# Государственное бюджетное общеобразовательное учреждение

средняя общеобразовательная школа №617

Приморского района Санкт-Петербурга

Проектная работа

«Голосовой помощник в Windows 10»



Выполнили:

Ученики 10ЛИ класса

Струков Артём

Левшенков Николай

Руководитель:

Преподаватель ITCube школы №617

Постаногова Елена Олеговна

Санкт-Петербург

2023

# **Оглавление**

[**Введение** 3](#_Toc134189752)

[**1.** **Понятие голосового помощника** 4](#_Toc134189753)

[**2.** **История развития голосового помощника** 5](#_Toc134189754)

[**3.** **Развитие голосовых помощников на примере Alexa** 7](#_Toc134189755)

[**4.** **Как голосовые помощники распознают речь** 8](#_Toc134189756)

[**5.** **Использование голосовых помощников** 9](#_Toc134189757)

[**6.** **Работа над проектом** 10](#_Toc134189758)

[**Заключение** 11](#_Toc134189759)

[**Список использованных источников** 12](#_Toc134189760)

[**Приложение** 13](#_Toc134189761)

# **Введение**

Как часто вы пользуетесь голосовыми поиском, либо голосовым ассистентом на своем смартфоне? С каждым годом количество поисковых запросов с помощью голоса растет в геометрической прогрессии, буквально каждый пятый владелец мобильного устройства не упускает возможность сэкономить время и «завойсить» интересующий вопрос в интернете. По прогнозам экспертов, скоро больше половины всех запросов в интернете будут выполняться с помощью технологий распознавания речи.

Мы хотим разработать голосового помощника, который будет помогать взаимодействовать с Windows и выполнять ваши команды, т.к при работе с компьютером у нас могут возникнуть разные проблемы, неудобства с управлением компьютерной мышью под некоторыми факторами). Из этого мы сделали вывод, что в Windows не хватает некоторых функций ГП[[1]](#footnote-1)

**Цель:** Разработать голосовой помощник для организации взаимодействия голоса с Windows 10.

**Задачи:**

1. Работа с информационными источниками.

2. Определение объектно-предметной области.

3. Анализ функций Windows.

4. Разработка ГП на основе полученной информации.

5. Апробация продукта.

6. Презентация и оценка результата.

7. Доработка продукта.

**Объект:** Организация взаимодействия голоса с Windows 10.

**Предмет:** ГП.

**Методы исследования:** Проанализировав информацию о ГП и на основе массового опроса среди пользователей ПК в возрасте от 16 до 75 лет, мы поняли, что ГП позволит людям эффективнее взаимодействовать с компьютером.

**Гипотеза:** На основе проведённого опроса, мы пришли к выводу, что добавление взаимодействия голоса с Windows 10 может помочь людям, у которых возникают проблемы или неудобства с управлением ПК мышью.

**Этапы работы над проектом:**

1. Выявление проблемы и области исследования.

2. Работа с информационными источниками

3. Определение объектно-предметной области

4. Описание теоретической части

5. Разработка проекта

6. Проведение анкетирования

7. Разработка проекта

8. Повторное анкетирование

9. Презентация проекта

# **Понятие голосового помощника**

ГП –Виртуальный ассистент с функциями голосового помощника.

Виртуальный ассистент (англ. Virtual assistant) — программный агент, который может выполнять задачи (или сервисы) для пользователя на основе информации, введенной пользователем, данных о его местонахождении, а также информации, полученной из различных интернет-ресурсов (погода, уличное движение, новости, курсы валют и ценных бумаг, розничные цены в магазинах и т. д.).

Програ́ммный аге́нт — это программа, которая вступает в отношение посредничества с пользователем или другой программой.

Голосовой помощник – это бот, работа которого построена на искусственном интеллекте, а именно - на механизме распознавания речи. Основной задачей такого бота является быстрое решение простых задач, таких как поиск информации в сети интернет или выполнение команд на смартфоне и других цифровых устройствах. Для работы голосового помощника компании-разработчику необходимы не только мощная аппаратная часть, но и крупная серверная сеть для обработки и хранения данных.

«Бот» (сокращение от «робот») – это программа, выполняющая автоматические заранее настроенные повторяющиеся задачи. Боты обычно имитируют поведение пользователя или заменяют его. Боты являются автоматизированными, потому они работают намного быстрее, чем пользователи. Они выполняют полезные функции, например, обслуживание клиентов или индексация поисковых систем.

Иску́сственный интелле́кт (ИИ; англ. artificial intelligence, AI) — свойство искусственных систем выполнять творческие функции, которые традиционно считаются прерогативой человека; наука и технология создания интеллектуальных машин, особенно интеллектуальных компьютерных программ.

Интеллектуа́льная систе́ма (ИС, англ. intelligent system) — это техническая или программная система, способная решать задачи, традиционно считающиеся творческими, принадлежащие конкретной предметной области, знания о которой хранятся в памяти такой системы.

Распознавание речи — автоматический процесс преобразования речевого сигнала в цифровую информацию.

СХД (Система хранения данных или Сервер хранения данных) — это устройство для хранения и управления данными, их резервного копирования. Она призвана решить типичные проблемы, связанные с растущими объемами информации в любой организации.

# **История развития голосового помощника**

С ростом технологий и возможностей, голосовые помощники медленно, но верно, становятся повседневной частью жизни. Однако эта идея не была всегда столь популярной. В этой статье мы постараемся рассказать о возникновении голосовых помощников, о том, как они развивались и что будет дальше.

История голосовых помощников начинается с исследователей, которые давно ищут способы, чтобы сделать взаимодействие человека с компьютером более естественным и удобным. Долгое время компьютеры были управляемы только через клавиатуру и мышь, что часто затрудняло работу в определенных условиях, таких как вождение автомобиля или путешествие на общественном транспорте.

Первым шагом в истории голосовых помощников был задуманный в 1960-х годах система “Shoebox”, которая использовала технологию распознавания ключевых слов. К сожалению, эта система не работала стабильно и не была широко распространена.

С тех пор технологии стали более совершенными, и к 1980-м годам был разработан компьютерный голосовой помощник. Однако, эта система использовалась только в стенах лаборатории - там, где носители языка могли работать со специальной аппаратурой, которая была необходима для корректной интерпретации звуков человеческого голоса и его правильного распознавания даже в шумных условиях.

Развитие голосовых помощников значительно ускорилось в последние годы. В 2010 году Apple выпустила Siri, голосового помощника для своих устройств. Siri был прообразом для многих последующих голосовых помощников, таких как Alexa от Amazon или Google Assistant.

Технологии голосовых помощников все еще совершенствуются, но уже сегодня они предоставляют функции для настройки будильников, выполнения поиска в Интернете, управления календарем, покупок онлайн и многих других. Безусловно, голосовые помощники с годами будут становиться еще более сложными, чтобы лучше соответствовать нашим потребностям и исполнить ежедневные задачи.

Развитие голосовых помощников происходило в сочетании с дополнительными технологиями, например, распознавание речи. Они были разработаны на основе глубоких нейронных сетей и множественных алгоритмов машинного обучения. В этом случае нейронные сети используются, чтобы обучить приложение распознавать звуки человеческой речи и осуществлять перевод в текстовый формат.

Возможности для использования голосовых помощников настолько широки, что они стали незаменимыми во всех сферах жизни. Голосовые помощники умеют контролировать степень освещенности в доме, отвечать на вопросы, управлять планированием вещей, производить платежи, выполнять заказы и многое другое.

Голосовые помощники основаны на распознавании речи и анализируют голос пользователя. Однако, с каждым днем количество голосовых помощников неуклонно растет, и каждый раз увеличивается их функциональность.

В заключении, можно сказать, что голосовые помощники находятся в поиске своей прикладной пользы и выхода на новые рынки, трансформирующиеся каждый день и развивающие свои услуги для пользователей.

в 1779 году было разработано первое в истории устройство для синтеза речи, созданное немецким инженером — Христианом Кратценштайном. Работая в России, он создал аппарат, который мог воспроизводить всего несколько гласных звуков (а, э, и, о, у)

Изобретение ученого, сотрудника Американской лаборатории Bell — Гомера Дадли — является самым первым электронным речевым синтезатором. Аппарат Voder был представлен в 1937 году.

В 1942 году, несмотря на войну и ужасы блокадного Ленинграда, советский физик Лев Мясников защитил свой проект, продемонстрировав фокус-группе аппарат, умеющий распознавать несколько гласных и согласных звуков.

В 1952 году был представлен первый голосовой распознаватель чисел от 1 до 9, разработанный в стенах лаборатории Bell

Спустя ровно десять лет компания IBM представила собственную технологию распознавания речи — Shoebox.Система могла распознавать кроме 16 заложенных слов и 6 цифр еще и 10 математических команд.

В период 1990-х появилось достаточно большое количество систем распознавания речи от разных издателей, и каждая из них становилась все лучше и лучше. Так, например, в 1990 году увидела свет одна из первых коммерческих программ по распознаванию речи — Dragon Dictate.

# **Как голосовые помощники распознают речь**

Распознавание речи - это процесс преобразования аудио-сигнала человеческой речи в текстовую форму, понятную компьютеру. Это позволяет использовать голосовые команды и интерфейсы в компьютерных программах, улучшает доступность к информации для людей с ограниченными возможностями и упрощает процессы взаимодействия человека с компьютером. Распознавание речи является сложным процессом, включающим в себя несколько этапов и алгоритмов обработки данных.  
Опознавание речи - это процесс идентификации голосовых характеристик человека, позволяющий установить личность говорящего. Это может быть полезным в различных областях, таких как безопасность, медицина и управление персоналом. Опознавание речи может осуществляться при помощи алгоритмов машинного обучения, которые анализируют спектральные характеристики голоса, такие как высота, длительность и интонация. Эти данные затем используются для создания уникального голосового профиля, который может быть использован для идентификации личности пользователя.

Основные компоненты программного обеспечения распознавания речи включают в себя акустическую модель, языковую модель и модель декодирования. Акустическая модель отвечает за определение звуков, содержащихся в аудио-сигнале, и их преобразование в последовательность чисел. Эти данные используются для создания матрицы параметров, которая отражает особенности звуков, содержащихся в записи.

Языковая модель определяет наиболее вероятные последовательности слов, вследствие чего возможно более точное преобразование звуковой информации в текст. Для этого используются словари со списком наиболее часто используемых слов в языке.

Модель декодирования принимает результаты, полученные из акустической и языковых моделей, и генерирует окончательную расшифровку речи. Это происходит путем применения алгоритмов, которые принимают во внимание как правильно сформулировано предложение, так и правильно произнесены отдельные слова.

Распознавание речи осуществляется в несколько этапов. Сначала аудио-сигнал разбивается на малые фрагменты, называемые кадрами. Затем каждый кадр подвергается цифровой обработке, что позволяет выделить характеристики звука. Эти характеристики затем используются для формирования матрицы параметров.

Для обработки этой матрицы применяются различные алгоритмы, такие как нейронные сети и скрытые марковские модели, которые используются для анализа и классификации звуковых сигналов.

После анализа каждого кадра создается последовательность звуков, которая затем используется для формирования предложения. Это происходит при помощи языковой модели, которая использует вероятностный подход, чтобы определить последовательность наиболее вероятных слов.

Наконец, модель декодирования используется для преобразования этой последовательности звуков в окончательный текст. Эта модель использует большую базу знаний о конкретном языке и интеллектуальные алгоритмы, что позволяет распознать даже трудные случаи, такие как шум или возможные искажения звука.

В заключение, распознавание речи и опознавание речи играют важную роль в современных технологиях, и тесно связаны друг с другом. Распознавание речи позволяет компьютерам понимать, что говорит пользователь, а опознавание речи может быть использовано для установления личности говорящего. Оба этих процесса выполняются при помощи комплексных алгоритмов и используются во многих областях, включая медицину, телекоммуникации, автомобильную промышленность и многие другие.

# **Использование голосовых помощников**

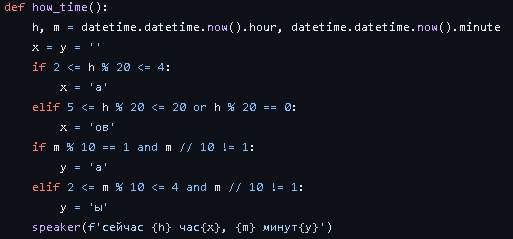
Сегодня, голосовые помощники стали незаменимым инструментом в ряде сфер нашей жизни:

* Домашняя автоматизация: голосовые помощники, такие как Apple Siri, Google Assistant и Amazon Alexa, используются для управления устройствами домашней автоматизации, такими как умные термостаты, устройства освещения и умные замки. Они могут помочь включить свет или вентилятор, открыть дверь или закрыть окно, включить телевизор или воспроизвести музыку.
* Здравоохранение: голосовые помощники используются в медицинских учреждениях для помощи пациентам и консультаций с врачами. Они могут предоставлять пациентам расписание приема медикаментов, напоминать о важных приемах и даже давать советы по лечению.
* Образование: голосовые помощники могут помочь студентам и преподавателям в учебном процессе. Они могут помочь с поиском информации для исследований, давать практические примеры и задания, предоставлять исторические факты и т.д.
* Персональная организация: голосовые помощники, такие как Siri и Google Assistant, могут помочь в личной организации. Они могут помочь составить список покупок, уведомить о важных заданиях и напомнить о прошлых событиях.
* Бизнес-сектор: голосовые помощники используются в бизнес-секторе для автоматизации усилий и повышения эффективности работы, путем управления элементами, такими как планирование встреч, проведение конференций и мониторинг состояния финансовой отчетности.
* Транспорт: голосовые помощники используются в автомобильной промышленности для управления элементами, такими как навигация, системы связи и инфотейнмент.
* Развлечения: голосовые помощники используются в развлекательной промышленности, чтобы сделать игровой процесс более интерактивным и универсальным. Они могут помочь участникам посмотреть телевизионную программу, заказать фильм или сделать заказ в интернет-магазине.

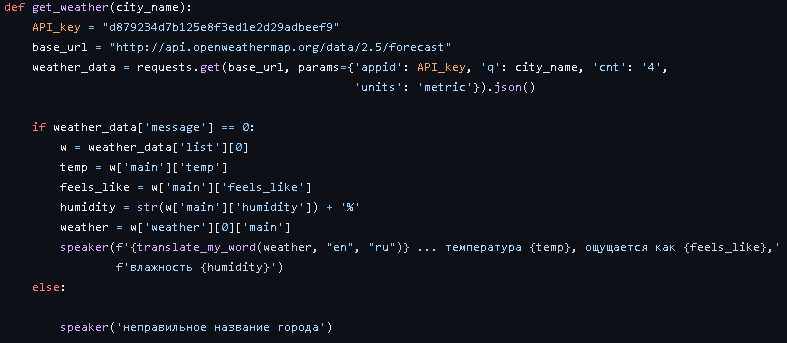
В целом, голосовые помощники широко применяются в разных сферах жизни и продолжают развиваться с каждым годом. Они помогают управляться с повседневными проблемами и упрощают рабочий процесс для тех, кто работает в других сферах.

# **Работа над проектом**

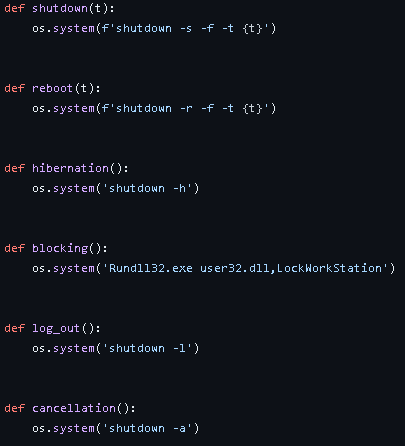
* Разработка некоторых, подфункций ГП таких как:
  + Настроить напоминания по времени
  + Узнать сколько времени в данный момент



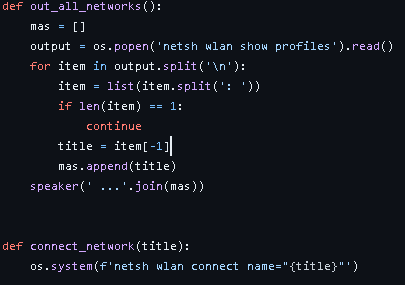
* + Поиск в Интернете с помощью Yandex
  + Выполнение вычислений и конвертация
  + Проверить погоду (с помощью API)



* + Базовые команды управления
    - Выключение
    - Перезапуск
    - Гибернация
    - Блокировка
    - Выход из пользователя
    - Отмена некоторых, из выше перечисленных действий



* + Переключение между сетями, что использовались ранее



* + Переключение языка раскладки клавиатуры
  + Пролистывание страницы
  + Регулировка громкости и яркости



* + Получить список команд / Справка
  + Функцию для шуток
* Разработка программы распознавания и воспроизведения речи
* Объединение в единую систему разработанных приложений
* Доработка разработанных подфункций под требования ГП

# **Заключение**

Мы реализовали некоторые функции необходимые для базового управления ПК, это частично облегчает взаимодействие с Windows 10, но мы далеки от того, чтобы полностью заменить мышь и клавиатуру голосовым помощником. Мы подтвердили поставленную гипотезу, а также узнали новые способы улучшения взаимодействия с ПК с помощью голоса. В будущем мы планируем улучшать и дорабатывать представленный проект новыми функциями, чтобы в какой-то момент он смог стать не голосовым помощником, а голосовым ассистентом, с помощью которого можно будет забыть о существовании таких вещей как мышь и клавиатура.

# **Список использованных источников**

1. [wikipedia.org](https://ru.wikipedia.org/wiki/Wi-Fi)

1. [vc.ru](https://vc.ru/)

1. [kaspersky.ru](https://www.kaspersky.ru/)

1. [reg.ru](https://www.reg.ru/)

1. [trashbox.ru](https://trashbox.ru)

1. [yandex.ru](https://yandex.ru)
2. forbes.com
3. nytimes.com
4. cnet.com
5. wired.com
6. businessinsider.com
7. technologyreview.com

1. Голосовой помощник [↑](#footnote-ref-1)