

ПРОГРАММИРОВАНИЕ НАВЫКОВ ГОЛОСОВОГО ПОМОЩНИКА НА ПРИМЕРЕ УМНОЙ КОЛОНКИ «АЛИСА»

А.Е. Булай, С.Б. Протосовицкая

Современные технологии развились до того уровня, когда у большинства людей в кармане находится средство управления обширным количеством информации и инструмент, способный решать задачи огромных масштабов – смартфон. Люди, не задумываясь, пользуются поиском, онлайн-сервисами, слушают музыку и смотрят

фильмы. Однако на бытовом уровне человеку все же привычнее общаться голосом, разговаривать. Темп и динамика жизни приводят к тому, что у людей нет времени читать длинные сообщения, у них даже нет времени писать текст. Эти и другие факторы стали причиной создания голосового помощника (иначе его называют голосовым ассистентом). В работе рассматривается задача программирования навыков для голосового ассистента Алиса. Под навыком будем понимать сервис, реализующий диалог, который запускается некоторой командой и расширяет возможности голосового ассистента. Объект исследования: голосовые ассистенты. Предмет исследования: программирование голосовых ассистентов. Методы исследования – анализ, обобщение, формализация, эксперимент.

Ведущие ИТ-компании мира (в частности, Apple, Amazon, Google, Microsoft, Alibaba, Baidu) настойчиво рекламируют и внедряют в свои продукты голосовых помощников на основе искусственного интеллекта. «Яндекс Алиса» – виртуальный голосовой помощник, который уникален своей человечностью. С ним, точнее с ней, можно общаться. Она распознает речь, поддерживает диалог, шутит, иногда вольничает и острит, дает ответы на вопросы и, благодаря запрограммированным навыкам, решает простые прикладные задачи.

Когда Алисе задают вопрос, она распознает голос пользователя и превращает его в текст. Эта технология называется SpeechKit. В Алисе собрано более миллиарда сообщений с разными голосами, произношениями и акцентами. После распознавания, помощник проецирует фразу в семантическое пространство широкой размерности. В этом пространстве находится множество точек. Каждая точка этого пространства – высказывание, которое имеет свою тему. Далее голосовой ассистент находит необходимое место и сравнивает запрос с подходящим ответом. Предугадать, что он ответит невозможно. Когда ответ выбран, Алиса его произносит с помощью технологии Text-to-Speech и нейросети, что придает голосу натуральность.

Ко всей полезности и красоте технологии, существуют пару недостатков.

1. На платформе Яндекс.Диалоги, бывает, что добавленный навык не проходит модерацию. Причиной может служить слишком общее имя навыка, например, «погода в Гродно», что схоже с «погода в Минске».

2. Документация платформы Яндекс.Диалоги не всегда соответствует действительности, так как платформа развивается слишком быстро

3. Конфиденциальность данных, используемых в диалоге с Алисой.

Последний пункт хочется выделить особенно: Алиса не слышит пользователя до момента ее активации голосом с помощью уникального слова для ее включения или кнопкой. Помимо этого, вся информация, которую сообщают Алисе, проходит сложное и надежное шифрование. Однако, при использовании голосовых помощников или умных колонок рекомендуется ограничить действия, которые они могут совершать, а также избавиться от случайной активации путем определения уникального слова для его включения [1]. Голосовые помощники облегчают повседневную жизнь людей, помогают людям с ограниченными возможностями, дают ресурсы для реализации потребностей своих пользователей.

Литература

1. Безопасны ли умные колонки? / Защитные решения для дома и бизнеса | Лаборатория Касперского [Электронный ресурс]. – Режим доступа: <https://www.kaspersky.ru/resource-center/threats/how-to-improve-your-smart-speaker-privacy>. – Дата доступа: 20.03.2021.