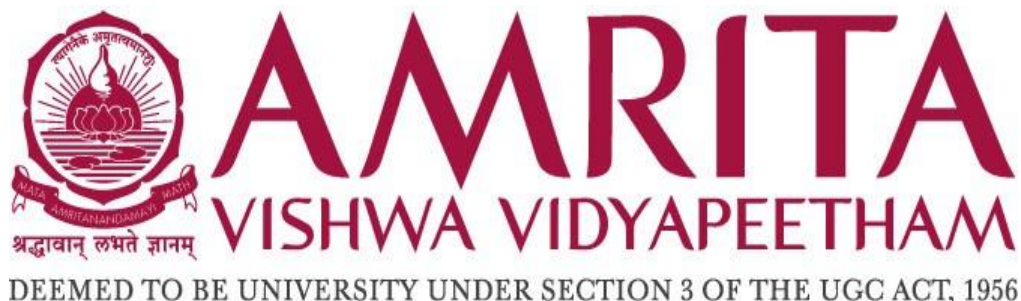


*As a part of the subject*

## Introduction To Cloud Computing 22AIE305



A project report submitted in partial fulfillment of the completion of the course 22AIE305 Introduction To Cloud Computing on **“CLIP CRUNCH – YouTube Video Summarizer using AWS”**

Centre for Computational Engineering and Networking  
**AMRITA SCHOOL OF ARTIFICIAL INTELLIGENCE**  
**AMRITA VISHWA VIDYAPEETHAM**

COIMBATORE - 641112

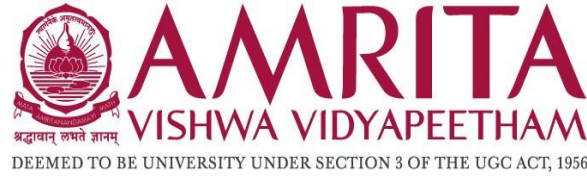
**NOVEMBER - 2024**

### **GROUP – 4**

#### **GROUP MEMBERS:**

S.NO	NAME	ROLL NO
1	ARIVANANTHAN M	CB.EN.U4AIE22007
2	SARAN DHARSHAN SP	CB.EN.U4AIE22045
3	A KATHIR	CB.EN.U4AIE22001
4	AIYYAPPAN MR	CB.EN.U4AIE22002

**AMRITA SCHOOL OF ARTIFICIAL INTELLIGENCE**  
**AMRITA VISHWA VIDYAPEETHAM**  
**COIMBATORE - 641 112**



**BONAFIDE CERTIFICATE**

This is to certify that the project entitled “Clip Crunch” submitted by Arivananthan M, Saran Dharsan, Kathir and Aiyyappan to Amrita Vishwa Vidyapeetham, Coimbatore in partial fulfillment for the award of the Degree of Bachelor of Technology in the “CSE(AI) ” is a bonafide record of the work carried out by him under our supervision at Amrita School of Artificial Intelligence, Coimbatore.

Dr. K.P.Soman  
Dean of AI

Mrs Prajisha C  
Project Incharge

**AMRITA SCHOOL OF ARTIFICIAL INTELLIGENCE**  
**AMRITA VISHWA VIDYAPEETHAM**  
**COIMBATORE - 641 112**

**DECLARATION**

We, Arivananthan M, Saran Dharsan, Kathir, and Aiyyappan, hereby declare that this project report titled " Clip Crunch " is a record of the original work carried out by us under the guidance of Mrs Prajisha, Assistant Professor, Centre for Computational Engineering and Networking, Amrita School of Artificial Intelligence, Coimbatore. To the best of our knowledge, this work has not formed the basis for the award of any degree, diploma, associateship, fellowship, or a similar award to any candidate in any University. In keeping with the ethical practice in reporting scientific information, due acknowledgments have been made wherever the findings of others have been cited.

Arivananthan M

Saran Dharsan SP

A Kathir

Aiyyappan MR

Place : Coimbatore

Date : 15-11-24

Signature of the students

## Table of Contents

Abstract.....	5
Introduction: .....	6
Related Works: .....	7
System Architecture .....	7
A. Frontend.....	7
B. Backend .....	8
Methodology.....	8
A. Website and User Interaction .....	8
B. Flask API and Backend Processing.....	8
C. YouTube Audio Extraction .....	8
D. Audio Storage in AWS S3.....	9
E. Transcription with AWS Transcribe.....	9
F. Summarization Using AWS Bedrock.....	9
G. Hosting on Azure App Service.....	9
Results and Discussion.....	10
Challenges and Limitations.....	11
Conclusion.....	11
Future Work.....	12
References .....	12

## Table of Figures

Figure 1:Overall Workflow .....	9
Figure 2: User Friendly Website .....	10
Figure 3:Logs coming in the source code .....	10
Figure 4: Summarization for the URL with video preview .....	11

## **Abstract**

In today's digital age, video platforms like YouTube host a vast array of educational, informational, and entertainment content. However, the sheer volume of available video material can make it challenging for users to consume information efficiently. Bardx.ai, an AI-driven web application, addresses this issue by providing concise, text-based summaries of YouTube videos. The platform leverages Amazon Web Services (AWS) to transform video content into succinct summaries, enabling users to gain key insights without investing significant time. When a user submits a YouTube link, Bardx.ai extracts the video's audio and processes it through a series of AWS services: audio is first converted to text using AWS Transcribe, and then AWS Bedrock generates a condensed summary from the transcription. This streamlined approach allows users to quickly capture essential content from lengthy videos.

Bardx.ai integrates a frontend interface built with HTML, CSS, and JavaScript, which facilitates a seamless user experience. The backend, developed with Python and Flask, handles all core operations and interactions with AWS, ensuring efficient processing and scalable performance. The application also incorporates several functionalities, such as URL validation, asynchronous processing for faster response times, and clear feedback to the user. By integrating advanced natural language processing (NLP) capabilities, Bardx.ai delivers summaries that retain the most relevant information, making it ideal for students, professionals, and casual users seeking fast content consumption.

In addition to detailing the application's architecture and key methodologies, this paper discusses the platform's potential to enhance content accessibility and usability. Key challenges, such as transcription accuracy and processing speed, are evaluated, along with potential areas for future development, including multilingual support and enhanced transcription quality. Bardx.ai represents a step forward in media consumption technology, demonstrating the power of AI in summarizing and personalizing content experiences on a large scale.

## **Introduction:**

The vast growth of video content on platforms like YouTube has created both opportunities and challenges for information seekers. While videos are an excellent medium for education, research, and entertainment, the time commitment required to watch lengthy content can be a significant limitation. In fields such as academia, business, and personal development, individuals often seek ways to quickly capture key insights from videos without watching the entire content. Recognizing this need, Bardx.ai provides a solution by delivering AI-powered text summaries of YouTube videos, thereby transforming video consumption into a more efficient and accessible experience.

With the rise of artificial intelligence and cloud computing, content summarization has become a viable tool for processing and understanding vast amounts of unstructured data, such as audio and video files. This application leverages AWS's suite of services to automatically extract, transcribe, and summarize YouTube video content, making complex information more digestible and time-efficient. Users can simply input a YouTube URL into Bardx.ai's interface, and within moments, they receive a summary that captures the primary points of the video. This feature is particularly beneficial for students, professionals, and researchers who need to quickly understand video content for decision-making or learning.

Bardx.ai's system architecture combines front-end simplicity with a robust back-end process using Flask and Python, interfacing with AWS services to automate and scale the summarization process. The backend framework uses AWS S3 for audio storage, AWS Transcribe for audio-to-text conversion, and AWS Bedrock for natural language summarization, ensuring that each video is processed accurately and efficiently. These technologies are seamlessly integrated to provide an intuitive user experience and deliver high-quality summaries that retain the essence of the original content.

The application of Bardx.ai extends beyond convenience; it has implications for education, content accessibility, and knowledge management in sectors where efficient content consumption is critical. This paper explores the methodologies, implementation, and effectiveness of Bardx.ai, and discusses potential improvements and future applications. By harnessing advanced AI and cloud services, Bardx.ai exemplifies the transformative potential of technology in optimizing digital media consumption.

## **Related Works:**

The demand for video summarization tools has grown with the increase in video content on platforms like YouTube. Numerous methodologies for automatic summarization, particularly for audio-visual content, have been proposed to assist users in obtaining concise, informative content.

Early approaches to video summarization typically relied on visual feature extraction and keyframe selection algorithms to highlight main scenes and events. However, these methods often missed important spoken information, limiting their effectiveness in content-focused summaries. With the rise of natural language processing (NLP) and machine learning, more advanced summarization techniques have been developed. Automatic speech recognition (ASR) systems, such as Google's Speech-to-Text API and AWS Transcribe, have become popular for transcribing audio content in video-based summaries. These transcriptions allow the application of NLP models for text summarization.

Several studies and projects have explored combining ASR with deep learning models to enhance summarization accuracy. For instance, BERT-based and transformer models have improved extractive and abstractive text summarization by enabling contextually aware summaries. Research has also highlighted the potential of using cloud-based platforms, such as AWS and Azure, for scalable ASR and NLP services, making them suitable for real-time applications.

Projects like Bardx.ai leverage these advancements by integrating ASR with cloud-based NLP solutions, such as AWS Bedrock, to create accessible, AI-powered video summarization tools. By storing and processing data on cloud platforms (AWS for storage, transcription, and summarization; Azure for hosting), Bardx.ai exemplifies how cloud infrastructure and NLP can combine to address modern video summarization needs efficiently.

## **System Architecture**

Bardx.ai's architecture includes a frontend user interface for user input and a backend for processing requests. The backend utilizes AWS for key operations such as audio storage, transcription, and summarization, ensuring scalability and efficiency.

### **A. Frontend**

The frontend, designed with HTML, CSS, and JavaScript, provides a simple user interface:

- YouTube URL Input: Users paste a YouTube URL to receive a video summary.
- Video Preview: Embedded video player displays a preview of the selected YouTube video.
- Summary Display: Once processed, the summary is displayed alongside the video preview for easy reading.

## **B. Backend**

The backend is built using Python and Flask to manage HTTP requests and interactions with AWS services. Flask handles the routing of requests, triggering the audio extraction, storage, transcription, and summarization processes.

## **Methodology**

Bardx.ai employs a multi-step approach to generate concise summaries of YouTube videos by integrating website development, cloud services, and AI models.

### **A. Website and User Interaction**

The website, built with HTML, CSS, and JavaScript, provides users a straightforward interface to input a YouTube URL and receive a summary. When a URL is submitted, the page sends the data to a Flask backend for processing. The site design also includes an embedded YouTube video preview and displays the generated summary alongside the video.

### **B. Flask API and Backend Processing**

A Flask server hosts the backend logic, receiving requests from the website. Flask validates the YouTube URL and orchestrates the processing workflow by sequentially calling functions for audio extraction, storage, transcription, and summarization. This backend also manages data transfer between the website and AWS services.

### **C. YouTube Audio Extraction**

Using `yt_dlp`, the backend extracts the audio from the submitted YouTube video in MP3 format. This audio file is stored temporarily before being uploaded to cloud storage.



### D. Audio Storage in AWS S3

The extracted audio is uploaded to an AWS S3 bucket via boto3, providing scalable and reliable storage. The S3 URL generated for the file facilitates access by other AWS services.

### E. Transcription with AWS Transcribe

AWS Transcribe processes the audio file to convert speech into text. The transcription job is initiated with the S3 URL and monitored for completion. Once the transcription finishes, the generated text is saved for further summarization.

### F. Summarization Using AWS Bedrock

The transcribed text is sent to AWS Bedrock, which summarizes it into a concise format. A prompt is designed to guide Bedrock's language model, ensuring the output is relevant and readable.

### G. Hosting on Azure App Service

The entire application, including the Flask backend and website, is hosted on Azure App Service, providing reliable and scalable web hosting. Azure App Service supports Flask natively, enabling Bardx.ai to handle user requests efficiently.

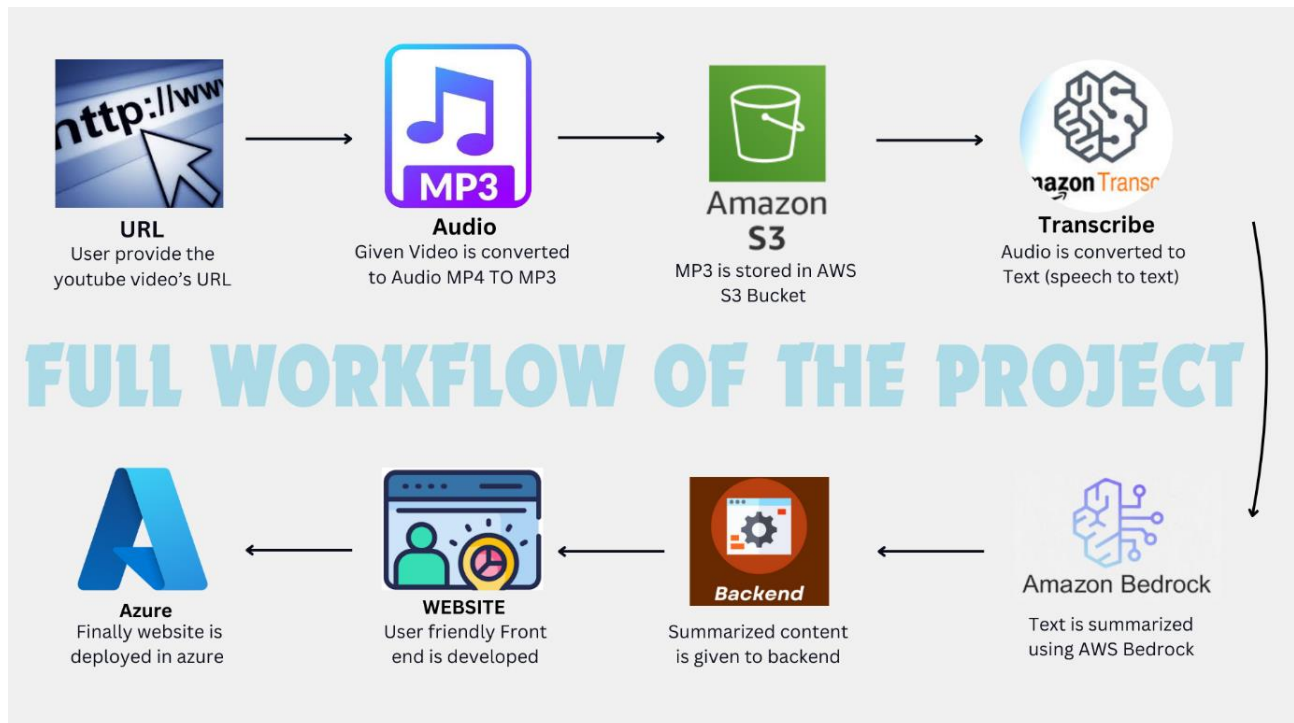


Figure 1: Overall Workflow

## Results and Discussion

The Bardx.ai platform demonstrates significant time-saving and efficiency improvements for users who need information from YouTube videos quickly. By using AWS services, the application ensures that audio transcription and summarization are scalable and effective. AWS Transcribe delivers accurate transcriptions, while AWS Bedrock produces meaningful, concise summaries. Initial testing on various types of YouTube videos, including educational content and news clips, shows that the summaries capture core details effectively, allowing users to bypass full video playback.

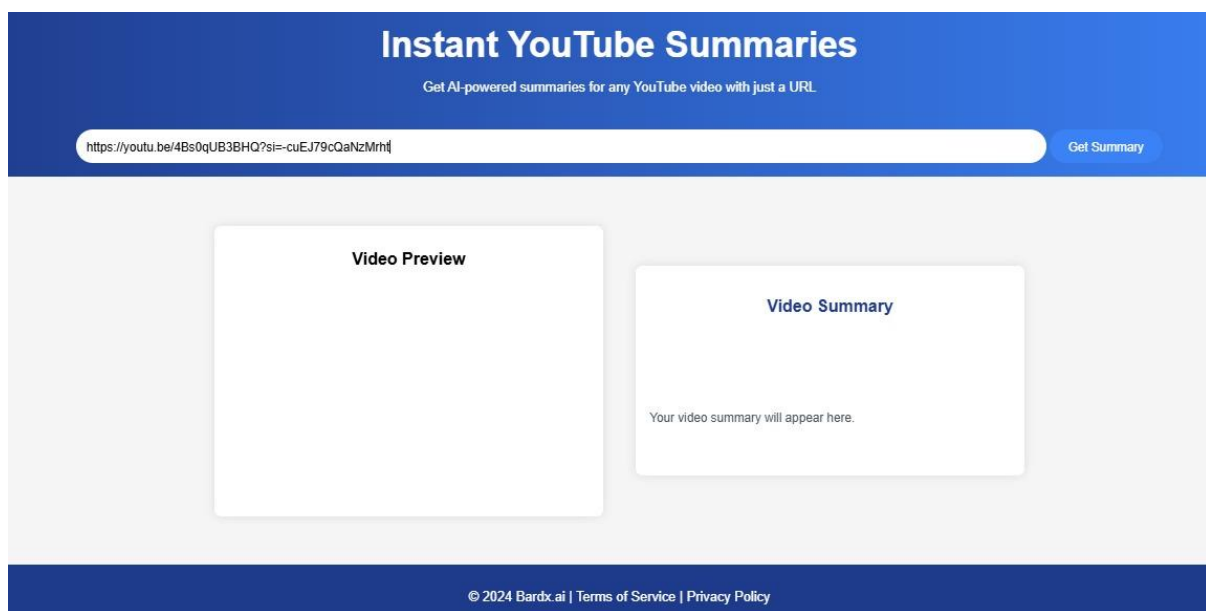


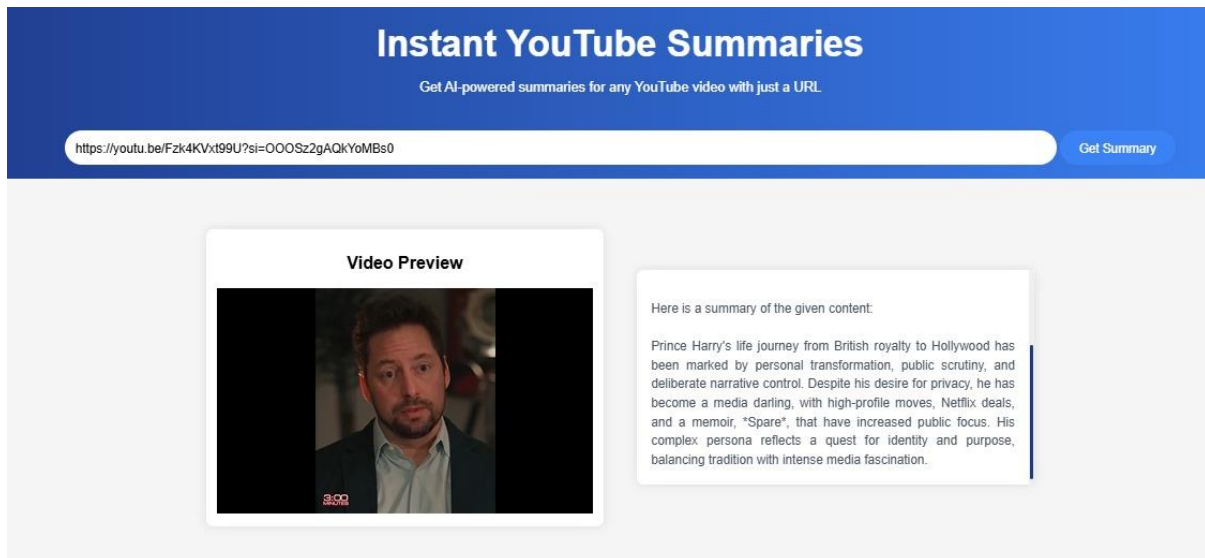
Figure 2: User Friendly Website

```
PROBLEMS 11 OUTPUT DEBUG CONSOLE TERMINAL PORTS JUPYTER
* Debugger: PID: 133-979-302
127.0.0.1 - - [17/Nov/2024 23:46:19] "GET / HTTP/1.1" 200 -
127.0.0.1 - - [17/Nov/2024 23:46:19] "GET /logo2.png HTTP/1.1" 404 -
127.0.0.1 - - [17/Nov/2024 23:46:20] "GET /static/script.js HTTP/1.1" 304 -
127.0.0.1 - - [17/Nov/2024 23:46:20] "GET /static/styles.css HTTP/1.1" 304 -
127.0.0.1 - - [17/Nov/2024 23:46:20] "GET / HTTP/1.1" 200 -
127.0.0.1 - - [17/Nov/2024 23:46:20] "GET /logo2.png HTTP/1.1" 404 -
127.0.0.1 - - [17/Nov/2024 23:46:20] "GET /static/styles.css HTTP/1.1" 304 -
127.0.0.1 - - [17/Nov/2024 23:46:20] "GET /static/script.js HTTP/1.1" 304 -

Video found: Deepak Chopra's Go-To 3-Minute Meditation To Stay Focused

Downloading audio to memory...
Audio download complete.
Uploading file to S3...
File uploaded to S3.
Transcription started...
Transcription in progress. Waiting...
Transcription in progress. Waiting...
Transcription in progress. Waiting...
Job SES4tFtyfghvbcvg is COMPLETED.
Transcription complete. Transcript saved.
Your Summary: Deepak Chopra's Go-To 3-Minute Meditation guides individuals toward focus and stress reduction through a structured process. Start with soft eyes, observing your surroundings without fixating, then shift to natural breath observation. Mentally repeat your name, progressively dropping parts of it to connect with your core self, ultimately reaching a state of pure awareness. Replace identity with sounds like aham or c
m to explore essence, and rest in mindful awareness. Conclude by observing bodily sensations, reconnecting to the present. This meditation promotes mindfulness, self-awareness, and relaxation in just a few minutes.
Your Summary: Here is a 3-4 line summary of the given content:
```

Figure 3: Logs coming in the source code



*Figure 4: Summarization for the URL with video preview*

## Challenges and Limitations

- **Latency:** Transcription and summarization add processing time, particularly for longer videos.
- **API Limits:** AWS service limits could constrain scalability if demand significantly increases.
- **Accuracy:** Transcription quality may vary based on audio clarity, impacting summary quality.
- **User Authentication and Security:** The project currently uses hardcoded AWS credentials, a potential security risk. Implementing a secure authentication system is essential for production deployment.

## Conclusion

Bardx.ai effectively summarizes YouTube video content by leveraging cloud-based AI services, making it a valuable tool for users seeking efficient content consumption. This project highlights the potential of integrating cloud services like AWS for video processing and summarization. Future work may focus on enhancing response times, improving transcription accuracy, and integrating secure user authentication.

## Future Work

- **Enhanced Summarization Techniques:** Experimenting with more advanced summarization models could improve output quality.
- **Multi-Language Support:** Expanding language options would broaden the application's usability.
- **Optimized Processing Pipeline:** Reducing latency in the transcription and summarization process could improve user experience.
- **Authentication and Security:** Implementing OAuth or AWS IAM roles for secure API access.

## References

1. M. Altschuler, et al., "Real-Time Summarization of Streaming Social Media," *IEEE Transactions on Multimedia*, vol. 20, no. 11, pp. 3026-3037, Nov. 2018.
2. Amazon Web Services, "AWS Transcribe - Developer Guide," [Online]. Available: <https://aws.amazon.com/transcribe>
3. "Amazon Bedrock Documentation," Amazon Web Services. [Online]. Available: <https://aws.amazon.com/bedrock/>
4. "AWS Security Best Practices," AWS Whitepapers & Guides, Amazon Web Services, 2022.
5. "AWS Transcribe Documentation," AWS. [Online]. Available: <https://docs.aws.amazon.com/transcribe/>. [Accessed: Nov. 14, 2024].
6. "Azure App Service Documentation," Microsoft Azure. [Online]. Available: <https://docs.microsoft.com/en-us/azure/app-service/>. [Accessed: Nov. 14, 2024].