DOI: 10.17117/na.2016.03.03.083

<http://ucom.ru/doc/na.2016.03.03.083.pdf>

Поступила (Received): 24.03.2016

Косач Д.И., Жидкова Л.О., Белехов А.Н. Виртуальные голосовые помощники с элементами искусственного интеллекта

**Kosach D.I., Zhidkova L.O., Belekhov A.N.
Virtual voice assistant with elements of artificial intelligence**

Голосовые ассистенты и умные помощники, наглядно показывают нам, на что именно будут способны персональные роботы в ближайшем будущем. Рассмотрен вариант разработки отечественного полноценного виртуального помощника с голосовым управлением, понимающего устные запросы на естественном, что самое главное, на русском языке

Voice assistants and smart assistants, clearly show us what it will be capable of personal robots in the near future. A variant of the development of domestic high-grade virtual assistant with voice control, understanding verbal requests to the natural, most importantly, in Russian

Ключевые слова: виртуальные голосовые помощники, Siri, Google Now, Cortana

Key words: Virtual voice assistant, Siri, Google Now, Cortana

Косач Дмитрий Иванович

Студент

Российский экономический университет им. Г.В. Плеханова

г. Москва, Стремянный пер., 36

Kosach Dmitry Ivanovich

Student

Russian university of economics named G.V. Plekhanov

Moscow, Stremyannyy lane, 36

Жидкова Лалита Олеговна

Студент

Российский экономический университет им. Г.В. Плеханова

г. Москва, Стремянный пер., 36

Zhidkova Lalita Olegovna

Student

Russian university of economics named G.V. Plekhanov

Moscow, Stremyannyy lane, 36

Белехов Александр Николаевич

Кандидат технических наук, доцент

Российский экономический университет им. Г.В. Плеханова

г. Москва, Стремянный пер., 36

Belekhov Aleksandr Nikolaevich

Candidate of Technical Sciences, Associate Professor

Russian university of economics named G.V. Plekhanov

Moscow, Stremyannyy lane, 36

По мере развития «виртуальных ассистентов» и «умных помощников», – следующим логическим этапом их развития будет интеграция с аппаратной частью. Их базы знаний, будут дополнены элементами естественной речи, позволяя нам искать во все возрастающих объёмах информации с меньшей затратой по времени и в более человеческой манере – разговаривая с системой [1].

Уже сейчас «голосовые ассистенты» и умные помощники, наглядно показывают нам, на что именно будут способны персональные роботы в ближайшем будущем.

Во-первых, голосовой интерфейс будет основным для взаимодействия со множеством функций персонального робота. Во-вторых, ваш персональный робот будет подключен к различным онлайн-сервисам, таким как календарь, новостной портал, интернет-магазин и так далее. Можно составить длинный список возможностей персонального робота, открывающихся при интеграции голосового интерфейса и онлайн-сервисов, таких как:

- понимание голосового ввода – персональный робот сможет реагировать на команды, отвечать на вопросы, предоставлять консультации – от медицинских персональных рекомендаций, до проведения обучающих занятий с вашими детьми;

- предложение личной помощи – планировать встречи, устанавливать напоминания, и контролировать их выполнение, проверять запасы, уведомлять о получении штрафа и напоминать о необходимости его оплатить и так далее;

- предоставление помощи в передвижениях – подсказывать направление движения, скоростной режим, прогнозировать пробки, найти стоянку автомобилей;

- взаимодействие с подключенными устройствами (интернет вещей) – управлять телевизором, гаражными воротами, фонарями, камерами наблюдения и так далее;

- возможность покупать вещи и отслеживать доставку приобретенных товаров;

- взаимодействовать с социальными медиа – отправлять фотографии, делать видеотрансляции, сообщать о происшествии в службу спасения, читать твиты и прочее;

- определять настроение – распознавать эмоции человека, язык тела [2].

Появление интеллектуальных персональных программных агентов, таких как: Siri, Cortana, Google Now во многом связано с тем, что большинство людей вообще не привыкли скачивать и использовать отдельные приложения для решения своих задач, всё больше программ станут поддерживать поиск по своему контенту через внешние сервисы и для этого их даже не придётся устанавливать.

Ключевой фактор роста популярности интеллектуальных помощников – возможность искать по контенту приложений. В свою очередь те будут использовать свои поисковые алгоритмы, чтобы забирать контент из приложений так, что пользователю даже не придётся устанавливать их на свой телефон [3].

Личные роботы будут широко использовать голосовой интерфейс, сродни тому, что мы видим сегодня в персональных ассистентах, потому что это естественная форма общения человека и она обеспечивает выполнение команд самым простым, быстрым и эффективным способом. Меню из графического интерфейса уйдут в прошлое, а личности новых цифровых помощников – подарят значительно большее удовольствие, чем симпатичные иконки [3].

Интеллектуальные персональные программные агенты также призваны предугадывать потребности своих пользователей. При желании, им можно дать доступ к своим личным данным, таким как электронная почта, адресная книга, история поисков в сети и т.п. – все эти данные она будет использовать для удовлетворения ваших нужд. Также некоторые виртуальные ассистенты не лишены чувства юмора: они могут поддерживать с вами беседу, петь песни и рассказывать анекдоты. Они заранее напомнят вам о запланированной встрече, дне рождения друга и других важных событиях. Помощники сообщают, если ваш авиарейс отменили, или на дорогах много пробок. Их интерфейс имеет очень гибкие настройки конфиденциальности, позволяющие пользователю самому определять, какого рода информацию предоставлять виртуальному ассистенту.

Учитывая опыт многих мировых компаний и используя доступность и открытость некоторого программного кода и алгоритмов, для российских программистов и разработчиков программного обеспечения должна быть поставлена задача о разработке отечественного полноценного виртуального помощника с голосовым управлением, понимающего устные запросы на естественном, что самое главное, русском языке. При задавании вопроса к которому, пользователь смог бы получить быстрый и релевантный ответ. Данная разработка позволила бы таким людям, как людям пенсионного возраста, которым в силу своих лет всё тяжелее освоить новые «гаджеты», людям со слабым зрением, людям с ограниченными возможностями, используя так называемых интеллектуальных персональных программных агентов, установленных в их смартфоны или планшеты диктовать смс сообщения, предоставлять информацию о последних событиях, происходящих в мире, давать маршруты по картам, находить, а также загружать понравившиеся музыкальные композиции, предоставлять прогнозы по погоде, рекомендовать рестораны, фильмы, а также бронировать билеты и столики, напомнить о дне рождения близкого человека и других важных событиях.

Ввести интеллектуального персонального программного агента в своих собственных проектах можно уже сейчас. В данной статье представлено краткое описание процедуры внедрения библиотеки с интеллектуальным персональным программным агентом в своё приложение. Изучение никаких новых языков программирования или технологий не требуется. Достаточно ограничиться для типовой разработки Андроид SDK (Android SDK) (см. табл.). Таким образом, программный агент представляет собой надстройку над обычными Android-сервисами. Он реализует интерфейс между ассистентом и логикой приложения.

После сборки приложения и установки на девайс ассистент загрузит информацию из вашего манифеста и модуль с перенаправлением всех подходящих запросов от пользователя к агенту, если «движок» ассистента посчитает, что фраза подходит лучше всего под паттерны команд модуля.

Таблица 1. Пошаговая программная процедура внедрения библиотеки

1. Запуск начального программного образующего модуля типа Андроид Студิโอ (Android Studio).

2. Для приложения пользователя представляется модуль с двумя командами (простой *xml*-файл, который нужно хранить в ресурсной директории *xml*), не содержащий какого-либо управляющего кода, который описывается в классе агента. Это отражает основной принцип подхода к голосовому агенту – декларативная часть, описывающая грамматику диалога, отделена от управляющего кода, который реализует логику обработки и полностью независим от языка.

3. Формирования быстрого ответа в одну строку с использованием оператора «Запрос» («*Request*»), содержащего всю необходимую информацию о запросе пользователя – идентификатор команды, контент запроса, сессию и т.д. На любой запрос агент должен сформировать ответ – «*Request*», содержащий контент ответа и, при необходимости, инструкции ассистенту о переключении контекста диалога.

4. Программный «ответ», содержащий кроме контента в виде строки информацию об изменении содержания диалога, обеспечивающую пользователю возможность доступа к командам только из модуля «*R.xml.name*», а после «озвучки» ассистентом ответа от агента автоматически включится микрофон – в так называемом «модальном режиме».

5. Описание в манифесте приложения «*AndroidManifest.xml*» сервиса каждого агента.

```
repositories {
    maven {
        url 'http://voiceassistant.mobi/m2/repository' } }
dependencies {
    compile 'mobi.voiceassistant:base:0.1.0-SNAPSHOT'
    compile 'mobi.voiceassistant:client:0.1.0-SNAPSHOT' }
public class HelloAgent extends AssistantAgent {
    @Override
    protected void onCommand(Request request) {
        switch (request.getDispatchId()) {
            case R.id.cmd_hello:
                onHello(request);
                break;
            case R.id.cmd_name:
                onName(request);
                break; } }
    request.addQuickResponse ("Привет!");
```

```
Response response = request.createResponse();
response.setContent(getString(R.string.hello_say_name));
response.enterModalQuestionScope(R.xml.name);
request.addResponse(response);
<service android:name=".HelloAgent">
    <intent-filter>
        <action android:name="mobi.voiceassistant.intent.action.COMMAND"/>
        <data android:scheme="assist" android:host="mobi.voiceassistant.ru"/>
    </intent-filter>
    <meta-data android:name="mobi.voiceassistant.MODULE" android:resource="@xml/hello"/>
</service>
```

Представленная разработкой пошаговой программной процедуры внедрения библиотеки с интеллектуальным персональным агентом в своё приложение позволяет повсеместно обеспечить внедрение и освоение новых «гаджетов» для пользователя с последующим совершенствованием доступа к функциональным возможностям «голосовых ассистентов» для решения широкого круга задач от управления промышленными объектами до применения в быту. Виртуальные голосовые помощники с элементами искусственного интеллекта должны оказать влияние на генерацию инновационных способов управления во всё ещё остающейся затратной российской экономике [4], предотвратить различного рода техническую утечку вербальной информации [5].

Список используемых источников:

1. Wikipedia: Интеллектуальный персональный программный агент.
URL: https://ru.wikipedia.org/wiki/Интеллектуальный_персональный_программный_агент
2. Пономарёв Ф. Голосовые ассистенты-предвестники персональных роботов.
URL: <https://robot-review.ru/2015/09/24/golosovyye-assistentyi-predvestniki-personalnyh-robotov/>
3. Лухачёв Н. Помощники вроде Siri и Google Now заменят собой приложения.
URL: <https://tjournal.ru/p/no-more-apps>
4. Мерзликин В.Г., Красс М.С., Кулаго А.Е. Худяков С.В. Затратность экономики России. Динамические модели развития // Научные исследования и разработки. Экономика фирмы. 2013. Т. 2. № 3-4. С. 30-34. DOI 10.12737/2493
5. Гришачев В.В., Халяпин Д.Б., Шевченко Н.А., Мерзликин В.Г. Новые каналы утечки конфиденциальной речевой информации через волоконно-оптические подсистемы СКС // Специальная техника. 2009. № 2. С. 2-9.

© 2016, Косач Д.И., Жидкова Л.О., Белехов А.Н.
Виртуальные голосовые помощники с элементами
искусственного интеллекта

© 2016, Kosach D.I., Zhidkova L.O., Belekhov A.N.
Virtual voice assistant with elements of artificial
intelligence