

Проектування та створення інформаційної системи голосового управління розумного будинку.

Спроековано інформаційний кросплатформенний додаток управління розумним будинком з голосовим інтерфейсом. Визначено структурну схему, Встановлено, що всі можливі переходи та зміни станів системи включають в себе голосове управління.

Домашня автоматизація, або розумний будинок – система домашніх пристроїв, здатних виконувати дії і вирішувати певні повсякденні завдання без участі людини. Домашня автоматизація розглядається як окремий випадок інтернету речей, вона включає доступні через інтернет домашні пристрої, в той час як інтернет речей включає будь-які пов'язані через інтернет пристрої в принципі. Системи голосового управління здатні полегшити та спростити взаємодію користувача з комп'ютерною системою. Особливо дана ідея розвивається у напрямку концепції так званих розумних пристроїв. Більше того, іноді голосовий інтерфейс є необхідним компонентом, наприклад, коли йде мова про людей з обмеженими можливостями. Задачею таких систем є виділення та розпізнавання із потоку звукового сигналу набору мовних команд. Причому при введенні якоїсь команди система не повинна реагувати на інші частини мовного сигналу.

Перш за все, необхідно спроектувати схему сайту відповідно до функцій, що має виконувати інформаційна система - це створення продуманої моделі ефективного веб-ресурсу на етапі підготовки до створення дизайну. Створення структури сайту, а саме - структури розділів і інтерфейсу, дозволяє оптимізувати процес розробки. Створений прототип дозволяє оцінити його відповідність завданням і при необхідності внести зміни в попередню модель. Таким чином, розробка дизайну і програмної частини відбуватиметься на основі прототипу, що дозволить уникнути несподіванок і виконати роботу максимально оперативно і ефективно [1].

Таким чином, структурна схема веб-додатку містить форму реєстрації, адміністративну панель та панель вибору кімнати з вікнами управління та тригерами. Панель головних налаштувань містить тригери управління глобальними функціями системи автоматизації. Панель управління кімнатами має доступ до форм керування, що були створені для сторінки вибору кімнати для управління. Панель перегляду зображень кімнат містить галерею із фото, що розташовані за принципом слайдера. Панель екстрених викликів – можливість швидкого доступу до гарячих ліній екстрених служб при виникненні непередбачуваних ситуацій. Усі можливі переходи та зміни станів системи включають в себе голосове управління.

Технологія голосового управління

Web Speech API дозволяє взаємодіяти з голосовими інтерфейсами у веб-додатках. Він складається з двох частин – SpeechSynthesis – перетворення тексту в мову та SpeechRecognition – асинхронне розпізнавання мови. Web Speech API дозволяє веб-програмам обробляти голосові дані. Доступ до мовлення здійснюється через інтерфейс SpeechRecognition, який забезпечує можливість розпізнавання контексту мовлення з аудіовходу (зазвичай за допомогою служби розпізнавання мови за умовчанням пристрою) і відповідати належним чином.

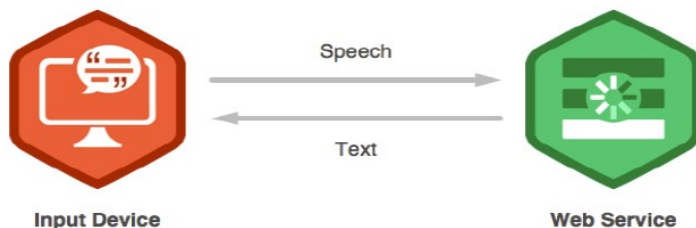


Рисунок 1 – Схема роботи Web Speech API

Як правило, використовується конструктор інтерфейсу для створення нового об'єкта SpeechRecognition, який має ряд оброблювачів подій для виявлення, коли мовлення здійснюється через мікрофон пристрою. Для створення голосового інтерфейсу будемо використовувати бібліотеку annyang.

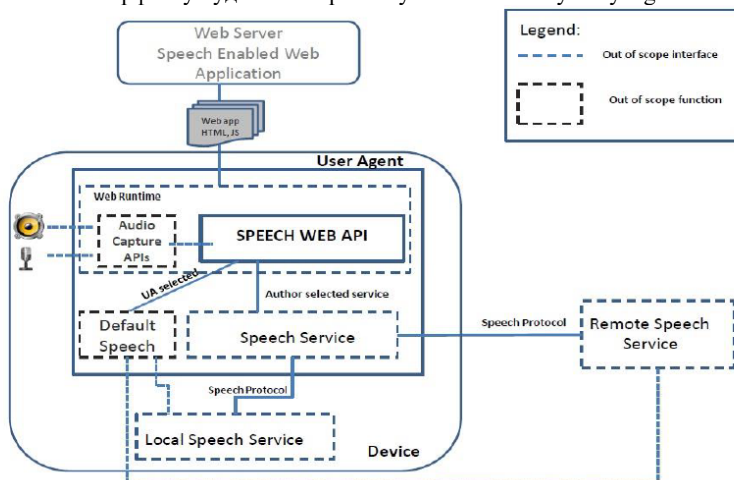


Рисунок 2. – Схема голосового розпізнавання Web Server

Інтерфейс SpeechGrammar представляє контейнер для певного набору граматики, який прилад повинен розпізнати. Граматика визначається за допомогою JSpeech Grammar Format (JSGF) [2].

Синтез мовлення здійснюється через інтерфейс SpeechSynthesis, компонент «текст-мовлення», який дозволяє програмам прочитати текстовий вміст (зазвичай за допомогою синтезатора рішень за замовчуванням пристрою). Різні типи голосів та текстів представлені об'єктами SpeechSynthesisVoice та SpeechSynthesisUtterance. Отримати голосове відображення можна, передаючи об'єкти методу SpeechSynthesis.speak() [3].

Голосове розпізнавання передбачає отримання мовлення через мікрофон пристрою, який потім перевіряється службою розпізнавання мов у списку граматики (в основному, словник, який окремо вказується в кожному додатку) Коли слово або фраза успішно розпізнаються, вони передаються як результат (або список результатів) у вигляді текстового рядку, з якими в результаті можуть бути ініційовані подальші дії. Web Speech API має основний інтерфейс контролера - SpeechRecognition – плюс ряд тісно пов'язаних інтерфейсів для представлення граматики, результатів тощо. Як правило, стандартна система розпізнавання мовлення, доступна на пристрої, буде використовуватися для розпізнавання мови - більшість сучасні ОС мають систему розпізнавання мовлення для випуску голосових команд. Наприклад, Mac OS X, Siri на iOS, Cortana в Windows 10, Android Speech, тощо [4].

Реалізувати голосове розпізнавання команд веб-додатком виконаємо за допомогою бібліотеки JavaScript Annyang. Annyang – невелика бібліотека із відкритим вихідним кодом, важить 2 кілобайта та знаходиться у вільному доступі для використання у проектах відповідно до MIT ліцензії. Розпізнавання мовлення відрізняється залежно від використовуваного протоколу:

<https://> Один раз просить про дозвіл і запам'ятовує вибір.

<http://> Просить багаторазового дозволу на кожну завантаження сторінки. Результати також повертаються значно повільніше в HTTP.

Фактична обробка мовлення та перетворення команд в текст виконується в хмарі, тому швидше з'єднання зумовлює оптимальнішу роботу додатку. Коли розпізнавання мови відбувається в безперервному режимі, результати повертаються повільніше (браузер чекає, коли людина закінчить розмовляти, щоб побачити, чи є щось ще, що людина хоче додати). Якщо вимкнути режим безперервного режиму, веб-переглядач повертає результат набагато швидше. Щоб запустити annyang у неперервному режимі, необхідно вказати continuous: false в об'єкті параметрів, який приймає annyang.start(). Також варто вимкнути autoRestart. Варто звернути увагу, що ці параметри вже є стандартними, якщо використовується HTTPS. Якщо використовується HTTP, за замовчуванням буде активовано безперервний режим (що призведе до повільного розпізнавання), щоб запобігти повторним повідомленням про безпеку [5]. Перш за все, перевіряємо чи розпочалося розпізнавання і чи працює бібліотека, для цього виконуємо перевірку if(annyang).

Далі створюємо функції, що додають логін і пароль до полів функції згідно з тим, що вимовив користувач. Створюємо перелік команд та ставимо їм у виконання функції відповідно до розпізнаної фрази. Таким чином, *name – означає параметр, що передається у функцію login(name) та буде встановлено у полі логін форми входу. Аналогічно виконується команда введення паролю. Якщо виконати команду 'sign in', вона запускає на виконання створену раніше

функцію signIN(), яка перевіряє наявність зареєстрованого користувача із введеними логіном та паролем. Для переходу між сторінками застосовуємо location.href. Створимо голосове розпізнавання для прокручування сторінки до початку, вгору та вниз. Для цього додаємо скрипт. Функції, що виконуються є звичайними у мові JavaScript. Для додавання голосового інтерфейсу варто лише створити відповідну команду, якій передати або існуючу функцію, або передати активність певного елементу чи створити нову функціональну одиницю, яка вбудована в псевдоеlementи CSS. Голосові функції веб-додатку є аналогічними з функціями, що здійснюються за допомогою допоміжних пристроїв: клавіатури, миші, тощо.

Опис створеного веб-додатку з можливістю голосового керування

Перша сторінка веб-додатку містить форму із полями для введення імені користувача – логіну та пароллю. Для перевірки введених даних та переходу далі до функціоналу веб-додатку існує кнопка для входу. Всі дані можна вводити вручну через клавіатуру або виконати голосові команди «username *name», де name – ім'я користувача. Додаток розпізнає слово чи фразу та вводить дані в поле «username». Аналогічно команда «password *pass» здійснює запис пароллю в поле «password». Для перевірки форми та переходу до головного функціоналу веб-додатку необхідно виконати голосову команду «sign in». Якщо логін і пароль відповідають певному зареєстрованому користувачу, веб-додаток переходить на сторінку з вибором кімнати.

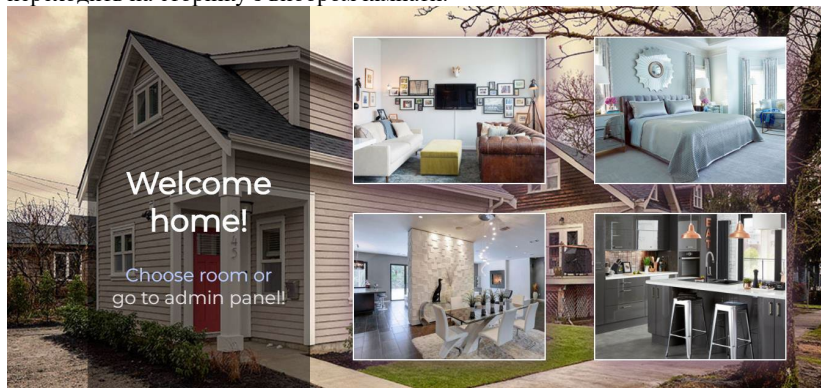


Рисунок 3 – Відображення сторінки із вибором кімнати

Для вибору кімнати необхідно в стандартному режимі навести мишею на певне зображення, а в голосовому режимі необхідно виконати команду із назвою кімнати. Наприклад, 'living room', 'bedroom', 'dining room', 'kitchen'. Для можливості керування елементами кімнати необхідно натиснути на кнопку на зображенні відповідної кімнати, або виконати голосову команду «show»+ назва кімнати. Наприклад, «show living room». З'являється модальне вікно із тригерами керування.

Щоб увімкнути певний елемент необхідно або натиснути на тригер навпроти назви, або виконати команду із назвою елементу та додати «on», наприклад,

«ceiling lights on». Для вимкнення тригера необхідно виконати команду аналогічну команду із словом «off», наприклад, «ceiling lights off».

Щоб закрити модальне вікно, можна натиснути на кнопку «хрестик», або виконати голосову команду «close room», або команду «close» + назва кімнати. Для переходу на адміністративну панель необхідно виконати команду «go to admin panel», «show admin panel», «admin panel» або натиснути на цю команду мишею. На адміністративній панелі є можливість перегляду поточної погоди відповідно до міста. Для перегляду необхідно виконати команду «set location *city», де city – місто, для якого потрібно дізнатися стан погоди.

Для встановлення головних налаштувань будинку необхідно виконати команду «назва елемента» + «on», для вимкнення тригера аналогічно з командою «off», або натиснути на тригер навпроти елементу. Для перегляду галереї зображень необхідно натиснути на панель «Rooms photos», або виконати голосову команду «show rooms» або «rooms photos». Для переходу між фото галереї необхідно виконати свайп або голосові команди «next photo» або «next room» - для переходу до наступного фото, «previous photo» або «previous room» - для переходу до попереднього фото. Щоб вийти з галереї треба клікнути на вільну зону екрану або виконати голосову команду «close view» або «close rooms». Для набору екстреного номеру треба просто виконати команду за назвою служби, або натиснути на номер. Додаток є кросплатформеним, що дає змогу використовувати його на різних пристроях з різними розширеннями екрану.

Таким чином, створений веб-додаток містить елементи управління, як через фізичний набір чи клік, так і через голосове управління. Створена інформаційна система має адаптивну структуру, тому є можливість кросплатформеного використання.

Список літератури

1. Шаблон інтерфейса для «умного дома». Режим доступу: <https://geektimes.ru/post/250858/>
2. Introducing the Web Speech API. Режим доступу: <https://www.sitepoint.com/introducing-web-speech-api/>
3. JavaScript: Web API чтение текста и распознавание голоса. Режим доступу: <https://archakov.im/post/javascript-web-api-recognition-and-speech-text.html>
4. Understand the Smart Home Skill API. Режим доступу: <https://developer.amazon.com/docs/smarthome/understand-the-smart-home-skill-api.html#how-the-smart-home-skill-api-works>
5. annyang! Tutorial. Режим доступу: <https://github.com/TalAter/annyang>