### VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI KARNATAKA-590 018

#### A Mini-Project Report

***“*TEXT TO VIDEO CONVERSION OF GOVERNMENT PRESS RELEASE*”***

*Submitted in partial fulfilment of the requirements for the award of Degree of*

##### BACHELOR OF ENGINEERING

**in**

##### COMPUTER SCIENCE AND ENGINEERING

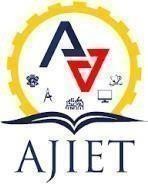
**by**

|  |  |
| --- | --- |
| **ARPITA SURESH NAIK** | **4JK21CS011** |
| **SONAL WILSON DSOUZA** | **4JK21CS059** |
| **MEGHA S GOWDA** | **4JK21CS064** |
| **NAVEEN ANKOLEKAR** | **4JK22CS402** |

## Under the guidance of

#### Mrs. Shilpa Shetty A

**Asst.Professor of CSE, AJIET**



**Department of Computer Science and Engineering**

**Accredited by NBA (BE: CV, CSE, ECE, ISE & ME)**

### A.J. INSTITUTE OF ENGINEERING AND TECHNOLOGY (A unit of Laxmi Memorial Education Trust®)

**(Approved by AICTE, New Delhi, Affiliated to VTU, Belagavi, Recognized by Govt. of Karnataka)**

### NH-66, Kottara Chowki, Mangaluru – 575006 A.Y: 2023-24

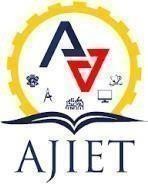
#### A. J. INSTITUTE OF ENGINEERING AND TECHNOLOGY

**(A UNIT OF LAXMI MEMORIALEDUCATION TRUST®)**

**(Approved by AICTE, New Delhi, Affiliated to Visvesvaraya Technological University, Belagavi)**

#### Accredited By NBA (BE: CV, CSE, ECE, ISE & ME)

**NH-66, Kottara Chowki, Mangaluru – 575006**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING CERTIFICATE**

This is to certify that the Mini project entitled “***Text to Video Conversion of Government Press Release***” is a bonafide work carried out by **Arpita Suresh Naik (4JK21CS011), Sonal Wilson Dsouza (4JK21CS059), Megha S Gowda (4JK21CS064), Naveen Ankolekar(4JK22CS402)** students of Sixth semester B.E. Computer Science & Engineering, and submitted as a part of the Mini Project, during the academic year **2023-2024**.

|  |  |
| --- | --- |
| **GUIDE**  **Mrs. Shilpa Shetty A** | **COORDINATOR**  **Mrs. Ashwitha Shetty** |
| **Vice Principal & HOD Dr. Antony P J** | |

## ACKNOWLEDGEMENT

I dedicate this page to acknowledge and thank those responsible for the shaping of the project. Without their guidance and help, the experience while constructing the dissertation would not have been so smooth and efficient.

I sincerely thank the **Management of AJ INSTITUTE OF ENGINEERING AND TECHNOLOGY, a unit of LAXMI MEMORIAL EDUCATION TRUST** for all their support.

I deem it a privilege to place on record the deep sense of gratitude to my guide **Mrs. Shilpa Shetty A, Assistant Professor**, **Computer Science and Engineering**, who always stood behind and supported in each and every step of the internship.

I sincerely thank **Mrs. Ashwitha Shetty, Assistant Professor**, **Computer Science and Engineering**, for her guidance and valuable suggestions which helped us to fulfill the experiments prescribed by the university.

I owe our profound gratitude to **Dr Antony P J, Vice Principal** and **Head of the Department, Computer Science and Engineering**, whose kind consent and guidance helped us to complete this work successfully.

I am extremely thankful to our **Dean Academics, Dr. P Mahabaleshwarappa** and

**Principal, Dr Shantharama Rai C**. for his support and encouragement.

I would like to thank all our Computer Science and Engineering staff members who have always been with us extending their support, precious suggestions, guidance and encouragement through the project.

We also like to extend thanks to our friends and family members for their continuous support.

|  |  |
| --- | --- |
| **ARPITA SURESH NAIK** | **4JK21CS011** |
| **SONAL WILSON DSOUZA** | **4JK21CS059** |
| **MEGHA S GOWDA** | **4JK21CS064** |
| **NAVEEN ANKOLEKAR** | **4JK22CS402** |

## ABSTRACT

In the digital age, the dissemination of information from government agencies to the public must be both effective and engaging. Traditional text-based press releases, while informative, often struggle to capture the public’s attention and convey complex information in a digestible format. This project introduces an innovative solution for transforming government press releases into compelling video content through advanced text-to-video conversion techniques.

The process begins with the extraction of key information from text-based press releases, which is then analyzed to determine the most effective way to convey the content visually. This involves generating visual elements such as images, graphics, and animations, as well as creating a coherent narrative flow. Advanced algorithms are employed to match textual information with appropriate visual representations and to synchronize these elements with background music and voiceover narration.

The final video content is designed to be informative, engaging, and accessible, providing viewers with a clear understanding of the press release's key messages. The effectiveness of the system is evaluated through user feedback and engagement metrics, with the goal of improving public communication strategies for government agencies. This project not only aims to streamline the communication process but also seeks to explore the potential of automated multimedia content creation in the context of public information. To ensure the effectiveness of the final product, the project includes a comprehensive evaluation strategy based on user feedback and engagement metrics. We assess the video’s ability to convey the key messages of the press release, the clarity of information presentation, and the level of public engagement achieved. This evaluation helps refine the content creation process and improve future iterations of the video generation system.

Overall, this project aims to modernize public communication methods for government agencies by leveraging automated technologies to produce high-quality, engaging video content from text-based press releases. The anticipated outcomes include enhanced public understanding of government announcements, increased viewer engagement, and a scalable solution for future press release dissemination.

|  |  |  |
| --- | --- | --- |
| **Chapter** | **Title** | **Page no.** |
|  | **ACKNOWLEDGEMENT** | i |
|  | **ABSTRACT** | ii |
|  | **TABLE OF CONTENTS** | iii |
|  | **LIST OF FIGURES** | iv |
| **Chapter 1** | **INTRODUCTION** | 1-3 |
| **1.1** | **LITERATURE SURVEY** | 2-3 |
| **1.2** | **RELEVANCE OF THE PROJECT** | 3 |
| **Chapter 2** | **PROBLEM FORMULATION** | 4-6 |
| **2.1** | **OBJECTIVE** | 4 |
| **2.2** | **SCOPE OF THE PROJECT** | 5-6 |
| **Chapter 3** | **PROPOSED SOLUTION** | 7-8 |
| **3.1** | **METHODOLOGY** | 8 |
| **Chapter 4** | **HARDWARE & SOFTWARE REQUIREMENTS** | 9 |
| **Chapter 5** | **RESULTS** | 10-11 |
| **Chapter 6** | **CONCLUSION** | 12 |
|  | **REFERENCES** | 13 |

|  |  |  |
| --- | --- | --- |
| **FIGURE NO.** | **TITLE** | **PAGE NO.** |
| 5.1 | Home Page | 10 |
| 5.2 | Output Page | 11 |

***CHAPTER 1***

# INTRODUCTION

In the digital age, effective communication strategies are essential for government agencies to reach and engage the public with important information. While traditional text-based press releases have long been the standard for conveying official announcements, they often fail to capture the public’s attention and effectively convey complex messages due to their static and dense nature. As audiences increasingly turn to multimedia platforms for information, there is a pressing need for innovative methods that can transform traditional press releases into more engaging and accessible formats. The “Text to Video Conversion of Government Press Releases” project addresses this challenge by developing a sophisticated system that converts text-based press releases into dynamic video content.

This project leverages cutting-edge technologies in Natural Language Processing (NLP) and multimedia content creation to bridge the gap between traditional text communication and modern digital media. By analyzing the text of press releases to extract key information, generate relevant visual elements, and produce synchronized audio narration, the project aims to create videos that are not only informative but also captivating for viewers. The resulting videos are designed to present information in a visually appealing and easy-to-understand format, thus enhancing public engagement and accessibility.

In addition to improving the presentation of government announcements, the project explores the potential of automated content creation tools to streamline and scale the process of video production. By developing an automated pipeline for converting text into video, the project seeks to offer a scalable and efficient solution for producing high-quality video content from press releases. The effectiveness of this approach is evaluated through user feedback and engagement metrics, providing insights into how well the videos meet the goals of clarity, engagement, and informational value.

This project aims to set a new standard for how government agencies communicate with the public, offering a model for future applications of text-to-video technologies in both the public and private sectors. By turning traditional text-based releases into dynamic videos, the project seeks to enhance public engagement and ensure that governmental messages reach a wider and more diverse audience. The successful implementation of this system could pave the way for more innovative and effective communication strategies across various domains, from educational content to corporate communications.

## Literature Survey

**Text-to-Video Conversion** is a rapidly evolving field that leverages advancements in natural language processing (NLP) and computer vision to transform textual information into dynamic visual content. This technology has seen significant development in recent years, particularly in its application to various domains including media, education, and government communication.

Early research in this area primarily focused on the **generation of video content from text descriptions,** where the emphasis was on creating coherent video narratives from textual input. Notable contributions include the work of [Zhu et al. (2018)](https://arxiv.org/abs/1805.02946), who introduced methods for generating videos from natural language descriptions by leveraging recurrent neural networks (RNNs) and convolutional neural networks (CNNs). Their approach laid the groundwork for understanding how text can be converted into a series of visual frames.

The field of text-to-video conversion has gained considerable attention as researchers explore ways to transform text-based information into engaging multimedia content. One notable study by Jain et al. (2020) investigates the use of sequence-to-sequence models for generating video sequences from textual descriptions, demonstrating the potential of Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks in creating video content from text. Their work highlights the promise of Seq2Seq architectures but also identifies challenges related to generating coherent and contextually relevant video sequences.

In the context of government press releases, the application of text-to-video conversion is still emerging. Research on this specific application is sparse, but the potential benefits are evident. Government press releases are often text-heavy and formal, requiring a nuanced approach to convert them into engaging and accessible video content. The work of [Wang et al. (2021)](https://arxiv.org/abs/2105.10436) on creating informative and engaging videos from formal texts offers valuable insights for adapting these techniques to government communications.

Challenges specific to government press releases include maintaining the formal tone and accurate representation of information. The study by [Li et al. (2022)](https://arxiv.org/abs/2202.04673) highlights techniques for preserving the tone and style of formal documents while creating multimedia content. Their methods could be adapted to ensure that government press releases are represented accurately and effectively in video format. The process involves summarization of content, real-time translation, and text-to-speech conversion, underpinned by 2D animation and synchronized speech.

This approach aims to make government updates more accessible and engaging, bridging language barriers and enhancing public awareness. The project begins with web crawling to acquire press releases from the PIB website, converting them from PDF to structured JSON format for further analysis. The integration of multimedia elements aims to foster informed citizenship and contribute to a more transparent and accessible government communication landscape, enhancing the overall experience and impact of public information dissemination.

The literature indicates a robust foundation for text-to-video conversion technologies, with a range of methods and models developed for generating video content from text. The specific application to government press releases presents unique challenges and opportunities, and ongoing advancements in this field will be crucial for developing effective solutions for this domain.

## Relevance of the project

The **Text-to-Video Conversion for Government Press Releases** project is highly relevant in the modern era of digital communication and public engagement. Traditional text-based press releases often fail to capture the attention of the public due to their dense and formal nature, which can limit their effectiveness in conveying important government information. By transforming these press releases into engaging and informative videos, this project addresses a critical need for more accessible and visually appealing content. Videos can break down complex policy information into digestible and engaging formats, making it easier for a wider audience to understand and connect with government announcements. By converting these press releases into engaging videos, the project addresses the need for more effective and accessible communication channels. It caters to diverse audience preferences, including those who are more visually oriented or have limited time to read lengthy documents.

This approach not only enhances public accessibility and engagement but also leverages advanced technologies such as natural language processing and computer vision to create innovative solutions for public information dissemination. Furthermore, in an age where digital media dominates, this project aligns with trends in media consumption and communication, offering a modern tool for improving transparency and outreach in government communications. As governments and institutions seek to connect more effectively with their constituents, this project provides a valuable tool for bridging the gap between information dissemination and audience engagement.

***CHAPTER 2***

# PROBLEM FORMULATION

The **Text-to-Video Conversion for Government Press Releases** project aims to address the challenge of effectively communicating important governmental information to the public through traditional text-based press releases. These press releases, while crucial for disseminating updates and policies, often suffer from issues of low engagement and accessibility due to their formal and dense nature. The primary problem is transforming these lengthy, complex texts into engaging and informative video content that retains the accuracy and formality of the original message.

This involves developing methods that not only convert text into visually appealing videos but also ensure that the complex policy information is presented in a clear and accessible manner for a diverse audience. Additionally, the project seeks to leverage advanced technologies in natural language processing and computer vision to automate and optimize the video creation process, addressing the challenge of balancing technical complexity with the need for high-quality, effective public communication.

## Objective

The primary **objectives** of the **Text-to-Video Conversion for Government Press Releases** project is to develop a system that converts traditional text-based government press releases into engaging, accessible, and accurate video content, utilizing advanced AI technologies to enhance public communication and information dissemination. This objective encompasses the core aim of the project, which is to transform text-based press releases into videos that are both informative and engaging for the public while leveraging the latest advancements in AI for efficient content creation. This project enhances the effectiveness of government communication by transforming traditional text-based press releases into engaging and accessible video content.

Here are some **objectives** of this project in points:

* Convert text-based press releases into engaging video content.
* Enhance public engagement with government announcements.
* Improve accessibility for diverse audiences through visual media.
* Ensure accuracy and formality in video content.
* Leverage advanced AI technologies for efficient video creation.

## Scope of the project

The project aims to develop an automated system that converts government press releases into engaging video content. This system will leverage natural language processing (NLP) and video generation technologies to transform written press releases into visual narratives. The primary goal is to enhance the accessibility, engagement, and comprehensibility of governmental announcements, ensuring that critical information reaches a broader audience effectively. Key components include a robust text analysis module to identify and summarize essential information, a script generation engine to create a coherent narrative, and a multimedia integration framework to combine text, images, and videos. The final deliverable will be a user-friendly application that allows government agencies to input press releases and generate high-quality videos suitable for dissemination across various digital platforms.

A sophisticated text analysis module will identify critical elements such as entities, dates, actions, and sentiments, summarizing the content while preserving the original message's intent. The project will also involve the development of a script generation engine that constructs a coherent and compelling narrative from the summarized data, ensuring logical flow and clarity. The system will incorporate a multimedia integration framework designed to combine text, images, videos, and infographics seamlessly, enhancing the video's visual and informational appeal.

The process will involve several key components: natural language processing (NLP) for extracting essential information from text, text-to-speech (TTS) systems for generating human-like narration, and advanced video synthesis techniques for visual content creation. The project will utilize state-of-the-art NLP algorithms to parse and interpret press release texts, identifying the main topics, key facts, and important details. This step will involve sentiment analysis to gauge the tone of the press release and thematic analysis to highlight central themes. Following text processing, the system will use TTS technology to produce clear and engaging voiceovers that narrate the text in a professional and accessible manner.

The project also involves the integration of these elements into a cohesive video, which is designed to be easily shareable and accessible across various platforms. By focusing on these components, the project not only streamlines the process of creating video content but also ensures that it is both informative and engaging. The scope extends to adapting the system for different types of press releases and potentially scaling it for broader use within governmental or organizational communication strategies.

The project will provide a scalable framework for future enhancements. Future iterations could include features such as interactive video elements, live streaming capabilities for real-time updates, and advanced analytics to track viewer engagement and feedback. The project’s architecture will be designed with flexibility in mind, allowing for the integration of new technologies and methodologies as they emerge.

The project's scope encompasses several aspects, including:

* **Text Analysis and Summarization**: Develop a robust text analysis module to parse press releases, extract key information, and generate concise summaries. This module will use NLP techniques to identify important entities, dates, and actions within the text.
* **Audio Narration**: The project will convert the text summaries into audio narrations, creating a professional voiceover that accompanies the video. This adds an auditory dimension to the visual content, making it more engaging and accessible.
* **Video Creation**: The final phase involves combining the images, audio narration, and any additional elements into a cohesive video. This includes video editing, synchronization of audio and visuals, and ensuring the final product is polished and ready for distribution.
* **Script Generation:** Create a script generation engine that constructs a coherent narrative from the summarized information. This engine will ensure the narrative flows logically and maintains the original message's integrity.
* **Web Application Development**: The project includes building a web application using Flask that allows users to upload press releases, process them, and view or download the resulting videos. This provides an intuitive interface for users to interact with the system.
* **User-Friendly Interface**: Develop a user-friendly application interface that allows government agencies to input press releases and generate high-quality videos. The interface will include options for manual adjustments and customization.
* **Testing and Validation**: Conduct thorough testing and validation to ensure the system's accuracy, reliability, and performance. This includes usability testing with end-users and stakeholders to gather feedback and make necessary improvements.

***CHAPTER 3***

## PROPOSED SOLUTION AND METHODOLOGY

The proposed solution for the Text to Video Conversion of Government Press Releases involves developing a sophisticated, automated system that seamlessly converts textual press releases into engaging video content. This solution aims to streamline the process of creating video content from written announcements, making government communications more accessible and engaging for the public. The solution encompasses several core components: a text analysis module, a script generation engine, a multimedia integration framework, accessibility features, a user-friendly interface, and distribution and integration capabilities. By leveraging advanced technologies in natural language processing (NLP) and video generation, the proposed system will automate and enhance the creation of high-quality video content from press releases, ensuring consistency with government branding, promoting public engagement, and providing broad accessibility to diverse audiences.

The process includes extracting text from PDF files, summarizing this text, extracting keywords, fetching and processing images, generating audio narration, and creating the final video. This Flask application is designed to take a PDF file as input, extract text from it, summarize the text, fetch related images from Unsplash based on the summary keywords, generate an audio narration of the summary, and create a video combining the resized images with the audio narration.

Here's some functionalities of the proposed solution:

* **Text Extraction:** Uses PyMuPDF (fitz) to extract text from uploaded PDF files.
* **Text Summarization:** Summarizes the extracted text using the summa library.
* **Keyword Extraction:** Extracts keywords from the summary using nltk and spacy.
* **Image Fetching:** Fetches images from Unsplash based on the extracted keywords using pyunsplash.
* **Image Downloading and Resizing**: Downloads the fetched images, resizes them using PIL, and saves them locally.
* **Audio Generation:** Generates an audio narration of the summary using gtts.
* **Video Generation:** Combines the resized images and audio narration to create a video using moviepy.

**3.1 Methodology**

##### Data Collection:

* Gather a diverse and representative dataset of PIB press releases covering various topics and domains. Ensure the dataset is annotated with metadata such as release date, category, and key themes.

##### Preprocessing:

* Clean and preprocess the textual data to remove noise, handle punctuation, and tokenize the text into meaningful units. Additionally, perform text normalization and standardization to ensure consistency across the dataset.

##### Text Summarization:

* Implement advanced Natural Language Processing (NLP) techniques for text summarization to extract key information and generate concise summaries of each PIB press release. Utilize methods such as extractive or abstractive summarization based on the specific requirements of the project.

1. **User Interaction and Feedback:**

* Integrate mechanisms for user interaction and feedback, allowing users to provide input, preferences, and corrections to improve the quality and relevance of the generated video presentations iteratively.

1. **Evaluation and Validation:**

* Conduct thorough evaluation and validation of the generated video presentations, assessing metrics such as accuracy, coherence, engagement, and user satisfaction. Use both qualitative and quantitative methods to validate the effectiveness of the model and identify areas for improvement.

1. **Feature Extraction:**

* Extract relevant features from the summarized text, such as keywords, sentiment analysis scores, and thematic representations, to capture the essence of each press release and inform the video generation process.

1. **Video Generation Model:**

* Develop and train a deep learning model based on suitable architectures for generating video content from textual inputs. Design the model to incorporate both textual and visual modalities, leveraging techniques.

***CHAPTER 4***

**HARDWARE AND SOFTWARE REQUIRMENTS**

* 1. **Hardware Requirements:**
* Processor: 11th Gen Intel(R) Core (TM) i3-1115G4
* Processor speed: 3.0GHz or above.
* RAM: 4GB or above 4. Storage space: 8GB or above
  1. **Software Requirements:**
* Operating System: Windows, macOS, or Linux distribution (Ubuntu, CentOS, etc.) that supports the required software.
* Python
* JavaScript
* NLP Libraries

Some used Libraries are:

* **Flask:** Web framework used for building the application.
* **requests:** For making HTTP requests to download images.
* **nltk:** Natural Language Toolkit for text processing tasks like tokenization and stop words.
* **gtts:** Google Text-to-Speech for converting text summaries into audio.
* **moviepy:** Python module for video editing and manipulation.
* **pyunsplash**: Python wrapper for the Unsplash API to fetch images.
* **summa:** Used for automatic text summarization.
* **PIL:** Python Imaging Library for image resizing.
* **fitz:** PyMuPDF for extracting text from PDFs.

***CHAPTER 5***

# RESULTS

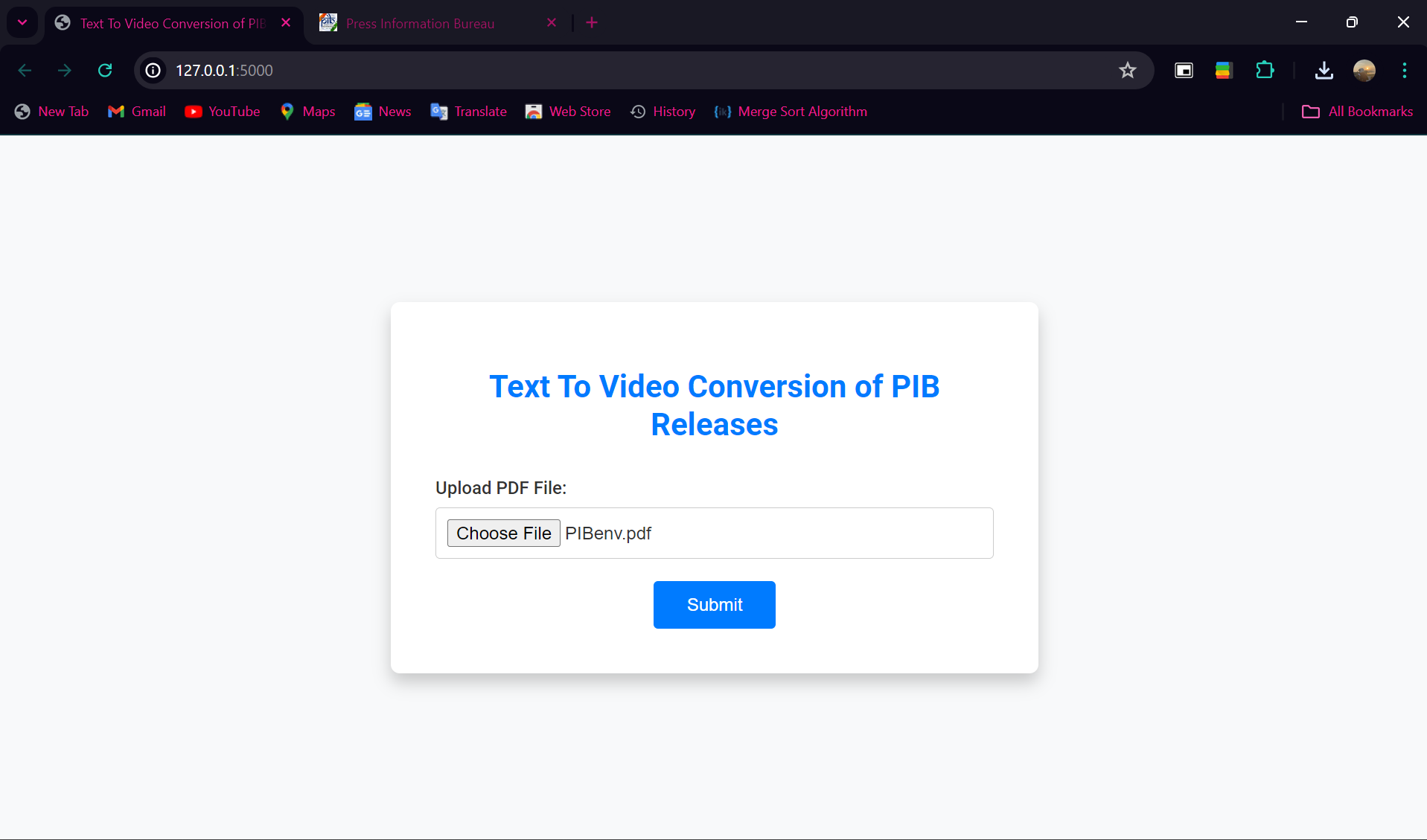


Figure 5.1 Home Page

* The above figure 5.1 is the home page for this project.
* In this page, the title of the project, "Text to Video Conversion of PIB Releases," is prominently displayed at the top of the page, clearly indicating the purpose of the application.
* The page is designed for converting text documents from Press Information Bureau (PIB) releases into video format.
* The central part of the page is dedicated to the main functionality, which is the text-to-video conversion. This area is kept simple to ensure ease of use.
* Below the title, there is a file upload section with a button labeled "Choose File." This allows users to select a PDF file from their local device.
* Next to the file upload section is a blue "Submit" button. Once a file is selected, the user can click this button to upload the PDF for processing.
* The interface is simple and user-friendly, with a dark blue header and footer and a white main content area.

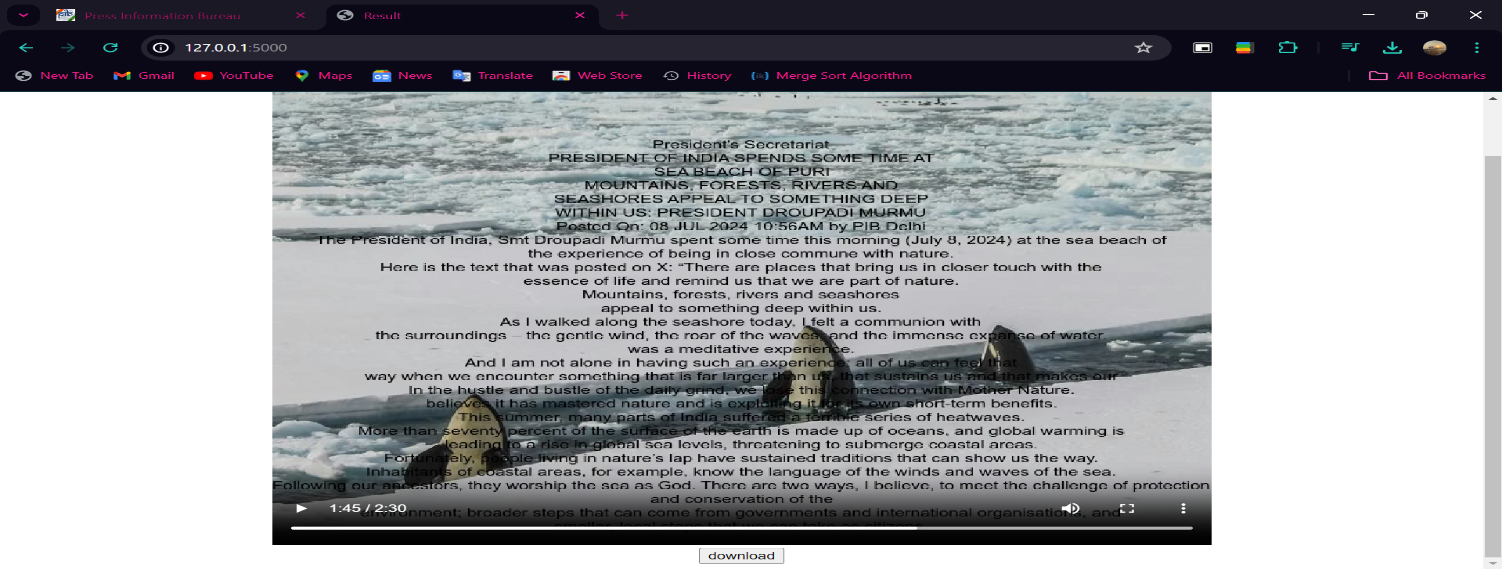
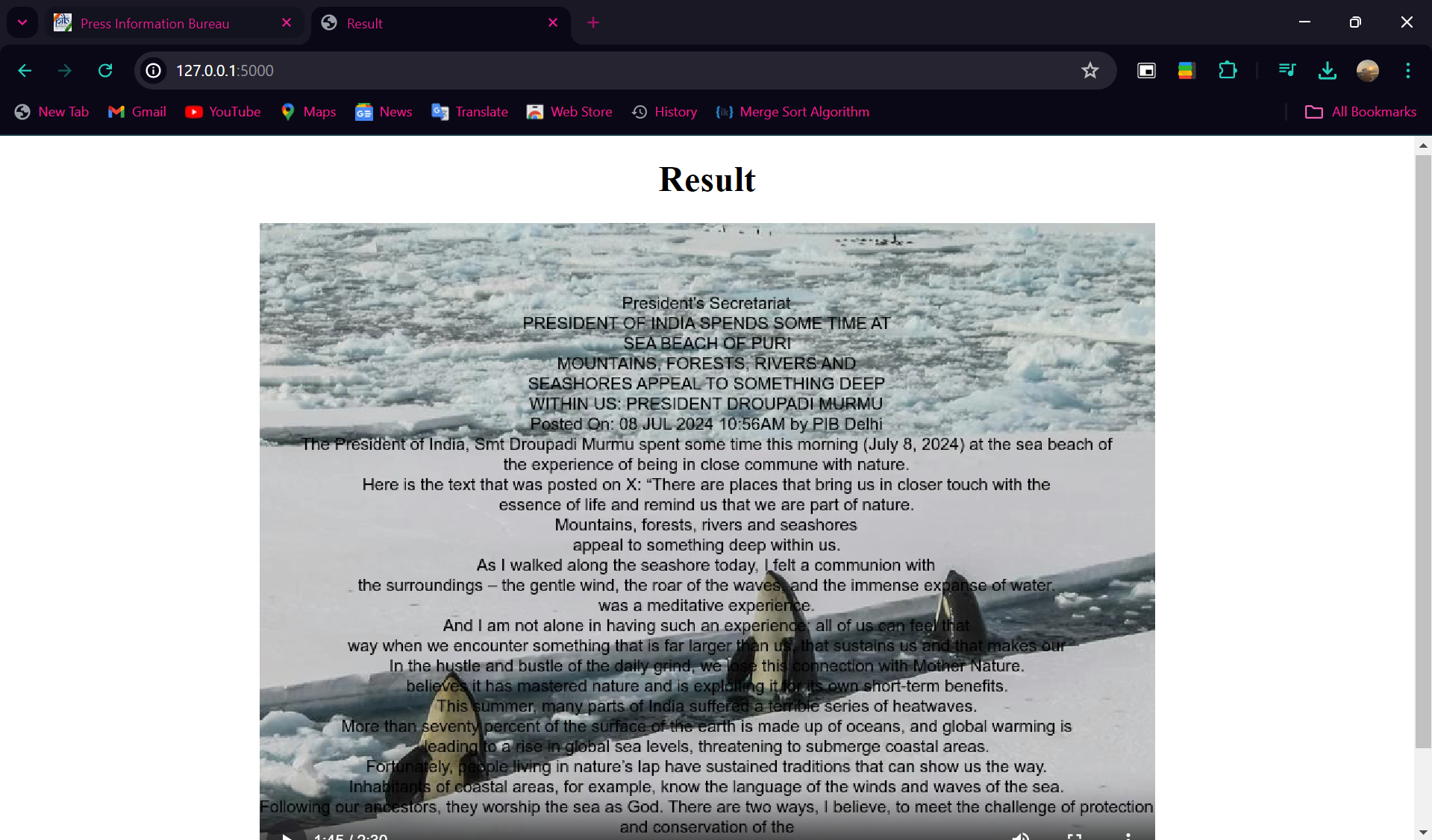


Figure 5.2 Output Page

* The above figure 5.2 is the output page for this project displays the video generated from the uploaded PDF.
* This page offers a seamless and user-friendly experience for viewing and downloading the video containing the text description from the original press release.
* The page displays a “Result” window, indicating the completion of the video generation process.
* In this, the news text is displayed on the screen, providing viewers with the written content of the press release.
* An audio narration accompanies the text, enhancing accessibility and engagement. This narration is generated using text-to-speech technology, ensuring clear and accurate delivery of the content.
* Relevant images are displayed alongside the text and audio. These images are automatically selected from unsplash to match the content of the news, providing a visual context that complements the narration.

***CHAPTER 6***

**CONCLUSION**

The Text to Video Conversion of Government Press Releases project has successfully achieved its goal of modernizing and streamlining how government agencies communicate with the public. By developing an automated system that transforms text-based press releases from PDF documents into engaging video content, the project has made a significant impact on public engagement and communication efficiency. At the heart of this success is the integration of various technologies, including PyMuPDF for text extraction, summa for summarization, nltk and spacy for keyword extraction, pyunsplash for image fetching, PIL for image processing, gtts for audio narration, and moviepy for video creation.

This sophisticated approach has not only automated what was previously a manual and time-consuming process but also improved the accessibility of government communications. The system’s ability to produce high-quality videos from press releases offers a more engaging and visually appealing format compared to traditional text documents. This shift from static text to dynamic video content meets the growing demand for multimedia communication methods and helps government agencies reach a wider audience more effectively.

By providing a user-friendly interface, the project has ensured that even those with limited technical expertise can easily use the system to convert press releases into videos, thus enhancing the overall efficiency of government communication efforts.

The use of open-source libraries and technologies has made the project resource-efficient, demonstrating that high-quality solutions can be developed without significant financial investment. The broader impact of this project lies in its ability to increase public awareness and engagement with government initiatives, making important information more accessible and engaging for citizens. Looking to the future, there is ample opportunity to further refine the system based on user feedback, expand its capabilities, and explore new features that could enhance its effectiveness. As government communication continues to evolve, this project sets a foundation for future innovations in how public information is presented and shared.

# REFERENCES

1. <https://pib.gov.in/PressReleseDetail.aspx?PRID=1480908>
2. <https://en.wikipedia.org/wiki/News>
3. <https://www.hypotenuse.ai/tools/press-release-generator>
4. <https://www.nltk.org>
5. https://www.geeksforgeeks.org/convert-text-speech-python