

Research Article

Innovation for Universal Voice Command

Atiela Amran¹, Nurin Sofea Rahim², Nazatul Najwa Mohd Zaini³, Nur Bilqis Dayini Amran⁴, Nur Hizzati Saliman⁵, and Farah Safiyah Abdul Basir⁶

¹ Universiti Teknologi MARA; atiela@uitm.edu.my;  ORCID ID (<https://orcid.org/0009-0000-8592-519X>)

² Universiti Teknologi MARA; nurinsofeara@gmail.com

³ Universiti Teknologi MARA; najwanazatul76@gmail.com

⁴ Universiti Teknologi MARA; bilqisdayini@gmail.com

⁵ Universiti Teknologi MARA; atieeleman31@gmail.com

⁶ Universiti Teknologi MARA; farahsafiyahbasir@gmail.com

* Correspondence: atiela@uitm.edu.my; 019-9344487.

Abstract: *The development of a universal voice command device involves problem identification, focusing on hands-free operation, voice assistance, and accessibility for elderly or disabled users. Key stages include ideation and market research, emphasizing user needs, ergonomic design, and competitor analysis. Research and design utilize a user-centered approach, integrating features like voice recognition, LED indicators, and tactile buttons for inclusivity. Engineering and development involve assembling hardware components (microphones, speakers) and implementing software for voice algorithms and connectivity (Wi-Fi, Bluetooth). Accessibility features like manual buttons, volume controls, and light cues are added for versatility. Manufacturing ensures durability and eco-friendliness while scaling production with quality control. Marketing and deployment target users with clear instructions and accessibility highlights. Feedback and iteration enable updates and feature enhancements to maintain relevance and effectiveness.*

Keywords: *Problem-solving; Accessibility; User-centered design; Voice recognition; Ergonomics; Manufacturing; Marketing; Feedback; Iteration.*



Copyright: © 2025 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. INTRODUCTION

Voice command technology, often known as voice recognition or speech recognition technology, is a creative approach that allows people to communicate with devices and systems using spoken language. By turning spoken words into machine-readable inputs, this technology eliminates the need for physical interfaces like keyboards, touchscreens and buttons. Instead, it offers a hands-free, intuitive and natural method for controlling devices. This technology interprets and understands voice instructions using Natural Language Processing (NLP), Artificial Intelligence (AI) and Machine Learning (ML) algorithms. NLP enables the system to understand human speech patterns, context and purpose. AI and machine learning allow the system to improve its accuracy and adapt to the user's preferences over time. These developments enable voice command systems to be far more advanced and user-friendly.

However, voice command technology faces a number of challenges that limit its full potential. Accuracy and understanding are significant issues since systems frequently struggle to recognise different accents, dialects and unclear speech which will result in dissatisfaction. Privacy and security

problems develop when devices constantly listen for activation words, possibly recording essential conversations. Environmental issues such as background noise and many voices may prevent correct detection. Furthermore, many systems cannot comprehend complicated or multi-step commands, which limits their efficiency. Finally, accessibility remains an issue for those with speech difficulties, making it difficult for them to utilise voice technologies fully. These limitations have an impact on the general use and effectiveness of voice command systems.

2. PROJECT OUTCOMES

The universal voice command initiative delivers an innovation aimed at simplifying and improving how people interact with their devices. At its core, the project aims to provide a single, universal voice-command interface that can operate a wide range of devices and systems, regardless of brand, platform, or operating system. This solution aims to revolutionise ease and accessibility in the innovative technology era by incorporating powerful voice recognition technology and exploiting cross-platform compatibility. The universal voice command technology will give many people effortless control over everyday chores like turning on appliances, operating home security systems, and accessing digital tools, making them faster and more intuitive.

Among the project's most important results is its capacity to improve accessibility. Designed with inclusion in mind, the technology provides a lifeline for those with physical limitations by allowing them to interface with devices using simple voice commands. Another significant advantage is increased efficiency; the technology consolidates many tasks onto a single platform, eliminating the clutter of devices and controllers while saving users vital time and effort. By focussing on cross-platform integration, the project assures that the voice-command system is versatile, adaptable to a variety of situations, and compatible with a wide range of technologies.

In addition, the universal voice command technology is designed to be scalable. This means it is built to develop and evolve, quickly adding new instructions and growing capabilities as technology progresses. The innovation also makes considerable progress on the Internet of Things (IoT) domain, helping to further the development of interconnected smart home ecosystems. Its ability to synchronize the control of several IoT-enabled devices is a significant step towards establishing a genuinely seamless and intelligent technological experience.

The universal voice command innovation intends to transform user-device interactions by focusing on accessibility, efficiency, and innovation, making them more intuitive, inclusive, and future-ready. This breakthrough, which addresses significant technological issues and offers scalable solutions, demonstrates the power of human-centred design in promoting connection and smart living.

3. PROJECT CHALLENGES

The concept of a universal voice command system represents a compelling breakthrough that could transform human interaction with technology. Nonetheless, implementing this goal is challenged by problems that encompass technical, social, and ethical domains. These challenges must be systematically addressed to guarantee the system's functionality, usefulness, and adoption.

A significant problem is accommodating the diversity of languages, dialects, and accents worldwide. An effective universal system must possess the ability to comprehend several languages while precisely translating regional accents and informal phrases. Accomplishing this needs comprehensive datasets and robust natural language processing (NLP) methods. The absence of these

elements may lead to user disengagement due to the system's inability to accurately interpret or identify orders, especially from individuals who do not utilize conventional or commonly accepted accents.

A significant concern is context awareness. Voice commands frequently exhibit confusion and are significantly influenced by the context in which they are spoken. A statement such as “Turn it on” may denote several devices, contingent upon the context and user intention. Developing a system that comprehends all these details requires advanced contextual analysis and machine learning, hence increasing the complexity of the development process.

Environmental issues, like ambient noise, provide considerable challenges. Voice command systems must operate consistently in noisy or unexpected environments, such as crowded public areas or industrial locations. Attaining this kind of durability requires complicated noise-cancellation technologies and adaptive algorithms, which can be resource-demanding to create and apply.

4. PROJECT SUCCESS INDICATORS

The success of the universal voice command system can be measured by its technical, social, and ethical performance, ensuring functionality, usability, security, and broad adoption. Key indicators include technical accuracy, language comprehension, and contextual awareness. The system should handle diverse languages, accents, and noisy environments while seamlessly integrating with multiple hardware and software platforms.

Usability and user experience are essential for widespread acceptance. The system must be intuitive and accessible for users of all skill levels, including the elderly and non-technical users. Commands should execute promptly, and the system must adapt to evolving language trends to stay relevant across demographics.

User privacy and data security are crucial for building trust. Strong encryption and compliance with global privacy standards should protect user data, while developers must maintain transparency in data handling. Balancing security with efficiency is vital to ensure trust without compromising performance.

Cultural and ethical considerations are key for global acceptability. The system must respect cultural norms, avoid biased judgments, and handle language nuances effectively. High adoption across diverse communities will indicate success in maintaining cultural sensitivity and inclusivity.

5. USER SCENARIO

5.1 Voice Command System Workflow (Step-by-Step)

5.1.1 Step 1: Wake Word Detection



Figure 1. Wake Word Detection
(Source: Raja Shekar, 2022)

Action: The voice assistant system (e.g. Zeno) for this voice command system is always listening for a specific "wake word" (e.g., "Hey Zeno," "Okay Zeno," or just call "Zeno").

Process: The system uses a microphone to listen for the wake word in its environment constantly. It remains in a passive, low-power state until it detects the wake word.

Example: The user says, "Hey Zeno."

5.1.2 Step 2: User Command Recognition



Figure 2. User Command Recognition
(Source: www.mathworks.com)

Action: Once the system detects the wake word, it activates and begins listening for the actual command from the user.

Process: The voice assistant starts recording the speech that follows the wake word and sends this audio to a cloud-based server (or sometimes processes it locally on the device if offline capabilities are supported).

Example: The user says, "Set the living room lights to 50%."

5.1.3 Step 3: Internet Recognition

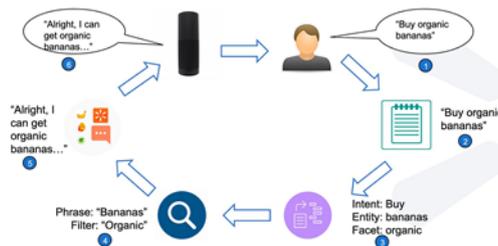


Figure 3. Internet Recognition
(Source: Mohan, 2019)

Action: The system analyzes the text for its intent and context to understand what action the user wants to perform.

Process: Using NLP, the system determines the intent (e.g., controlling lights) and identifies parameters (e.g., "living room lights," "50%" brightness).

Example: The system recognizes the intent to control smart lights and identifies the target as the "living room" and the desired brightness level as "50%."

5.1.4 Step 4: Command Processing

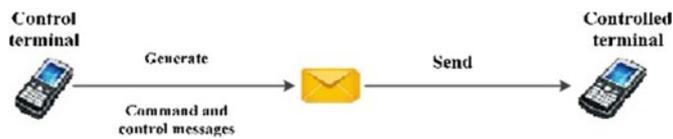


Figure 4. Command Processing
(Source: Tao Liu & Lai Zhu, 2015)

Action: Once the intent is recognized, the system processes the command by linking it to the appropriate action.

Process: The system translates the request into an executable action. If it requires integration with a smart home device (e.g., controlling lights), it sends the command to that device via the home automation hub or a cloud-based service.

Example: The system sends a signal to the bright light in the living room to adjust the brightness to 50%.

5.1.5 Step 5: Device Action Execution

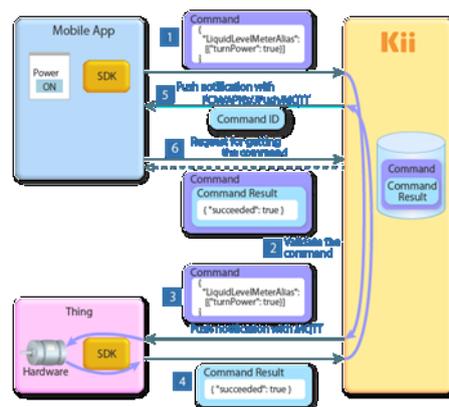


Figure 5. Command Processing
(Source: Kii Cloud, 2022)

Action: The smart device (e.g., lights, thermostat, music player) receives the command and acts.

Process: The device receives the data from the voice assistant system, which could involve adjusting settings (e.g., dimming lights, changing the temperature, playing music, etc.).

Example: The living room lights are adjusted to 50% brightness.

5.1.6 Step 6: Feedback to the User



Figure 6. Feedback to the User
(Source: Cheeky Munkey, 2017)

Action: After the action is performed, the system provides feedback to the user (either through audio or visual confirmation).

Process: The system may speak back to the user or provide visual indicators on connected devices to confirm that the action has been completed.

Example: The voice assistant might respond, “The living room lights are now set to 50%,” or simply show the visual change (e.g., dimmed lights).

5.1.7 Step 7: User Confirmation (Optional)

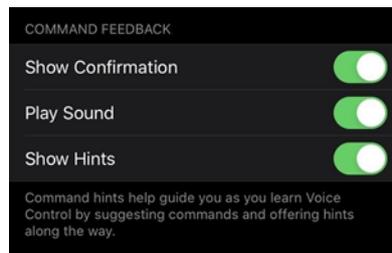


Figure 7. User Confirmation
(Source: Gadget Hacks)

Action: Sometimes, the system may ask for user confirmation or perform a follow-up action.

Process: Depending on the complexity of the command, the system may follow up to ensure the action was successful or to ask for further instructions.

Example: If the user wants to adjust the lighting further, they may continue the conversation: “Increase brightness to 70%.”

5.1.8 Step 8: System Idle

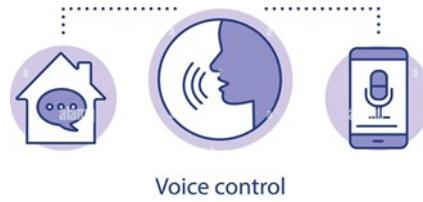


Figure 8. System Idle
(Source: Alamy Limited, 2025)

Action: After the task is completed and the conversation ends, the system returns to a passive listening state, awaiting the next wake word or command.

Process: The system re-enters a low-power mode or continues listening for the next user interaction.

Example: The system waits for the next wake word (“Hey Zeno” or “Zeno”).

6. RAPID PROTOTYPING



Figure 9. Prototyping
(Source: Xvkdfashion.shop, 2023)

Table 1. Features Description

Features	Description
Volume Up (+ Button)	Adjust the speaker's volume to increase audio levels
Volume Down (- Button)	Adjust the speaker's volume to decrease audio levels
Action Button	<ul style="list-style-type: none"> i. Allows manual activation of the voice assistant without the need for a wake word. ii. Perform specific functions such as stopping an alarm.
Mic Off Button	Disables the device's microphone for privacy
Light Ring	<ul style="list-style-type: none"> i. Provides a visual indicator, making it easy to recognize the device's status or activation, even for those with hearing disability. ii. It makes it easier for users to locate the device even in complete darkness
Power Port	Connects the device to an external power source
3.5 mm Audio Output	Allows connection to external speakers or audio systems for enhanced sound output
Height (1.7")	Compact and unobtrusive, making it suitable for any space
Diameter (3.9")	Ensures a table base for better placement

7. CONCLUSION

In conclusion, the universal voice command system represents a game-changing breakthrough in human-technology interaction, seeking to alter how people interact with their devices in a variety of settings. This breakthrough, which allows for hands-free operation, promises to simplify daily tasks, increase productivity, and offer more inclusivity for users with various linguistic, cultural, and physical needs. However, to realise this ambitious vision, a number of technical, social, and ethical difficulties must be addressed. The detailed concept visualisation process identified critical elements that will ensure the system's success, such as strong voice recognition capabilities, context-aware processing, seamless integration with various devices, noise resilience, and comprehensive security and privacy measures.

Acknowledgments: First and foremost, we thank Allah, the Almighty, for His countless blessings, granting us the health, courage, and ability to complete this report without significant obstacles. We are deeply grateful to Dr. Ayu Kamareena Abdullah Thani, our esteemed lecturer for ASM657 Innovative Thinking Strategies, for her unwavering guidance, patience, and encouragement throughout the preparation of this report. Her dedication and valuable insights have been a source of inspiration and motivation for us to work diligently on this collaborative effort. We also express our heartfelt gratitude to our parents for their unwavering support, encouragement, and sacrifices, which have been our greatest source of strength. Additionally, we are thankful to our classmates for their collaboration, friendship, and shared ideas, making this journey both enriching and memorable. Lastly, we extend our appreciation to everyone who has contributed directly or indirectly to the successful completion of this assignment, acknowledging their effort and dedication, which made this report possible.

References

- Alamy Limited. (2025). *Voice control concept icon. Speech recognition. Voice command device for home automation system. Virtual assistant. VUI idea thin line illustration*. V. Alamy.com; Alamy images. Retrieved from <https://www.alamy.com/voice-control-concept-icon-speech-recognition-voice-command-device-for-home-automation-system-virtual-assistant-vui-idea-thin-line-illustration-v-image341489773.html>.
- Cheeky Munkey. (2017, June 29). *Voice Activated Computers - the Future?* Retrieved from <https://cheekymunkey.co.uk/blog/voice-activated-computers-future/>
- Gadget Hacks. (n.d.). *24 voice control features in iOS 13 that let you use your iPhone totally hands-free*. Retrieved from <https://ios.gadgethacks.com/how-to/24-voice-control-features-ios-13-let-you-use-your-iphone-totally-hands-free-0206364/>
- Harshasri, M., Durga, B. M., & Ravikanth, M. (2021, June 10). *Voice based email for blind*. Ssrn.com. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3901972
- Iliev, Y., & Ilieva, G. (2022). A framework for Smart Home System with voice control using NLP methods. *Electronics*, 12(1), 116–116. Retrieved from <https://doi.org/10.3390/electronics12010116>
- Khusnul Khotimah, Santoso, A. B., Miftahul Ma'arif, Alfiantin Noor Azhiimah, Bambang Suprianto, Meini Sondang Sumbawati, & Tri Rijanto. (2020). Validation of voice recognition in various Google voice languages using Voice Recognition Module V3 based on microcontroller. 1–6. Retrieved from <https://doi.org/10.1109/icvee50212.2020.9243184>
- Kii Cloud. (2022). *Action and Command Execution*. Retrieved from https://docs.kii.com/en/functions/thingifsdk/thingifsdk/model/actions_commands/
- Liu, T., & Zhu, K. (2015). The research of control mechanism in Mobile Botnet. *2015 3rd International Conference on Machinery, Materials and Information Technology Applications*. Retrieved from <https://doi.org/10.2991/icmmita-15.2015.279>
- Mathworks.com. (n.d.). *Speech command recognition using deep learning - MATLAB & Simulink*. www.mathworks.com. Retrieved from <https://www.mathworks.com/help/deeplearning/ug/deep-learning-speech-recognition.html>
- Mohan, D. (2019, December 12). *Joint intent classification and entity recognition for conversational commerce*. Medium. Retrieved from <https://medium.com/walmartglobaltech/joint-intent-classification-and-entity-recognition-for-conversational-commerce-35bf69195176>
- Mohsen Bakouri, Alsehaimi, M., Ismail, H. F., Khaled Alshareef, Ganoun, A., Alqahtani, A., & Alharbi, Y. (2022). Steering a robotic wheelchair based on voice recognition system using convolutional neural networks. *Electronics*, 11(1), 168–168. Retrieved from <https://doi.org/10.3390/electronics11010168>
- Pervaiz, A., Hussain, F., Huma Israr, Tahir, M. A., Raja, F. R., Baloch, N. K., Farruh Ishmanov, & Yousaf Bin Zikria. (2020). Incorporating noise robustness in speech command recognition by noise augmentation of training data. *Sensors*, 20(8), 2326–2326. Retrieved from <https://doi.org/10.3390/s20082326>
- Raja Shekar. (2022). *How to build your own wake word detector*. Retrieved from <https://www.rajashekar.org/wake-word/>
- Xvkdfashion.shop. (2023). *Amazon Echo Dot*. Retrieved from <https://www.xvkdfashion.shop/?path=page/ggitem&ggpid=2841482>