

USING SPEECH RECOGNITION TO CREATE MEDICAL REPORTS

MADALINA-ELENA DINEA(COCEAN)^{1,2}, ANDREEA GRIPARIS², INGE GAVAT²

Abstract. Doctors spend about half of their time doing the paperwork and writing medical reports. The use of speech recognition in the creation of medical transcription applications aims to ease the work of medical professionals, freeing them from writing reports. Moreover, these reports contain a larger volume of information, with more details included, the information being added while the doctor examines the patient in the case of a medical consultation, respectively while analyzing the images obtained in the case of a medical investigation. Medical transcription solutions reduce the time needed to complete the reports related to a medical examination or investigation, giving the medical staff more time to focus on the patient.

Keywords: transcription, speech recognition, medical report.

1. INTRODUCTION

Speech technology has several applications in the health domain [1] that use speech recognition, synthesis, or processing, such as fixation of speech and hearing impairments, speech interfaces for doctors and patients, speech evaluation for psychological disorders, home care, or speech biomarkers [2] used to detect diseases.

This article aims for one application from the speech recognition branch that has a wide range of applications, among them the transformation of spoken phrases into written text, the applications that have this purpose being part of the "Speech to Text" category.

One application of speech recognition that is used in the medical field is known as medical transcription. It aims to create medical reports from audio recordings or in real-time, transforming the doctor's observations during the consultation and/or the doctor-patient conversation into written text.

¹ "Politehnica" University of Bucharest, Romania

² Luxoft Romania

Some of the important advantages of these applications are:

- Easier handling of a large volume of information – especially in the field of digital imaging where large reports with many specialized terms are created after analyzing investigations such as MRI, CT, mammograms and X-rays.

- Saving time: the doctor does not have to write the medical report during or at the end of the consultation.

- Enhanced communication between the doctor and the patient: the medical staff can focus more on the consultation and diagnosis, without the need to write the medical report.

- Detailed reports: handwritten reports include notes/summaries related to the patient's diagnosis and symptoms, while reports created with speech recognition applications can contain more details, as the report is created while the doctor consults with the patient.

A comparison of the performance in noisy environments of three of the most well-known and performing speech recognition systems was made: Google Cloud, Mozilla Deep Speech and IBM Watson. The purpose of this study [3] was to determine how these systems would perform in environments such as the doctor's office or emergency room, situations that will result in recordings where the useful signal is mixed with noise. This evaluation concluded that the chosen recognition systems need to be improved to be largely adopted in the medical field.

The results of the study [4] indicate that clinicians are interested in using speech recognition software for medical documentation because, even in the case of transcription errors, when corrections are needed to be done manually, the time spent on writing a report is shortened.

Currently, many medical transcription applications are available for the English language, but none are specifically designed for the Romanian language. Among the applications available on the market are: Nuance Dragon Medical [5], VTEX Voice Solutions SayIt [6], PrognoCIS [7], VoiceboxMD [8], WebChartMD [9], Amazon Transcribe Medical [10] and Augmedix [11].

2. MEDICAL TRANSCRIPTION SOLUTION

The solution proposed for using speech recognition applied in the medical field that will be described in this paper is a medical transcription application for the Romanian language. This solution targets doctors and it is a web application that has the following main functionalities:

- Creation of medical reports for investigations/consultations
- Viewing the report created
- Sharing the report.

The medical report contains the following sections:

- Information about the patient: name, surname, age, CNP, phone number, email address, known allergies, and medical history.
- Information about the doctor: name, surname, initial code, field, medical unit.
- Consultation/investigation information
- Date and time of report creation.

The doctor can select the type of report he wants to create. Thus, he will be able to create reports for:

- Medical investigations (useful in the field of medical imaging)
- Medical consultations (useful for consultations conducted by specialists in the polyclinic/emergency room).

The information contained in the medical report for an investigation is:

- Type of investigation conducted
- Observations of the doctor based on the images obtained during the procedure
- Conclusions of the doctor
- Recommendations of the doctor.

The information contained in the medical report for a medical consultation:

- Reason for the patient's presentation
- Observations of the doctor during the consultation
- Diagnosis
- Recommendations of the doctor, which will include the prescribed treatment and referrals to other specialist doctors if needed.

The date and time of the consultation are filled in automatically when the medical report is saved. A significant part of the fields of the medical report form could be populated in two ways:

- Written – by typing the information associated with the current section.
- Verbal – by pressing the record button and dictating the information, with the possibility to edit the transcription result from the interface.

The medical report will be saved after verifying the entered data by pressing a button. In order to protect the patient in case the doctor prescribes an erroneous treatment or makes a wrong diagnosis, a report that has been saved cannot be edited. After saving, the complete report will appear in web format with the possibility of being converted to PDF format for sending by email or for printing on physical media, thus realizing its sharing.

3. INTERFACE DESCRIPTION

The web interface of the application has 3 types of pages:

- Home page

- Creating a medical report
- Viewing and sharing the report created.

The initial page of the application contains a simple menu that allows the user to select the type of medical report. The elements of this page are the text with the indication for the user and the two buttons corresponding to the two categories of medical reports, each of them opening the form specific to the named type.

Selectați tipul de raport medical pe care doriți să îl creați



Fig. 1 – Home page capture.

Once the type of medical report is selected, the content of the page viewed by the user turns into a form. This form has 3 sections: patient data, doctor data, investigation / medical consultation data, the third section being determined by the type of report initially selected.

The patient data section contains fields for the following information:

- Last Name: mandatory field, used to mention the patient's last name
- First name: mandatory field, used to mention the patient's first name
- Age(years): field that supports only numerical values containing the declared age of the patient
- CNP: mandatory field, used to mention the personal numerical code; this field in the form has additional restrictions declared, so the user has to enter 13 numeric characters for the information to be valid
- Telephone number: field intended for the patient's contact number; this field is only validated if numeric characters, the "+" character and space have been added, the rest of the characters lead to its invalidation and return an error
- Email address: mandatory field, intended for the email address to be used in sharing the medical report; this field is invalid and prevents the completion of the medical report if it does not contain a valid email address
- Known allergies: field intended for medical information that the patient declares regarding his allergies
- Medical history: field used to list interventions, previous investigations as well as previously diagnosed conditions.

The doctor data section contains 5 mandatory fields, marked with*:

- Last Name: used to mention the last name of the doctor
- First name: used to mention the doctor's first name
- Identification code: used to mention the doctor's identification code
- Specialty: used to mention the field of the doctor who consults the patient or analyzes the images obtained following a medical investigation
- Medical unit: used to mention the medical unit where the patient was evaluated.

The content of the third section varies depending on the type of medical report selected. Thus, the third section of the form can be called “Investigation data” or “Medical consultation data”.

The fields of the "Investigation data" section are:

- Imaging investigation performed
- Observations: the doctor includes the analysis resulting from the analysis of the images obtained from the investigation mentioned in the previous field
- The conclusions of the investigation: the doctor highlights the significant observations in this section, observations that can help the referring doctor in making a diagnosis
- Recommendations: here are included the recommendations of the specialist doctor in case the patient is recommended to resume the investigation after a time or if other investigations/analyses are required.

The fields of the “Medical consultation data” section are:

- The reason for the patient's presentation
- Observations: the doctor includes the resulting analysis during the consultation
- Diagnosis: here is included the diagnosis made following the consultation, which justifies the recommendations in the next section
- Recommendations: here are included the recommendations of the specialist doctor in case the patient is recommended to take treatment, follow a regimen, carry out a set of analyses or a medical investigation or consult a doctor from another field.

The button to save the report is initially disabled when not all mandatory fields are filled with data. In addition to the error that warns of the existence of unfilled mandatory fields, the errors associated with the special validators added for the CNP, phone number and email address fields in the patient data section can be displayed. If an error is listed under the save report button in the form, the fields causing the error are highlighted in red as seen in Fig. 2.

Date pacient

Nume*	Prenume*	Vârsta(ani)
CNP*	Număr de telefon	Adresă de email*

0 / 13

Fig. 2 – Fields with errors.

The text area type fields can be completed both by writing and by dictating the information, and this is highlighted by the existence of a button with a microphone next to it, but also by the helper text existing in its area before adding the information. This text tells the user how to add the content, the existing message being: “Press the button to start transcribing the content or type...”. After starting the recording, the blue button with the microphone is replaced by the red button with a muted microphone which allows the recording to stop. Once the audio recording is started for a field in the form, typing text into it is temporarily disabled as the real-time transcription is available. After turning off the microphone, this field becomes accessible again for editing the existing text. If the user turns on the microphone for another field even though it already has transcription active, it stops recording in the field it was active in and starts transcribing in the newly selected field.

Informații medicale



Alergii cunoscute pacientul este alergic la praf	
Istoric medical pacientul a fost diagnosticat anterior cu	

Fig. 3 – Recording data for text area fields.

After completing and verifying the data in the form, the doctor can save the medical report by pressing the button at the end of the form. The date and time of the report will be given when it is saved. Immediately after pressing the button to save the report, it will be automatically downloaded, in PDF format, and can be easily redirected to the clinic's database. The name of the file containing the filled

data is given by the patient's first and last name, the date of the consultation/investigation and the field of the doctor. After saving the document, the user is redirected to the page that allows viewing and sharing the report.

In the page that appears after saving the created report, the user will have the report's data in its content, followed by 3 buttons:

- Print Report
- Send the patient report by email
- Create a new report.

By pressing the “Send the patient report by email” button, the generated medical report is sent to the address specified on the patient's contact details. The email sent has the subject “Medical report [date]”, and in its content is specified the name of the patient and instructions to access the attached document. The PDF document containing the medical report sent is secured with a password, which is represented by the last 6 digits of the patient's CNP. The name of the attachment contains the doctor's field and the date when the report was made. An example of an email sent through the application is available in Fig. 4.

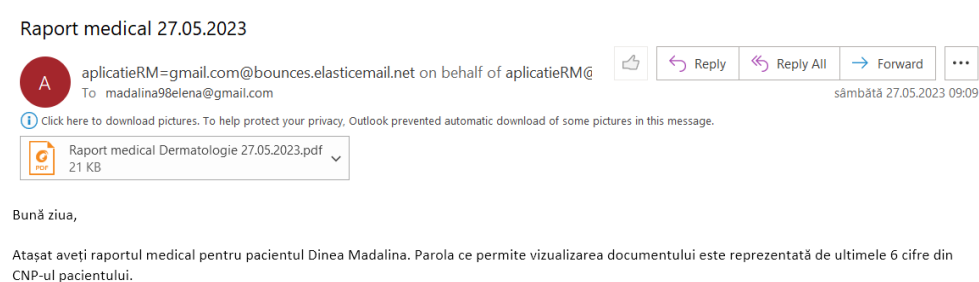


Fig. 4 – Email sent using medical transcription application.

Pressing the “Print Report” button opens a new window in the browser, through which the report can be printed after selecting the printer and pressing the “Print” button.

If the user clicks on the third button, “Create new report”, he is redirected to the initial page of the application, which allows him to select the type of report to be created, and the data corresponding to the previous report is deleted from the application's memory.

3. TECHNOLOGIES USED IN THE DEVELOPMENT OF THE MEDICAL TRANSCRIPTION SOLUTION

The following technologies were used in the realization of this project:

- Angular 15 [12]: for the development of the web application

- Angular Material [13]: for design elements within the interface
- Web Speech API [14]: for speech recognition integration
- pdfMake [15]: for generating PDF files manipulated through the application
- Smtplib [16]: for sending email with the report to patient
- Elastic Email [17]: for the server through which emails are sent.

Considering the chosen technologies and the application field, TypeScript programming language was used in the .ts files of the project, HTML markup language in the files. html and the SCSS formatting language in .scss files.

The working environment used in writing the code was Visual Studio Code. The command line where Angular CLI was installed was used to compile the code, install dependencies and start the application. The application was tested and viewed using the Google Chrome browser.

4. PROJECT ARCHITECTURE

The application is made up of the basic component of a project developed using Angular technology, 4 components and 3 services, each having methods dedicated to a specific visual part, in the case of components, respectively functionalities, in the case of services.

The services that constitute the functional part of the application are:

1. `SpeechToTextService`: used to configure the speech recognition system, start and stop recording using the microphone, and pass the obtained real-time transcription of the spoken content to the component using the speech-to-text function.

2. `MedicalReportService`: used for memorizing and manipulating the type of medical report and the data in its content, as well as monitoring the stage in which the user is, thus controlling the indicators that determine the visibility of the components in the graphic interface.

3. `SharingService`: used for generating, downloading and printing the PDF document containing the data of the saved medical report, as well as sharing it by email.

The Angular components corresponding to the graphic elements visible to the user are:

1. `SpeechToTextElementComponent`: defines the appearance and functions of the elements in the medical report form that allow entering data both by typing and by voice recording, each element defined by this component is an instance of this class and uses the functions from `SpeechToTextService` service.

2. `HomePageComponent`: defines the appearance and functions of the graphic elements from the initial page of the application, setting through the `MedicalReportService` the type of form to be displayed for the medical report data.

3. **MedicalReportFormComponent**: defines the appearance and functions of the graphic elements on the page that contains the form used to fill in the medical report data, those that enable voice recognition being instances of the **SpeechToTextElementComponent** class and uses all 3 services previously defined to allow saving data from the report and downloading PDF document after saving it.

4. **MedicalReportDisplayComponent**: defines the appearance of the page that allows viewing and sharing the created report, acquiring the data saved through the **MedicalReportService** service and allowing it to be printed and sent by email using the **SharingService** service.

The basic component of the project, **AppComponent**, instantiates the 3 components corresponding to the pages in the application and controls which of them is visible at a given moment with the help of the indicators that define the stage in which the medical report is.

In Fig. 5, the application architecture represented by instances of the 5 classes corresponding to the existing Angular components in the project is available. The way the user can move from one page to another is illustrated by arrows.

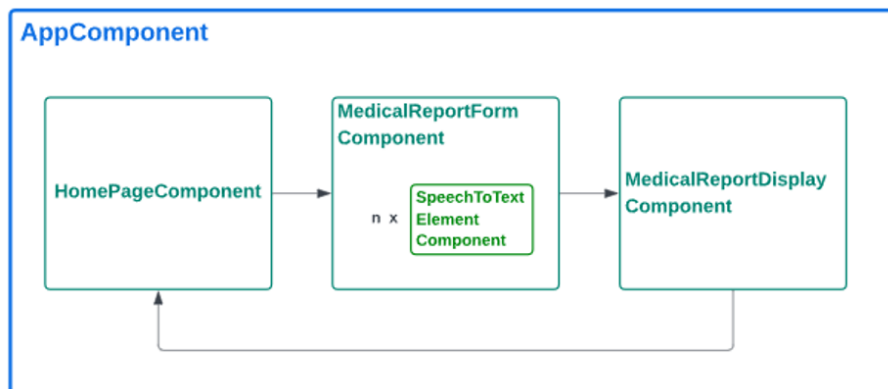


Fig. 5– Solution architecture.

5. CONCLUSION

The use of speech recognition in the creation of medical transcription applications aims to ease the work of medical professionals, freeing them from writing reports. Moreover, these reports contain a larger volume of information, with more details included, the information being added while the doctor examines the patient in the case of a medical consultation, respectively, while analyzing the images obtained in the case of a medical investigation.

The proposed and implemented solution for reducing the time allocated for writing a medical report is represented by a medical transcription application for

the Romanian language that also allows viewing and sharing the created medical report. This solution reduces the time needed to complete the reports related to a medical consultation or investigation, giving the medical staff more time to focus on the patient or the images acquired during an investigation.

This web application can be integrated into complex solutions that offer more flexibility in the medical field and the way medical reports are managed. One of the future directions in the development of the solution would be the integration of the existing interface into a system that stores the created reports and offers functionality to both doctors and patients. Storing the reports in the electronic environment would facilitate access to them, making it much easier to follow the medical history of a patient in this way.

Received on November, 2024

REFERENCES

1. LATIF, S., QADIR, J., QAYYUM, A., USAMA, M., YOUNIS, S., *Speech Technology for Healthcare: Opportunities, Challenges, and State of the Art*, IEEE Reviews in Biomedical Engineering, **14**, pp. 342–356, 2020.
2. RAMANARAYANAN, V., LAMMERT, A., C., ROWE, H., P., QUATIERI, T. F., GREEN, J., R., *Speech as a Biomarker: Opportunities, Interpretability, and Challenges*, Perspectives of the ASHA Special Interest Groups, **7**, 1, pp. 276–283, 2022.
3. MOHAN, M., S., MATHEW, S., G., PAUL, S., NANDARAJ, V., A., ANIYAN, A., *Analysis of Different Speech Recognition Toolkit for EMR*, EasyChair Preprint 6316, 2021.
4. ZUCHOWSKI, M., GOLLER, A., *Speech recognition for medical documentation: an analysis of time, cost efficiency and acceptance in a clinical setting*, British Journal of Healthcare Management, **28**, 1, 2022.
5. Nuance Dragon Medical One, <https://www.nuance.com/healthcare/provider-solutions/speech-recognition/dragon-medical-one.html>, accessed on 12.11.2024.
6. SayIt, <https://www.vtexvsi.com/speech-recognition/sayit/>, accessed on 12.11.2024.
7. prognocIS, <https://prognocis.com/medical-speech-recognition/>, accessed on 12.11.2024.
8. VoiceboxMD, <https://voiceboxmd.com>, accessed on 12.11.2024.
9. WebChartM, <https://www.webchartmd.org/>, accessed on 12.11.2024.
10. Amazon Transcribe Medical, <https://aws.amazon.com/transcribe/medical/>, accessed on 12.11.2024.
11. AUGMEDIX, <https://augmedix.com/>, accessed on 12.11.2024.
12. Introduction to Angular concepts, <https://angular.io/guide/architecture>, accessed on 12.11.2024.
13. Angular Material, <https://material.angular.io/>, accessed on 12.11.2024.
14. Web Speech API, https://developer.mozilla.org/en-US/docs/Web/API/Web_Speech_API, accessed on 12.11.2024.
15. pdfmake, <http://pdfmake.org/#/>, accessed on 12.11.2024.
16. SmtplibJS.com, <https://smtpjs.com/>, accessed on 12.11.2024.
17. Elastic Email, <https://elasticemail.com/>, accessed on 12.11.2024.