


## Article

# The Use of Voice Control in 3D Medical Data Visualization Implementation, Legal, and Ethical Issues

Miklos Vincze<sup>1,\*</sup>, Bela Molnar<sup>2</sup> and Miklos Kozlovsky<sup>3,4</sup> <sup>1</sup> AR/VR Laboratory, John von Neumann Faculty of Informatics, Obuda University, 1034 Budapest, Hungary<sup>2</sup> Image Analysis Department, 3DHISTECH Ltd., 1141 Budapest, Hungary; bela.molnar@3dhitech.com<sup>3</sup> Applied Cyber-Medical Systems Research Team, Laboratory of Parallel and Distributed Systems, Institute for Computer Science and Control (SZTAKI), Hungarian Research Network (HUN-REN), 1518 Budapest, Hungary; kozlovsky.miklos@sztaki.hu<sup>4</sup> BioTech Research Center, Obuda University, 1034 Budapest, Hungary; kozlovsky.miklos@nik.uni-obuda.hu

\* Correspondence: miklos.vincze@uni-obuda.hu

**Abstract:** Voice-controlled devices are becoming increasingly common in our everyday lives as well as in medicine. Whether it is our smartphones, with voice assistants that make it easier to access functions, or IoT (Internet of Things) devices that let us control certain areas of our home with voice commands using sensors and different communication networks, or even medical robots that can be controlled by a doctor with voice instructions. Over the last decade, systems using voice control have made great progress, both in terms of accuracy of voice processing and usability. The topic of voice control is intertwined with the application of artificial intelligence (AI), as the mapping of spoken commands into written text and their understanding is mostly conducted by some kind of trained AI model. Our research had two objectives. The first was to design and develop a system that enables doctors to evaluate medical data in 3D using voice control. The second was to describe the legal and ethical issues involved in using AI-based solutions for voice control. During our research, we created a voice control module for an existing software called PathoVR, using a model taught by Google to interpret the voice commands given by the user. Our research, presented in this paper, can be divided into two parts. In the first, we have designed and developed a system that allows the user to evaluate 3D pathological medical serial sections using voice commands. In contrast, in the second part of our research, we investigated the legal and ethical issues that may arise when using voice control in the medical field. In our research, we have identified legal and ethical barriers to the use of artificial intelligence in voice control, which need to be answered in order to make this technology part of everyday medicine.

**Keywords:** medical voice controlling; voice assistant; ethical voice assistant; legal voice assistant; voice control in 3D environment; voice control in VR



Academic Editor: Haridimos Kondylakis

Received: 20 November 2024

Revised: 18 December 2024

Accepted: 27 December 2024

Published: 30 December 2024

**Citation:** Vincze, M.; Molnar, B.; Kozlovsky, M. The Use of Voice Control in 3D Medical Data Visualization Implementation, Legal, and Ethical Issues. *Information* **2025**, *16*, 12. <https://doi.org/10.3390/info16010012>

**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

### 1.1. The State of Art of the Voice Assistant Systems

Nowadays, the use of voice assistants and various systems based on voice control is almost commonplace. More and more Internet of Things (IoT) systems are offering the functionality for the user to control various actions with their voice. A good example is the spread of devices that are entirely voice-controlled. These include Amazon Echo, Google Home, and Apple HomePod. In addition to these solutions, most phones now have some form of voice assistant. These include Siri, Cortana, and Google Now. In this area, users can access almost any functionality of their phone just by using their voice.

The most common part of any voice-based program is the dictionary. This determines how well the software will understand the words, phrases, and sentences that the user says. Today, a large proportion of voice control systems are language-independent, meaning that the user can communicate with the software in several languages. Different languages are understood and processed by so-called voice models. These models are trained by different service providers using a huge sample base to be able to provide Automatic Speech Recognition (ASR).

### Application of Voice Control in Various Fields

Today, users use voice control in a wide variety of areas [1,2]. One of the biggest areas of use for voice control is smart phones and communication with various Internet of Things devices.

With voice-based control solutions, it is possible to control devices in our home wirelessly. With the help of various actuators, controllers, and sensors, such solutions can provide voice-based control of our home. In their paper [3], the authors present a solution that is based on the so-called Google Assistant system, and with the help of which the user can turn on and off a ventilation system with voice commands. These and similar solutions enable voice control of our home, which can make life easier for people who cannot do this otherwise due to their physical limitations.

Another prominent area of application of voice control is the various voice assistants in smartphones. With this software, users can use their voice to access almost any function provided by their mobile phone. Nowadays, solutions such as Google Assistant, Amazon Alexa, or Apple Siri can be found in almost every smartphone. Many scientific studies deal with the accuracy, usability, and usefulness of these solutions [4,5]. The usability of so-called Smart Personal Assistant (SPA) systems has many influencing factors, such as processing time and accuracy. Based on the authors' research [5], the Google Assistant system proved to be the best in both accuracy (95%) and speed (0.62 s).

How voice control can be used in medicine is an area of active research [2,6,7]. Several studies [8–10] have shown that productivity can be improved with voice-based control, voice processing, or even an artificial intelligence assistant. Medical surgeries are also an area in which active research is being conducted regarding the use of voice control. In the case of robotic surgery, the use of voice control allows the surgeon to have his hands-free during surgery, so he can focus on principal tasks. The authors' research reveals that several requirements must be met to use voice control in robotic surgery. Examples include a low delay between issuing the command and its execution, as short as possible specific commands, and providing the surgeon with the right image [11].

### 1.2. Guidelines to Follow When Designing Voice Control

#### 1.2.1. The ISO 9242-11 Standard

The usability of the interface between the user and the software is described in ISO 9242-11 [12]. By applying this standard, it is possible to enable users to effectively use software that includes voice control. The standard emphasizes the importance of usability, both for software solutions that are used frequently and for software solutions that are used infrequently. In both cases, it is important that the user can perform his or her task quickly and efficiently, to which the use of the current software contributes. In addition, by improving usability, software errors that may result from the inappropriate use of the software can be avoided. This standard is used in many cases in the literature to define the usability of software [13–15].

#### 1.2.2. The ISO/IEC 30122 Standard

The ISO 30122 [16] family of standards addresses the principal voice commands commonly used with various information and communication technology (ICT) devices.

Due to the rapid development of speech recognition, it was necessary to further develop and extend the existing standards. Thus ISO/IEC 30122-2 was born, which focuses on technical criteria and test procedures for voice control systems [17].

### 1.2.3. The Guidelines of the European Union

The European Data Protection Board (EDPB) has published the second version of the guideline on virtual voice assistants (VVA) in 2021. The document contains the technological background that a system offering voice assistants must have. It also defines the aspects of data protection and the limits of data availability from the point of view of the VVA system. According to the document, VVA systems must primarily comply with the GDPR [18] regarding the storage and handling of user data. The document also highlights the importance of applying the e-Privacy Directive.

### 1.2.4. The Guidelines of the Health Insurance Portability and Accountability Act (HIPAA)

The HIPAA rules include strict privacy standards that all medical apps, including AI and voice-controlled systems, must follow [19]. The focus is on privacy of health care information, controlling access rights, third-party responsibilities, and transparent data management procedures. AI systems in the medical environment must operate in compliance with HIPAA regulations to ensure patient privacy and regulatory compliance.

One of the main purposes of the HIPAA Privacy Rule is to ensure that an individual's health information is adequately protected while allowing the flow of information necessary to promote high-quality health care [20].

## 1.3. Problem Definition and Motivation

Our research in this paper can be broken down into two major parts. The first is the design, implementation, and testing part, in which we developed a voice control solution for an existing virtual reality (VR) medical data visualization software. In the second part of our research, we investigated the legal and ethical implications of using voice control. We summarized the issues that may arise when implementing voice control. The main objectives of our research are

- **Research Objective 1:** Defining the state of the art of voice control systems.
  - People are increasingly using different voice control systems in their daily lives. Currently, there are many different manufacturers offering several different solutions for voice control, which differ in both functionality and implementation.
- **Research Objective 2:** Design and development of a voice control system to control the visualization of digitalized medical samples in a virtual reality environment.
  - The goal of our research was to implement a system that can speed up the processing of medical samples using voice control. By using voice control, it would be possible in the future to design a system where the user would not need to have any other input device than the microphone built into the virtual reality headset.
- **Research Objective 3:** Formulation and summary of legal and ethical issues arising from the design and development of voice control.
  - When using software solutions that use voice control, there are many details that need to be considered that are not present when using traditional input peripherals. Examples include the continuous processing of the speech heard or the saving of words spoken by the user. In our research, we examined several legal and ethical issues that may affect the use of this technology in medical devices.

## 2. Materials and Methods

In our research, we used anonymized, digitized pathological samples to visualize the results of voice commands. These samples built up a complete pathology serial section. The pathological samples used in our research were digitized using a p1000 hardware device manufactured by 3DHISTECH in Budapest, Hungary. The serial section used in this study is made up of 122 pathological samples, each with a resolution of 26,880 pixels  $\times$  32,256 pixels. Our voice control solution, designed and developed in PathoVR software [21], is based on a free model [22] taught by Google. We did not have access to further teaching of the model or the teaching dataset during our research.

In order to create the designed system, we used a graphics engine called Godot. With the help of this graphics engine, we were able to create a 3D environment in which the user can use the implemented voice control system. In our research, we used an Oculus Quest Pro VR device. We used the microphone of the device as the input peripheral for the voice control and the speaker of the device to play different audio feedback.

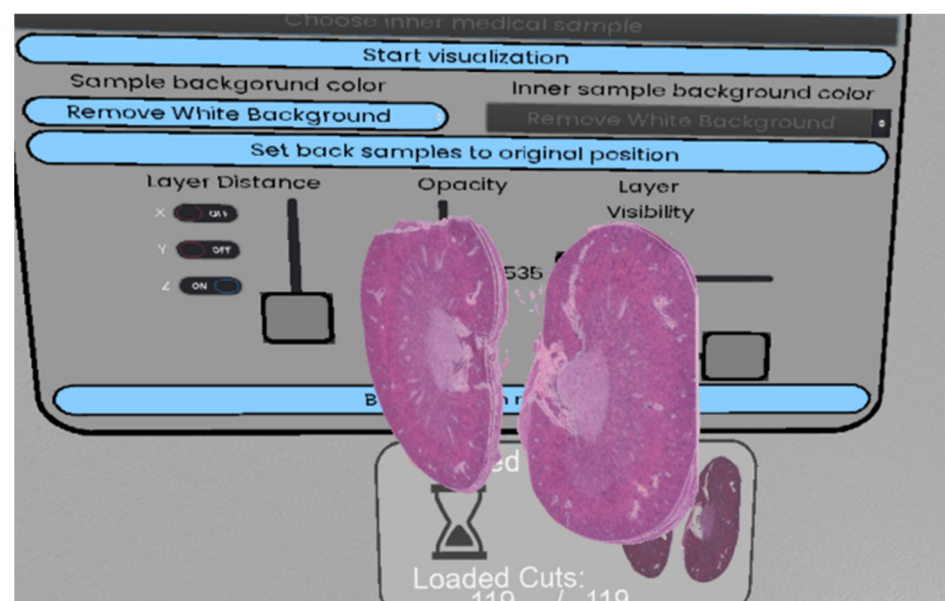
The volunteers in the test were all first-time users, with no prior learning phase. The testers all speak English daily, so they had no difficulty in giving voice commands in English.

## 3. Results

### 3.1. The PathoVR Software

PathoVR [23] is a software solution that can use VR technology to provide a new 3D visualization for the evaluation of medical samples. The software developed by us can process medical samples from different imaging modalities (CT, PET, MRI) and visualize them in a 3D VR environment as a tomogram. Figure 1 shows an example of a 3D visualization of the samples. By using our software, the user can examine the visualized medical sample from any viewpoint in 3D space. In addition, the use of VR technology allows the user to examine the internal structure of a 3D sample with a single head movement. Software that implements 2D visualization in everyday medical practice provides users with a range of functionalities aimed at making diagnoses more accurate and faster. Such functionalities include

- Creating voice memo,
- Place markers,
- Zooming.



**Figure 1.** Three-dimensional Visualization of a 3D kidney sample made by PathoVR software.

One of the most recent results of our research was the implementation of voice control, the results of which are presented in this paper.

### 3.2. The Voice Control System of the PathoVR Software

By providing voice control, we can make our software easier and more comfortable to use, which should result in users being able to spend more time in virtual reality. By using voice control, it is possible to replace the use of controllers in virtual reality software so that the user does not have to learn new control methodologies. This also makes the learning curve associated with proper use of the software less steep. Before implementing voice control, we need to answer important questions such as:

- Does the expected system understand only specific commands, or can it also process whole conversations?
- Does the system execute the command immediately after the command is spoken and understood, or does it wait for a confirmation from the user?
- Will the system use an online database that has already been validated and taught, or will it use its own local database created specifically for this purpose?
- Do we want to cover full control of the software with voice control or only certain functionalities?

The answer to the questions in the list will determine the voice control developed. The answers to these questions, in relation to the development presented in the paper, are discussed in Section 3.2.1, Section 3.2.2, and Section 3.2.3, respectively.

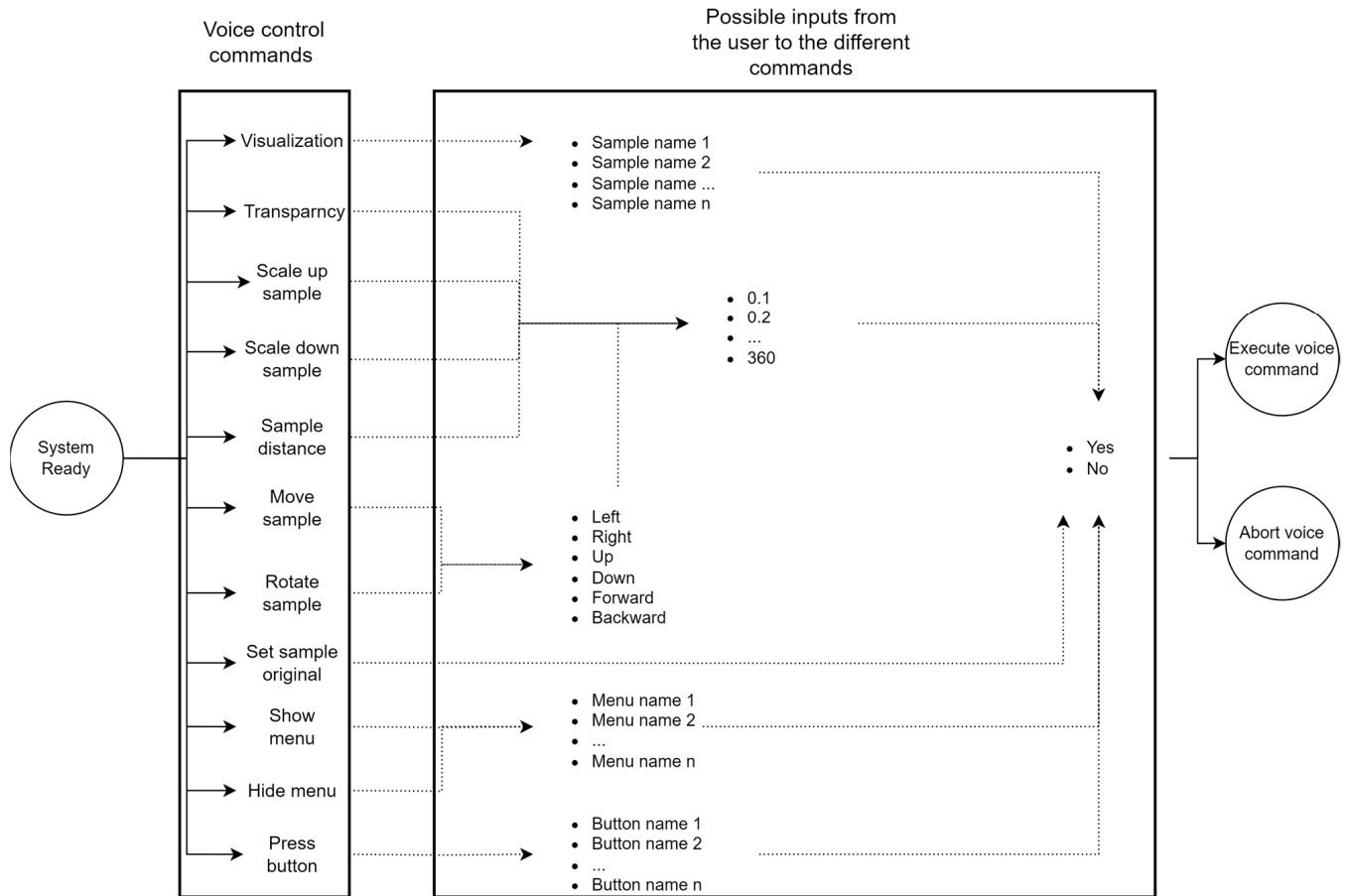
#### 3.2.1. The Implemented Voice Control System

In the current PathoVR system, a solution based on an online database has been developed. An important detail about our solution is that only certain functionalities are available in the PathoVR software using voice control. Such functionalities are the following:

- Display various 3D medical samples,
- Moving a 3D medical model,
- Change transparency of 3D medical sample,
- Rotate sample in the 3D space,
- Move sample back to its original position,
- Show/hide menu system,
- Scaling up/down a 3D medical model.

Figure 2 shows the commands understood by the developed voice control and their possible input parameters. Figure 2 also shows that the user is always able to abort the execution process.

As shown in Figure 2, the user can currently access 11 functionalities within the PathoVR system without the use of controllers, using only the voice control system. As shown in Figure 2, each functionality can have different input parameters that the user must specify when using voice control. An example of this can be the rotation of the displayed medical sample in 3D space, as shown in the figure. In this case, the user has to say the command for the rotation, the direction of rotation, and the angle of rotation. If these are correctly understood by the system, the user will issue the “Yes” command, giving the final permission to the renderer to execute the command. From this example, in the system shown, it is always up to the user to decide whether the system has understood the correct command with the correct input parameters. The issued command is executed only if the user gives final permission for it. This security step is to filter out wrong executions that could interfere with the result of an evaluation performed in the PathoVR software.



**Figure 2.** The possible voice control commands interpreted by the PathoVR system and the input parameters required for them.

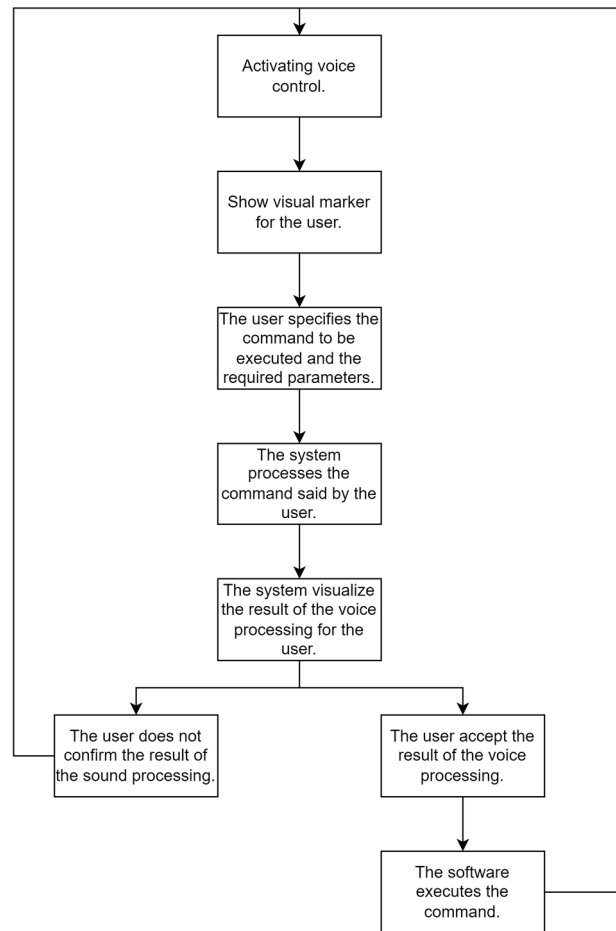
The presented solution works by using an online database, but it has many advantages and some negative features. For example:

- Only available with an internet connection.
- The user will not be notified that a change has been made to the database used for the model.

The second element of the list can raise serious safety issues if voice control is implemented in a medical device.

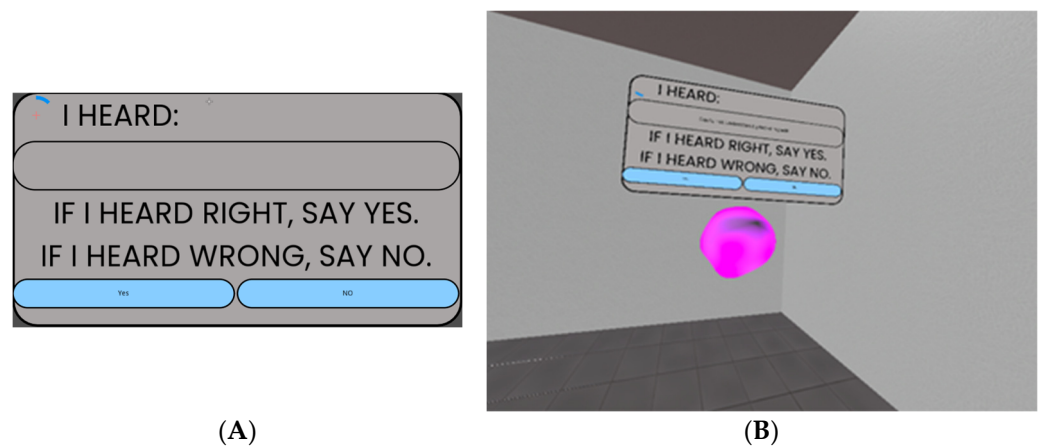
The voice control system implemented in the PathoVR software does not continuously check the audio data provided by the user. This is only carried out if the user gives permission by pressing a specific button. Figure 3 shows the steps in the operation of the implemented voice control.

As can be seen in Figure 3, the voice control implemented in our research can only be used with bound steps. This ensures that the user is always informed about the current state of the voice data processing and minimizes the number of errors due to improper use.



**Figure 3.** Basic steps for using voice control.

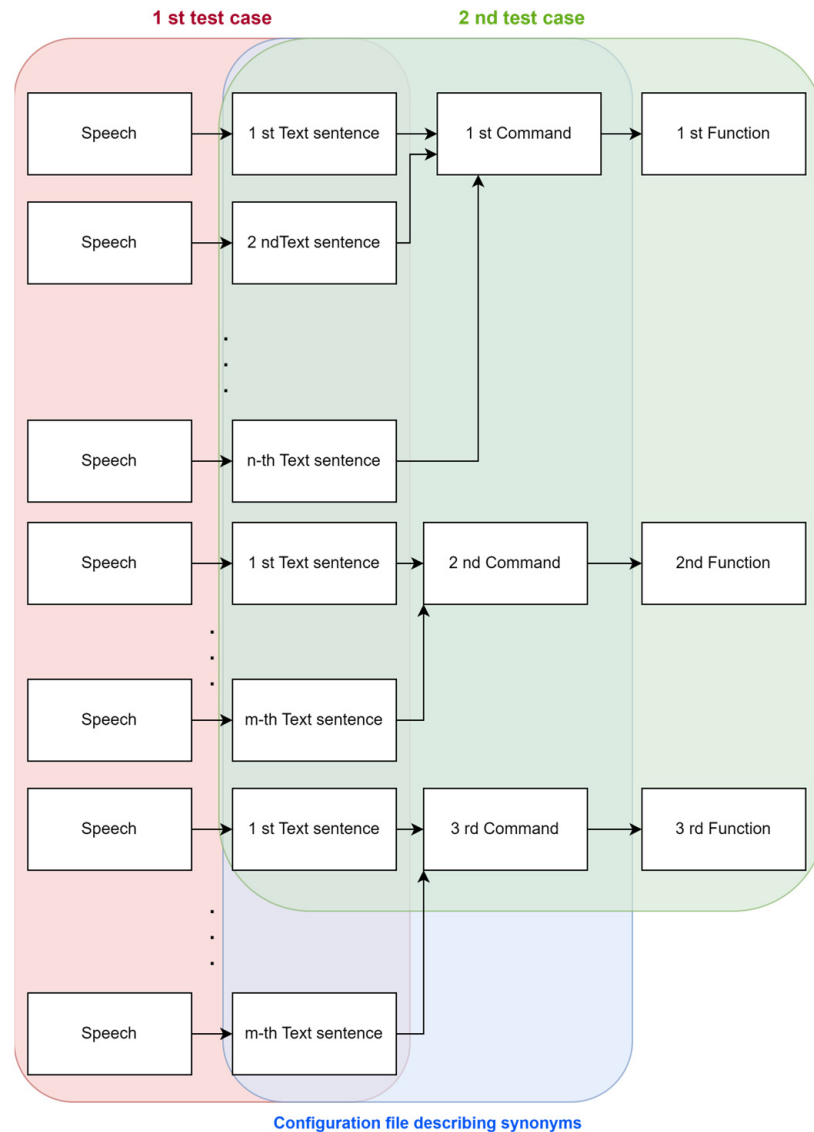
Our solution provides the user with continuous audio-visual feedback from our system to ensure that they always know the status of the voice control subsystem. In the PathoVR software, we have implemented a separate menu system to display only the status of the voice control and its result. In addition, an animated 3D sphere model is placed below the menu system to provide visual feedback to the user that the voice control is active. This can be seen in Figure 4.



**Figure 4.** The menu system, that connects to voice control in the PathoVR software. **(A)** Two-dimensional plan of the PathoVR software voice control menu. **(B)** The PathoVR software voice control menu in use.

### 3.2.2. Handling Synonyms in the Implemented Voice Control System

In order to allow the user to issue commands not only from a predefined set of voice commands, we have created a synonym system. In this solution, we have organized the synonyms for each instruction in a descriptor file, which from now on point to the same method to be executed. With the solution described, any number of synonyms can be assigned to a method to be executed, allowing the user to customize the voice control operation. With our solution, the user can remove or add new synonyms to the system at any time, and this new data will be taken into account by the voice control the next time the software is started. The synonym system is described in Figure 5.



**Figure 5.** Linking synonyms with their associated command and method.

As shown in Figure 5, the user can associate any number of synonyms with a given voice command. By using these synonyms, the user does not need to learn new software-specific commands and can associate any number of voice commands with each method executed. We split our voice control system into two parts for the testing presented in the paper.

In the first part, we tested the accuracy of the system in translating verbal commands issued by the user into written form. This part of the system is shown in red in Figure 5, and the corresponding test results are given in Section 3.2.3. In the second test case, we tested the assignment of each synonym to the corresponding command and then the execution



of the corresponding function. This part of the presented system is marked in green in Figure 5, and the corresponding test results are given in Section 3.2.4.

### 3.2.3. Testing the Accuracy of the Implemented Voice Control System

To test the accuracy and the usability of our implementation, we have created a testing environment. In this environment, test subjects were asked to issue predefined verbal commands, and then we tested the accuracy with which the system understood the issued instructions. The system was tested by three users, each issuing all the commands three times. The commands used during testing could be issued to the system in any order. The accuracy of six commands was tested. The number of commands understood by the system is shown in Table 1.

**Table 1.** The number of commands tested and their correct/incorrect detection.

Command	Number of Correct Understandings of Issued Command	Total Number of Times the Command Has Been Issued
Visualize sample mouse embryo.	6	9
Change transparency to 0.5.	9	9
Scale up sample.	7	9
Scale down sample.	5	9
Rotate sample backward.	3	9
Set sample original.	7	9

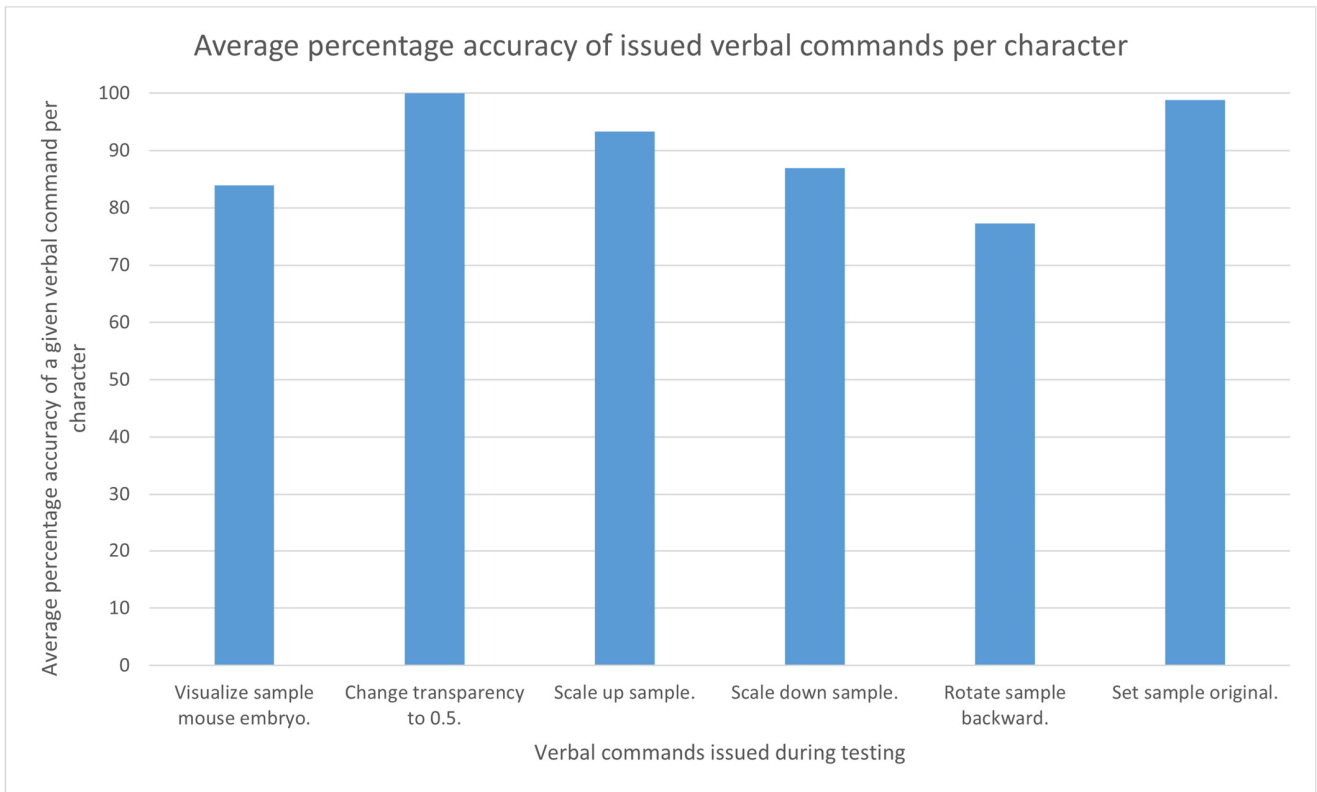
To test the per-character accuracy of the system as interpreted, each character was compared to the character in the same position in the original command. This gave us the percentage of characters in the correct position for the entire command. This data were used to produce a summary showing which commands were interpreted with what accuracy by the system. Equation (1) was used in our research to determine the accuracy of the commands issued.

$$C_H = \frac{C_C}{C_T} \quad (1)$$

The notations in the equation mean the following:

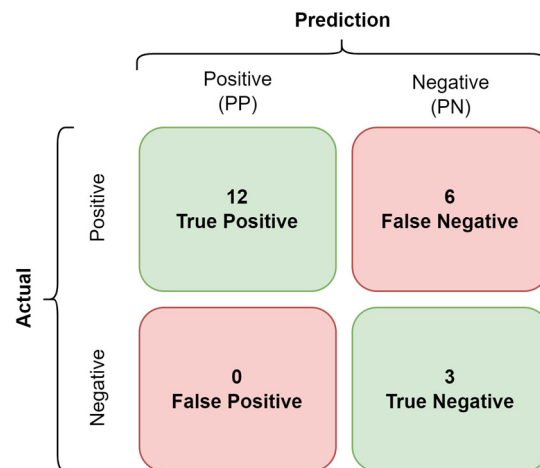
- $C_H$ : The percentage accuracy of a given command.
- $C_C$ : The proportion of characters that appear in the correct position in the given command.
- $C_T$ : The number of characters in the given command.

The average per-character accuracy of the commands tested is shown in Figure 6. As we can see in Figure 6, the accuracy per character did not drop below 70 percent for any of the commands tested. Regarding the testing results, it should be noted that in our solution, we only accept a command issued by the user if it is 100 percent identical to a predefined command. This constraint was introduced because it completely avoids incorrect modification of 3D medical samples.



**Figure 6.** The average accuracy per character of the verbal commands tested.

To describe how our system works and to perform statistical calculations on our tests, we have created a confusion matrix for our system [24]. Using this solution, we determined the number of cases in our tests where our solution worked as expected and the number of cases where we received an incorrect result. The values achieved during testing are shown in Figure 7.



**Figure 7.** Confusion matrix of the presented voice control system.

To test the correct/incorrect perception of our voice controls and the resulting operation, we have distinguished 4 classes. These are as follows:

- True Positive (TP): In the test, we have included cases where the goal was for the system to detect the issued command, and it did.
- True Negative (TN): In the test, we included cases where the goal was to ensure that the system did not detect the issued command, and it did.

- False Positive (FP): In the test, we have included cases where the goal was not to detect an issued command, but the system did.
- False Negative (FN): In the test, we included cases where the goal was for the system to detect an issued command, but it was not detected.

The confusion matrix was used to determine the accuracy, the balanced accuracy, and the  $F_1$ -score of our system. The equation of the confusion matrix-based accuracy measurement can be seen in Equation (2).

$$ACC = \frac{TP + TN}{P + N} \quad (2)$$

The notations in the equation mean the following:

- ACC: The accuracy of the voice control system.
- TP: The number of true positive detections.
- TN: The number of true negative detections.
- P: Number of positive measurements during testing.
- N: Number of negative measurements during testing.

Using these equations, the accuracy of our system is 0.71.

Since one target class occurs more often than the other in our measurements, we also determined the balanced accuracy during testing. The equation used to calculate balanced accuracy is shown in Equation (3).

$$BA = \frac{TPR + TNR}{2} \quad (3)$$

The notations in the equation mean the following:

- BA: The balanced accuracy of the system.
- TPR: True positive rate. Shows how many of the total number of positive results the system returned correctly.
- TNR: True negative rate. Specifies the ratio of the outputs correctly evaluated as negative by the system compared to the total negative outputs.

Using this equation, the balanced accuracy of our system is 0.83.

In our research, we determined the  $F_1$  score of our system to give the statistical accuracy of a test. Equation (4) was used to determine the  $F_1$  score.

$$F_1 = \frac{2 TP}{2 TP + FP + FN} \quad (4)$$

The notations in the equation mean the following:

- TP: The number of true positive detections
- FP: The number of false positive detections.
- FN: The number of false negative detections.

Based on the measurements, the  $F_1$  score of our system is 0.8.

In our tests, the model we used achieved 71% accuracy, 83% balanced accuracy, and an 80%  $F_1$  score. Our future goals include using a more accurate model.

#### 3.2.4. The Testing of the Implemented Synonym System

In the voice control system described in this paper, 5 synonyms are currently available to the user to start a new 3D pathology sequence visualization. During testing, we examined whether the correct command is executed for the different synonyms spoken by the user, i.e., whether the 3D visualization is started or the transparency value is set. The test results for the synonyms are shown in Table 2.

**Table 2.** The results of the synonym testing.

Test ID	The Synonym Said by the Tester	Expected Launched Command (Ground Truth)	Observed Launched Command
1	start sample visualization	start visualization	start visualization
2	sample visualization	start visualization	start visualization
3	visualize	start visualization	start visualization
4	show sample	start visualization	start visualization
5	add 3D sample	-	-
6	show me sample	start visualization	start visualization
7	transparent	set transparency	set transparency
8	make transparent	set transparency	set transparency
9	hide white parts	set transparency	set transparency
10	make white parts disappear	-	-

As we can see in the table, in the case of synonym testing, we tested synonyms for two different functions. The functions tested and the expected result of the test are shown in the third column of Table 2. The first command tested was start visualization, which allows the user to start a 3D visualization of medical samples in a VR environment. The second command is set transparency, which allows to set the transparency of already visualized 3D medical samples.

The table shows that for the synonyms that were placed in the configuration file, the expected result was achieved after testing. This is shown in rows 1–4 and rows 6–9 of Table 2. During testing, we also tested synonyms that were not placed in the configuration file. These test cases are shown in rows 5 and 10 of Table 2. In these cases, we can see that our solution did not perform any functions as expected.

### 3.3. Legal Views on the Use of Voice Control

The use of voice control software can bring many benefits to the user. Such advantages may include ease of use, a less steep learning curve, or the elimination of the need to use different input peripherals. On the other hand, voice control raises legal issues that, without being addressed, make it impossible to use this technology in systems such as a medical technique device [25–28].

#### 3.3.1. The Doctor or the AI Is the Final Decision Maker?

An important question about voice control is who makes the final decision, the doctor or the software that interprets the instructions and carries them out. Nowadays, artificial intelligence is gaining ground in decision support, but the literature recognizes that the final decision should remain in the hands of the physician [29,30]. This is no different for voice guidance. When designing artificial intelligence, transparency is an important aspect. This is also particularly important for doctors. Several scientific articles [31–33] mention that even though a doctor uses AI, he/she cannot hide behind it. In other words, both the doctor and your decision-making must remain transparent even when using AI.

#### 3.3.2. The Question of Feedback from the System

An important issue is the form and method of confirmation provided to the user. If the system does not provide the user with information about what it has understood from the voice-based instructions, the user should have full confidence in the software. One possible solution to this problem is the so-called Read Back systems, which read back to the user what the software has understood. With this solution, the user can obtain accurate confirmation of what the system has understood from the instructions given [34,35].

If the client receives data from the software about the processing of commands, it can take several forms. It can be audio, visual, or a combination of audio-visual feedback. The form this takes in the case of medical devices is subject to heavy regulation. In the case of audio-only feedback, a specific audio frequency or playback length must be used to convey the signal. Visual feedback should use predefined colors, depending on the seriousness of the event for the user. When using combined audio-visual feedback, the system must meet all these requirements [36].

### 3.3.3. The Question of Responsibility When Using AI

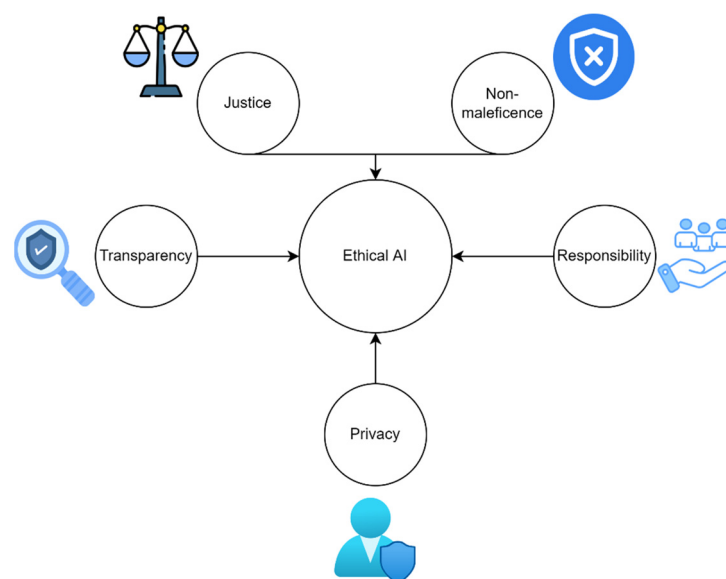
The use of artificial intelligence is challenging our society at many points. Just think of issues such as access to personal data or machine decision-making in acute situations. The use of AI in medicine is a hotly debated area. A decision or execution error in healthcare can have fatal consequences for the patient, so the utmost care must be taken when designing medical devices.

On the subject under consideration, the European Parliament's resolution is based on research commissioned, monitored, and published by the policy unit "Citizens' Rights and Constitutional Affairs" [37].

Unlike those working in the health sector, software/hardware systems cannot be held liable if they have an unforeseen impact on the patient during their operation. In principle, the currently accepted approach is that new frameworks and approaches are needed for the use of different systems using artificial intelligence. However, it is also important that the hospital staff who use them should be further trained in the system, as the ultimate responsibility lies with the person who uses the software. This is basically because AI tools help users to make decisions, speeding up their work but not replacing them. There is currently little literature on the regulation of the use of AI in the medical field [38,39], which is why it is important for policymakers to define the framework within which the benefits of AI can be safely exploited.

### 3.4. Ethical Issues in the Use of Voice Control Systems

The use of software systems that are constantly active and monitor what the user has said raises ethical concerns [40,41]. This is no different with voice control [42]. Ethical concerns about the use of voice control are an area of active research [41–43]. The 5 principals of an ethical AI can be seen in Figure 8.



**Figure 8.** The 5 principals of an ethical AI software.

### 3.4.1. Ethical Views on the Use of Artificial Intelligence in Medicine

Today, the use of artificial intelligence is gaining ground in both research and everyday use. This is no different in medicine [44–46]. The application of AI is leading to advances in many areas, such as image processing, analysis of huge amounts of data, or decision support. But beyond its unquestionable benefits, the ethical issues surrounding the use of the technology cannot be ignored [47–49].

Today, in the case of so-called Artificial Intelligent Systems (AIS), these solutions may be able to hide the actual operating mechanisms behind the system output. This creates a system whose output is virtually uncontrollable. In the future, solutions that use AI to process patient medical history, laboratory results, and complaints and then use this information to support decisions in patient management will become increasingly common. In the case of such systems, the use of so-called “black box” systems is questionable, as researchers, doctors, and patients do not know how the system arrived at the conclusions it produces [50,51].

### 3.4.2. The 5 Principles of an Ethical Voice Assistant System

Systems implementing voice control can be static in terms of their knowledge base, operating on an existing data set. In addition to this type, there is another type of voice control that can learn continuously. However, for continuous learning to take place, it is essential that the software’s data set is augmented. One solution to this is for the software to continuously save the words, phrases, and expressions used by the user and then use them to train itself. With this and similar solutions, the more and more people use the software, the more accurate the voice control system’s learning can be. This method is demonstrably efficient in terms of extending the sample system used for teaching, but the user must be constantly alert to what the software is saying as it is used.

In an article published in 2019 [49], the authors present 5 basic points that can be used to develop an ethical artificial intelligence (AI). Although not all software that offers voice control uses AI, even those that implement the functionality without using AI can benefit from adhering to these 5 points. The points mentioned in the article are

- Transparency,
- Justice, fairness and equity,
- Non-maleficence,
- Responsibility and accountability,
- Privacy.

When considering transparency in AI, attention should be paid to aspects such as the interpretability of the model used or the disclosure of training data, model bias, or performance. In the case of virtual assistants or voice control, there are several steps to improve transparency. For example, informing the user. The user should be informed about who the software is communicating with, and it is important to inform them about the way in which the data are stored and uploaded. By following these steps, the user can be aware of the consequences of using the software and can give his/her consent to the system using the data generated by the voice control for further teaching. Designing software that is transparent can improve the user experience, making it easier for users to adapt to the software.

Justice in AI means that they are not biased towards any group, product, or topic. This point is basically an issue for voice assistant solutions, where the user can initiate a continuous conversation with an AI model. For voice control systems where the user can issue commands in a much narrower system, the role of justice as a consideration point in development is minor.

The issue of responsibility and accountability is an important and unavoidable detail for voice-based control systems. This is where details such as the ownership of the data generated using the software should be considered. In the case of voice assistant systems,

this means acting fairly towards users, i.e., informing them about the ownership of the data generated. There are currently several different solutions for voice control, based on different business models. These different business models can, of course, also affect the way the data generated is handled and stored. The question arises whether the data generated when using a voice assistant system is shared by the company with, for example, advertising companies. In any case, the user should be informed about this activity and the handling of the data.

Privacy means the management of personal data, i.e., that the user's data are treated confidentially and is only shared with third parties with the user's consent. Many countries have laws requiring service providers to handle user data appropriately. In the case of data protection, several issues need to be addressed. These are, for example, the following:

- Where and how does the system store the data generated during voice control?
- Can the user permanently delete the data generated during the conversation?
- Is the communication channel encrypted, if so in what form?
- Who has access to the data generated during the conversation?
- Will the system use the new data to improve itself?
- Is the data stored locally or globally?

If service providers manage user data in a secure, transparent, and ethical way, this can improve user satisfaction with the software. Providing users with the right information can also help them to have confidence in the virtual assistant, so that they can feel more comfortable sharing their data if they know it is being handled in the right way.

#### 4. Discussion

The idea behind our research was to replace the use of controllers when displaying medical image data in 3D in a VR environment. To solve this problem, we applied voice control, which allows the user to display and modify medical data in 3D according to specific commands. Our research has also focused on the ethical and legal issues surrounding the use of voice control systems. In our work, we have highlighted standards and decisions that may help researchers/developers who wish to design similar systems in the future.

In Section 1.3, we have identified three key points along which the paper is built. The first of these was aimed at presenting and describing current solutions. The second is a presentation of our own research/development results, and the third presents the main ethical and legal constraints.

In our research, we presented the current state of voice control and its various applications. In addition, our own development results and test cases of these development results are presented in the paper. In the third part of our paper, we have presented the reader with the main ethical and legal issues in the application of voice control systems.

We plan to further develop our current solution in the future. The solution presented in this paper provides the possibility to replace the use of controllers with voice control. As a result, the learning curve of the software can be less steep, which can speed up the learning of the PathoVR software or any VR software. In the future, we plan to extend our voice control system with functionalities such as handling different synonyms, handling extended command set, and handling different languages.

The relationships between the claims made in the literature used in our research and the results of the presented study are shown in Table 3.

**Table 3.** Summary of the connection between the previous research on which the study is based and the results of the present study.

Findings from the Research Used in the Current Study	Results of the Actual Study
The HIPAA rules include strict privacy standards that all medical apps, including AI and voice-controlled systems, must follow [19]. The focus is on privacy of health care information, controlling access rights, third-party responsibilities, and transparent data management procedures.	Even though the research presented in this paper deals with 3D visualization of medical samples, it uses an online model that requires a continuous internet connection to use the speech-to-text function. This feature is one of the main weaknesses of the presented research, and its correction is part of future research.
The algorithmic predictive capability of AI is an input into the decision-making process, and the human expert's final decision (judgment) remains critical [30].	The article agrees with the findings of the cited research [30]. To achieve this, we have designed the presented solution in such a way that the user must always contribute to the execution of the issued command.
Pronunciation tests with Google Assistant are the most successful of the smart assistants [5].	Although we used a Google pre-trained model to interpret the spoken commands in our research, we achieved 71% accuracy and an 80% $F_1$ score. We plan to further improve the accuracy of our solution in the future.
The voice commands must be as short as possible [11].	We agree with this statement by the authors [11]. To minimize the number of errors in the application of voice control, the commands that can be issued and their parameters have been specified. In addition, however, it is possible to add synonyms to the configuration file to allow the user to personalize the system. The length and complexity of these synonyms is not limited.

In summary, the use of voice control in healthcare depends on several key points. One of these is the proper management of user data. Both the data generated during voice control and the medical data that the software works with must be handled in accordance with the standards described in this paper. A further key factor is the accuracy and usability of the application. To support these parameters, we added, for example, a freely extensible synonym system to our solution.

## 5. Conclusions

Today, software that uses some form of artificial intelligence is becoming more and more common. This is no different in medicine; just think of the various image-processing software. Nowadays, solutions that not only speed up the work of doctors but also make it more comfortable are also gaining ground. One such solution could be a voice assistant, where the user can issue various commands that the artificial intelligence then executes after understanding them. For such solutions, reliability is an important issue, as is the correct understanding of the heard text. In all cases, the software must execute the command or sequence of commands given by the doctor and nothing else.

In our research, we have designed and implemented a voice-based control system that allows the user to control 3D VR software. In this paper, we have presented the architecture of our solution and the functionalities that the user can currently access via voice commands.

In the second half of our research, we looked at the ethical and legal aspects of voice control and the sub-application of voice assistants. We examined the most important factors to consider from an ethical point of view in case we need to implement voice control in our system. We have reviewed the current literature on the use of systems that use artificial intelligence to understand auditory text. The literature reviewed for this article suggests that few academic articles address the legal implications of using artificial intelligence for speech recognition.



During our research, we were limited in the number of test participants we could include, so we would like to include more volunteers in the further development of the system presented.

The application of voice control is expected to remain an area of intense research in the future. Today, solutions that increase user efficiency and software usability are in high demand. The field of science presented in this paper is an actively developing sector, so every day new solutions are published by researchers. The number of VR software applications using artificial intelligence is constantly increasing. It is expected that solutions such as the voice assistants presented in this research will become part of everyday medical research in the future.

As a future continuation of our research, we aim to further investigate possible implementations of voice control in the medical field. An important goal in our research is to be able to create a virtual voice-based assistant in the future that can improve the usability of software but at the same time provide a safe alternative for the medical use of artificial intelligence.

**Author Contributions:** Conceptualization, M.V.; methodology, M.V.; software, M.V.; validation, M.V.; formal analysis, M.V. and M.K.; investigation, M.V.; resources, B.M. and M.K.; data curation, M.V.; writing—original draft preparation, M.V.; writing—review and editing, M.V.; visualization, M.V.; supervision, B.M. and M.K.; project administration, M.K.; funding acquisition, M.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research was supported by the 2024-2.1.1 University Research Scholarship Program of the Ministry for Culture and Innovation from the source of the National Research, Development, and Innovation Fund. The authors would like to thank the 2019-1.3.1-KK-2019-00007 “Innovációs szolgáltató bázis létrehozása diagnosztikai, terápiás és kutatási célú kiberorvosi rendszerek fejlesztésére” national project for the financial support.

**Institutional Review Board Statement:** Not applicable; the dataset was anonymized and did not include patient data or personal information.

**Informed Consent Statement:** Not applicable; the dataset was anonymized and did not include patient data or personal information.

**Data Availability Statement:** The medical serial sections and images used in this study cannot be made available at any request.

**Acknowledgments:** The authors would like to thank the AIAM (Applied Informatics and Applied Mathematics) doctoral school of Obuda University, Budapest, Hungary, for their support in this research. The research was supported by the 2024-2.1.1 University Research Scholarship Program of the Ministry for Culture and Innovation from the source of the National Research, Development, and Innovation Fund.

**Conflicts of Interest:** Author Bela Molnar was employed by the company 3DHISTECH Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Srinivasan, S.; Brown, E. Is speech recognition becoming mainstream? *Computer* **2002**, *35*, 38–41. [[CrossRef](#)]
2. Johnson, M.; Lapkin, S.; Long, V.; Sanchez, P.; Suominen, H.; Basilakis, J.; Dawson, L. A systematic review of speech recognition technology in health care. *BMC Med. Inf. Decis. Mak.* **2014**, *14*, 94. [[CrossRef](#)]
3. Yadav, S.; Kumar, P.; Kumar, S.; Singh, S. Google Assistant Controlled Home Automation with Voice Recognition. In Proceedings of the 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 17–18 December 2021. [[CrossRef](#)]
4. Brill, T.; Munoz, L.; Miller, R. Siri, Alexa, and other digital assistants: A study of customer satisfaction with artificial intelligence applications. In *The Role of Smart Technologies in Decision Making*; Routledge: London, UK, 2022. [[CrossRef](#)]

5. Isyanto, H.; Arifin, A.; Suryanegara, M. Performance of Smart Personal Assistant Applications Based on Speech Recognition Technology using IoT-based Voice Commands. In Proceedings of the 2020 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Republic of Korea, 21–23 October 2020. [CrossRef]
6. Mohr, D.; Turner, D.; Pond, G.; Kamath, J.; Vos, C.; Carpenter, P. Speech Recognition as a Transcription Aid: A Randomized Comparison with Standard Transcription. *Journal of the American Medical Informatics Association. JAMIA* **2003**, *10*, 85–93. [CrossRef]
7. Kauppinen, T.; Koivikko, M.; Ahovuo, J. Improvement of Report Workflow and Productivity Using Speech Recognition—A Follow-up Study. *J. Digit. Imaging Off. J. Soc. Comput. Appl. Radiol.* **2008**, *21*, 378–382. [CrossRef]
8. Fernandes, J.; Brunton, I.; Strudwick, G.; Banik, S.; Strauss, J. Physician experience with speech recognition software in psychiatry: Usage and perspective. *BMC Res. Notes* **2018**, *11*, 690. [CrossRef] [PubMed]
9. Zhao, Y. Speech-recognition technology in health care and special-needs assistance [Life Sciences]. *Signal Process. Mag. IEEE* **2009**, *26*, 87–90. [CrossRef]
10. Matarneh, R.; Maksymova, S.; Zeleniy, O.; Lyashenko, V. Voice Control for Flexible Medicine Robot. *Int. J. Comput. Trends Technol.* **2018**, *56*, 1–5. [CrossRef]
11. Rogowski, A. Scenario-Based Programming of Voice-Controlled Medical Robotic Systems. *Sensors* **2022**, *22*, 9520. [CrossRef]
12. ISO 9241-11:2018(en); Ergonomics of Human-System Interaction—Part 11: Usability: Definitions and Concepts. Available online: <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en> (accessed on 28 October 2024).
13. Faruk, L.; Pal, D.; Funilkul, S.; Chan, J. A Systematic Review of Voice Assistant Usability: An ISO 9241-11 Approach. *SN Comput. Sci.* **2022**, *3*, 267. [CrossRef]
14. van Pinxteren, M.; Pluymaekers, M.; Lemmink, J. Human-like communication in conversational agents: A literature review and research agenda. *J. Serv. Manag.* **2020**, *31*, 203–225. [CrossRef]
15. Weichbroth, P. Usability attributes revisited: A time-framed knowledge map. In Proceedings of the 2018 Federated Conference on Computer Science and Information Systems (FedCSIS), Poznan, Poland, 9–12 September 2018. [CrossRef]
16. ISO/IEC 30122-1:2016(en); Information Technology—User interfaces—Voice commands—Part 1: Framework and General Guidance. Available online: <https://www.iso.org/obp/ui/en/#iso:std:iso-iec:30122:-1:ed-1:v1:en> (accessed on 28 October 2024).
17. Janačković, G.; Vasović, D.; Vasović, B. On standardization in the field of user interfaces. In Proceedings of the XII International Symposium Engineering management and Competitiveness 2022 (EMC 2022), Zrenjanin, Serbia, 17–18 June 2022; pp. 272–277.
18. General Data Protection Regulation (GDPR)—Official Legal Text. General Data Protection Regulation (GDPR). Available online: <https://gdpr-info.eu/> (accessed on 18 October 2023).
19. Nass, S.J.; Levit, L.A.; Gostin, L.O. Institute of Medicine (US) Committee on Health Research and the Privacy of Health Information: The HIPAA Privacy Rule. In *Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research*; Nass, S.J., Levit, L.A., Gostin, L.O., Eds.; National Academies Press (US): Washington, DC, USA, 2009. [PubMed]
20. Ness, R.B.; Joint Policy Committee, Societies of Epidemiology. Influence of the HIPAA Privacy Rule on health research. *JAMA* **2007**, *298*, 2164–2170. [CrossRef] [PubMed]
21. The Official Website of the PathoVR Software. Available online: <https://www.pathovr.com/> (accessed on 13 December 2024).
22. The Official Website for the Google Speech-to-Text Model Descriptions. Available online: <https://cloud.google.com/speech-to-text/docs/transcription-model> (accessed on 13 December 2024).
23. Vincze, M.; Molnar, B.; Kozlovsky, M. 3D Visualization in Digital Medicine Using XR Technology. *Future Internet* **2023**, *15*, 284. [CrossRef]
24. Khare, P.; Srivastava, S. Enhancing Security with Voice: A Comprehensive Review of AI-Based Biometric Authentication Systems. *Int. J. Res. Anal. Rev.* **2023**, *10*, 398–403.
25. Gong, Y.; Poellabauer, C. An Overview of Vulnerabilities of Voice Controlled Systems. *arXiv* **2018**, arXiv:1803.09156.
26. Lupton, M. Some ethical and legal consequences of the application of artificial intelligence in the field of medicine. *Trends Med.* **2018**, *18*. [CrossRef]
27. Gerke, S.; Minssen, T.; Cohen, G. Ethical and legal challenges of artificial intelligence-driven healthcare. *Artif. Intell. Healthc.* **2020**, 295–336. [CrossRef]
28. Laptev, V.A.; Ershova, I.V.; Feyzrakhmanova, D.R. Medical Applications of Artificial Intelligence (Legal Aspects and Future Prospects). *Laws* **2022**, *11*, 3. [CrossRef]
29. Funer, F.; Liedtke, W.; Tinnemeyer, S.; Klausen, A.D.; Schneider, D.; Zacharias, H.U.; Langanke, M.; Salloch, S. Responsibility and decision-making authority in using clinical decision support systems: An empirical-ethical exploration of German prospective professionals’ preferences and concerns. *J. Med. Ethics* **2024**, *50*, 6–11. [CrossRef]
30. Göndöcs, D.; Dörfler, V. AI in Medical Diagnosis: AI Prediction & Human Judgement. *Artif. Intell. Med.* **2024**, *149*, 102769. [CrossRef] [PubMed]
31. Lysaght, T.; Lim, H.Y.; Xafis, V.; Ngiam, K.Y. AI-Assisted Decision-making in Healthcare: The Application of an Ethics Framework for Big Data in Health and Research. *Asian Bioeth. Rev.* **2019**, *11*, 299–314. [CrossRef]

32. Lorenzini, G.; Ossa, L.A.; Shaw, D.M.; Elger, B.S. Artificial intelligence and the doctor-patient relationship expanding the paradigm of shared decision making. *Bioethics* **2023**, *37*, 424–429. [CrossRef]
33. Khosravi, M.; Zare, Z.; Mojtabaean, S.; Izadi, R. Artificial Intelligence and Decision-Making in Healthcare: A Thematic Analysis of a Systematic Review of Reviews. *Health Serv. Res. Manag. Epidemiol.* **2024**, *11*, 1–15. [CrossRef]
34. Boyd, M.; Cumin, D.; Lombard, B.; Torrie, J.; Civil, N.; Weller, J. Read-back improves information transfer in simulated clinical crises. *BMJ Qual. Saf.* **2014**, *23*, 989–993. [CrossRef]
35. Brown, J.P. Closing the Communication Loop: Using Readback/Hearback to Support Patient Safety. *Jt. Comm. J. Qual. Saf.* **2004**, *30*, 460–464. [CrossRef] [PubMed]
36. Miller, A.C.; Scissum, K.; McConnell, L.; East, N.; Vahedian-Azimi, A.; Sewell, K.A.; Zehtabchi, S. Real-time audio-visual feedback with handheld nonautomated external defibrillator devices during cardiopulmonary resuscitation for in-hospital cardiac arrest: A meta-analysis. *Int. J. Crit. Illn. Inj. Sci.* **2020**, *10*, 109–122. [CrossRef] [PubMed]
37. Report on the Proposal for a Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation). Available online: [https://www.europarl.europa.eu/doceo/document/A-7-2013-0402\\_EN.html](https://www.europarl.europa.eu/doceo/document/A-7-2013-0402_EN.html) (accessed on 19 October 2023).
38. Pesapane, F.; Volonté, C.; Codari, M.; Sardanelli, F. Artificial intelligence as a medical device in radiology: Ethical and regulatory issues in Europe and the United States. *Insights Imaging* **2018**, *9*, 745–753. [CrossRef] [PubMed]
39. Palaniappan, K.; Lin, E.Y.T.; Vogel, S. Global Regulatory Frameworks for the Use of Artificial Intelligence (AI) in the Healthcare Services Sector. *Healthcare* **2024**, *12*, 562. [CrossRef] [PubMed]
40. Li, J.; Chen, C.; Pan, L.; Azghadi, M.R.; Ghodosi, H.; Zhang, J. Security and Privacy Problems in Voice Assistant Applications: A Survey. *Comput. Secur.* **2023**, *134*, 103448. [CrossRef]
41. Seymour, W.; Zhan, X.; Cote, M.; Such, J. A Systematic Review of Ethical Concerns with Voice Assistants. In Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society, Montréal, QC, Canada, 8–10 August 2023; pp. 131–145. [CrossRef]
42. Artificial Intelligence in Health Care: Medical, Legal and Ethical Challenges Ahead. Available online: [https://www.eerstekamer.nl/bijlage/20201105/artificial\\_intelligence\\_in\\_health/document3/f=/vldiey9antv3.pdf](https://www.eerstekamer.nl/bijlage/20201105/artificial_intelligence_in_health/document3/f=/vldiey9antv3.pdf) (accessed on 20 October 2023).
43. Legal Obligation and Ethical Best Practice: Towards Meaningful Verbal Consent for Voice Assistants | Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. Available online: <https://dl.acm.org/doi/abs/10.1145/3544548.3580967> (accessed on 13 December 2024).
44. Zhang, J.; Zhang, Z. Ethics and governance of trustworthy medical artificial intelligence. *BMC Med. Inform. Decis. Mak.* **2023**, *23*, 7. [CrossRef] [PubMed]
45. Möllmann, N.R.; Mirbabaie, M.; Stieglitz, S. Is it alright to use artificial intelligence in digital health? A systematic literature review on ethical considerations. *Health Inform. J.* **2021**, *27*, 14604582211052391. [CrossRef]
46. Rosemann, A.; Zhang, X. Exploring the social, ethical, legal, and responsibility dimensions of artificial intelligence for health—A new column in Intelligent Medicine. *Intell. Med.* **2021**, *2*, 103–109. [CrossRef]
47. Cheng, P.; Roedig, U. Personal Voice Assistant Security and Privacy—A Survey. *Proc. IEEE* **2022**, *110*, 476–507. [CrossRef]
48. Kaul, A. *Virtual Assistants and Ethical Implications*; IntechOpen: London, UK, 2021. [CrossRef]
49. Naik, N.; Hameed, B.M.Z.; Shetty, D.K.; Swain, D.; Shah, M.; Paul, R.; Aggarwal, K.; Ibrahim, S.; Patil, V.; Smriti, K.; et al. Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility? *Front. Surg.* **2022**, *9*, 862322. [CrossRef] [PubMed]
50. Jobin, A.; Ienca, M.; Vayena, E. The global landscape of AI ethics guidelines. *Nat. Mach. Intell.* **2019**, *1*, 389–399. [CrossRef]
51. Durán, J.; Jongsma, K. Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI. *J. Med. Ethics* **2021**, *47*, 329–335. [CrossRef] [PubMed]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.