

РАЗРАБОТКА И СОЗДАНИЕ ДЕСКТОПНОГО ГОЛОСОВОГО АССИСТЕНТА

В настоящее время огромную популярность набирают голосовые ассистенты и различные системы голосового управления, однако подобные системы не являются локальными, а производят обработку данных на серверах и для большинства задач достаточно небольшого словаря.

Обычный способ распознавания речи заключается в следующем: берутся аудиоданные, разделяются по высказываниям, по молчаниям, а затем распознается, что говорится в каждом высказывании. Для этого берутся все возможные комбинации слов и сопоставляются их с аудио, затем выбирается лучшая подходящая комбинация.

В соответствии со структурой речи для её распознавания используются три модели: акустическая модель, фонетический словарь, языковая модель. В рамках решаемой задачи необходимо было реализовать ещё один уменьшенный словарь, содержащий в себе только команды, которые может выполнять голосовой ассистент. Такой словарь представляет собой ветвящуюся структуру, именно на данном этапе создается логика взаимодействия с пользователем.

Следующим этапом в разработке голосового ассистента является реализация аудио модели голоса человека, что представляет собой звуки и слова в битовом формате и соответствующие им захэшированное символьное представление звука или слова.

Для реализации задачи использовался объектно-ориентированный язык программирования Java 15 и интегрированная среда раз-

Современные информационные технологии
Прикладные программно-аппаратные системы

работки JetBrains IntelliJ IDEA 2020.1. В качестве основы для распознавания речи – переделанная библиотека с открытым исходным кодом Sphinx4.

В результате был разработан голосовой ассистент, который считывает аудиоданные с микрофона, переводит их в понятные ему команды, а затем выполняет их. Так же в функционал разработанного голосового ассистента входит синтез речи.