



ЛЕКСИ — ГАДЖЕТ С ХАРАКТЕРОМ

ДМИТРИЙ СУВОРОВ
sales@lexybot.com

В последние десятилетия в информационных технологиях четко прослеживаются тренды упрощения доступа к информации, автоматизации рутинной работы, упрощения интерфейсов доступа к функциям бытовой техники. Интерфейсы смартфонов и компьютеров становятся интуитивнее, Интернет — дешевле, быстрее и доступнее. Гаджеты действительно делают нашу жизнь комфортной. Тем интереснее, что ожидать от их разработчиков в ближайшем будущем. О домашнем роботе, который все видит, слышит и может пообщаться с хозяином, рассказал сооснователь и генеральный директор ООО «ЛЕКСИ» Дмитрий Суворов.

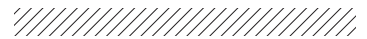
Системы, основанные на анализе больших данных и методах искусственного интеллекта, все плотнее входят в нашу жизнь. Многие люди уже не могут представить свой быт без компьютера, смартфона и интернета. Конечно, во многом такое мнение навязано маркетологами, но современные гаджеты и интернет-сервисы действительно сильно упрощают жизнь и решают множество проблем. В мире сложились целые индустрии по разработке потребительских гаджетов, программного обеспечения к ним и интернет-сервисов. На первый взгляд, эти индустрии развиваются очень динамично: каждый месяц корпорации выпускают новые модели смартфонов и планшетов, каждый год выходят новые версии операционных систем, каждый день появляются новые ИТ-стартапы. Но если приглядеться повнимательнее, станет видно, что новая версия ОС отличается от предыдущей только дизайном, у нового смартфона просто чуть лучше камера и чуть больше экран,

интернет-сервисы часто конкурируют друг с другом только дизайном. Многие новые аппаратные продукты не предоставляют своим покупателям качественно новых возможностей. Производители лишь немного улучшают качество своих предыдущих версий или моделей.

На этом фоне команда проекта «Лекси» пришла к решению начать работу над совершенно новым продуктом, который даст своим пользователям принципиально новые возможности и благодаря этому будет конкурентен на мировом рынке. Обзор нашей разработки начнем с проблем, которые она решает.

Текущие технологии предполагают использование визуального канала для получения информации и взаимодействия с техникой — мониторы компьютеров, экраны телефонов и планшетов, видеопроекторы на конференциях и совещаниях, дисплеи на стиральных машинах, роботах-пылесосах, микроволновых печах... Визуальный канал восприятия информации понятен, удобен и при-

Команда «Лекси» состоит из опытных разработчиков, занимающихся программированием, проектированием электроники, твердотельным моделированием и подготовкой производственных процессов. Ключевые члены команды, Дмитрий Суворов и Роман Жуков, ранее успешно реализовали проект робота удаленного присутствия Webot. Сооснователь, инвестор и ментор проекта — Игорь Ашманов, известный в России специалист в области искусственного интеллекта и интернет-бизнеса.
<http://lexybot.com>



вычен, но человек должен находиться в непосредственной близости от экрана. Насколько бы ни был интуитивен визуальный интерфейс, он требует от пользователя его предварительного изучения, что может вызвать проблемы у детей, пожилых людей, технически малограмотных лиц. Также использование только визуального канала не оставляет возможностей для взаимодействия с техникой людям с ограниченными возможностями, у которых отсутствуют необходимые функции: людям с нарушениями зрения (не могут воспринимать информацию), людям с серьезными травмами конечностей (не могут вводить информацию).

В то же время голосовой канал общения является полноценной альтернативой. Он более естественен для человека: дети сначала осваивают

речь, пожилые люди больше к нему привыкли. Часто человеку быстрее и проще воспринимать информацию на слух, а голосовой ввод информации значительно быстрее текстового. Текущие голосовые интерфейсы для техники (голосовое управление автомобилем, «умный дом», Google Now, Siri) являются узкоспециализированными интерфейсами для конкретных устройств, подразумевают ввод информации в виде команд (не позволяют запрашивать информацию на естественном языке), требуют нахождения в непосредственной близости от устройства, предоставляющего этот интерфейс.

Наш проект интеллектуально-домашнего помощника Лекси направлен на создание гаджета следующего поколения, который предоставит голосовой доступ к информации, голосовое взаимодействие с умной техникой, будет значительно удобнее компьютера или смартфона при доступе к Интернету из дома.

Лекси — это настольное устройство, с которым можно общаться при помощи голоса, естественной речью. Мы хотим дать человеку возможность использовать голосовой канал общения, а не визуальный, как в случае с телефоном или компьютером. Взаимодействие с Лекси очень простое: человек спрашивает его о чем-нибудь, а он отвечает. Также Лекси может проявить инициативу и начать разговор самостоятельно. Одна из ключевых особенностей Лекси — то, что это не просто «вопрос-ответная» система, а устройство с характером, в отличие от Siri или Google Now. Лекси может веселиться или скучать. Его можно обидеть. Он хочет человеческого внимания. Лекси умеет взаимодействовать с окружающим миром, реагировать на внешние события. Спектр его возможностей обширен, о них речь пойдет далее.

Итак, с Лeksi можно просто общаться, это ключевой момент. Но Лeksi, кроме того, является и интеллектуаль-

ным помощником: его можно спросить, например, о погоде или пробках на дорогах города. Если человеку скучно, Лeksi развлечет его анекдотом или предложит сыграть в какую-нибудь словесную игру. Можно попросить Лeksi поставить будильник или напоминание, а также он сможет рассказать последние новости и многое другое.

Поскольку Лeksi — это устройство для дома, он способен управлять простыми устройствами: выключать лампочку или чайник, запускать вентилятор...

Проект «Лeksi» решает указанные выше задачи не только связыванием трех давно известных технологий виртуального собеседника, распознавания и синтеза речи в одном устройстве. Решение гораздо глубже. Оно лежит на границе нескольких областей: программирования, инженерии, прикладной лингвистики и психологии. Цель проекта — создать устройство, которое станет естественной частью домашней обстановки, будет вырабатывать у людей привычку использовать его, станет незаменимо полезным для них.

Мы наделили Лeksi личностью, чтобы людям было комфортнее с ним общаться. Замечено, что людям свойственно одушевлять интеллектуальные устройства. Им так проще принять их интеллектуальность. А Лeksi — это не только интеллектуальное устройство из концепции «Интернета вещей», но и гаджет, с которым необходимо разговаривать. Чтобы люди, купившие Лeksi, не забыли про него, поиграв с ним несколько недель, он периодически напоминает о себе, предлагает пообщаться, рассказывает свежие новости, комментирует изменение внешних условий.

Лeksi — это устройство для семейного использования. Для удобного общения со всеми членами семьи ему необходимо различать их. Поэтому Лeksi наделен средствами идентификации диктора по голосу и алгоритмом идентификации пользователей по видео. Узнав человека, с которым сейчас идет общение, Лeksi начинает собирать о нем различную информацию (модель пользователя), которая потом будет использоваться в процессе общения. Модель пользователя составляется на основе двух типов данных — основных, доступных всегда, и дополнительных.

Основные данные — это данные, полученные в результате обработки

Никто не знает, что должен делать робот. И рано или поздно кто-нибудь сделает для роботов то, что Стив Джобс сделал для смартфонов: придумает абсолютно убедительный приятный способ использования существующих технологий...

Например, способ использования домашних роботов пока не придуман: то ли это полезный помощник, секретарь, дворецкий, то ли это собеседник, а может быть, это «умный» дом, который управляется голосом.

Игорь Ашманов

голоса собеседника, с видеокамеры, установленной в Лекси, и извлеченные при помощи виртуального собеседника. Дополнительные данные — это данные, полученные из социальной сети, если пользователь разрешит Лекси их использовать.

Обработывая голос, Лекси может определить возрастную группу и пол своего собеседника; научиться распознавать его эмоции (смех, плач...) и физиологическое состояние собеседника (кашель, чихание...).

При помощи виртуального собеседника Лекси может попытаться узнать возраст и имя собеседника; пол; день его рождения; родственные связи собеседника с остальными пользователями; интересы собеседника (любимые книги, фильмы, спортивные увлечения, интересующую тематику новостей...); текущее настроение (плохое, нейтральное или хорошее); кем работает собеседник и др.

Обработывая видео с камеры, Лекси сможет обнаруживать наличие рядом человека и узнавать текущего собеседника.

Из социальной сети Лекси может узнать, например, дни рождения друзей или собеседника, пол и возраст, семейное положение, а также интересы собеседника.

Некоторые данные о пользователе Лекси получает несколькими способами, поэтому будет возможность проверить адекватность данных. Извлеченные данные передаются рекомендательной системе, информация от которой используется в процессе общения.

При разработке аппаратной составляющей проекта наша команда уделила большое внимание мощной вычислительной базе. Нам удалось запустить распознавание и синтез речи, а также виртуального собеседника прямо «на борту» устройства. Для этого мы разработали собственную материнскую плату на базе четырехъядерного процессора архитектуры ARM7. Лекси может работать без подключения к Интернету. Это важно для начала работы с устройством, так как пользователи сразу достают из коробки работающего Лекси, в отличие от конкурентов, которых еще нужно подключать к Wi-Fi в течение 10–15 минут. Благодаря отсутствию необходимости посылать в Интернет звук для распознавания речи, Лекси не требует быстрого и стабильного интернет-соединения,



что является значительным преимуществом на рынке развивающихся стран (в том числе и России).

Основной технической сложностью в проекте стала реализация качественного распознавания речи. Для решения данной задачи наша команда разработала специализированную звуковую карту, которая захватывает звук на частоте 16 кГц и занимается его предобработкой — шумоподавлением, акустическим эквализацией, благодаря которому Лекси не будет пытаться распознать, что говорит сам, и аналоговым автоматическим регулированием усиления. Кроме того, Лекси можно перебить — например, если он рассказывает сказку, не обязательно слушать ее от начала и до конца. Чтобы добиться таких возможностей, необходима разработка собственного «железа». Для повышения качества распознавания и идентификации по голосу мы создаем собственную акустическую модель нашей звуковой карты. Схема обучения учитывает условия работы при реверберациях.

Лекси — это устройство, включенное дома 24 часа в сутки. Он слышит и видит все, что происходит вокруг. Мы посчитали важной задачей защитить частную жизнь от вмешательства через наши устройства. Разместив все ключевые вычисления (распознавание и синтез речи, искусственный интеллект) на борту устройства, мы лишили возмож-

ных злоумышленников возможности перехватить данные на пути от устройства к серверу. Наш Лекси — один из самых защищенных домашних роботов.

Работа над проектом и первыми прототипами началась летом 2013 г., осенью 2014 г. проект получил резидентство в инновационном центре «Сколково». На данный момент мы находимся на этапе подготовки предсерийного производства, а в конце марта 2015 г. планируем начать сбор предзаказов через систему краудфандинга. ●

