

ОБРАБОТКА ГОЛОСОВЫХ ЗАПРОСОВ С ИСПОЛЬЗОВАНИЕМ ОБЛАЧНЫХ ТЕХНОЛОГИЙ

Попов И. А. (федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»)
**Научный руководитель – доцент факультета инфокоммуникационных технологий
Добриборщ Д. Э.**

(федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»)

Целью доклада является демонстрация прикладного применения технологии распознавания речи с использованием услуг облачной инфраструктуры для обработки голосовых запросов в качестве команд, предназначенных к исполнению браузером. Ключевой составляющей проекта является процесс интеграции различных приложений и сервисов, а также особенности их взаимодействия между собой.

Введение. В последние годы все большую актуальность обретает возможность голосового управления различными процессами как в промышленной, так и в социальной сферах. Голосовые помощники и ассистенты помогают человеку как, например, на высоконагруженных складах, автоматически отбирая нужные товары, так и управляют светом и температурой в «умных домах».

Помимо технологичных устройств с удобными голосовыми интерфейсами и уникальных приложений, популярнее и доступнее становятся сервисы, предоставляющие возможность сторонней обработки речи по требованию. Подобные сервисы, обеспечивающие использование программного обеспечения как услуги, базируются в облачных инфраструктурах и являются мощным инструментом, позволяющим решать широкий спектр различных задач от машинного перевода потоковой речи до синтеза текста.

Основная часть. Основными преимуществами использования облачных сервисов обработки голосовых запросов является отсутствие необходимости развертывать сложную архитектуру и самостоятельно заниматься машинным обучением.

Основная цель данной работы заключается в создании системы, демонстрирующей современный потенциал облачных технологий по распознаванию речи, а также универсальность их программных интерфейсов, обеспечивающих адаптивность в применении и простоту интеграции в проекты, использующие и сочетающие в себе разноплановые технологии.

Ключевой идеей, заложенной в основу разработки, является объединение функционала имеющихся приложений и сервисов для последовательного решения задач приема голосовых команд от пользователя, их обработки, и трансформации в перечень команд, позволяющих управлять браузерными событиями на удаленном компьютере.

Выводы. В ходе работы над проектом было разработана система, использующая мессенджер Telegram в качестве платформы для приема голосовых сообщений и облачного сервиса, используемого для их обработки и распознавания.

За счет использования облачных технологий, предложенное решение имеет легковесную архитектуру, основанную на внутреннем обмене обработанными данными, и предоставляет возможность пользователю осуществлять удаленное управление веб-браузером компьютера с помощью голосовых сообщений, отправляемых в мессенджере.

Попов И. А. (автор)

Добриборщ Д. Э. (научный руководитель)

