

**БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**  
**Факультет радиофизики и компьютерных технологий**  
**Кафедра интеллектуальных систем**

Аннотация к дипломной работе

**«Голосовые ассистенты.  
Построение прототипа смарт-спикера»**

Баркун Инна Владимировна

Научный руководитель: кандидат физико-математических наук, доцент  
Е.И. Козлова

## РЕФЕРАТ

Дипломная работа: 52 страницы, 19 рисунков, 6 таблиц, 22 использованных источника.

ГОЛОСОВОЕ УПРАВЛЕНИЕ, ГОЛОСОВЫЕ АССИСТЕНТЫ, РАЗРАБОТКА СМАРТ-СПИКЕРА, ДЕТЕКТИРОВАНИЕ РЕЧЕВОГО СИГНАЛА, РАСПОЗНАВАНИЕ РЕЧИ, DEEPSPEECH, СИНТЕЗ РЕЧИ.

*Объект исследования* – голосовые помощники.

*Цель работы* – разработка требований к аппаратному и программному обеспечению для проектирования и последующего создания голосового помощника.

*Методы исследования* – сравнительный анализ, методы моделирования систем обработки информации.

Рассмотрены основные сферы и аспекты использования голосовых ассистентов, описан лидирующий на основе представленных статистических данных в своей области смарт-спикер Amazon Echo от Amazon, уступающий по базе знаний устройству Google Home от Google.

При разработке программного обеспечения используются алгоритмы для детектирования, распознавания и синтеза речи – WebRTC-VAD, DeepSpeech, Amazon Polly соответственно, позволяющие выполнить взаимодействие «вопрос-ответ» пользователя и устройства в режиме реального времени при помощи голосового ввода без предварительного взаимодействия с интерфейсом.

В результате выполнения работы реализован собственный вариант смарт-спикера, способного консультировать пользователя в вопросах оказания первой доврачебной помощи пострадавшему.

## РЭФЕРАТ

Дыпломная праца: 52 старонкі, 19 малюнкаў, 6 табліц, 22 выкарыстанныя крыніцы.

ГАЛАСАВОЕ КІРАВАННЕ, ГАЛАСАВЫЯ АСІСТЭНТЫ, РАСПРАЦОЎКА СМАРТ-СПІКЕРА, ДЭТЭКТАВАННЕ МАЎЛЕНЧЫХ СІГНАЛАЎ, РАСПАЗНАВАННЕ МАЎЛЕННЯ, DEEPSPEECH, СІНТЭЗ МАЎЛЕННЯ.

*Аб'ект даследвання* – галасавыя памочнікі.

*Мэта* – распрацоўка патрабаванняў да апаратнага і праграмнага забеспячэння для праектавання і наступнага стварэння галасавога памочніка.

*Метады даследавання* – параўнальны аналіз, метады мадэлявання сістэм апрацоўкі інфармацыі.

Разгледжаны асноўныя сферы і аспекты выкарыстання галасавых асістэнтаў, апісаны смарт-спікер Amazon Echo ад Amazon, які лідыруе на аснове прадстаўленых статыстычных даных у сваёй вобласці, але саступае па базе ведаў прыладзе Google Home ад Google.

Пры распрацоўцы праграмнага забеспячэння выкарыстоўваюцца алгарытмы для дэтэктавання, распазнавання і сінтэзу маўлення – WebRTC-VAD, DeepSpeech, Amazon Polly адпаведна, якія дазваляюць выканаць узаемадзеянне «пытанне-адказ» карыстальніка і прылады ў рэжыме рэальнага часу пры дапамозе галасавога ўводу без папярэдняга ўзаемадзеяння з інтэрфейсам.

У выніку выканання работы рэалізаваны ўласны варыянт смарт-спікера, здольнага кансультаваць карыстальніка ў пытаннях аказання першай даўрачэбнай дапамогі пацярпеламу.

## ABSTRACT

Thesis: 52 pages, 19 figures, 6 tables, 22 sources.

VOICE CONTROL, VOICE ASSISTANTS, SMART SPEAKER DEVELOPMENT, VOICE ACTIVITY DETECTION, SPEECH RECOGNITION, DEEPSPEECH, SPEECH SYNTHESIS.

*The object of research* – voice assistants.

*Objective* – determination of the requirements for voice assistant hardware and software for prototyping and future voice assistant development.

*The methods* – comparative analysis, information processing system modelling techniques.

We examine primary areas and aspects of application of voice assistants. We also describe the leading (according to provided statistical data) smart speaker Amazon Echo by Amazon, which is surpassed by Google Home by Google.

Development of such an assistant involves use of voice activity detection, speech recognition and speech synthesis algorithms – WebRTC-VAD, DeepSpeech, Amazon Polly. Applying these algorithms, we implement a real-time speech-controlled «question and answer» system, which doesn't require any preliminary interaction with physical interface.

Finally, using mentioned algorithms we develop own smart speaker prototype, which is able to advise a user in terms of rendering first aid to a victim.