

УДК 004. 934.1

ВЫБОР ТЕХНОЛОГИИ РАСПОЗНАВАНИЯ РЕЧИ ДЛЯ РАЗРАБОТКИ КРОСС-ПЛАТФОРМЕННОГО МОБИЛЬНОГО ПРИЛОЖЕНИЯ

Г. О. КОРЕННОЙ, Д. И. ЕРОХОВА
Научный руководитель О. В. СЕРГИЕНКО
Белорусско-Российский университет
Могилев, Беларусь

В повседневной жизни современный человек использует множество различных мобильных приложений, помогающих ему решать текущие бытовые или жизненные вопросы и задачи. Среди таких приложений можно выделить планировщики, напоминания, программы для формирования списков. Использование таких приложений помогает распланировать время, не забыть принять лекарство, заранее спланировать поход в магазин. Большинство таких приложений обладают схожими функциональными возможностями, к которым можно отнести управление данными и их редактирование.

Но подобные приложения должны отвечать еще ряду требований, к которым можно отнести кросс-платформенность и наличие голосового ввода. Очень часто человек не имеет возможности ввести в приложение новые данные из-за спешки, при этом возникает риск просто о них забыть. Голосовой ввод позволил бы обеспечить быстрый ввод в данной ситуации. Также отдельную категорию, для которой необходима данная возможность, составляют пользователи с ограниченными возможностями, например с нарушением зрения.

Одной из проблем решения задачи добавления возможности голосового ввода является выбор и применение технологии преобразования или имплементации речи в текст.

Если мобильное приложение создается только для определенной мобильной операционной системы, т. е. не должно быть кросс-платформенным, то оно строится на нативных фреймворках, имеющих собственные технологии распознавания речи. Они достаточно просты в применении. Но такой вариант построения может существенно сузить круг его пользователей. Для кросс-платформенных фреймворков можно использовать три основных подхода:

- 1) применение сторонних библиотек с функциями распознавания речи;
- 2) имплементация при помощи алгоритмов обработки естественного языка, тренировка нейронных сетей под распознавание голоса;
- 3) использование комплектов для разработки ПО (Software development kit, SDK), содержащих необходимые библиотеки.

После анализа всех вариантов использование SDK признано наиболее оптимальным по функциональным возможностям. Крупные компании уже предоставляют свои библиотеки для Android и iOS. Они отлично распознают десятки языков, поддерживают все основные форматы звука и пользоваться ими можно абсолютно бесплатно.