УДК 004.934

VOICE ASSISTANT FOR WEB APPLICATIONS

Nesmiian D.
Науковий керівник – доц. Голян Н.В.
Харківський національний університет радіоелектроніки, каф. ПІ
м. Харків, Україна
e-mail: dmytro.nesmiian@nure.ua

The current work is dedicated to comparing user interaction algorithms with web interfaces of web applications using voice, particularly in educational platforms. The main principles of voice technology operation at different levels of hardware and software systems are considered. A comparison of the most popular voice assistants and the technologies used to create them has been conducted. It has been determined that optimizing additional functionality for controlling a web page using voice improves the user experience and contributes to user's return to the web application.

Investments in the field of information technology are increasing, because more and more people are using digital solutions. The transition of a large number of buyers to the Internet environment requires developers to create convenient navigation. Convenient navigation is one of the most important parameters of a commercial website interface. It must be simple, usable and understandable. Neglecting the requirements for navigation design will lead to loss of income [1]. If websites and online applications are designed and designed appropriately, they can be used by people with special needs [2].

Language is the main form of communication of thoughts, ideas and feelings in the human environment. The ability to control the environment by voice has prompted scientists and experts of advanced corporations to create such software applications as Assistant from Google, Cortana from Microsoft and others.

Solving the problem of voice recognition allows users to switch from using the command line and WIMP (Window, Image, Menu, Pointer) to the SILK (Speech, Image, Language, Knowledge) interface of software systems [3]. In contrast to user interaction with a computer system using a keyboard, mouse, joystick and display, a voice interface has the following advantages:

– to communicate with a computer, a person does not need to have special skills or abilities in the field of information technologies;

– language reduces the psychological and physical distance between a person and a computer and can be connected to it through communication systems, for example, a telephone;

– the voice interface provides efficiency and mobility of communication, frees hands and relieves the visual channel when receiving information.

Speech recognition (speech recognition, speech to text (STT)) is an automated process of converting a speech signal into a text stream. Voice

recognition technology has two modes of interaction: voice command mode and voice search mode. Voice search works in the same way as a normal search, but it is controlled by voice. Both are important to consider when drawing attention to interface elements. A significant part of the voice commands is search-oriented.

During a voice request, the technology converts the audio signal into text. The search engine or applications then process this and display the appropriate result.

The voice request is processed in four stages:

– filtering – the required phrase is extracted from the cloud of noise. This is necessary to eliminate external obstacles.

– digitization – sound waves are converted into a code that can be processed by computing.

– analysis – the received data is processed, compressed and prepared for identification.

– pattern data identification – the query is compared with examples from the database and search history to reveal specific words and phrases.

Analysis of voice recognition technology and existing voice assistants for web browsers revealed their certain shortcomings and the urgency of developing a new solution. The most interesting examples are VoxPow [4], Handsfree for Web [5], as well as the Web Speech API. Their approaches to the translation of user commands, the possibility of integration into a web page, and the support of several languages were taken into account during the development of the software product. The development of a new solution is intended to eliminate the shortcomings of the listed analogues.

References:

1. Interfax-Ukraine – news from Ukraine and the world. Retrieved February 20, 2022, from https://interfax.com.ua/news/blog/755659.html

2. Introduction to Web Accessibility // Web Accessibility Initiative (WAI) | W3C. Retrieved October 5, 2019 from https://www.w3.org/WAI/ fundamentals/accessibility-intro

3. Tetyana Tarnavska. Classification of Information Technology in Education. International Journal of Scientific & Engineering Research Volume 4, Issue3, March-2013 1

4. Speech Recognition and Voice Commands for websites – Voxpow.com. Retrieved February 23, 2022, from https://voxpow.com/

5. Handsfree for Web – Voice Control – Browse using your voice. Retrieved February 24, 2022, from https://www.handsfreeforweb.com/en/