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

2022-2023

# **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**

Первая стадия в работе голосового ассистента — это обработка звука.

Если сказать помощнику «Лев Толстой», то он услышит не два слова, а звуковой сигнал, в котором гласные и согласные плавно перетекают друг в друга.

Процесс распознавания осложняется тем, что одни и те же фразы, произнесённые разными людьми в разной обстановке (например, в вагоне метро или в безлюдном парке), будут давать непохожие друг на друга сигналы. Поэтому для распознавания речи голосовой помощник делит слова на звуковые фрагменты — фонемы. Носители языка узнают именно их: например, первые гласные в словах «дом» и «домашний» мы воспринимаем одинаково, хотя звуки разные. В Алисе различают всего 48 фонем — в соответствии с подходом Ленинградской фонологической школы.

Для того, чтобы распознавать фонемы, помощник работает с их числовым представлением. Так, когда вы произносите запрос, смартфон записывает его и отправляет на сервер компании, которая разрабатывала голосового помощника. Там запись разделяется на множество маленьких фреймов — фрагментов речи длиной в сотые доли секунды.

Затем каждый фрейм подвергают математическим преобразованиям и в результате получают коэффициенты, описывающие его частотные характеристики. А на основании этих данных ассистент может предположить, к какой фонеме принадлежит фрейм.

Как и человек, голосовой помощник не всегда хорошо распознаёт звуки, и ему приходится достраивать слова исходя из контекста и статистики. У него есть таблица вероятностей переходов, которая указывает, что после «к» с одной вероятностью будет «о», с другой — «т» и так далее. Это позволяет определить варианты последовательности фонем и потом варианты слов, которые вы могли сказать.

Голосовому помощнику нужно учитывать множество акцентов, в которых фонемы звучат по-разному. Например, в Москве чаще будет произноситься гласная «а», а в Архангельске — «о».

# **Использование голосовых помощников**

Одним из самых ярких примеров использования голосовых помощников являются умные колонки и устройства домашней автоматизации – например, Amazon Echo и Google Home. С их помощью пользователи могут не только слушать музыку, но и управлять своим домом, задавать вопросы в Интернете и выполнять другие задачи.

Голосовые помощники облегчают повседневную жизнь людей. С их помощью пользователи могут находить нужную информацию, создавать списки покупок, просматривать графики работы компаний, читать и отправлять электронные письма, заказывать пиццу и многое другое.

Однако, это только начало развития. В будущем голосовые помощники могут даже заменить настоящих помощников, будь то личный помощник в офисе или домашний помощник. С развитием искусственного интеллекта голосовые помощники становятся все более умными и полезными.

# **Работа над проектом**

* Разработка некоторых, подфункций ГП таких как:
  + Открытие\поиск папок\файлов\приложений
  + Голосовой ввод
  + Настроить напоминания по времени
  + Поиск в Интернете с помощью Google (или другой поисковой системы)
  + Выполнение вычислений и конвертация
  + Проверить погоду
  + Создание событий календаря
  + Получить список команд / Справка
  + Базовые команды управления (выход, сон и т.д.)
  + Переключение между сетями
  + Регулировка громкости и яркости
* Разработка программы распознавания и воспроизведения речи
* Объединение в единую систему разработанных приложений
* Доработка разработанных подфункций под требования ГП

# **Заключение**

Голосовые помощники стали тесно связаны с нашими ежедневными делами. Люди настолько привыкли к ним, что уже не могут обойтись без них. Среди компаний, производящих и создающих голосовых ассистентов, с каждым днем возрастает конкуренция. Как показывают исследования , на популярность продукта в первую очередь влияет практичность и простота в использовании. Новые технологии настолько быстро вникают в нашу повседневную жизнь, что уже и самим людям приходится обучаться правилам использования гаджетов. Но как быть со старшим поколением, которому сложно свыкнуться с миром быстроразвивающихся технологий? Данный фактор имеет весомое влияние в гонке между создателями . По показаниям исследований можно сказать, что популярность голосовых помощников напрямую зависит от следующих факторов: территориальный, языковой и человеческий. И самое главное: данные, которые мы можем предоставить сегодня, завтра будут неверными, потому что, как и было сказано ранее, мир новых технологий ежесекундно растет и людям ежедневно нужно познавать что-либо новое.

# **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)