structured presentation slide decks. The process begins when a user uploads a voice recording or enters raw text through the application interface. If the input is audio, a Speech-to-Text module converts the recording into text while preserving the important spoken information. The NLP processing module cleans and analyzes the text to identify the main topic, keywords, concepts, sections, and important statements. A content-understanding module summarizes lengthy information and removes unnecessary repetition. The AI planning module creates a logical presentation outline based on the topic, audience, requested slide count, and presentation objective. A Generative AI engine produces slide titles, bullet points, explanations, examples, and summary content. A visual recommendation module determines suitable images, icons, diagrams, charts, timelines, or other visual elements for individual slides. A presentation-design module applies templates, layouts, fonts, spacing, and other design rules to organize the generated content. A validation module checks the generated deck for content consistency, slide structure, readability, and possible unsupported

information. Users can preview the presentation and edit, rearrange, regenerate, or remove individual slides. The final presentation can be exported into formats such as PPTX or PDF. The system can also generate speaker notes to assist the presenter during delivery. This workflow enables users to move from unstructured ideas to a structured presentation with significantly less manual effort.

Existing System

Existing presentation-creation systems generally require users to manually convert their ideas or source material into presentation slides. Users may begin with voice notes, documents, meeting notes, or raw text and then manually extract the important information. When voice recordings are involved, transcription may be performed separately before the content can be used in a presentation. Users must manually determine the presentation structure, slide sequence, titles, and key points. Selecting suitable images, diagrams, charts, and layouts also requires additional effort. Conventional presentation software provides templates and design tools but generally does not understand the complete meaning of the user's raw content. Basic AI presentation features may generate slides from text but can provide limited control over complex content structures. Long paragraphs may be transferred directly into slides without sufficient summarization. Visual selection may also be irrelevant or inconsistent with the slide topic. Users often need to manually check the generated presentation for factual and structural problems. Maintaining consistent formatting across many slides can require considerable work. Different presentation types may require different narrative structures that conventional systems do not automatically identify. These limitations create a need for an intelligent platform capable of understanding voice or text content and automatically transforming it into a coherent and visually organized presentation.

Disadvantages of Existing System

- Requires significant manual slide creation.
- Voice notes usually require separate transcription.
- Manual content summarization is time-consuming.
- Users must manually create the slide outline.
- Manual selection of important points is required.
- Visual selection can be difficult.
- Requires manual formatting and layout adjustment.
- Large amounts of text can result in overcrowded slides.
- Limited understanding of presentation context.
- Basic automation may generate generic slide structures.
- Manual consistency checking is required.
- Creating diagrams and charts can require additional effort.
- Speaker notes often need to be created separately.
- Difficult to rapidly generate multiple presentation versions.
- Limited personalization for different audiences.

Proposed System

The proposed system introduces an AI-powered platform that automatically transforms voice notes and text into complete presentation slide decks. Users can upload an audio

recording or enter raw text through the system interface. Voice input is processed by a Speech-to-Text module that converts spoken content into an editable transcript. The NLP engine analyzes the transcript or text and extracts topics, keywords, important concepts, sections, and relationships between ideas. A content-processing module summarizes lengthy information and removes unnecessary repetition while preserving important details. The AI planning engine creates a logical slide structure based on the presentation objective, audience, topic, and requested slide count. The Generative AI module produces suitable slide titles, bullet points, descriptions, examples, summaries, and conclusions. A visual intelligence module recommends appropriate images, icons, diagrams, charts, and other visual elements for each slide. The presentation-generation engine applies suitable templates and layouts while maintaining consistent typography, spacing, and visual hierarchy. The system can automatically generate speaker notes from the slide content. A validation layer checks the presentation for logical consistency, readability, duplication, and potentially unsupported claims. Users can preview the generated deck and edit or regenerate selected slides through the interface. The final presentation can be exported in formats such as PPTX and PDF. This integrated approach allows users to convert raw ideas into professional presentations quickly while retaining control over the final output.

Advantages of Proposed System

- Converts voice notes directly into presentations.
- Supports both audio and text input.
- Automatically performs speech-to-text conversion.
- Extracts important information using NLP.
- Automatically summarizes lengthy content.
- Creates logical slide structures.
- Generates slide titles and key points.
- Uses Generative AI for content creation.
- Recommends relevant visuals.
- Automatically applies presentation layouts.
- Maintains consistent slide formatting.
- Generates speaker notes.
- Allows users to edit and regenerate slides.
- Reduces presentation-creation time.
- Supports different audiences and presentation styles.
- Enables rapid export and sharing of completed presentations.

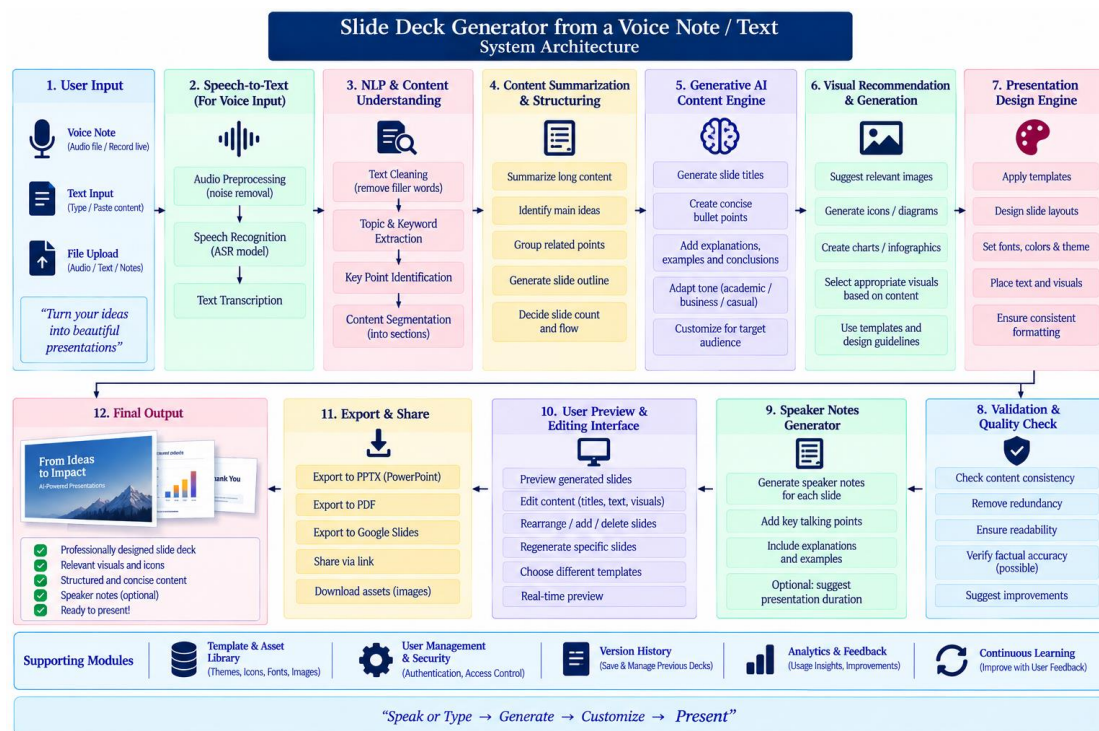
IV. Methodology

The methodology of the proposed system consists of several stages beginning with voice or text input and ending with a complete presentation. The first stage is **input collection**, where the user uploads a voice note or enters raw text through the application interface. If audio is provided, the Speech-to-Text module converts the recording into an editable textual transcript. The preprocessing module removes unnecessary repetitions, filler words, and irrelevant content while preserving meaningful information. The NLP engine analyzes the content to identify the topic, keywords, important concepts, sections, entities, and relationships. A summarization module condenses lengthy information into concise points suitable for presentation slides. The AI planning module determines the number and sequence of slides

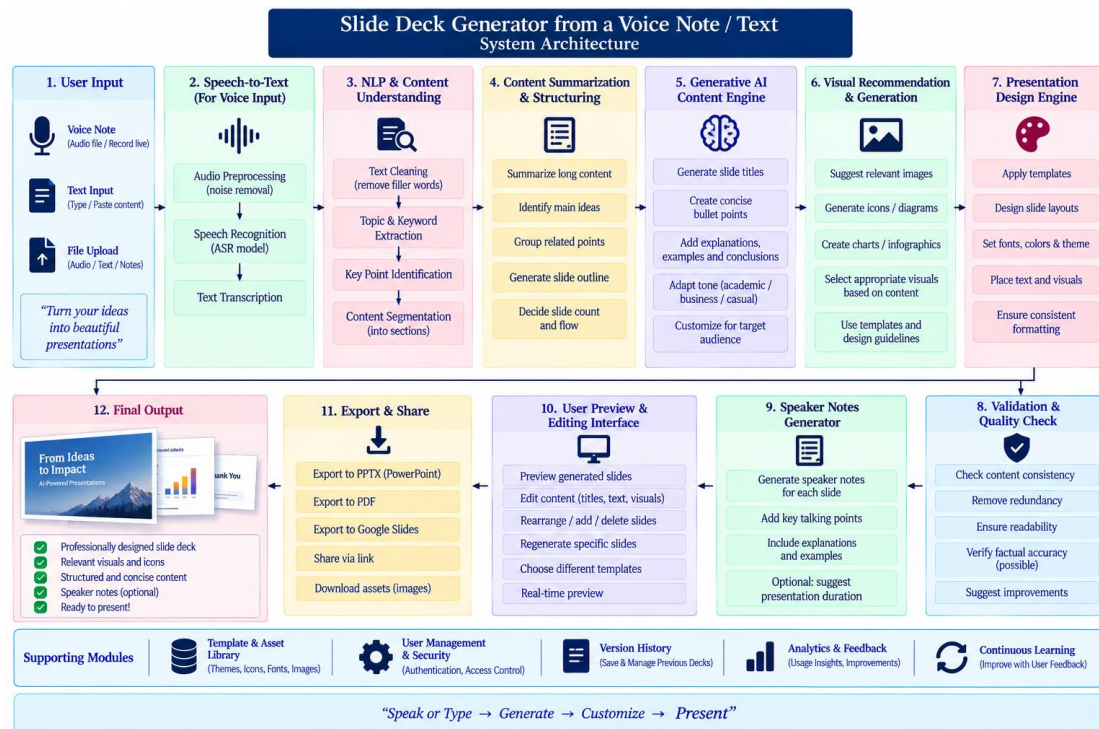
according to the user's requirements. The Generative AI engine creates slide titles, bullet points, explanations, examples, and conclusions based on the analyzed content. A visual-generation or recommendation module selects appropriate images, icons, diagrams, charts, and other visual elements for individual slides. The presentation-design engine combines the content and visuals using predefined templates and layout rules. A quality-control module checks the generated presentation for consistency, readability, duplication, and potential unsupported statements. The system generates optional speaker notes to help the presenter explain each slide. Users can preview, edit, rearrange, regenerate, or remove slides before finalization. Finally, the presentation is exported into formats such as PPTX or PDF and can be shared or presented directly.

System Architecture

The system architecture consists of several interconnected components including the User Interface, Voice/Text Input Module, Speech-to-Text Engine, NLP Processing Layer, Content Summarization Module, AI Slide Planner, Generative AI Engine, Visual Recommendation Module, Presentation Design Engine, Validation Module, Speaker Notes Generator, and Export Module. The process begins when a user uploads a voice note or provides text through the application interface. Audio input is passed to the Speech-to-Text engine, which generates an editable transcript. Both transcripts and direct text inputs are processed by the NLP layer to identify topics, concepts, keywords, entities, and important information. The summarization module condenses the content into concise presentation-ready information. The AI Slide Planner determines the logical sequence, number, and purpose of individual slides. The Generative AI engine creates titles, bullet points, explanations, examples, and conclusions for each slide. The Visual Recommendation module identifies suitable images, icons, charts, diagrams, and other visual components based on slide content. The Presentation Design Engine applies templates, layouts, typography, spacing, and visual hierarchy to create the complete slide deck.



The Validation Module checks content consistency, readability, structure, duplication, and potential unsupported information. The complete architecture follows the flow **Voice Note/Text → Speech-to-Text (for audio) → NLP & Content Understanding → Summarization → AI Slide Planning → Generative AI Content → Visual Recommendation → Presentation Design → Validation → Human Review → Speaker Notes → PPTX/PDF Export**.



V. Result and Output

SlideGen AI
Turn your ideas into stunning presentations

Home Templates My Decks Help R Ranjitha

1. Input Your Content

Voice Note Text Input

Recording... Speak about your topic

Or upload an audio file

Drag & drop audio file here (MP3, WAV, M4A)

Browse Files

OR

Paste your text

Create a presentation on renewable energy covering introduction, types, benefits, challenges and future scope. Keep it simple and suitable for a college audience.

124/5000

Generate Slides

2. AI Processing

- Converting speech to text Completed
- Analyzing content & extracting key points Completed
- Generating slide outline Completed
- Creating slide content with AI Completed
- Selecting visuals and design In Progress
- Finalizing presentation Pending
- Preparing download files Pending

Generating your slides... 80%

Tip: You can edit the outline or specify the number of slides, audience and tone for better results.

3. Slide Outline (Editable)

Edit Outline

- Title Slide: Renewable Energy: A Cleaner, Brighter Future
- Introduction: What is Renewable Energy?
- Types of Renewable Energy: Solar, Wind, Hydro, Biomass, Geothermal
- Benefits of Renewable Energy: Environmental, Economic, Social
- Challenges: Limitations and Current Issues
- Future Scope: Trends and Opportunities
- Conclusion: Key Takeaways
- Thank You: Questions?

4. Slide Preview

Slide 1 of 8

Preview Full Deck Download

Final Generated Slides (Sample Output)

1. Title Slide

Renewable Energy
A Cleaner, Brighter Future

Sustainable Today for a Better Tomorrow

2. Introduction

What is Renewable Energy?

- Energy from naturally replenishing sources
- Sustainable and eco-friendly
- Reduces dependence on fossil fuels
- Helps in combating climate change

3. Types of Renewable Energy

Solar Energy, Wind Energy, Hydropower, Biomass, Geothermal Energy

4. Benefits

Benefits of Renewable Energy

- Cleaner environment
- Lower greenhouse gas emissions
- Long-term cost savings
- Energy security
- Creates new job opportunities

5. Challenges

Challenges

- High initial cost
- Weather dependency
- Energy storage issues
- Infrastructure limitations
- Policy and regulatory challenges

6. Future Scope

Future Scope

- Advancements in storage technology
- Improved efficiency
- Integration with smart grids
- Increased global investments
- Greater adoption in industries and homes

7. Conclusion

Conclusion

- Renewable energy is essential for a sustainable future
- Offers long-term environmental and economic benefits
- Continued innovation and support are needed
- Together we can build a cleaner, greener tomorrow

8. Thank You

Thank You!
Questions?

"Clean Energy. Bright Future."

VI. Conclusion

The Slide Deck Generator from a Voice Note/Text provides an efficient and intelligent way to create presentations from simple voice recordings or text input. The system reduces the effort and time required to manually prepare slide content, organize information, and design presentations.

By using Speech-to-Text, Natural Language Processing, summarization, and Generative AI, the system can understand the user's ideas, identify important concepts, generate suitable slide titles and content, and organize them into a meaningful presentation structure. It can also recommend visuals, apply appropriate templates, and generate speaker notes to improve the overall presentation quality.

The system also provides users with a preview and editing facility, allowing them to modify the generated content before finalizing the presentation. This human-in-the-loop approach helps users maintain control over the final output while benefiting from AI-based automation.

Overall, the proposed system makes presentation creation faster, easier, and more accessible, especially for students, teachers, professionals, and users who prefer explaining their ideas through voice rather than manually creating slides. In the future, the system can be enhanced with multilingual voice support, real-time collaboration, advanced visual generation, automatic charts, and personalized presentation styles.

References

- [1] Kumar, R. D., Prudhviraaj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine learning algorithm. In Handbook of Artificial Intelligence and Wearables (pp. 145-158). CRC Press.
- [2] Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In The International Conference on Artificial Intelligence and Smart Environment (pp. 557-564). Cham: Springer Nature Switzerland.
- [3] Sv satyakrishna, shirisha rangi ,bhargavi nalacheruve.(2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm
- [4] Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.
- [5] V. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, "Real-Time Object Detection in Drone Surveillance Using YOLOv5," in Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.
- [6] B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, "Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks," in Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart Environment, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.
- [7] R. D. Kumar, V. N. S. Manaswini, "Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology," in

Blockchain for Smart Cities, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.

[8] Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, “An advanced movie recommender using collaborative filtering and sentiment analysis,” International Research Journal of Modernization in Engineering Technology and Science, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.

[9] Ravi Kumar Banoth, Ramana Murthy B V, “Automatic crop recommendation system using LightGBM and decision tree machine learning models,” Journal of Machine and Computing, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.

[10] Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, “Smart agriculture through IoT and machine learning for analyzing carbon footprints,” in Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE), Apr. 2025.

[11] Ravi Kumar Banoth, B. V. Ramana Murthy, “Soil image classification using transfer learning approach: MobileNetV2 with CNN,” SN Computer Science, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.