IJAS

A Smart and Reliable Virtual Voice Assistant Using Machine Learning to Ease Our Task Efficiency

Renuka Palsodkar[1], Roshni Devkate[2], Ruchita Ingole[3], Nikhil Lade[4], Nikhil Bawane[5], Rushi More[6], Prof. F. N. Mawale[7]

[1,2,3,4,5,6]*Student, Dept of Computer Engineering, Bapurao Deshmukh College of Engineering Sevagram, Wardha.*
[7]*Professor, Dept of Computer Engineering, Bapurao Deshmukh College of Engineering Sevagram, Wardha.*

Email: [1]*renukapalsodkar2002@gmail.com,* [2]*roshnidevkate67@gmail.com,* [3]*ruchitaingole03@gmail.com,* [4]*nikhil.lade32@gmailcom,* [5]*nikhilbawane1999@gmail.com,* [6]*rushimore155@gmail.com*

Abstract: Digitization brings new possibilities to ease our daily life activities by the means of assistive technology. Amazon Alexa, Apple Siri, Microsoft Cortana, Samsung Bixby, to name only a few were successful in the age of smart personal assistants (spas). A voice assistant is defined a digital assistant that combines artificial intelligence, machine learning Speech Recognition, Natural Language Processing (NLP), Speech Synthesis and various actuation mechanisms to sense and influence the environment. We use different NLP techniques to convert Speech to text (STT), then process the text, convert Text to Speech (TTS), add various functionalities. However, SPA research seems to be highly fragmented among different disciplines, such as computer science, human-computer-interaction and information systems, which leads to 'reinventing the wheel approaches' and thus impede progress and conceptual clarity. In this paper, we present an exhaustive, integrative literature review to build a solid basis for future research. Hence, we contribute by providing a consolidated, integrated view on prior research and lay the foundation for an SPA classification scheme. Instead of pattern recognition we use NLP techniques to recognize the text which is context based. Operates online as well as offline. Data is Stored in Application itself, reduces Time and Space Complexity. We are developing a system which recognizes the voice on the basis of dataset, and then the corresponding response will be given on the basis of input voice received by the user. There are three stages on which it prominently works: -pre-processing, classification and data extraction. Desktop voice assistant are programmes that use integrated voice systems to recognize and respond human voice.

Keywords: Voice Assistant, Machine Learning, Speech Recognition, Natural Language Processing, Voice Recognition.

1. INTRODUCTION

Voice assistants (VAs) which are also called intelligent personal assistants are computer programs capable of understanding and responding to users using synthetic voices. Voice assistants have been integrated into different technological devices, including smartphones

and smart speakers. The voice modality is the central mode of communication used by these devices, rendering the graphic user interface (GUI) inapplicable or less meaningful. People use VA technology in different aspects of their lives, such as for simple tasks like getting the weather report or managing emails. In addition, the VA can perform complex tasks like client representative tasks and controllers in autonomous vehicles. In other words, VA's can revolutionize the way people interact with computing systems currently, there is a massive global adoption of voice assistants.

Usability is a critical factor in the adoption of voice assistants. A study by Zwakman et al. highlighted the importance of usability in voice assistants. An additional study by Coronado et al. reiterated the importance of usability in human–computer interaction tools. Numerous studies have been carried out on the usability heuristics used in a VA, each study adopting a unique approach. Virtual voice assistant is the manifestation of artificial intelligence by using data of human conversation as basics. it uses machine learning as well as natural language processing to process data and respond to the instruction of user. This virtual voice assistant provides response on the command or instruction of user. Early there were devices which used to work on single command of the user since 1950s, the first idea of virtual personal assistant came in 1996 from Anastacia Brice. Through various Virtual personal assistant can be seen in the last century, the first modern virtual assistant was developed by apple for iphone4S. the modern virtual personal assistant refers to software with features of current standard. the virtual assistant not only provides many advance and convenient features for devices like a smartphone. These not only developed the artificial intelligence field but also took devices to the next level.
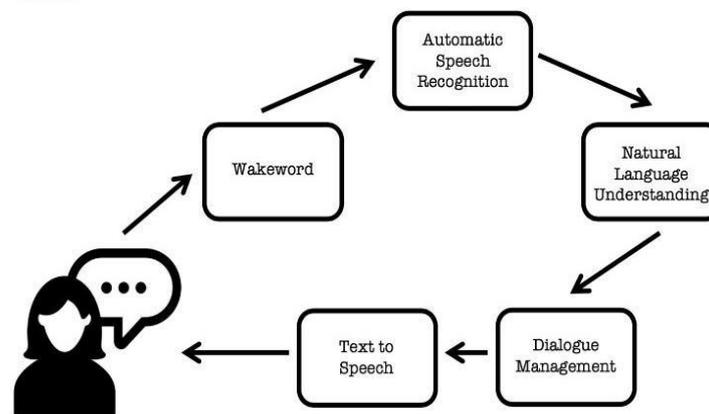

Fig. 1. Diagram of Virtual Voice Assistants

2. LITERATURE SURVEY

Voice assistant has a long history with several waves of major innovations. Voice assistant for dictation, search, and voice commands has become a standard feature on smartphones and wearable devices. The study stems from an overlooking literature review in order to present generic knowledge (theory and concepts) about voice control, virtual assistants, fields of use and more. When looking at a number of currently available intelligent programs with natural language processing capabilities, many examples can be found in everyday life filling a variety of roles.

The first speech recognition system, named Audrey, was created by Bell Laboratories in 1952. Audrey was rather rudimental and limited technology wise, understanding only ten digits -

spoken by particular people. About 10 years later, IBM developed and demonstrated their Shoebox Machine. The device recognized and responded to 16 different spoken words, including all ten digits "0" to "9" as well as calculating commands such as "plus" or "minus" (IBM, 2018). Shoebox Machine recognized and responded to 16 spoken words, including the ten digits from "0" through "9", only in English by a designated speaker. In today's world we train our machine to think like humans and do their task by themselves and what human being can do are being replaced by machines. Based on this situation there comes concept of voice assistant capable of completing various task for the humans based on their voice. Specific commands given by the user to virtual assistant is capable of filtering out the command and return relevant information. People in the whole world are transforming their digital experience using upcoming technologies like virtual reality, augmented reality, voice interaction etc. Voice control is emerging as new evolution in Human and Machine interaction where analog signal is converted by speech signal to digital wave. In Last few years huge increase in the use of smart phones led to the great use of voice assistant like Apple's Siri, Google's Assistant, Microsoft's Cortana and Amazon's Alexa etc. Voice assistants are built using technologies like voice recognition, speech synthesis, and Natural Language Processing (NLP) to provide indefinite applications to the users to make their life easy and comfortable.

Voice assistants have several interesting services for their users such as:
- Answer to questions asked by users.
- Play music from streaming music services and Playing YouTube videos.
- Set timers or alarms.
- Send WhatsApp, email messages.
- Provide information about the weather.
- Control other   smart devices (lights, locks, thermostats, vacuum cleaners, switches).
- The capabilities of voice assistants are continuously extending according to the users need.

According to Deepak Shende, Ria Umabiya, the AIVA (Microsoft, Google Assistant from Google, and the recently
Appeared intelligent assistant under the name "AIVA" 2018) aimed at developing a voice-controlled personal assistant which is doing many things such as to search the Internet. It has some new features like posting comments on the social media websites such as Facebook, Twitter, etc. By just few simple commands. You can also know the weather around you and can get the climate conditions in your region. Tulshan explained that because of continuous typing there may be possibility of injuries to the fingers of the user. To avoid such problems, we need to design a system in which we can get our work done through our voice commands. The voice will be recognized by the system and that recognized words will be synthesized and if they are appropriate or makes some sense then that will be printed on screen and after this again by recognizing the specific keywords the program will be compiled and executed.

3. PROPOSED METHOD

The proposed system of voice assistant will solve some issues of existing system as well introduce new features for better quality and usage. So, let's have a brief of the new updated version of the voice assistant. Instead of pattern recognition technique which has been used in previous models, we use Natural Language Processing (NLP) techniques to recognize the text which is context based rather the usual pattern based. This Operates in online as well as offline mode.

System application runs on offline mode, whereas web-based operations run on online mode. Data is Stored in Application itself, rather than cloud which reduces Time and Space Complexity. It even reduces the economic cost due to reducing high bundles of data usage.
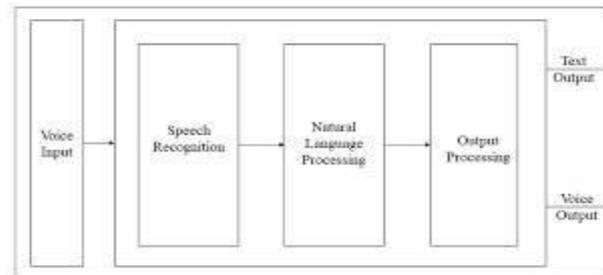

Fig. 2. Block Diagram of Virtual Voice Assistant

The Voice Assistant performs basic operations such as controlling computer tasks and operations, asking for temperature, humidity, date, time, and year. Adding, reading and deleting notes using voice commands and playing YouTube videos on demand. The above tasks can be performed using certain methodologies in which each technique has its own functionality and different operations to be performed. Each technique has different process logic to be executed

Techniques to be implemented
- Speech Recognition
- Text-to-Speech
- Natural Language Processing
- Command Processing

1. Speech Recognition: Utilize a library like Speech Recognition to convert spoken words into text. This will be the foundation for your assistant to understand voice commands.
2. Text-to-Speech (TTS): Implement TTS using libraries like g TTS (Google Text-to-Speech) to enable your assistant to respond audibly to user queries.
3. Natural Language Processing (NLP): Integrate NLP libraries such as NLTK or spaCy to understand and interpret user inputs. This will allow your assistant to comprehend more complex commands and questions.
4. Command Processing: Create a set of predefined commands or actions that your assistant can perform, such as checking the weather, opening applications, setting reminders, or searching the web.

Use-Case Diagram
The use case diagram shows the main interactions between the user and the personal voice assistant. The user initiates the interactions by speaking to the assistant, which responds with the requested information or action. The assistant is responsible for recognizing the user's speech, interpreting their intent, executing the relevant action, and generating an appropriate response.

IJAS
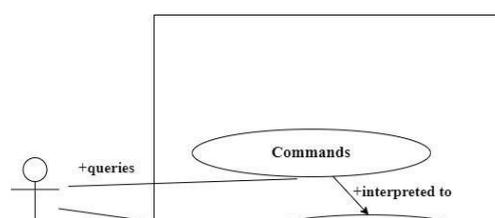
Some of the common use cases for a personal voice desktop assistant include asking for information, setting reminders or appointments, sending emails, controlling home automation devices, playing music, performing web searches, and checking the weather or news updates.
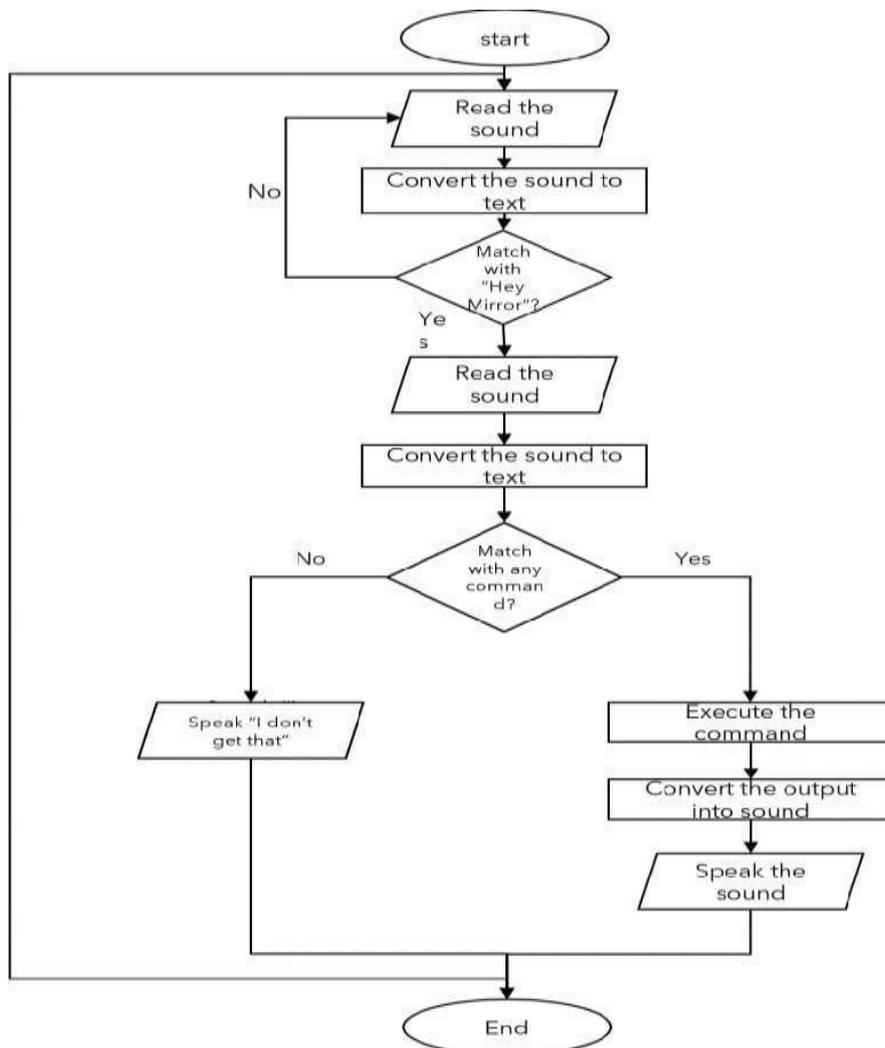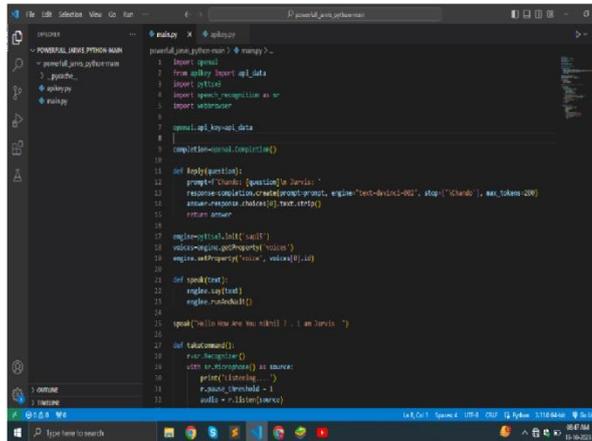
Flowchart

Fig. 3. Flowchart of Virtual Voice Assistant
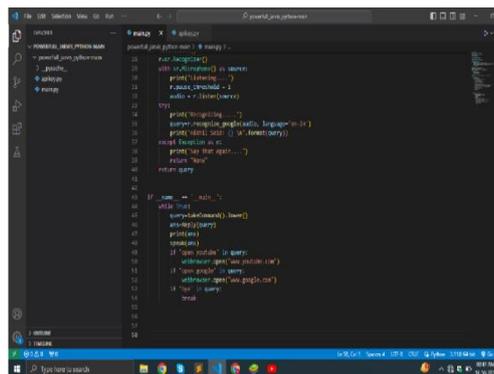
4. PRAGRAM AND DISCUSSION

Fig. 4.1. Program of Virtual Voice Assistant

Which uses open ai, speech recognition, pyttsx3, speechrecognition, web browser (functions used are: -define function, reply, speak, take command).
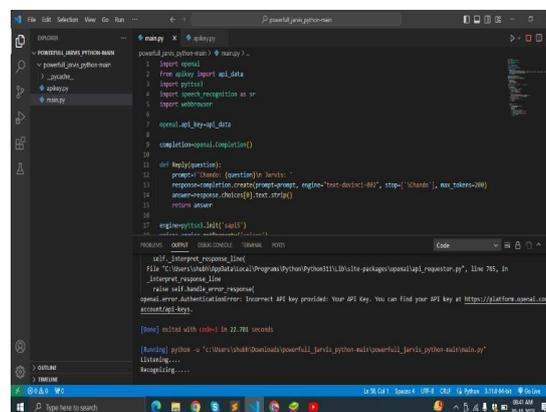

Fig. 4.2. Program of Virtual Voice Assistant

Which uses open ai, speech recognition, pyttsx3, speech recognition, web browser (functions used are: -define function,reply, speak, take command).

5. OUTPUT AND ANALYSIS

Fig. 5.1. Screenshot of Virtual Voice Assistant

In this above fig of output after listening the voice command of user the words / instructions of user are been recognized by our voice assistant.
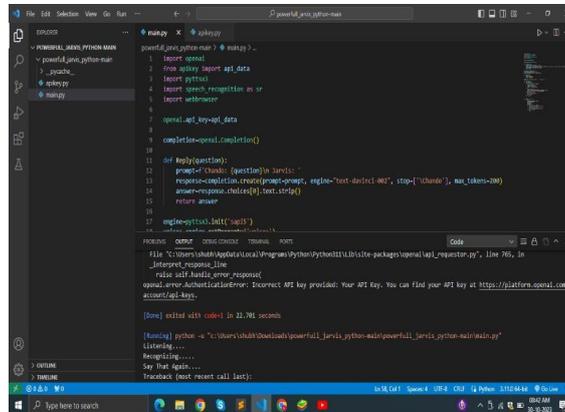
Fig. 5.2. Screenshot of Virtual Voice Assistant

In this above fig the since the assistant was not able to hear the instructions given by user properly so it provided the output as say it again.
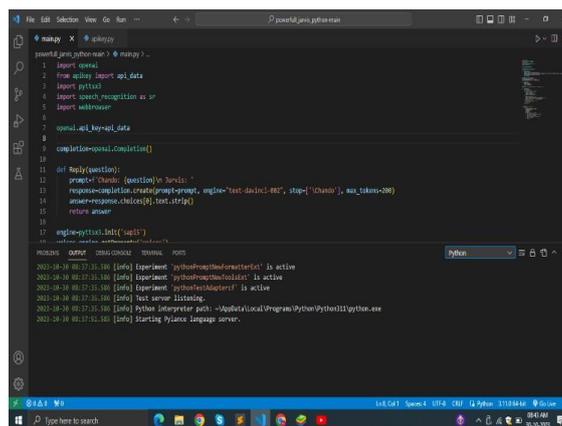
Fig. 5.3. Screenshot of Virtual Voice Assistant

In this above fig of output, our virtual voice assistant becomes active and listen the voice command of user.

Based on the survey we recommend that the application should be developed which accomplishes the desire of different users. The main reason that the user wants to use the voice assistant is to make their life easier, so by implementing the below mentioned features the user can be facilitated.

- Developing for different languages and different accents.
- Portability for any environment.
- Voice authentication technology can be implemented for more security.
- Chatbot implementation requires corpus.
- Dialogue flow needs stack with neurals.
- NLP features such as finding entities, topic modelling.

6. CONCLUSION AND FUTURE WORK

Voice Controlled Personal Assistant System will use the Natural language processing and can be integrated with Machine learning techniques to achieve a smart assistant that can perform action on various applications and will make human life comfortable. The system will have the following phases: Data collection in the form of voice; Voice analysis and conversion to text; Data storage and processing; generating speech from the processed text output. This application will also make life easier for those who are physically disabled and every common user who is fascinated by voice recognition. Academically, raising awareness for systems like this for students can give them better understanding of topics like Artificial Intelligence, Neural Networks, Natural Language Processing, Machine Learning and Human Computer Interaction and also how to improve user experience in application development. The formulated solution is able to process voice commands offline allowing users to cut down on the cost of data bundles. This also helps to make it faster in comparison to alternative applications like Apple's Siri, Google assistant, etc. Moreover, the solution is capable of carrying out a variety of tasks with ease such as telling the date and time, playing music/videos, making phone calls, finding weather, temperature, googling information etc. This paper can also act as a prototype for many advanced applications. The software has been created in such a way that the user will easily deal with it. Our proposed methodology, "Desktop Assistant"— an artificial intelligence smart voice assistant, is often built using face recognition and speech recognition modules, making the operations more effective and resilient. The Voice Assistant delivers two kinds of services. Firstly, the facial recognition technology makes it safer to use. Second, the voice or text-controlled application. With hands-free voice management of their system, the Desktop Assistant can assist a user. It will be able to automate a number of activities with the help of this voice assistant using single-line commands. Virtual assistants are a secret weapon for start-up heads for success. Virtual assistants not only save money but also boost productivity. Unlike a regular employee, you will pay your VA on an hourly basis. What's more, you won't have to spend any money on training if you take a VA on board. So, hiring a VA is a sure-fire way to success.

7. REFERENCE

[1]   Lasi, H.; Fettke, P.; Kemper, H.G.; Feld, T.; Hoffmann, M. Industry 4.0. Bus. Inf. Syst. Eng. 2014, 6, 239–242. [Google Scholar] [CrossRef]

[2]   Reis, A.; Rocha, T.; Barroso, J.; Carvalho, D. The Ambient Assisted Working (AAW) concept: Assistance according to I4.0 Technical Assistance Design Principle. In Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments, Corfu, Greece, 29 June–1 July 2022; pp. 316–318. [Google Scholar]

[3]   Hermann, M.; Pentek, T.; Otto, B. Design Principles for industrie 4.0 scenarios. In Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS), Koloa, HI, USA, 5–8 January 2016; pp. 3928–3937. [Google Scholar]

[4]   Gregor, S.D.; Hart, D.N. Information Systems Foundations: Building the Theoretical Base; Australian National University: Canberra, Australia, 2002; pp. 1–20. [Google Scholar]

[5]   Gregor, S. Building theory in the sciences of the artificial. In Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology, Philadelphia, PA, USA, 7–8 May 2009; pp. 1–10. [Google Scholar]

[6]   Nguyen, H.N.; Lasa, G.; Iriarte, I.; Atxa, A.; Unamuno, G.; Galfarsoro, G. Human-

IJAS

centered design for advanced services: A multidimensional design methodology. Adv. Eng. Inform. 2022, 53, 101720. [Google Scholar] [CrossRef]

[7] Sung, T.K. Industry 4.0: A Korea perspective. Technol. Forecast. Soc. Chang. 2018, 132, 40–45. [Google Scholar] [CrossRef]

[8] Soofastaei, A. Introductory Chapter: Virtual Assistants. In Virtual Assistant; Soofastaei, A., Ed.; IntechOpen: London, UK, 2021; pp. 1–9. [Google Scholar]

[9] Sarikaya, R. The technology behind personal digital assistants: An overview of the system architecture and key components. IEEE Signal Process. Mag. 2017, 34, 67–81. [Google Scholar] [CrossRef]

[10] Piñeiro-Martín, A.; Garcia-Mateo, C.; Docío-Fernández, L.; López-Pérez, M.d.C. Ethical Challenges in the Development of Virtual Assistants Powered by Large Language Models. Electronics 2023, 12, 3170. [Google Scholar] [CrossRef]

[11] Bolton, T.; Dargahi, T.; Belguith, S.; Al-Rakhami, M.S.; Sodhro, A.H. On the security and privacy challenges of Virtual Assistants. Sensors 2021, 21, 2312. [Google Scholar] [CrossRef] [PubMed]

[12] Gubareva, R.; Lopes, R.P. Virtual Assistants for Learning: A Systematic Literature Review. CSEDU 2020, 1, 97–103. [Google Scholar]

[13] De Cock, C.; Milne-Ives, M.; van Velthoven, M.H.; Alturkistani, A.; Lam, C.; Meinert, E. Effectiveness of conversational agents (Virtual Assistants) in health care: Protocol for a systematic review. JMIR Res. Protoc. 2020, 9, e16934. [Google Scholar] [CrossRef] [PubMed]

[14] Lugano, G. Virtual Assistants and self-driving cars. In Proceedings of the 2017 15th International Conference on ITS Telecommunications (ITST), Warsaw, Poland, 29–31 May 2017; pp. 1–5. [Google Scholar]

[15] Kagermann, H.; Wahlster, W.; Helbig, J.; Hellinger, A. Recommendations for implementing the strategic initiative INDUSTRIE 4.0. Final Rep. Ind. 2013, 4, 82. [Google Scholar]

[16] Drath, R.; Horch, A. Industrie 4.0: Hit or hype? [Industry forum]. IEEE Ind. Electron. Mag. 2014, 8, 56–58. [Google Scholar] [CrossRef]

[17] Zhou, K.; Liu, T.; Zhou, L. Industry 4.0: Towards future industrial opportunities and challenges. In Proceedings of the 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Zhangjiajie, China, 15–17 August 2015; pp. 2147–2152. [Google Scholar]

[18] Olsen, T.L.; Tomlin, B. Industry 4.0: Opportunities and challenges for operations management. Manuf. Serv. Oper. Manag. 2020, 22, 113–122. [Google Scholar]

[19] Ghobakhloo, M. Industry 4.0, digitization, and opportunities for sustainability. J. Clean. Prod. 2020, 252, 119869. [Google Scholar] [CrossRef]

[20] Schmidt, B.; Borrison, R.; Cohen, A.; Dix, M.; Gärtler, M.; Hollender, M.; Klöpper, B.; Maczey, S.; Siddharthan, S. Industrial Virtual Assistants: Challenges and opportunities. In Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, Singapore, 8–12 October 2018; pp. 794–801. [Google Scholar]

[21] Vaidya, S.; Ambad, P.; Bhosle, S. Industry 4.0—A glimpse. Procedia Manuf. 2018, 20, 233–238. [Google Scholar] [CrossRef]

[22] Mohamed, M. Challenges and benefits of industry 4.0: An overview. Int. J. Supply Oper. Manag. 2018, 5, 256–265. [Google Scholar]

[23] Margherita, E.G.; Bua, I. The role of human resource practices for the development of Operator 4.0 in Industry

[24] 4.0 organisations: A literature review and a research agenda. Businesses 2021, 1, 18–33. [Google Scholar] [CrossRef]

[25] Romero, D.; Stahre, J.; Wuest, T.; Noran, O.; Bernus, P.; Fast-Berglund, Å.; Gorecky, D. Towards an operator 4.0 typology: A human-centric perspective on the fourth industrial revolution technologies. In Proceedings of the proceedings of the international conference on computers and industrial engineering (CIE46), Tianjin, China, 29–31 October 2016; pp. 29–31. [Google Scholar]

[26] Gorecky, D.; Schmitt, M.; Loskyll, M.; Zühlke, D. Human-machine-interaction in the industry 4.0 era. In Proceedings of the 2014 12th IEEE international conference on industrial informatics (INDIN), Porto Alegre, Brazil, 27–30 July 2014; pp. 289–294. [Google Scholar]