Development of a Speech-to-Text Program for People with Haring Impairments

Diana Koshtura, Vasyl Andrunyk, Tetiana Shestakevych

Lviv Polytechnic National University, Stepana Bandery Street, 12, Lviv, 79000, Ukraine

Abstract

The process of inclusion of a person with special needs into society can be improved with different information technologies. For people with hearing impairments, it should be based on speech-to-text converter. The main functions of such an application should allow the user to see the text on the screen of the device. To create such an application, its main functions were chosen based on the analyzed analogs, so the designed application met all the necessary demands. The UML diagrams were chosen to model the designed application, called CLON. The software was tested with different single words, phrases, and sentences.

Keywords [1]

Speech-to-text, hearing impairment, risk low, speech recognition, quick answer, convert speech, voice message

1 Introduction

The effective development of society can be determined by how government measures are aimed at human well-being. This development can be considered at a high level if the state directs its forces to help the least protected and most vulnerable members of society. Issues of social assistance and protection of the rights of persons with disabilities are considered relevant throughout the civilized world. Having hearing impairments causes troubles for not only persons with such disability, but also for their families, friends, all the other members of society, and creating a mean to improve such communication, will be of greater help to the society of hearing people. Taking into account the level of the electronic devices spread, such helpful mean should be designed and developed as an information technology. The object of research is the process of communication between people with hearing impairments. The subject of the research is a development of an application as a mean of interaction between people. The practical value of the results is to develop and test new tools and methods for converting speech into text, saving human resources and money, and building convenient, useful, and important for people with hearing impairments software and systems. The data that will be extracted and directed to the development of such a science as speech recognition, other systems that use the human voice to convert to text will be very useful and will improve these systems.

2 State of art

To support the communication processes, when some of the participants have hearing impairments, the methodology of *speech-to-text* is used. Such methodology, risen from interdisciplinary connections of computer science and computational linguistics, is many various fields. In *Scopus* database, the *speech-to-text* keyword was first mentioned in 1982, in a paper, where the author considered the unlimited-vocabulary speech-to-text system as the key to future phonemic typewriter [1]. In May 2021

MoMLeT+DS 2021: 3rd International Workshop on Modern Machine Learning Technologies and Data Science, June 5, 2021, Lviv-Shatsk, Ukraine

EMAIL: Diana.Koshtura.sa.2017@lpnu.ua (D. Koshtura); Vasyl.A.Andrunyk@lpnu.ua (V. Andrunyk); Tetiana.V.Shestakevych@lpnu.ua (T. Shestakevych)

ORCID: 0000-0003-0697-7384 (V. Andrunyk); 0000-0002-4898-6927 (T. Shestakevych)

with the *speech-to-text* keyword, the Scopus database as a result gives 953 documents, 470 of those papers were published in 2017-2020. The word cloud from the abstracts of these documents is in fig. 1 (numbers and common words were removed).

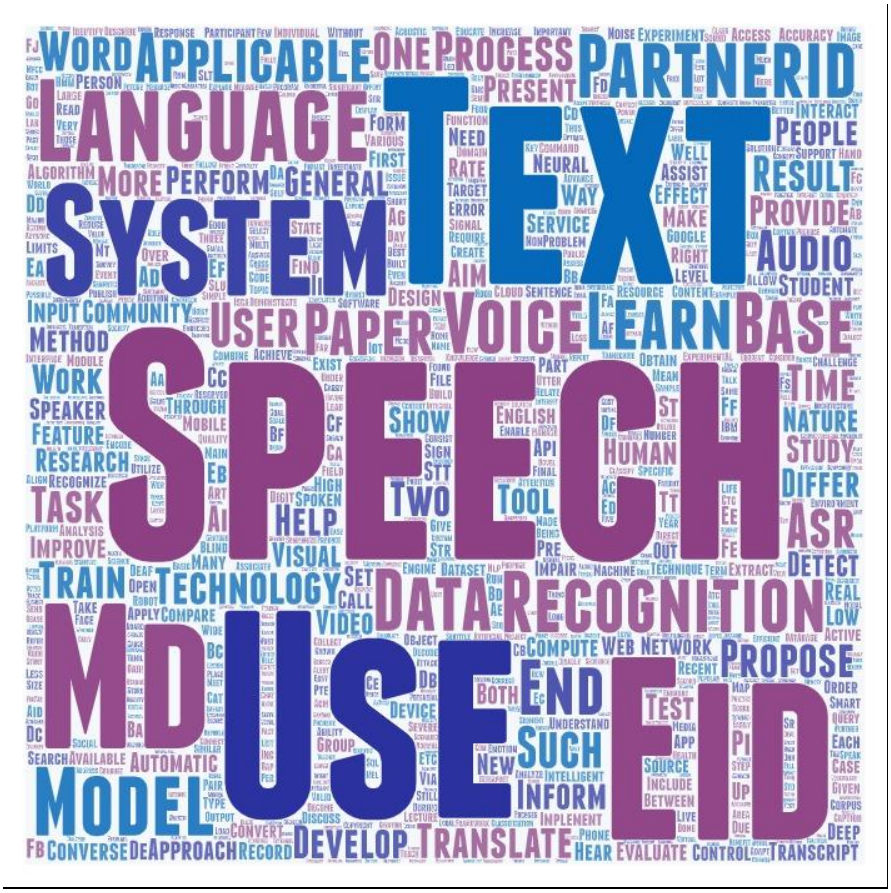


Figure 1: The word cloud from the abstracts of papers, indexed in Scopus, with speech-to-text keyword

The results of the implementation of speech-to-text methodology can be used in various fields, such as advertisment, automatic keywords extraction, education, linguistics, military, data analysis, etc. [2-12]. The peculiarities of translation of spoken language into sign language were investigated in [13, 14]. Silent speech recognition is actual in a noisy place can be a part of a wider system of human body image recognition [15-24].

The speech-to-text method was used to assess the risk of Alzheimer's disease when patients` texts were analyzed to find special markers [25-28]. Also, the speech-to-text method is used to treat a motor speech disorder resulting from neurological injuries [29-30].

In [31], scientists, while examining the language and literacy needs among students in youth detention, declared the need for consistent text-level language assessment. To better identify functional difficulties within their language, authors used speech-to-text methods [25-28]. To control electrical appliances and door, the authors [32] constructed a prototype with a speech recognition system.

The effects of assistive technology for students with severe disabilities (reading and writing) were investigated in [33]. The researchers concluded such technology to be supportive, motivating, but not with obstacles in implementation [34].

The peculiarities of stemming as a part of non-formal conversation converting into the text were investigated by authors who concluded, that the non-formal sentences should be normalized to formal ones to be used for text classification [35-47], hate speech detection and content analisys, etc [48-61].

Let us analyze the capabilities of the developed software product in comparison with modern analogs. Programs similar to those present on the market are products of foreign production: RogerVoice, Live Transcribe, App MyEar, AVA, Voxsci.

2.1 RogerVoice

RogerVoice (https://rogervoice.com/) is the world's first mobile application for creating subtitles for telephone conversations, ie converts mobile voice calls into a much more accessible text format. Developed by deaf engineer Olivier Jeannel. The application uses speech recognition technology to convert voice to text, so deaf and deaf people can read what the other person is telling them.

The application was created thanks to a successful campaign on Kickstarter, a beta version of the application is currently available, which can be downloaded from the RogerVoice website (Fig. 2*a*).

2.2 Live Transcribe

The application reads human language and translates it into text. To do this, the application uses a smartphone microphone and Google's language API (Fig. 2b), which supports more than 70 languages. To ensure that the transcription is as error-free as possible, machine learning technologies are responsible for recognition. But Live Transcribe can not only decrypt the voice, but also notify the user that someone wants to talk to him, and also allows you to participate directly in the dialogue, giving answers using the built-in keyboard. The spoken text is perceived by the phone's microphone and delivered to the screen of the Android phone via Wi-Fi or other network connection. This can be useful for people who do not hear and attend conferences or lectures, for example. Spoken words will appear on the phone of the person who has the applications.

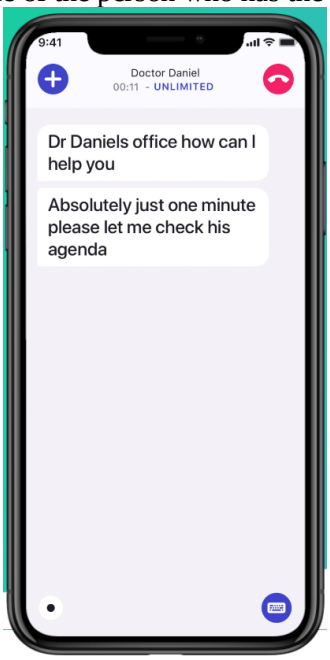


Figure 2a: View of the RogerVoice application

Figure 2b: View of the application from Google

2.3 MyEar app

The MyEar app was developed by Gerald Isobe, a deaf golfer and his son Brandon.

Gerald grew up reading his lips, but was disappointed to understand only 30% of what he said, and tired of asking people around him, "What are you talking about?"

This program was created based on these frustrations, and it is a program that he and others use to communicate with their hearing colleagues, friends, and family.

Direct pricing. It costs a one-time $ 9.99 with free updates when we launch a new software update.

For emergencies. If a police officer stopped you, would you be able to effectively understand what he was saying, for example, at night? Most likely not. Here the MyEar application can help. It will

record what the officer is saying so that you do not guess or try to understand what the officer is saying at night and at any other time. Last minutes of a meeting or presentation, and no time to call an interpreter. Sometimes last-minute meetings can happen and then just open the MyEar program and you can use it for long presentations and you will understand what it is about.

Influence on new words and topics. TV subtitles and ASL translators do not translate 100% word for word what people say on TV or on the phone. However, you may not realize this until you try the MyEar app. The MyEar app will translate every word a person says to you.

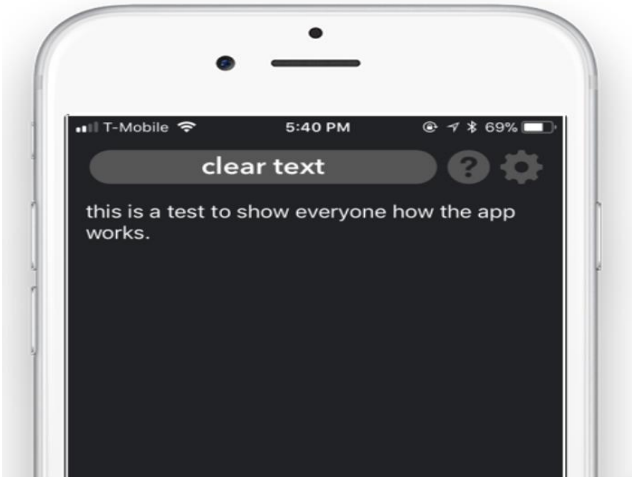


Figure 3: View of the application from Apple

2.4 AVA

Lip reading can be more difficult in a group of people, and this is one of the main reasons for creating AVA. If a person who is deaf or suffers from hearing problems is with a group of friends, he or she can force those friends to join the program - then the person (s) with hearing impairments will see the live broadcast of the group conversation. The language is perceived by the phone's microphone, and the name of the speaker is displayed on the screen before the person speaks.

AVA works with employers, teachers, event organizers and other accessibility professionals who seek to fully engage their deaf and hard of hearing members.

How can Ava fit into your daily life? See how users use Ava → Ava gives you a whole new level of autonomy for many situations in your daily life.

2.5 Voxsci

Voxsci (https://www.voxsci.com/) is a language-to-text program that converts voicemail messages into texts and emails that you can store, search and share. Expenses start at £ 5 a month for 30 voicemails or emails.

Listening to voice messages can be very inconvenient. VoxSciences provides a paradigm shift by transcribing voice messages into text messages. This gives voice messages a quantum leap for joining email, SMS, and instant messaging on the same basis with all the inherent benefits such as texture search. Our VERBS engine (virtual engine for basic language recognition) converts voice messages into text messages and delivers them as e-mail, SMS, or via the API. Voicemail to Text (SMS) is ideal for personal or corporate voicemail systems.

Voice messages, transcribed and delivered by e-mail, are mainly used by call centers, comment or contest lines, and corporate voicemail systems.

The Client's Voice is a market research technique that provides a detailed set of needs and customer needs. It includes analysis of feedback from various sources, such as e-mail surveys, the Internet, and IVR. VoxSciences provides a key component to facilitate the analysis of audio feedback for analysis, providing transcription of transcripts almost in real-time in the business class through our API.

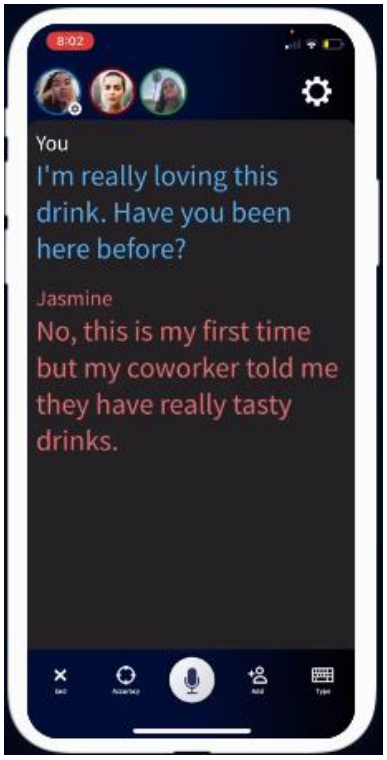


Figure 4: View of the AVA application

Figure 5: View of the application from VoxScience

Now we shall compare the RogerVoice, Live transcribe, AVA, VoxScience, MyEar, and the CLON application will be compared. All evaluations will be conducted on a scale from 0 to 5 points, where 0 points means that the application does not meet the purpose, conditions, or characteristics, and 5 points when it meets all requirements. The comparison will be based on the following characteristics or requirements:

- Compliance with the purpose. Analogs will be evaluated for such purposes as helping people with hearing impairments to communicate.
- Ease of use. This item will take into account items such as availability from different operating systems.
- Pricing policy, tariff. Here, analogs will be evaluated by their price categories, free products, paid.
- Practical use. Are analogs used or have they found value among people with hearing impairments?
- Simplicity of a design. Are analogs difficult to use?
- Network requirements. Analogs will be judged by whether a network connection is required when using the application.
- The complexity of manufacturing, development. What technologies were used to create the application, resources?
- Design. It is not possible to objectively assess whether the design is successful or not, so this assessment will be based on feedback.
- Control panel. Is a convenient and complete control panel for users?
- Parameters. How many parameters can be configured in analogs-applications, and what is their width (variety)?
- Support. Do applications inflate support in case of problems when interacting with the application?

The application being developed will be evaluated against the requirements it will meet. Other characteristics will be evaluated according to the developed mockaps or prototypes. So, we can conclude that only applications similar to Roger Voice and VoxSci do not meet all the requirements - according to the purpose, they do not fully satisfy it.

Table 1
The comparison the analogs

Characteristic	Roger Voice	Live Transcribe	MyEar	AVA	VoxSci	CLON
Purpose compliance	3	5	5	5	2	5
Ease of use, available on	All OS, beta version 4	All operating systems 5	iOS 3	All OSs 5	All OSs 5	Android 3
Fee	Free 5	Free 5	$ 10 one-time purchase 4	Free 5	£ 5 a month - 30 posts 3	Free 5
Practical use	Beta version 4	5	4	4	3	Under development, expectations 5
Ease of design	4	4	4	4	3	Under development, expectations 5
Network requirements	Yes 4	Yes 4	Yes 4	Yes 4	Yes 4	Yes 4
Complexity of development	Existing APIs are used 3	Custom APIs are used 4	(own) 4	(available) 3	(own + MN technologies) 5	(available) 3
Design	4	5 (reviews)	4 (reviews)	4	4	5
The complexity of control panel	Medium 4	Low 5	Low 5	Low 5	High 3	Low 5
Additional options	Yes 4	No 3	No 3	Yes 5	Yes 4	Yes 5
Support	3	4	5	3	4	4

Therefore, we can conclude that the availability of different operating systems does not depend on the price of the application. Therefore, applications that are free offer the use of the application on different operating systems. From the ratings, we can conclude that the applications are not difficult to use, and they are popular among users, and often used. Applications were evaluated according to reviews on official websites, Play Market, App Store (Fig.6 presents the feedback about MyEar, with an overall rating 4.1 and positive text reviews; Fig.7 presents a review of Live Transcribe. Reviews are mostly positive, often mentioned the usefulness of the application). In general, there are no negative reviews, there are simply lower ratings, where people are dissatisfied with the quality of translation from language to text, dissatisfied with the pricing policy, or have additional wishes, such as quick answers. Therefore, it can be concluded that the complexity of development depends on the technologies used to develop the application and whether the application uses a network connection.

Thus, we can conclude that all analogs are competitive, have many advantages, but also have disadvantages. The disadvantages that emerged during the comparison of analogs are:

- Difficulty of use - complex interface.

- High fee for the application - people like the free application more.
- A small number of settings - users would like to be able to give short answers, adjust call filters.
- Some operating systems are not supported.
- Complexity of development - many companies create their own technologies for language recognition, which makes the application more difficult to use and more expensive.

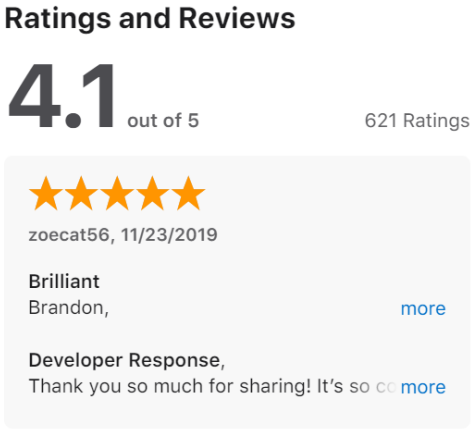


Figure 6: MyEar reviews

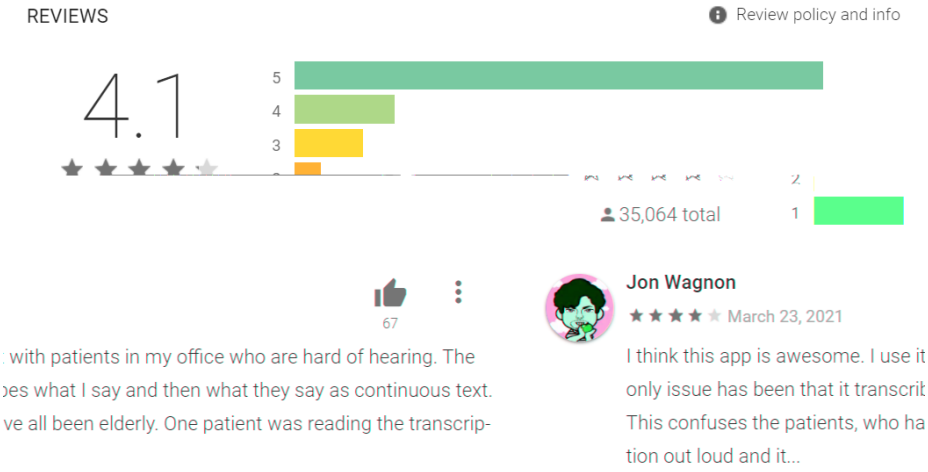


Figure 7: Feedback Live Transcribe

Compared to analogs, the developed application will have many advantages and will include and correct those disadvantages that have analogs. The main advantage of the application will be that it will allow you to configure call filters, the first time will not charge for the use of services, and will allow you to provide quick answers.

3 Materials and methods

The target audience of the application includes the following groups of people:
1) Developers, business analysts.
2) Users (people with hearing impairments).
3) Sponsors.

Application languages:

Two language versions of the application: Ukrainian and English. One language will be implemented at the beginning.

Basic Success Scenario: This section of the specification describes a "success scenario". That is, actions that lead to the successful completion of events in the main process, for example:

- The user installs the CLON application.
- The user logs in / registers (or skips this step) in the application.
- The user gets acquainted with the settings and interface of the application.
- The user selects a live broadcast recording point.
- The application shows the user the text that has just been spoken.
- The application analyzes the text and offers quick answers.
- The user selects one of the suggested answers or selects a typing item independently.

Baseline extensions or alternate streams: This section indicates all other possible scenarios that lead to the successful completion of the main scenario or alternative scenarios that lead to incorrect completion of the precedent. In this case, after processing all possible extensions of the precedent, the aircraft must ensure the return of the user to the main scenario, if the aircraft does not provide an alternative course of events. For example, baseline extension:

- The user has logged in to the application.
- The user makes settings for the application - language and color settings.
- The user makes conversation filters.

4 Terms of reference

4.1 The structure and functionality of the application

The application will develop a system that will perceive speech from a person and convert that speech into text. This feature is gaining popularity not only for people with hearing impairments but also in different systems, such as various applications or smart homes or devices.

Consider the activities of the application. The application asks the user to convert speech into text. The system does not produce components, but only converts speech into text. The system should allow centralized management of all processes related to text analysis.

The main types of work in the application:

- The system receives speech from a person.
- The program processes the converted speech into text.

At the stage of development of **functional requirements,** the following list of functions and their attributes was created. It was decided to follow the principles of flexible software development methodologies, which is why the first version of the product will implement a minority of the following functions.

In fig. 8 we used the following notation for functions of the compared applications:

- User graphical interface is the interface through which the user will interact with the system.
- Registration in the system – for the system to analyze new users.
- Login – so that the system stores filters and user settings.
- Speech recording – buttons on the home page that allow you to record speech.
- Typing – buttons on the main page to select the typing function.
- Speech to text conversion - API to convert speech to text.
- Language – choice of language country for analysis, perception, settings, filters.
- Password recovery – recover the user's password if the user forgot it.
- Text auto-correction - text correction functions - auto-correction based on context.
- Filtering – filters that will perceive the speech and process it according to the theme.

Non-functional requirements

- You need to provide user support when working with the application.
- You need to ensure the future use of text-to-speech.
- You need to provide offline support for the program.
- List of required technologies and additional devices.
- The application must be designed as an application for the phone.

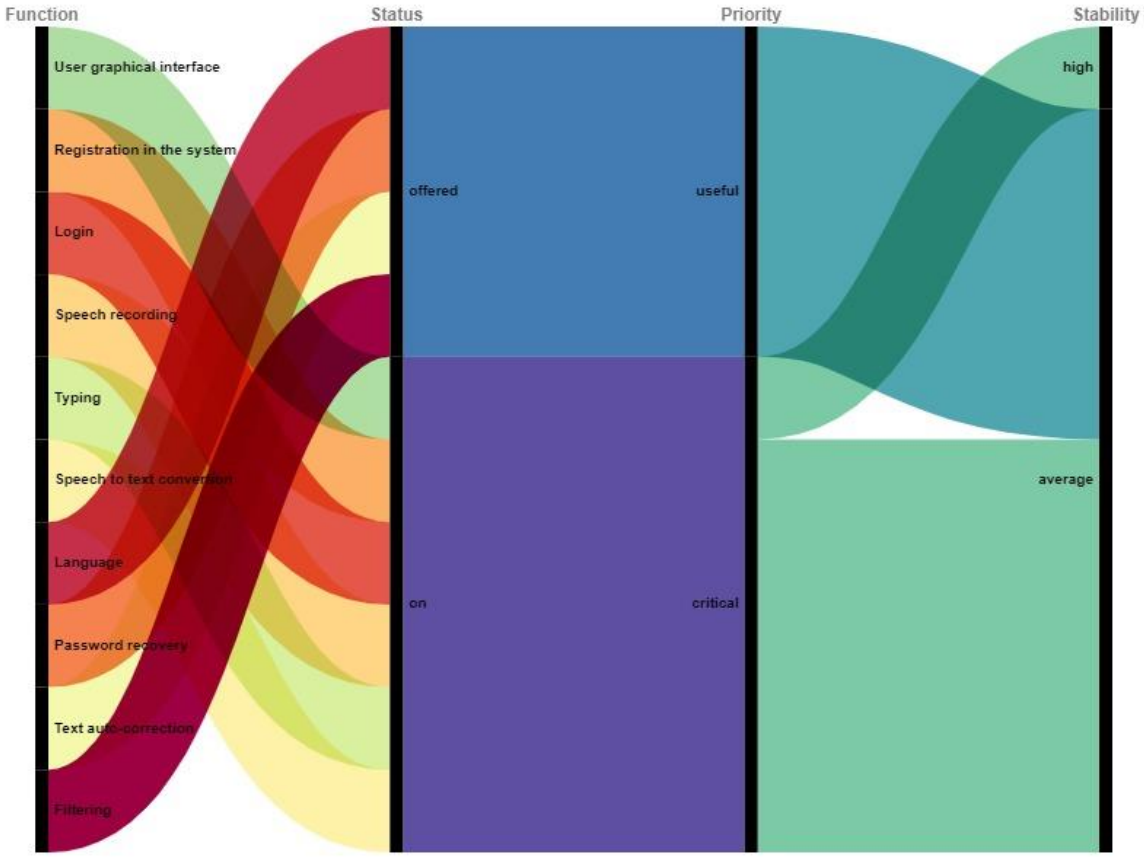


Figure 8: Comparing the functions of the application

Application architecture requirements. The application architecture must provide scalability and expansion of the system. Adding additional functionality should be done by adding additional modules without significant modernization of existing modules. The architecture of the application should provide for the independence of the information display system module from the information storage and management module. The application architecture should provide for the independence of the system implementation from the hardware platform and the server operating system.

The application architecture includes the following components:

- Graphical interface for user interaction.
- Authentication for users.
- Buttons for recording speech through a microphone.
- Access to the user's microphone.
- Functions for recording speech in text.

Application access requirements. All users, including the developer, have access to the application.

Requirements for data transmission to users. The application can transmit the following data to users:

- Application information.
- Team information.
- Information about filters and settings.

The user can transfer data. Data to be read through the microphone.

Finally, the structural model of the user`s interaction with an application is in fig. 9.

Class application CLON. For the class Convert speech to text: public: Convert recorded speech to text, Filter text by topic. For the Text Input class, add the following operations: public: use the keyboard to enter. For the class Using a microphone, add the following operations: protected, provide access to the microphone, record speech through the microphone. For the User Interface class, add the following

operations: public: grant access to application functions, provide access to filters, provide access to application settings.

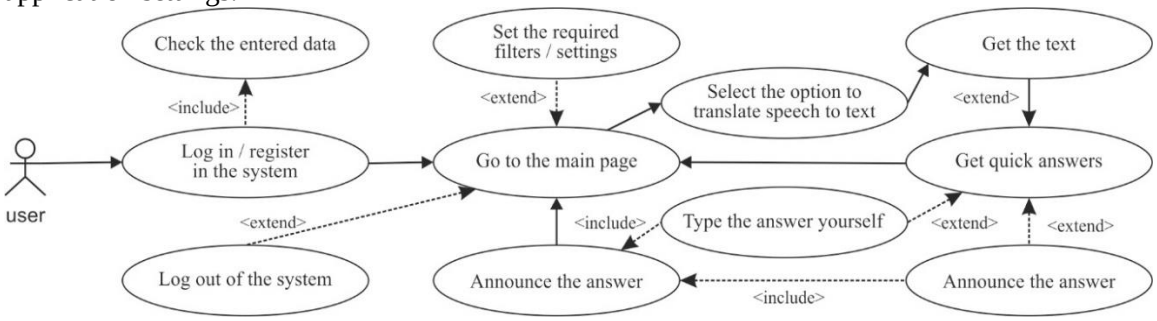


Figure 9: Diagram of use cases

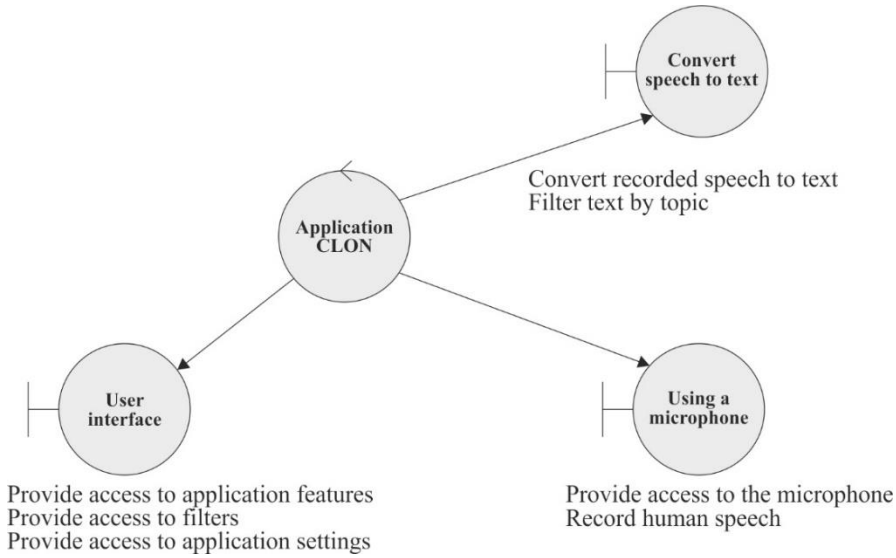


Figure 10: Diagram of classes

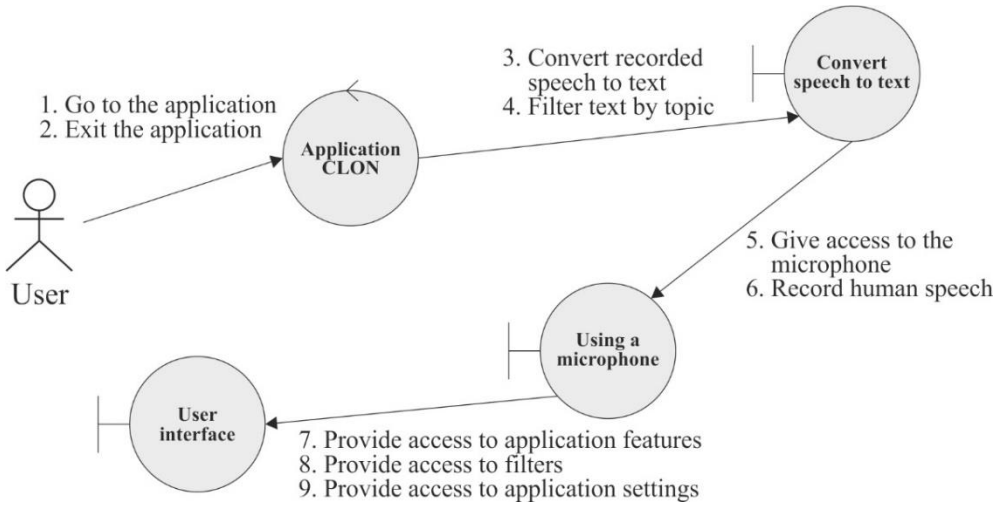


Figure 11: Diagram of cooperation

Add class objects from the class chart with the following names: System User, Application, Convert Speech to Text, User Interface, Use Microphone.

Add links that connect class objects to names: System User, Application, Convert Speech to Text, User Interface, Use Microphone. Add Message 1 - 14:

- Log in / register in the system.
- Configure the application.
- Exit the application.
- Convert recorded speech into text, provide access to a microphone.
- Record speech.

The development of the activity diagram shows how the program will work at the stage of speech perception from the speaker. The activity diagram assumes that after the user selects a text item, the system will record the speech in the device memory. After this step, the program uses API, which was specially developed for text recognition. API performs speech recognition. Then the program will create text based on the speech and pass it to the program interface. Then the user will be able to see the text.

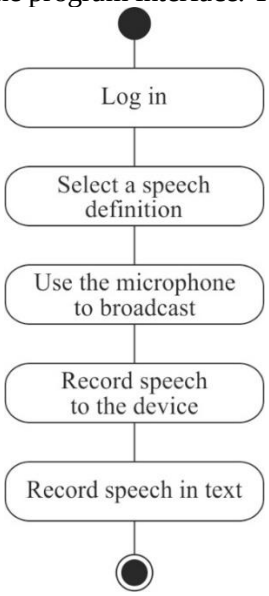


Figure 12: Diagram of activity

The sequence diagram is constructed from the classes of the class diagram in chronological order.

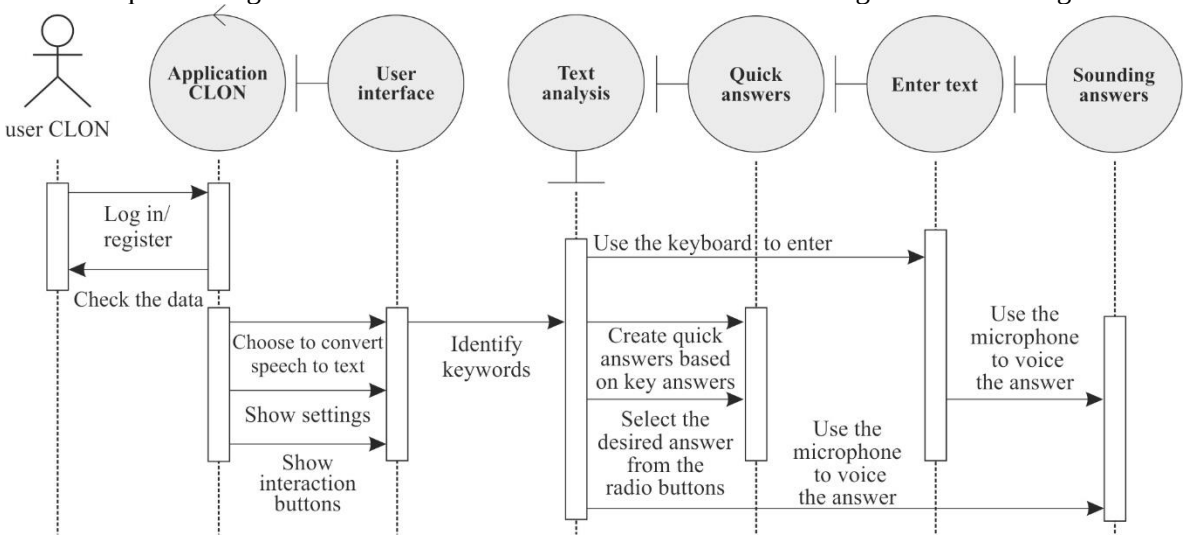


Figure 13: Sequence diagram

The deployment diagram, that visualizes the hardware and software of the designed system, is in fig. 14. The nodes are CLONE application, API and text analysis Servers, and a connector between server and host.

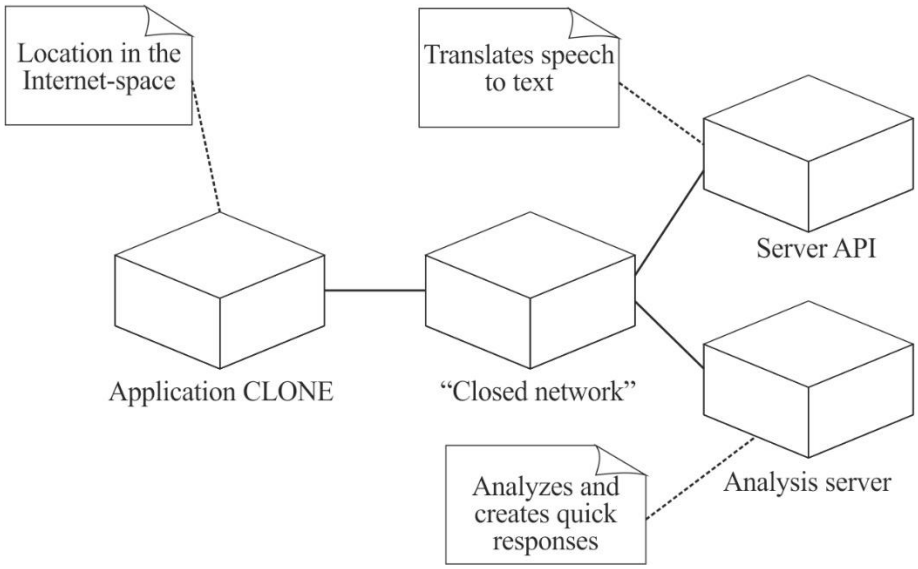


Figure 14: Deployment diagram

5 Experiment

Type of application (mockap, prototype). The application provides login or registration so that the user's settings are saved, his statistics are processed and improvements are provided.

If the user is not yet registered in the system, one can register.

You can also sign in with a social network, such as Twitter or Facebook.

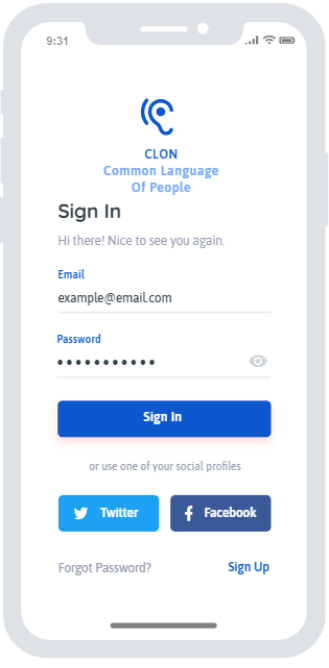


Figure 15a: Prototype of the entrance

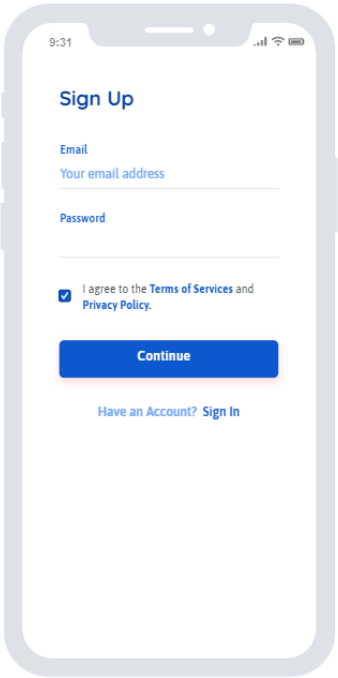
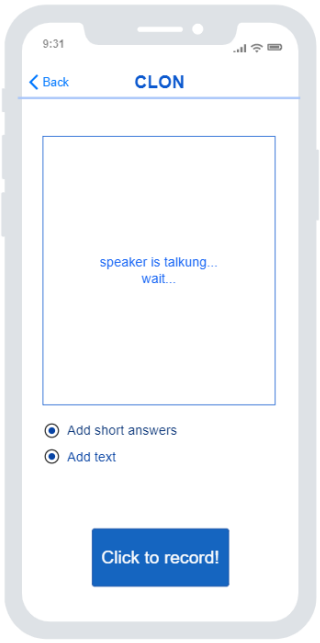


Figure 15b: Prototype of registration

The following prototype shows how the user will interact with the person. When a person says something, it will be automatically written as text in the user's application.

The following mockap shows how a person will interact with the application. To get the text from the person speaking, the user needs to click on the button to start writing the text. The user can also go back, select options from the radio button, such as add text yourself, use quick answers.

Next, the program must analyze the text so that the user can provide a quick response from the proposed options, which will be voiced. If the user does not find the right answer, he can type the text and it will be sounded through the phone speaker.



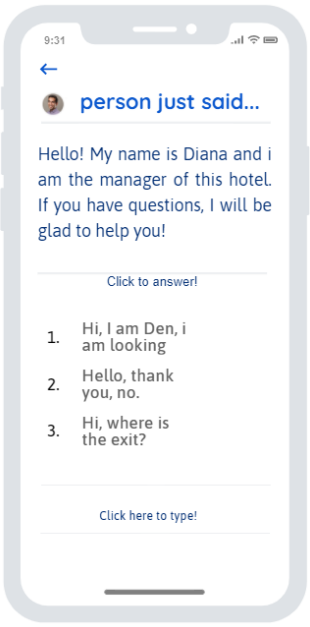


Figure 16: The main program window **Figure 17:** Prototype of the user's work

The first version will only receive speech from the user and translate it into text.

The first example will look like this. The user will be asked to say something.

```
Say Something!
```

Figure 18: To record speech

After that, there will be a pause of 3 seconds after which the speech will be interrupted. After that, the user of the application will see the text that was said by another person.

Text to be spoken by the user: *Hello it's my first program speech recognition app*.

```
CLON speech recognition thinks you said: hello it's my first problem speech recognition app
```

Figure 19: The result

You can see that the recognition was almost accurate. One word did not match - program - CLON interpreted it as a problem. Other results are satisfactory.

The following versions are needed in order to improve the results and implement all these features for the mobile application.

6 Results

Testing will follow these steps.

1. Single words recognition. At this stage, a certain vocabulary of some words will be checked to see if there are any difficulties with a certain category of words, or whether the recognition error does not depend on the category of the word.

2. Phrases recognition. At this stage, certain phrases consisting of two or more words will be checked. Here you can see how recognition occurs when more words flow.
3. Sentence recognition. Check whole sentences. Check more than two sentences.

6.1 Single words recognition

The first check is a check of certain words, pronounce for example. Fig. 20 shows the results for the personal pronoun *I* recognition.

```
CLON speech recognition thinks you said: bye
```

Figure 20: Pronoun *I* recognition

The result (*bye)* is not satisfactory. The next pronoun is *you*.

```
CLON speech recognition thinks you said: yo
```

Figure 21: Pronoun *you* recognition

The result is just as disappointing. The recognized word is *yo*. The next pronoun is *he*.

```
Clon Speech Recognition could not understand audio!
```

Figure 22: Pronoun *he* recognition

As the result, the application did not recognize the speech. The next pronoun is *she*.

```
CLON speech recognition thinks you said: skip
```

Figure 23: Pronoun s*he* recognition

The result is also not successful. Instead of the pronoun, *skip* is recognized. Now you can try to recognize some other words, much longer. For example, you can start with greetings. The first greeting is *hello*.

```
CLON speech recognition thinks you said: hello
```

Figure 24: Greetings *hello* recognition

The result is positive. The next word is *bye*.

```
CLON speech recognition thinks you said: bye
```

Figure 25: Greetings *bye* recognition

The result is positive. The next word is *goodnight*.

```
CLON speech recognition thinks you said: goodnight
```

Figure 26: Greetings *goodnight* recognition

Then other words were checked, both complex and easy to pronounce and to perceive.
The first word is *kernel*.

```
CLON speech recognition thinks you said: kernel
```

Figure 27: Word *kernel* recognition

Then the word is *linguistics*.

```
CLON speech recognition thinks you said: linguistic
```

Figure 28: Word *linguistic* recognition

The next word is *help*.

```
CLON speech recognition thinks you said: help
```

Figure 29: Word *help* recognition

The last word is *name*.

```
CLON speech recognition thinks you said: name
```

Figure 30: Word *name* recognition

It can be concluded that very short single words are not very well received by the application. And words that are much more commonly used, or that are longer than 4-5 letters are better recognized.

6.2 Phrases recognition

The first phrase is *my name*.

```
CLON speech recognition thinks you said: my name
```

Figure 31: The phrase *my name* recognition

Another phrase is *help me*.

```
CLON speech recognition thinks you said: help me
```

Figure 32: The phrase *help me* recognition

The third phrase is *nice day*.

```
CLON speech recognition thinks you said: nice day
```

Figure 33: The phrase *nice day* recognition

The fourth phrase is a *computer program.*

```
CLON speech recognition thinks you said: computer program
```

Figure 34: The phrase *computer program* recognition

Thus, we can conclude that the more words, the better the recognition. The text is self-correcting and it is possible that the artificial intelligence itself can correct or add the right words.

6.3 Sentences recognition

The first sentence is *Today is Monday*.

```
CLON speech recognition thinks you said: hello today is Monday
```

Figure 35: The sentence *Today is Monday* recognition

The next sentence is *How tall are you*.

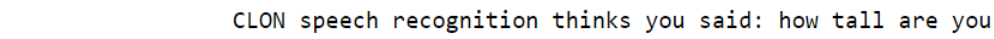


Figure 36: The sentence *How tall are you* recognition

The third sentence is *Computational linguistics is the scientific and engineering discipline concerned with understanding written and spoken language from a computational perspective, and building artifacts that usefully process and produce language, either in bulk or in a dialogue setting*.

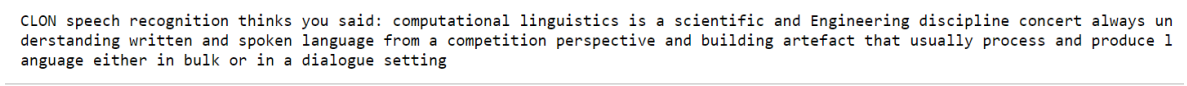


Figure 37: Third sentence recognition

7 Conclusions

Many language recognition problems are constantly being studied, improved, and understood. One such problem is modeling the process of translating speech into text. Translating speech into text is one of the problems that the science of speech recognition is trying to solve.

Speech recognition is the process of converting a speech signal into a text stream. This technology is used to solve such tasks as computer control, voice information, dictation, phonogram transcription. This technology is also used for speech recognition in Google translators, as well as in applications that help people with hearing impairments to communicate.

In the last few years, the demand for communication applications between people has been growing, because it is a fast way to communicate. Modern messengers are used for messaging, but their idea prompted people around the world to create a similar application for people with hearing impairments. This is one of the first and defining steps to solving the global problem because people with disabilities live among us and also want to communicate. Not everyone understands sign language and not everyone can immediately understand that a person has hearing problems, so this application is very necessary.

The system developed a system that recognizes the speaker's text and translates speech into text - yes, people with hearing impairments will be able to "hear" their interlocutor and understand him if he does not know sign language. Therefore, the developed application can and should be improved by expanding its functionality for further use in commercial companies seeking to help people with hearing impairments, as well as those entrepreneurs who intend to monetize such applications to increase their profits. You can provide the following options:

- adding filters for speech perception - that is, the conversion of speech into text will be more accurate if you specify the topic of conversation;
- integration with modern platforms;
- adding advertising banners of interactive format, namely mini-games based on the canvas model of the user's Internet browser.

References

1. E. T. Johnson, WILL TYPEWRITERS EVER TAKE DICTATION? Speech Technology, 1982, 1 (3), pp. 35-42.
2. G. Szymanski, P. Lipinski, Model of the effectiveness of Google Adwords advertising activities, 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018 - Proceedings, 2, art. no. 8526633, 2019, pp. 98–101.
3. S. Popova, G. Skitalinskaya, Extended list of stop words: Does it work for keyphrase extraction from short texts? Proceedings of the 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2017, 1, 2017, 8098815, pp. 401-404.

4. T. Shestakevych, Modeling the process of analysis of statistical characteristics of student digital text, CEUR Workshop Proceedings, 2021, 2870, pp. 657-669.
5. I. Dilai, M. Dilai, Automatic Extraction of Keywords in Political Speeches, 2020 IEEE 15th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2020 - Proceedings, 2020, 1, art. no. 9322011, pp. 291-294.
6. N. Kunanets, O. Levchenko, A. Hadzalo, The application of AntConc concordanger in linguistic researches, 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018 - Proceedings, 2, 2018, 8526591, pp. 144-147.
7. V. Vysotska, O. Kanishcheva, Y. Hlavcheva, Authorship Identification of the Scientific Text in Ukrainian with Using the Lingvometry Methods, 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018 - Proceedings, 2, 2018, art. no. 8526735, pp. 34-38.
8. O. Boreiko, V. Teslyuk, N. Kryvinska, M. Logoyda, Structure model and means of a smart public transport system, Procedia Computer Science, 2019, 155, pp. 75-82.
9. N. Hrytsiv, T. Shestakevych, J. Shyyka, Quantitative parameters of Lucy Montgomery's literary style, CEUR Workshop Proceedings,2021, 2870, pp. 670-684.
10. N. Shakhovska, T. Cherna, The method of automatic summarization from different sources, Econtechmod, 2016, Vol. 5, No. 1, 103–109.
11. V. Larin, N. Chichikalo, A.M. Kardo, K. Larina, Integrated Intellectual Approach to the Diagnostics of Defects of Operations of Induction Motors,, 2020 IEEE 15th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2020 - Proceedings, 2020, 1, art. no. 9322004, pp. 352-356.
12. Y. Kryvenchuk, I. Helzynskyy, T. Helzhynska, N. Boyko, R. Danel, Synthesis control system physiological state of a soldier on the battlefield, CEUR Workshop Proceedings, 2019, 2488, pp. 297-306.
13. R.A. Ramadani, I.K.G.D. Putra, M. Sudarma, I.A.D. Giriantari, A new technology on translating Indonesian spoken language into Indonesian sign language system, International Journal of Electrical and Computer Engineering, 2021, 11 (4), pp. 3338-3346.
14. Speech-to-text is used as a part of speech-to-sign communication [D.D. Chakladar, P. Kumar, S. Mandal, P.P. Roy, M. Iwamura, B.-G. Kim, 3d avatar approach for continuous sign movement using speech/text, Applied Sciences (Switzerland), 2021, 11 (8), art. no. 3439, .
15. L. Pandey, A.S. Arif, Liptype: A silent speech recognizer augmented with an independent repair model, Conference on Human Factors in Computing Systems – Proceedings, 2021.
16. K.Fornalczyk, A. Wojciechowski, Robust face model based approach to head pose estimation, Proceedings of the 2017 Federated Conference on Computer Science and Information Systems, FedCSIS 2017, 2017, art. no. 8104720, pp. 1291-1295.
17. G. Glonek, A. Wojciechowski, Hybrid orientation based human limbs motion tracking method, Sensors (Switzerland), 2017, 17 (12), art. no. 2857, .
18. A. Wojciechowski, Hand's poses recognition as a mean of communication within natural user interfaces, Bulletin of the Polish Academy of Sciences: Technical Sciences, 2012, 60, pp. 331-336.
19. P. Napieralski, M. Kowalczyk, Detection of vertical disparity in three-dimensional visualizations, Open Physics, 2017, 15 (1), pp. 1028-1033.
20. V. Khavalko, I. Tsmots, A. Kostyniuk, C. Strauss, Classification and recognition of medical images based on the SGTM neuroparadigm, Workshop Proceedings, 2019, 2488, pp. 234-245.
21. M. Nazarkevych, M. Logoyda, S. Dmytruk, Y. Voznyi, O. Smotr, Identification of biometric images using latent elements, CEUR Workshop Proceedings,2019, 2488, pp. 99-108.
22. D. Uchkin, T. Korotyeyeva, T. Shestakevych, Bitmap Image Recognition with Neural Networks, Econtechmod, 2020, Vol. 09, No. 1, 30 – 35
23. S. Kudubayeva, D. Ryumin, A. Sndetbayeva, Y. Krak, Computing of hands gestures' informative video features, Computer Sciences and Information Technologies - Proceedings of the 11th International Scientific and Technical Conference, CSIT 2016, 2016, art. no. 7589867, pp. 55-58.
24. N. Jaworski, I. Farmaha, U. Marikutsa, T. Farmaha, V.Savchyn, Implementation features of wounds visual comparison subsystem, 2018 14th International Conference on Perspective Technologies and Methods in MEMS Design, MEMSTECH - Proceedings, 2018, pp. 114-117.

25. A. Roshanzamir, H. Aghajan, M. Soleymani Baghshah, Transformer-based deep neural network language models for Alzheimer's disease risk assessment from targeted speech, BMC Medical Informatics and Decision Making, 2021, 21 (1), art. no. 92.
26. M. Sazhok, V. Robeiko, R. Seliukh, D. Fedoryn, O. Yukhymenko, Written Form Extraction of Spoken Numeric Sequences in Speech-to-Text Conversion for Ukrainian, CEUR workshop proceedings, 2020, 2604, 442-451.
27. N. Shakhovska, O. Basystiuk, K. Shakhovska, Development of the Speech-to-Text Chatbot Interface Based on Google API, CEUR Workshop Proceedings, 2019, Vol-2386, 212-221.
28. K. Tymoshenko, V. Vysotska, O. Kovtun, R. Holoshchuk, S. Holoshchuk, Real-time Ukrainian text recognition and voicing, CEUR Workshop Proceedings. 2021. Vol-2870, pp. 357-387.
29. M.S. Börjesson, L. Hartelius, K. Laakso, Communicative Participation in People with Amyotrophic Lateral Sclerosis, Folia Phoniatrica et Logopaedica, 2021, 73 (2), pp. 101-108.
30. N. Sharonova, V. Lytvyn, O. Cherednichenko, Y. Kupriianov, O. Kanishcheva, T. Hamon, N. Grabar, V. Vysotska, A. Kowalska-Styczen, I. Jonek-Kowalska, Preface: Computational Linguistics and Intelligent Systems, CEUR Workshop Proceedings, Vol-2870, 2021.
31. N.R. Kippin, S. Leitao, A. Finlay-Jones, J. Baker, R. Watkins, The oral and written narrative language skills of adolescent students in youth detention and the impact of language disorder, Journal of Communication Disorders, 2021, 90, art. no. 106088.
32. A. Abdulkareem, T.E. Somefun, O.K. Chinedum, F. Agbetuyi, Design and implementation of speech recognition system integrated with internet of things, International Journal of Electrical and Computer Engineering, 2021, 11 (2), pp. 1796-1803.
33. I. Svensson, T. Nordström, E. Lindeblad, S. Gustafson, M. Björn, C. Sand, G. Almgren/Bäck, S. Nilsson, Effects of assistive technology for students with reading and writing disabilities, Disability and Rehabilitation: Assistive Technology, 2021, 16 (2), pp. 196-208.
34. S. Ahmad, S. Asmai, S. Zaid, N. Kama, Shopping Assistant App For People With Visual Impairment: An Acceptance Evaluation, International Journal of Computing, 2019, 18, 285-292.
35. I. Khomytska, V. Teslyuk, Modelling of phonostatistical structures of the colloquial and newspaper styles in English sonorant phoneme group, in: Proceedings of the 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies, 2017, pp. 67 70.
36. I. Khomytska, V. Teslyuk, Modelling of phonostatistical structures of English backlingual phoneme group in style system, in: 14th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics, CADSM - Proceedings, 2017, pp. 324-327.
37. I. Khomytska, V. Teslyuk, Authorship Attribution by Differentiation of Phonostatistical Structures of Styles, in: IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT - Proceedings, 2018, 2, pp. 5-8.
38. I. Khomytska, V. Teslyuk, A. Holovatyy, O. Morushko, Development of methods, models, and means for the author attribution of a text, volume 3(2-93) of Eastern-European Journal of Enterprise Technologies, 2018, pp. 41-46.
39. I. Khomytska, V. Teslyuk, Authorship and Style Attribution by Statistical Methods of Style Differentiation on the Phonological Level, volume 871 of Advances in Intelligent Systems and Computing, 2019, pp. 105-118.
40. B. Rusyn, L. Pohreliuk, A. Rzheuskyi, R. Kubik, Y. Ryshkovets, L. Chyrun, S. Chyrun, A. Vysotskyi, V.B. Fernandes, The mobile application development based on online music library for socializing in the world of bard songs and scouts' bonfires, volume 1080 of Advances in Intelligent Systems and Computing, 2020, pp. 734-756. DOI: 10.1007/978-3-030-33695-0_49
41. A. Batyuk, V. Voityshyn, Apache storm based on topology for real-time processing of streaming data from social networks, in: Proceedings of the 2016 IEEE 1st International Conference on Data Stream Mining and Processing, DSMP, 016, pp. 345-349. DOI: 10.1109/DSMP.2016.7583573
42. J. Su, V. Vysotska, A. Sachenko, V. Lytvyn, Y. Burov, Information resources processing using linguistic analysis of textual content, in: Proceedings of the International Conference on Intelligent Data Acquisition and Advanced Computing Systems Technology and Applications, Romania, 2017, pp. 573-578. DOI: 10.1109/IDAACS.2017.8095038
43. V. Lytvyn, V. Vysotska, O. Veres, I. Rishnyak, H. Rishnyak, Content linguistic analysis methods for textual documents classification, in: Proceedings of the International Conference on Computer

Sciences and Information Technologies, CSIT, 2016, pp. 190-192. DOI: 10.1109/STC-CSIT.2016.7589903
44. V. Vysotska, Linguistic Analysis of Textual Commercial Content for Information Resources Processing, in: Modern Problems of Radio Engineering, Telecommunications and Computer Science, TCSET, 2016, pp. 709-713. DOI: 10.1109/TCSET.2016.7452160
45. V. Vysotska, V. Lytvyn, Y. Burov, P. Berezin, M. Emmerich, V. B. Fernandes, Development of Information System for Textual Content Categorizing Based on Ontology, volume Vol-2362 of CEUR Workshop Proceedings, 2019, pp. 53-70.
46. R. Bekesh, L. Chyrun, P. Kravets, A. Demchuk, Y. Matseliukh, T. Batiuk, I. Peleshchak, R. Bigun, I. Maiba, Structural modeling of technical text analysis and synthesis processes, CEUR Workshop Proceedings, 2020, 2604, pp. 562-589
47. A. Yarovyi, D. Kudriavtsev, Method of Multi-Purpose Text Analysis Based on a Combination of Knowledge Bases for Intelligent Chatbot, CEUR Workshop Proceedings, 2021, 2870, 1238-1248.
48. Y. Bodnia, M. Kozulia, Web Application System of Handwritten Text Recognition, CEUR Workshop Proceedings, 2021, Vol-2870, pp. 1323-1337.
49. A. B. Rianto, Mutiara, E.P. Wibowo, P.I. Santosa, Improving the accuracy of text classification using stemming method, a case of non-formal Indonesian conversation, Journal of Big Data, 2021, 8 (1), art. no. 26.
50. O. Kuropiatnyk, V. Shynkarenko, Text Borrowings Detection System for Natural Language Structured Digital Documents, CEUR workshop proceedings, 2020, Vol-2604, 294-305.
51. O. Cherednichenko, N. Babkova, O. Kanishcheva, Complex Term Identification for Ukrainian Medical Texts, CEUR Workshop Proceedings, 2018, Vol-2255, 146-154.
52. A. Berko, V. Andrunyk, L. Chyrun, M. Sorokovskyy, O. Oborska, O. Oryshchyn, M. Luchkevych, O. Brodovska, The Content Analysis Method for the Information Resources Formation in Electronic Content Commerce Systems, CEUR Workshop Proceedings, 2021, 2870, pp 1632-1651
53. V. Kuchkovskiy, V. Andrunyk, M. Krylyshyn, L. Chyrun, A. Vysotskyi, S. Chyrun, N. Sokulska, I. Brodovska, Application of Online Marketing Methods and SEO Technologies for Web Resources Analysis within the Region, CEUR Workshop Proceedings, 2021, 2870, pp 1652-1693.
54. N. Antonyuk, L. Chyrun, V. Andrunyk, A. Vasevych, S. Chyrun, A. Gozhyj, I. Kalinina, Y. Borzov, Medical news aggregation and ranking of taking into account the user needs, CEUR Workshop Proceedings, 2019, 2488, pp. 369-382.
55. V. Andrunyk, A. Vasevych, L. Chyrun, N. Chernovol, N. Antonyuk, A. Gozhyj, V. Gozhyj, I. Kalinina, M. Korobchynskyi, Development of information system for aggregation and ranking of news taking into account the user needs, CEUR Workshop Proceedings, 2020, 2604, 1127-1171.
56. V. Andrunyk, L. Chyrun, V. Vysotska, Electronic content commerce system development, in: Proceedings of 13th International Conference: The Experience of Designing and Application of CAD Systems in Microelectronics, CADSM, 2015.
57. A. Demchuk, B. Rusyn, L. Pohreliuk, A. Gozhyj, I. Kalinina, L. Chyrun, N. Antonyuk, Commercial content distribution system based on neural network and machine learning, CEUR Workshop Proceedings, 2019, 2516, pp. 40-57.
58. L. Chyrun, V. Andrunyk, L. Chyrun, A. Gozhyj, A. Vysotskyi, O. Tereshchuk, N. Shykh, V. Schuchmann, The Electronic Digests Formation and Categorization for Textual Commercial Content, CEUR Workshop Proceedings, 2021, Vol-2870, pp. 1816-1831.
59. A. Berko, I. Pelekh, L. Chyrun, M. Bublyk, I. Bobyk, Y. Matseliukh, L. Chyrun, Application of ontologies and meta-models for dynamic integration of weakly structured data, in: Proceedings of the IEEE 3rd International Conference on Data Stream Mining and Processing, DSMP, 2020, pp. 432-437. DOI: 10.1109/DSMP47368.2020.9204321
60. A. Berko, I. Pelekh, L. Chyrun, I. Dyyak, Information resources analysis system of dynamic integration semi-structured data in a web environment, in: Proceedings of the IEEE 3rd International Conference on Data Stream Mining and Processing, DSMP, 2020, pp. 414-419. DOI: 10.1109/DSMP47368.2020.9204101
61. I. Pelekh, A. Berko, V. Andrunyk, L. Chyrun, I. Dyyak, Design of a system for dynamic integration of weakly structured data based on mash-up technology, in: Proceedings of the IEEE 3rd International Conference on Data Stream Mining and Processing, DSMP, 2020, pp. 420-425. DOI: 10.1109/DSMP47368.2020.9204160