|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **FORM 1**  **THE PATENTS ACT 1970 (39 OF 1970)**  **&**  **The Patents Rules, 2003 APPLICATION FOR GRANT OF PATENT**  **(See section 7,54 & 135 and rule 20 (1))** | | | **(FOR OFFICE USE ONLY)**  **Application No.:  Filing Date:**  **Amount of Fee Paid:  CBR No.: Signature:** | | | | | |
| **1. APPLICANT** | | | | | | | | |
| Name | | Nationality | | | | | Address | |
| Prof.Chanchla Tripathi  Dr.Lalit B. Damahe | | Indian | | | | | **Department of Computer Science Engineering, Y.C.C.E.,**  **Hingna Road, Wanadongri, Nagpur-441110** | |
|  | |  | | | | |  | |
|  | |  | | | | |  | |
| **2. INVENTOR ( S )** | | | | | | | | |
| Name | | Nationality | | | | | Address | |
| Hemanshu Waghmare Rishabh Jain  Dhruv Dalvi Sanket Asole Yuvraj Chavan | | Indian | | | | | **Department of Computer Science Engineering, Y.C.C.E., Hingna Road, Wanadongri, Nagpur-441110** | |
| **3. TITLE OF THE INVENTION**  **Speak2Summarize : Daily Recap** | | | | | | | | |
| **4. ADDRESS FOR CORRESPONDENCE OF APPLICANT**  **Prof.Chanchla Tripathi, Department of Computer Science Engineering, Y.C.C.E., Hingna Road, Wanadongri, Nagpur-441110**  **Dr. Lalit B. Damahe, HoD, Department of Computer Science Engineering, Y.C.C.E., Hingna Road, Wanadongri, Nagpur- 441110** | | | | | | Telephone No. :  E-mail : | | |
| **5. PRIORITY PARTICULARS OF THE APPLICATION (S) FILED IN CONVENTION COUNTRY** | | | | | | | | |
| Country | Application Number | Filing Date | | Name of the Applicant | | | | Title of the Invention |
| N.A. |  |  | |  | | | |  |
| **6. PARTICULARS FOR FILING PATENT COOPERATION TREATY (PCT) NATIONAL PHASE** | | | | | | | | |
| International application number | | | | | N.A. | | | |
| **9. DECLARATIONS :** | | | | | | | | |
| 1. **Declaration by the Inventor (s)**   We, the above named inventor(s) are the true & first inventor(s) for this invention and declare that the applicant(s) herein is our assignee.   * 1. Date   2. Signature (s) Signature **Signature**   3. Name(s) **Prof.Chanchla Tripathi**   **Dr. Lalit B. Damahe** | | | | | | | | |

|  |
| --- |
| 1. **Declaration by the applicant (s):**   **I/We, the applicant(s) hereby declare(s) that:**   * + We are in possession of the above – mentioned invention.   + The provisional/complete specification relating to the invention is filed with this application.   + The invention as disclosed in the specification uses the biological material from India and the necessary permission from the competent authority shall be submitted by me/us before the grant of patent to me/us.   + There is no lawful ground of objection to the grant of the Patent to me/us.   + I am/We are the assignee or legal representative of true & first inventors.   + The application or each of the applications, particulars of which are given in Para – 5 was first application in convention country/countries in respect of my/our invention.   + I/We claim the priority from the above mentioned application(s) filed in convention country/countries and state that no application for protection in respect of the invention had been made in a convention country before that date by me/us or by any person from which I/We derive the title.   + My/our application in India is based on international application under Patent Cooperation Treaty (PCT) as mentioned in Para – 6.   + The application is divided out of my/our application particulars of which are given in Para – 7 and pray that this application may be treated as deemed to have been filed on under sec. 16 of the Act.   + The said invention is an improvement in or modification of the invention particulars of which are given in Para – 8 |
| 1. **Following are the attachments with the application:**    1. Complete specification.    2. Complete specification (in conformation with the international application) / as amended before the International Preliminary Examination Authority (IPEA), as applicable (2 copies ).   No. of pages No. of claims 04   * 1. Drawings (in conformation with the international application) / as amended before the International Preliminary Examination Authority (IPEA), as applicable (2 copies). No. of Sheets   2. Priority documents   3. Translation of priority document / specification/ International Search Report   4. Statement and undertaking on Form 3   5. Power of Authority   6. Declaration of inventorship on Form 5   7. Sequence listing in electronic form (floppy disc)   8. ………………………………   Fee Rs. ………………………in Cash/Cheque/Bank Draft bearing no. ……………….. Date ………. On Bank.  I/We hereby declare that to the best of my/our knowledge, information and belief the fact and matters stated herein are correct and I/We request that a patent may be granted to me/us for the said invention.  Dated this day of September 2005.  Signature :-  Name :  To,  The Controller of Patent  The Patent Office, at Mumbai |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **FORM 3**  **THE PATENT ACT, 1970 ( 39 OF 1970 )**  **&**  **The Patents Rules, 2003**  **STATEMENT AND UNDERTAKING UNDER SECTION 8**  **( See section 8, rule 12 )** | | | | | |
| We **YESHWANTRAO CHAVAN COLLEGE OF ENGINEERING** hereby declare  (i) that /We who have made this application No. 00000000000 Dated 00000000000 alone, made for the same/substantially same invention, application(s) for patent in the other countries, the particulars of which are given below : | | | | | |
| Name of the country | Date of Application | Application No. | Status of the application | Date of publication | Date of grant |
| Not Applicable |  |  |  |  |  |
| (iii) That the rights in the application(s) have been assigned to us and that We undertake that up to the date of grant of the patent, by the Controller. We would keep him informed in writing the details regarding corresponding applications for patents filed outside India within three months from the date of filing of such application. | | | | | |
| Dated this day of September 2005  For **YESHWANTRAO CHAVAN COLLEGE OF ENGINEERING** | | | | | |
| To,  The Controller of Patent  **The Patent Office, Mumbai** | | | | | |

# F O R M 2

## THE PATENTS ACT, 1970 (39 of 1970)

COMPLETE SPECIFICATION

## (See section 10 and rule 13)

|  |
| --- |
| **1. TITLE OF THE INVENTION:**  **Speak2Summarize : Daily Recap** |
| 1. **APPLICANT(S)**    1. **NAME**   Hemanshu Waghmare  Rishabh Jain  Dhruv Dalvi Sanket Asole Yuvraj Chavan   * 1. **NATIONALITY** Indian   2. **ADDRESS** |
| **Prof. Chanchla Tripathi, Department of Computer Science Engineering, Y.C.C.E., Hingna Road, Wanadongri, Nagpur-441110**  **Dr. Lalit B. Damahe, HoD, Department of Computer Science Engineering, Y.C.C.E., Hingna Road, Wanadongri, Nagpur-441110**  **3. PREAMBLE TO THE DESCRIPTION**  The growing reliance on AI (artificial Intelligence) technology in the domain of Natural Language Processing (NLP) has brought forward noticeable improvement in both Automatic Speech Recognition (ASR) and text summarization. These technologies are fundamental in fostering inclusivity, providing better communication, and facilitating data management around the world. Speech-to-text transformation involves the conversion of voice into text form through sophisticated Automatic Speech Recognition (ASR) models hoping from older statistical Hidden Markov Models (HMM) to newer deep learning structural Recurrent Neural Networks (RNN) and Transformers. On the other hand, summarization of text employs abstractive procedures to transform the provided text into brief and useful summaries. The advancement of these branches has led to sophisticated uses of virtual assistants, meeting transcribers, and auto content generation systems. This review aims to critically analyze the effectiveness of all existing approaches and the methodologies that have been developed in STT and text summarization systems of a new primary focus. |

|  |
| --- |
| **Type of Application: COMPLETE**  The following specification particularly describes the invention and the manner in which it is to be performed. |
| **4. DESCRIPTION** (Description shall start from next page) |
| **5. CLAIMS** (written on separate page) |
| **6. DATE AND SIGNATURE** (given on the last page of specification) |
| **7. ABSTRACT OF THE INVENTION (given along with complete specification on the separate page)** |
| **Note:**  \*Repeat boxes in case of more than one entry  \*To be signed by the applicant(s)or the authorized registered patent agent  \*Name of the applicant should be given in full, family name in the beginning  \*Complete address of the applicant should be given stating with postal index no. / code, state and country  \*Strike out the column which is/are not applicable |

### DESCRIPTION OF THE PATENT

1. **Field of Invention**

This invention relates to the Machine Learning and Natural Language Processing fields concerning Speech-To-Text (STT) transformations and Text Summarization. It includes the recent techniques of Automatic Speech Recognition (ASR) in the context of deep learning architectural models and open source software . Also, it incorporates methods of summarization by use of these open source models and techniques aimed at enhancing data process automation, accessibility and content creation. The invention is useful for many sectors such as systems for transcription of protocols of meetings, automatic text summarization systems, and voice writing systems and daily voice based journalling .

1. **Background of the Invention**

The improvement of AI and NLP technologies has created a new demand for efficient speech-to-text transformation and text summarization. Earlier speech recognition systems used statistical methods like Hidden Markov Models (HMM) and Gaussian Mixture Models (GMM) because of their effectiveness. These systems, however, failed in dealing with accent variations and noisy background systems. The accuracy and efficiency of speech to text systems with deep learning methods has advanced with the introduction of various forms of RNNs, LSTMs and transformer architectures such as BERT and GPT. These systems now recognize speech with greater speed and precision in multiple languages and dialects.

At the same time, large volumes of spontaneously generated digital content requires advanced text summarization tools that can efficiently extract relevant information. The previously existing rule-based and statistical methods for extracting summaries of a text have slowly been transitioned to incorporate machine learning and transformer techniques which are far more capable of producing accurate and concise summaries.

The enhancement of speech to text tools through the combination of speech recognition and automated summarization has benefits for virtual assistants, meeting transcription software, podasts, content analysis, and speech accessibility applications. Multi-speaker environments, background noise, context understanding, and handling specific vocabularies still pose a challenge.

This invention seeks to address these issues through the application of advanced deep learning and NLP models aimed at improving the power and precision of speech recognition, text comprehension, and, ultimately, human-machine communication, information retrieval, and content consumption

### OBJECT OF INVENTION / DISTINGUISHING FEATURES OF THE INVENTION WITH PRIOR ART:

* + To make speech-to-text and summarization tools more accessible to **visually impaired users, professionals, students, and organizations** requiring fast and reliable transcription and summarization services.
  + To enable individuals with memory related problems to **record, transcribe, and summarize conversations**, providing a textual reference for later recall.

### STATEMENT OF INVENTION

The present invention relates to a S**peech-to-Text (STT) and Text Summarization system** designed to convert spoken language into as **accurate as possible textual representation** and generate summary of that textual representation which is **concise and context-aware**. Utilizing **Natural Language Processing (NLP) and Open Source Machine Learning and Deep Learning models,** the invention aims to utilize **speech recognition accuracy, memory efficiency, and contextual understanding of these models to produce meaningful summaries**.

A key aspect of this invention is its ability to support individuals experiencing **memory-related challenges and the individuals working in any working or learning space where conversations are frequent**, by providing **real-time transcription, conversation recall, and automated summarization**. Through **context-aware information retrieval and transcripts** the system allows users to retain, review, and recall crucial information with minimal cognitive effort.

The invention integrates **hardware components** to enable efficient speech processing and real-time summarization, including:

* **MAX9814 Electret Microphone Module** – Records high-quality audio input with **automatic gain control (AGC)** to enhance speech clarity.
* **Arduino Nano (ATmega328P) Microcontroller** – Handles **audio signal pre-processing, real-time data acquisition, and integration with external modules**.
* **DS1302 Real-Time Clock (RTC) Module** – Ensures **accurate timestamping** of recorded speech data, aiding in recall and organization.
* **Micro SD Card Reader & Micro SD Storage (8GB - 32GB, FAT32)** – Stores transcriptions and summaries locally for offline access.
* **USB Power Bank & Power Management System** – Provides **portable, uninterrupted power supply** for real-time speech recognition and processing.
* **Tactile Push Buttons & LED Indicators** – Enables **manual control for starting/stopping recordings** and provides **visual feedback on system status**.

**A brief description of drawing is as follows:**

### 1. Setup Phase

* + RTC is initialized. If invalid, it's set to the current time from the computer (\_\_DATE\_\_ & \_\_TIME\_\_ macros).
  + SD Card is initialized.
  + Buttons and LEDs are configured.

### 2. Recording Phase

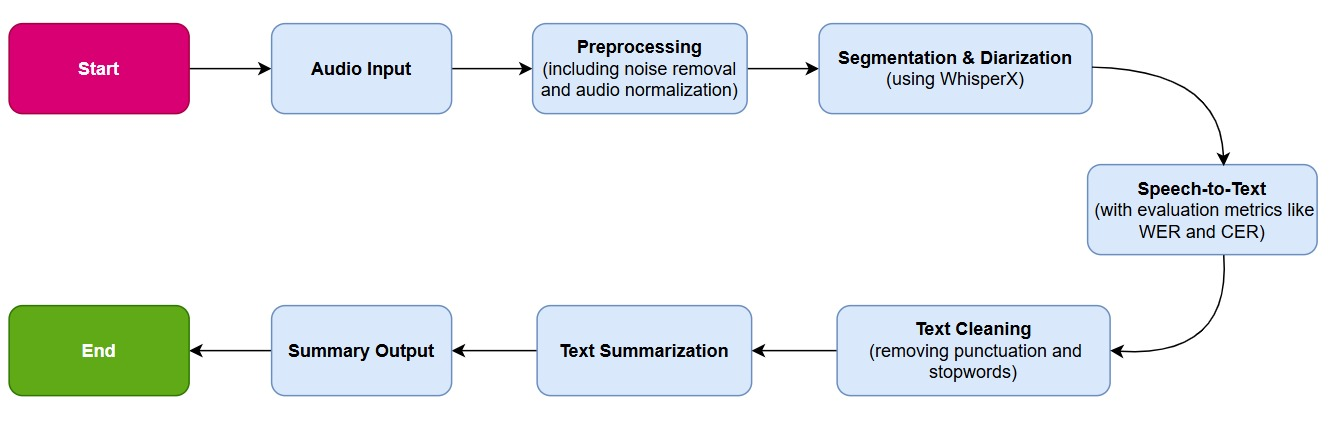
* + When the **Start button** is pressed:
    - The RTC fetches the **current date and time**.
    - A unique .mp3 filename is generated using the format:  
      **YYYYMMDD\_HHMMSS.mp3**
    - The SD card creates a new file with that name.
    - **LED1 turns ON**, indicating recording has started.
    - (You would need to interface the MAX9814 with a proper ADC or external recorder for actual audio.)

### 3. Stop Phase

* + When the **Stop button** is pressed:
    - The file is closed to ensure it's saved correctly.
    - **LED1 turns OFF**, and **LED2 turns ON** briefly to indicate successful saving of the audio

A diagram of a computer

AI-generated content may be incorrect.  
Fig 1. Circuit diagram of the audio recording apparatus

Fig 2. Flow Diagram of Data Processing Pipeling of Audio data into Summarised text

### DETAILED DESCRIPTION OF THE INVENTION

After observing these problems faced by individuals with memory and the individuals like professials and students who has to consumer large amount of audio based data everyday, we decided to develop a system for effective processing of the audio data and generate text summaries so that the individuals and easily get gist of the happenings of the day. The Arduino serves as the central processing unit of the recording system. It receives the audio signals from the **MAX9814 Electret Microphone Module**. The Arduino processes this audio signal and saves the audio file along with the timestamp generated by the RTC into the Micro SD Card which is accessed thorough Micro SD Card Reader module connected to the **Arduino Nano (ATmega328P)**. The user then put this Micro SD Card into their system and the software we have developed, using Open Source Maching learning and NLP model processes the audio present in the Micro SD Card. The Micro SD Card is FAT32 formatted which give it abilty to work with any operating system.

The software as mentioned consists of three main models which are Speaker Diarization, Speech Recognition Model and Text Summarization Model. These models works sequentially in the aforementioned order. The Speaker Diarization Model Breaks the audio into the segments if there are more than one speaker present in audio data (this is possible in sceanario such as meetings and interviews); the process is done in memory. The Speech Recognition Model then transcribe this segmented data in the sequential format based on timestamping. The transcribed data which is currently in text format is stored for the user to read in detail, along with this the text data is also fed into the Text Summarization Model to generate a concise and meaningful summaries while maitaining the context awareness. This summarised data is then presented to user with the aim of helping them remember their daily conversations to improve their productivity.

## WE CLAIM

### ****1. Independent Claims****

1. **A Audio Recording and Storage System, comprising:**

### **A microphone module (MAX9814) for capturing speech input with automatic gain control (AGC);**

### **A microcontroller (Arduino Nano) for processing audio signals and managing data acquisition;**

### **A real-time clock (DS1302 RTC) for timestamping speech recordings;**

### **A storage module (Micro SD card) for storing transcriptions and summaries;**

1. **A method for converting speech to text and generating text summaries, comprising:**

### **Capturing speech input via a microphone module;**

### **Preprocessing audio using noise reduction and segmentation techniques;**

### **Converting speech to text using an ASR model based on Transformers;**

### **Applying text summarization models to generate summaries;**

### ****2. Dependent Claims****

### **The system of claim 1, wherein the system processes audio signals to digital audio and stores the audio file in Micro SD Card**

### **The system of claim 1, wherein users can manually start and stop recording along with storing the timestamp of when the recording has started.**

### **The system of claim 2, wherein the summarization process includes abstractive techniques, utilizing pre-trained models fine-tuned for speech-based text summarization.**

### **The system of claim 2, wherein transcriptions are automatically segmented and summarized based on number of speakers.**

### **The system of claim 2, wherein the audio data is processed to a .wav file format so that pre-processing can be performed efficiently.**

### **The system of claim 2, wherein the speech recognition system filters background noise and enhances speech clarity using noise reduction techniques.**

### ABSTRACT OF THE INVENTION

The present invention relates to an **Speech-to-Text (STT) and Text Summarization System** that integrates **hardware and software** to efficiently convert speech into accurate text and generate concise, context-aware summaries. The system is designed to assist individuals to get a record of their daily summaries as well as individuals with **memory-related challenges as** by providing **transcription, conversation recall, and automated summarization**. The invention comprises a **hardware module** that includes a **MAX9814 electret microphone for speech input**, an **Arduino Nano (ATmega328P) for audio signal processing**, a **DS1302 Real-Time Clock (RTC) for timestamping**, and an **SD card storage unit for saving transcriptions and summaries**. The **power management system** ensures seamless operation, making it suitable for both **portable and embedded applications**. On the software side, the system utilizes **Natural Language Processing (NLP) and Deep Learning models**, such as **Transformer-based Automatic Speech Recognition (ASR) models (WhisperX, Wav2Vec 2.0)** for **highly accurate speech-to-text conversion**, and a**bstractive summarization techniques** to generate meaningful summaries. The system is currently adaptable to Englishmaking it suitable for applications in **assistive technologies, virtual assistants, business transcription services, and education**. By combining **AI-driven text summarization, and memory-assistive features**, this invention provides a **scalable, accessible, and efficient solution for speech recognition and content retrieval**, enhancing accessibility and information retention for a wide range of users. The system aims to help with memory retention by keeping a log of summarised day to day interactions of an individual.