electronics

MDPI

Article

Edge Computing Robot Interface for Automatic Elderly Mental Health Care Based on Voice

Camille Yvanoff-Frenchin [1,*], Vitor Ramos [2], Tarek Belabed [2] and Carlos Valderrama [2]

[1] ESIEE, 93162 Noisy-le-Grand, France

[2] Department of Electronics and Microelectronics (SEMi), University of Mons, 7000 Mons, Belgium; vitor.ramos@umons.ac.be (V.R.); tarek.Belabed@umons.ac.be (T.B.); carlos.valderrama@umons.ac.be (C.V.)

* Correspondence: camille.yvanoff-frechin@edu.esiee.fr

Received: 31 January 2020; Accepted: 21 February 2020; Published: 29 February 2020

Abstract: We need open platforms driven by specialists, in which queries can be created and collected for long periods and the diagnosis made based on a rigorous clinical follow-up. In this work, we developed a multi-language robot interface helping to evaluate the mental health of seniors by interacting through questions. Through the voice interface, the specialist can propose questions, as well as receive users' answers, in text form. The robot can automatically interact with the user using the appropriate language. It can process the answers and under the guidance of a specialist, questions and answers can be oriented towards the desired therapy direction. The prototype was implemented on an embedded device meant for edge computing, thus it was able to filter environmental noise and can be placed anywhere at home. The proposed platform allows the integration of well-known open source and commercial data flow processing frameworks. The experience is now available for specialists to create queries and answers through a Web-based interface.

Keywords: edge computing; mental health care; voice-enabled techniques; embedded systems

1. Introduction

Mental health care and diagnosis are today migrating towards mobile solutions [1,2]. Indeed, mobile applications provide more accessible support [3]. This becomes particularly interesting, knowing that people dealing with mood, stress or anxiety do not always seek professional help or get care when it is really needed [4]. On the other hand, care or help is not always available when needed, for reasons such as location, financial averages or for societal reasons [5].

A plethora of mobile applications is available for healthcare. In the context of our work, we can cite MIMOSYS [6] and CHADmon [7]. MIMOSYS [6] is a smartphone app that monitors mental health by analyzing the human voice to detect diseases or disorders from emotional changes. The authors in [7] present CHADMon, a dedicated mobile application for voice analysis and monitoring of mental state and phase change detection. Their interest and the required techniques were already under study, taking into account multiple aspects, from acceptability to clinical efficacy, including targeted therapies and clinical benefits [2]. Regarding applications, care must be taken with diagnosis, which can be harmful and stigmatizing without specialized intervention [1]. Moreover, evaluation and experimental testing mechanisms are fundamental for clinical and appropriate validation [8]. Indeed, we need open platforms driven by specialists, in which queries can be created and collected for long periods and the diagnosis made based on a rigorous clinical follow-up.

Today, smartphones are popular and available for private usage. Some applications are attractive to users for the same reasons, in particular young adults and users looking for self-help support. However, that is not always the case with seniors not so familiar with the technology but still interested in hands free interaction such as a robot or voice.

Voice-enabled technologies are leading multiple domains, from automotive to home automation [9]. According to [10], 50 percent of searches will be based on voice by 2020, idem for smart speakers by 2022 [11]. Voice search tends to be more mobile and locally targeted because it is integrated with many mobile apps and devices. Many digital assistants are integrated with products that are part of our everyday life [10]; Microsoft integrated Cortana into Windows 10 for text and voice search. Amazon's Echo is ready to answer questions as well as to control other home devices. Voice assistants such as Amazon, Google Home or Sonos One are free; several requests are product searches that also offer placements to advertisers. Although these commercial products are not always open to customization, they are supported by development platforms, as in the case of Amazon AWS [12], Google [13] and IBM Watson [14], for example.

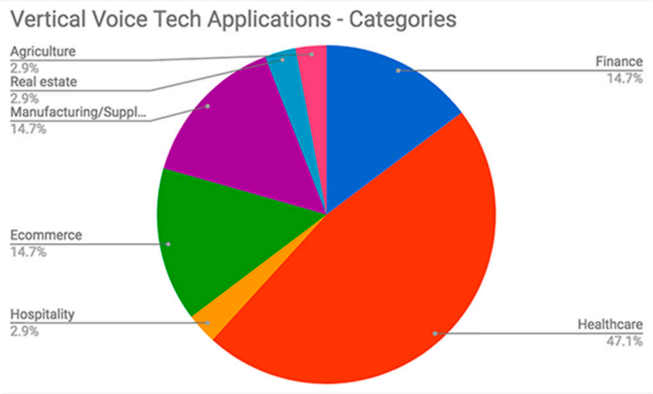


Figure 1. Vertical applications in voice tech [15].

According to [15], the healthcare sector is the most popular category (47.1%) (Figure 1) for vertical voice-based applications. Telemedicine encourages conversation applications in the health field, particularly where hospitals have a strong incentive to provide high-quality follow-up care. However, restrictions such as the confidentiality of the data involved, and low error tolerance make it difficult to grow quickly. Thus, the high cost of physicians and caregivers is spent on hours of data collection in electronic health records. The voice health sector also extends to seniors who wish to stay at home, especially those who refuse mobile or smart technologies requiring dexterity or good vision [9]. Aging at home implies socializing, AI-based activity-oriented interfaces and daily monitoring services. Robot-based patient-caregiver communication saves time and therefore increases the productivity of already planned tasks such as reminders and appointments. Physician notes, such as the Electronic Health Record (EHR) and patient feedback, now use voice technology and AI-based natural language scribes [16] on multiple platforms (PC, smartphones), including new microphones and wearable voice interfaces [17].

This work followed several objectives: to provide a multilingual voice interaction platform, facilitating the specialist's intervention by creating protocols and text queries, as well as text forms for advice and collected results, to integrate and experiment with existing technologies able to provide an automatic assessment of emotions providing graphical views of the evolution of the results, to pay attention to non-response situations, and to integrate the platform to a robot or a voice interface.

The article is organized as follows: Section 2 presents the state-of-the-art in-home health care voice-based products. We are particularly interested in embedded voice interfaces and devices, as well as the development tools available. Section 3 describes the system implemented. Section 4 presents the results. We conclude this document in Section 5.

2. State of Art

Healthcare home products are evolving thanks to AI-based platforms and on-line technologies (Figure 2). AI-based natural language systems capture patient-physician interaction, prepare real-time patient notes in the exam room, and produce text-form EHRs [16,17]. Several healthcare platforms are working with Amazon Alexa and Google Assistant smart speakers, for instance, Cuida Health LISA [18], a friendly voice assistant and companion who remembers medicines, appointments, and monitors wellness status daily, RemindMeCare [19], Memory Lane [20] or Senter [21]. Some are AI-based social robots, such as ElliQ [22] and Senter [21], encouraging daily personalized activities. On-line technologies, such as LifePod [23], enable schedules and voice services, providing valuable data to professionals and caregivers. Many are actually hands-free voice devices, such as Rosie Reminder [24] and ElliQ [22], or wearable, such as Notable [17].

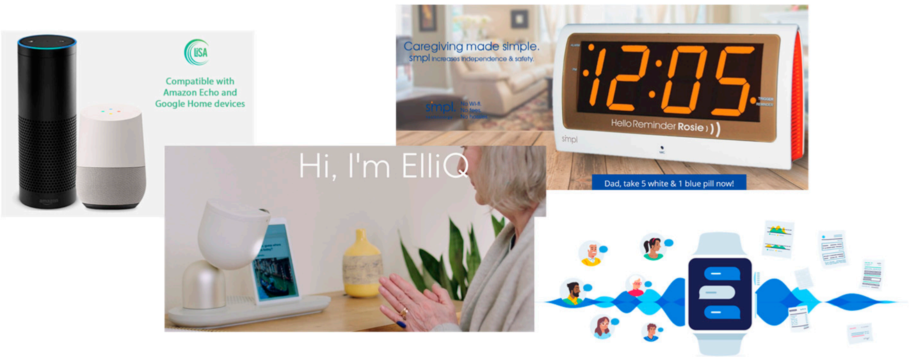

Compatible with Amazon Echo and Google Home devices
Caregiving made simple.
12:05
Hello Reminder Rosie
Dad, take 5 white & 1 blue pill now!
Hi, I'm ElliQ

Figure 2. Healthcare home products and hands-free voice devices. Healthcare platforms, such as Cuida Health LISA [18] (left), work with Amazon Alexa and Google Assistant smart speakers. Many are actually hands-free voice devices, such as Rosie Reminder [24] (right) and ElliQ [22] (center), others are wearable, as Notable [17] (bottom-right).

There are several tools and options available for processing, storing, indexing and managing streaming data, making it difficult for practitioners to choose the right combination of tools and platforms to build applications for data flow analysis [25]. In addition, healthcare analysis and recommendation systems must process continuous data streams within very short timeframes [26]. In [27], the authors present a comparative study of distributed data stream processing and analysis frameworks. In this study open source and commercial frameworks were examined regarding their ability to implement real-time distributed data stream processing. In [26], the authors survey state of the art architectures proposed to use edge computing, steam processing engines and mechanisms for data stream processing. Their results helped us identify the needs of specialists/users, data infrastructure and voice interface capabilities. Indeed, our internal data is text-based for the purposes of reporting and analysis.

Dialog systems, powered by artificial intelligence, are interactive virtual conversational agents used in a wide range of applications, including healthcare. Interactive and multilingual voice systems identify personalized needs in order to respond effectively to users' moods, tones, and languages. Voice assistants such as Amazon, Google or IBM Watson provide some libraries and APIs (Application programming interface). These commercial products are not always open to customization but are supported by development platforms. Indeed, due to their popularity and ease of use, we chose Google's and IBM's APIs as the first candidates to integrate into our open platform. IBM Watson [14] proposes tools for speech (convert text and speech with the ability to customize models), language (analyze text and extract meta-data from unstructured content) and empathy (understand tone, personality, and emotional state). As with Google, they also provide a language translator for documents, which can be enhanced with the Natural Language Classifier, a machine learning component to analyze text and labels by organizing data into custom categories, the Tone Analyzer, designed to understand the emotions and communication styles in the text and the Personality Insights, which predicts the characteristics, needs and values of the personality through written text. However, some of the APIs are only available for the English language, so tools such as tone analysis may lose their value in the context of a multilingual voice interface.

In [28], the authors investigated such popular APIs, evaluating in particular IBM Watson and Google. Their evaluation use case aimed to meet user needs regarding exam stress, based on university student survey data generated using Google Forms. Their results of the measurement of the effectiveness to analyze the responses concerning the stress related to the exams indicated that the APIs respond in an appropriate manner to the queries of the users concerning what they think of the exams at 76.5%. We are interested in a platform managed by a specialist, so the results of an automatic analysis can only be considered as nice-to-have complementary information.

3. Implementation Architecture

This work aims at a hands-free voice device suitable for edge computing. Therefore, the system consists of a programmable embedded device that can be placed anywhere at home, assisted by specialized hardware for audio processing and environmental noise filtering (Figure 3).

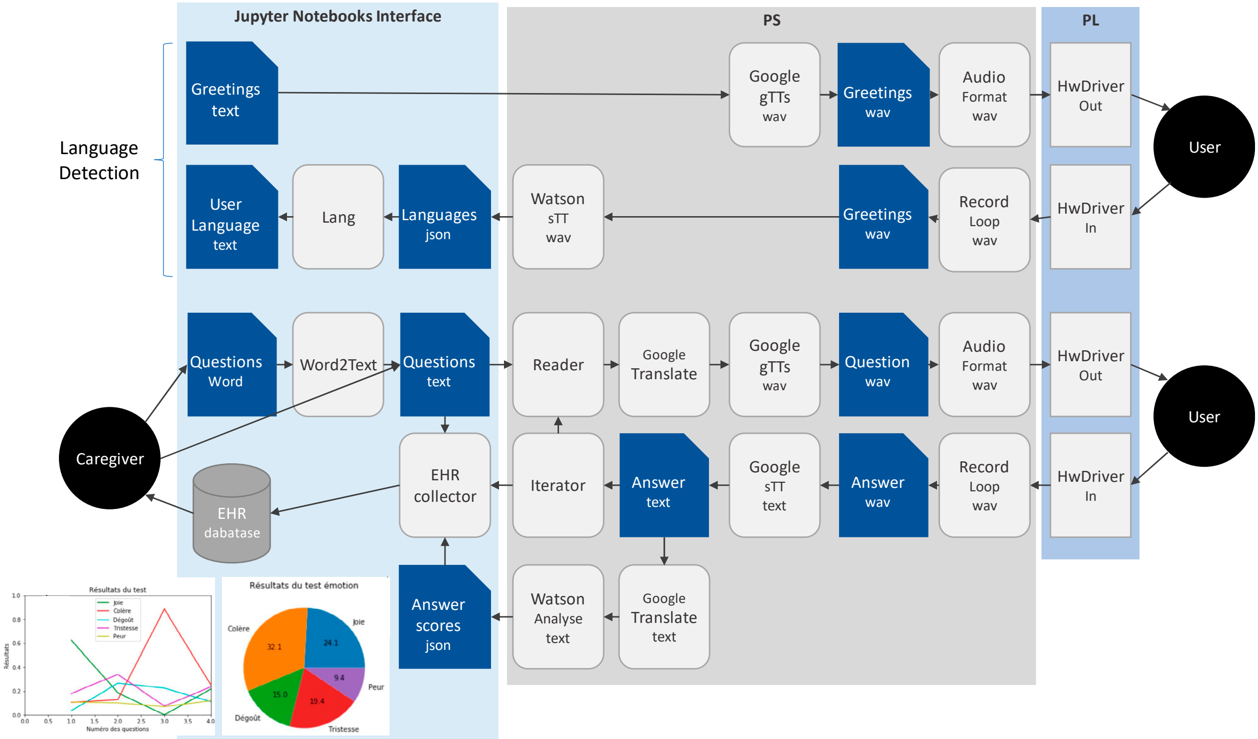


Figure 3. Healthcare home hands-free voice device system architecture. The edge-computing embedded system is composed of three sections: Jupyter Notebook WEB interface, Programmable Software PS and Programmable Logic PL. The PL part includes the headset user interface and real-time audio processing. The WEB interface allows retrieving the results in text and graphical form, in addition to entering queries and advice. The PS part recognizes the user's language and processes the questions and answers stored in the Electronic Health Record (EHR) database. Note: the screenshot of the WEB interface shows the graphic and circular views in the language chosen by the specialist (therefore automatically translated by the tool into French).

The proposed platform allows the integration of well-known open source and commercial data flow processing frameworks [27]. The programmable part runs Python on an ARM A9 CPU. The specialist (caregiver) can access the request and response database via a web interface. The prototype was implemented on a Xilinx PYNQ-Z1 board [29], designed to be used as an open-source framework, enabling embedded programmers to exploit the capabilities of reconfigurable hardware on the APSoC (All programmable System-on-Chip) Zynq family.

The software part (Programmable Software PS in Figure 3) of the APSoC is programmed using Python libraries from Google Cloud [13] (translate) and IBM Watson [14] (speech, empathy and natural language), all in a Jupyter Notebook [30] development environment. The hardware part (Programmable Logic PL in Figure 3) is a real-time audio processing programmable logic circuit, imported as a hardware library and programmed through an API, in the same way as the software. The platform can be accessed through a Web server hosting the Jupyter Notebooks design environment that includes the IPython kernel and packages running on a Linux OS.

3.1. Record a Complete User Response

The hardware API Pynq.Record is used to record the microphone input into an audio file. The audio driver (HwDriver in Figure 3) continually generates audio, but the API can only record a time interval. For this reason, the Python program continuously records 4 seconds each time, until there is no more incoming data (Record Loop in Figure 3). In this manner, we can record a complete answer to be sent to the speech-to-text conversion module Google.SpeechToText. This last operation can be intertwined, so we can build the response in text format during audio recording. We can also provide a playback of the answer by using the audio output and the recorded file.

3.2. Queries and Answers: Audio and Text Formats

Queries or answers to the user are entered by the specialist (Caregiver) as text. This is a shortcut for the professional, he can write it using his own language and be translated according to the user. He can also use a Word document, with each question ended with a question mark. In that case, the questions will be added to a list in a text-format file. We use Google.TextToSpeech, to create the equivalent mp3 audio file for each question. Alternatively, the professional can use the microphone to prepare his sets. The resulting audio must be converted to a wav file and adjusted to the driver parameters (24-bit, 48-kHz, 2-channel) using Subprocess (Audio Format in Figure 3). One option was to use Audacity [24], but then, it had to be integrated into the Python program and follow 2 conversion steps: from mono to stereo, then to 24 bits. The same can be realized by using PyPI.PyDub.AudioSegment [31] conversion from mono to stereo and PyPI.SoundFile [32] conversion from 16-bit 44 KHz to 24-bit 48 KHz.

3.3. Text from User Responses in the Appropriate Language: Language Detection

Based on queries, we can collect user answers through the audio interface. They can be recorded as an audio file. Instead of simply creating a multilingual multimedia EHR, anothersolution is to obtain user responses in text form and in the appropriate language. This format facilitates the search for keywords and features. By using Google.SpeechToText.recognize_google along with the audio file and language as input parameters we can get the text form. The file can then be translated into the language desired by the specialist using Google.Translator.translate for further processing of text-form recordings. In a similar manner, we can recover an audio file. This is especially important when using tools only available for English input, as we will show later. The library Google.TextToSpeech.gTTs will produce a wav format audio file at the desired speech rate.

3.4. Artificial Intelligence and Emotions on the Spot

The libraries offered by IBM Watson [14] for speech, language and empathy are based on AI and machine learning engines. Some libraries are available in Python; however, most have paid access. We, therefore, limited the experience to only language translation (LanguageTranslator) and tone analysis (ToneAnalyzer). The ToneAnalyzer library is intended to understand the emotions and style of communication. The Analyze library processes a text document based on emotion and sense parameters to focus on and provide a json (an open-standard file format) answer with a confidence score based on 6 alternative results: joy, anger, disgust, fear, sadness and positivity/negativity. In this case, the experiment consisted of detecting the language used, converting any information into that language, initiating the audio exchange between the specialist (questions and advice) and the user (answers) and finally of providing an emotional score.

4. Implementation Results

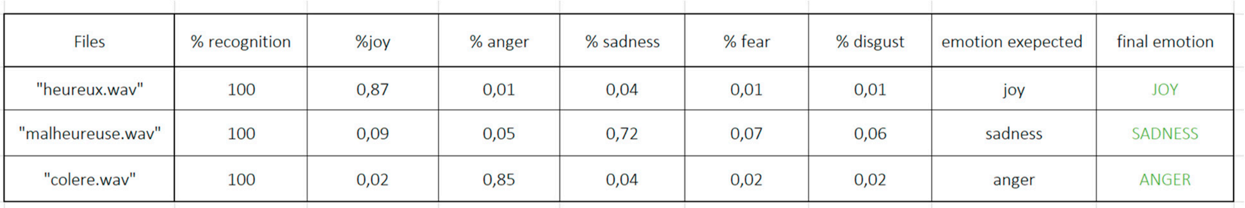
The full system was implemented on a PC and on the Xilinx PYNQ-Z1 board [29], for evaluation purposes. The prototype uses a Jupyter Notebook Web interface for the full process—in this manner we can make changes during the development process. However, the final version just provides the tools necessary for the specialist to enter queries and advice, receive user responses and graphical results.

The language detection (Language Detection in Figure 3) consists of a welcome text sentence that is transcribed into audio using Google.TextToSpeech, then we detect the language of the user's answer with IBM Watson SpeechToText which returns a json file containing several evaluated languages, the highest scored language is selected for the rest of the process. The answers collection process (Iterator in Figure 3) associates a list of questions (each question is transformed and sent to the audio output) to the user's responses (each answer received on the device audio input is transformed to a text-format). The process also detects when the user does not answer a question, in this case, the question is repeated, if not, it continues with the next question until the end of the list.

Considering that the platform implemented on a PYNQ card is less powerful than a PC, we carried out some evaluation tests in this direction. We use SpeechRecognition (pip3 install speech-recognition) [33], Google Text to Speech gTTs (pip3 install gTTs) [33], json and IBM Watson Natural Language Understanding (pip3 install NaturalLanguageUnderstandingV1) [34] libraries for information processing. On the PC we also used TempFile (pip3 install tempfile) [35] and PyGame (pip3 install pygame) for audio files processing. On Pynq we use SoundFile (pip3 install soundfile) [32] and PyDub (pip3 install pydub) [31] to manipulate audio files. In addition, we use Time to create pauses during execution and Numpy for the graphics. We first evaluated the processing speed compared to an I5 processor. The results showed that, in the worst case, the board spent 21 seconds per question compared to the 37 seconds on the PC.

The aim of this work is to provide a platform managed by a specialist, given that the results of an automatic empathy analysis are not sufficiently precise. Moreover, the Analyzer, used to identify emotions and communication styles is only applied to the text response, not to the audio input capable of containing contextual intonations specific to the user and the language used. However, this additional information could, at first, be used to help the specialist to choose the set of follow up questions to better identify the emotions. For this reason, we translated the user's response into English (the only language accepted by IBM Watson) and sent it to the Analyzer (Watson Analyze in Figure 3) which returned a json with a confidence score based on 5 alternative results. Figure 4 shows the results for three input audio files (initially recorded in French, then internally translated into English by the platform): heureux.wav "I'm so happy to live here", malheureuse.wav "I hate this world" and colère.wav "I can't tolerate this. I don't understand why people do that". As the example shows, the sentences were created using words specific to the expected emotion. The table at the top of the Figure 4 shows the input file and the associated resulting json, the emotion scores (joy, anger, sadness, fear, disgust) produced by the tool as well as the expected emotion. Note that, even for very precise answers, the maximum scores have never been higher than 87%. The recognition percentage is indicated at 100% when the expected score corresponds to the maximum score. We performed additional tests with short recorded sentences (wav audio format) in the French language. The results showed a maximum estimation accuracy of 87%, corroborating the results provided in [28].

In addition to using a translated text, the results of the Analyzer are not satisfactory enough for short answers or without the right set of directed questions. For that reason, the WEB interface (Figure 4. bottom) has been extended to provide the results for each set of questions. It is important to mention that the interface is automatically personalized according to the language chosen by the user (the example shows a screenshot of the page seen by a French-speaking specialist). We attach the scores associated with each answer to a chart (Figure 4. Bottom-left) and show the computed average score for the set of questions (Figure 4. Bottom-right) in the form of a pie chart. As this work can be extended in the future with supervised learning modules, we provide a multimedia HER including audio and text responses, as well as a graphical view of the result obtained with the Analyzer.



Files	% recognition	%joy	% anger	% sadness	% fear	% disgust	emotion exepected	final emotion
"heureux.wav"	100	0,87	0,01	0,04	0,01	0,01	joy	JOY
"malheureuse.wav"	100	0,09	0,05	0,72	0,07	0,06	sadness	SADNESS
"colere.wav"	100	0,02	0,85	0,04	0,02	0,02	anger	ANGER

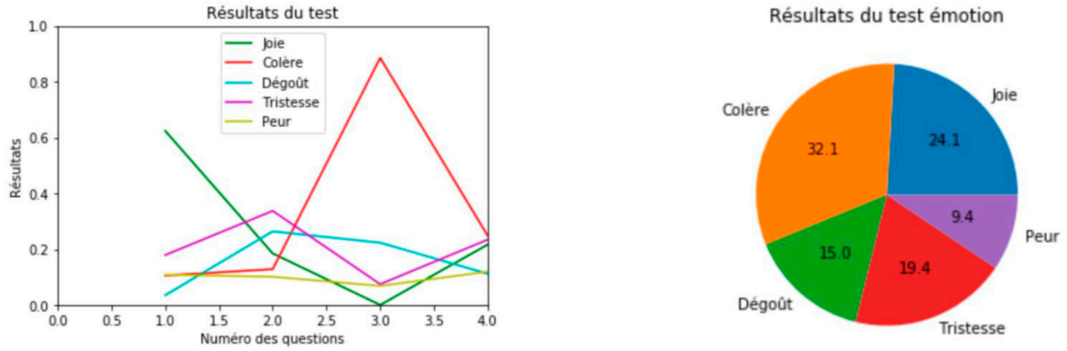


Figure 4. Results including the IBM Watson analyzer. The table (top) shows the results of three answers (audio files in French) according to the emotion score produced by the analyzer, the WEB Jupyter interface shows the evolution of the scores by the number of answers (bottom left screenshot) and the score graph for the answers set (bottom right screenshot). Note: the screenshot of the WEB interface shows the graphic and circular views in the language chosen by the specialist (therefore automatically translated by the tool into French).

5. Conclusions

In this work, we developed a multi-language robot interface helping to evaluate the mental health of seniors by interacting through questions. The prototype, implemented on an embedded device, is meant for edge computing. The platform is able to process text form queries from the caregiver and collect user answers. The device can also filter environmental noise and be placed anywhere at home. The experience is now available for specialists to create queries and answers through a Web-based interface. Queries can be created and collected for long periods and the diagnosis made, based on a rigorous clinical follow-up. The specialist can propose questions, as well as receive users' answers, in text form. The robot can automatically interact with the user using the appropriate language. It can process the answers and under the guidance of a specialist, questions and answers can be oriented towards the desired therapy direction. To fully exploit the advantages of the prototype, a set of questions used for the user follow up must be created and organized by a specialist according to the user (patient) and pathology. For that reason, generic tools such as Tone Analysis, in addition to the language barrier, can only be used as complementary information. As this work can be extended in the future with supervised learning modules, we provide a multimedia EHR including audio and text responses, as well as a graphical view of the result obtained with IBM Watson. The platform is now available to specialists to build a EHR database per patient. We expect that with this platform, clinical tests will be created with the help of specialists and patients, which will allow for the improvement of the platform to develop automatic voice-based care and pathologies protocols.

Author Contributions: All authors conceived and planned the experiments, C. Y.-F. contributed to the implementation and test. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding**.**

Acknowledgments: The authors would like to acknowledge the contribution of the COST Action CA16226 Indoor living space improvement: Smart Habitat for the Elderly.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bakker, D.; Kazantzis, N.; Rickwood, D.; Rickard, N. Mental Health Smartphone Apps: Review and Evidence-Based Recommendations for Future Developments. *JMIR Ment. Health* **2016**, *3*, e7.
2. Simon, G.E.; Ludman, E.J. It's time for disruptive innovation in psychotherapy. *Lancet* **2009**, *374*, 594–595.
3. Watts, S.E.; Andrews, G. Internet access is NOT restricted globally to high income countries: So why are evidenced based prevention and treatment programs for mental disorders so rare? *Asian J. Psychiatr.* **2014**, *10*, 71–74.
4. Mojtabai, R.; Olfson, M.; Mechanic, D. Perceived Need and Help-Seeking in Adults with Mood, Anxiety, or Substance Use Disorders. *Arch. Gen. Psychiatry* **2002**, *59*, 77.
5. Collin, P.J.; Metcalf, A.T.; Stephens-Reicher, J.C.; Blanchard, M.; E Herrman, H.; Rahilly, K.; Burns, J.M. ReachOut.com: The role of an online service for promoting help-seeking in young people. *Adv. Ment. Health* **2011**, *10*, 39–51.
6. Medical-Pst. MIMOSYS: Voice Analysis of Pathophysiology. 2019. Available online: https://medical-pst.com/en/ (accessed on 1 September 2019).
7. Antosik-Wójcinska, A.; Chojnacka, M.; Dominiak, M.; Święcicki, Ł. The use of smartphones in the management of bipolar disorder- mobile apps and voice analysis in monitoring of mental state and phase change detection. *Eur. Neuropsychopharmacol.* **2019**, *29*, S528–S529.
8. American Psychological Association (APA). Evidence-based practice in psychology. *Am. Psychol.* **2006**, *61*, 271–285.
9. Brownstein, J.; Lannon, J.; Lindenauer, S. 37 Startups Building Voice Applications for Healthcare. MobiHealthNews. Available online: https://www.mobihealthnews.com/content/37-startups-building-voice-applications-healthcare (accessed on 31 August 2019).
10. Olson, C. Voice Search Is Mobile-And Part of Your Everyday Life. Campaign. Available online: https://www.campaignlive.co.uk/article/just-say-it-future-search-voice-personal-digital-assistants/1392459 (accessed on 31 August 2019).
11. Perez, S. Voice-Enabled Smart Speakers to Reach 55% of U.S. Households by 2022, Says Report. TechCrunch. Available online: https://techcrunch.com/2017/11/08/voice-enabled-smart-speakers-to-reach-55-of-u-s-households-by-2022-says-report/?guccounter=1 (accessed on 31 August 2019).
12. AWS. Amazon Polly. 2019. Available online: https://docs.aws.amazon.com/polly/index.html (accessed on 1 September 2019).
13. Cloud, G. Google Cloud: Cloud Translation : Translation Client Libraries. 2019. Available online: https://cloud.google.com/translate/docs/reference/libraries (accessed on 1 September 2019).
14. Watson, I.B.M. IBM Watson Products and Services. 2019. Available online: https://www.ibm.com/watson/products-services/ (accessed on 1 September 2019).
15. Van der Straten, S. Voice Tech Landscape: 150+ Infrastructure, Horizontal and Vertical Startups Mapped and Analysed. Medium. Available online: https://medium.com/point-nine-news/voice-tech-landscape-150-startups-mapped-and-analysed-82c5adaf710 (accessed on 31 August 2019).
16. Kiroku. Kiroku: Automated Clinical Record Keeping. 2019. Available online: https://trykiroku.com/ (accessed on 31 August 2019).
17. Notable. Notable: Autopilot for Health Care. 2019. Available online: https://notablehealth.com/ (accessed on 31 August 2019).
18. CuidaHealth. You Can Make It Fun for Older Adults. 2019. Available online: https://cuidahealth.com/ (accessed on 31 August 2019).
19. RemindMeCare. RemindMeCare: Dementia & Elderly Care Technology. 2019. Available online: https://www.remindmecare.com/ (accessed on 31 August 2019).
20. MemoryLane. Your Memories Powered by Voice. 2019. Available online: http://memorylane.ai/ (accessed on 31 August 2019).
21. Senter. SENTER: Smarter Health at Home. 2019. Available online: http://senter.io/ (accessed on 31 August 2019).
22. Elliq. ElliQ, the Sidekick for Happier Aging-Intuition Robotics. 2019. Available online: https://elliq.com/ (accessed on 3 August 2019).
23. Lifepod. LifePod: Personalized Caregiving Powered by Proactive-Voice. 2019. Available online: https://lifepod.com/ (accessed on 31 August 2019).

24. Smpltec. Reminder Rosie: SMPL Technology. 2019. Available online: Https://smpltec.com/reminder-rosie#prodmenu (accessed on 31 August 2019).
25. Psaltis, A.G. *Streaming Data: Understanding the Real-Time Pipeline*, 1st ed., Manning Publications: Shelter Island, NY, USA, 2017.
26. Assuncao, M.; Veith, A.D.S.; Buyya, R. Distributed data stream processing and edge computing: A survey on resource elasticity and future directions. *J. Netw. Comput. Appl.* **2018**, *103*, 1–17.
27. Isah, H.; Abughofa, T.; Mahfuz, S.; Ajerla, D.; Zulkernine, F.; Khan, S. A Survey of Distributed Data Stream Processing Frameworks. *IEEE Access* **2019**, *7*, 154300–154316.
28. Ralston, K.; Chen, Y.; Isah, H.; Zulkernine, F. A Voice Interactive Multilingual Student Support System using IBM Watson. In Proceedings of the 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), Boca Raton, FL, USA, 2019, pp. 1924–1929.
29. Xilinx. PYNQ: Python Productivity for Zynq. 2019. Available online: http://www.pynq.io/ (accessed on 1 September 2019).
30. Jupyter.Org. Project Jupyter Notebook. 2019. Available online: https://jupyter.org/ (accessed on 1 September 2019).
31. Pypi.Org. PyPI: PyDub. 2019. Available online: https://pypi.org/project/pydub/ (accessed on 1 September 2019).
32. Pypi.Org. PyPI: SoundFile. 2019. Available online: https://pypi.org/project/SoundFile/ (accessed on 1 September 2019).
33. PyPI.Org. PyPI: SpeechRecognition. 2019. Available online: https://pypi.org/project/SpeechRecognition/ (accessed on 1 September 2019).
34. Watson, I.B.M. Ibm-Watson: NaturalLanguageUnderstandingV1. 2019. Available online: http://watson-developer-cloud.github.io/node-sdk/master/classes/naturallanguageunderstandingv1.html (accessed on 1 September 2019).
35. Python.Org. Tempfile: Generate Temporary Files and Directories. 2019. Available online: https://docs.python.org/3/library/tempfile.html (accessed on 1 September 2019).