

(19)



**Евразийское  
патентное  
ведомство**

(11) **047172**

(13) **B1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ**

(45) Дата публикации и выдачи патента  
**2024.06.17**

(21) Номер заявки  
**202391961**

(22) Дата подачи заявки  
**2023.08.03**

(51) Int. Cl. **H04M 3/493** (2006.01)  
**G10L 15/06** (2013.01)  
**G10L 15/22** (2006.01)  
**G10L 15/30** (2013.01)

---

(54) **СПОСОБ ОБУЧЕНИЯ ГОЛОСОВОГО РОБОТА**

---

(43) **2024.06.13**

(96) **2023000130 (RU) 2023.08.03**

(71)(73) Заявитель и патентовладелец:  
**ОБЩЕСТВО С ОГРАНИЧЕННОЙ  
ОТВЕТСТВЕННОСТЬЮ  
"ТОМОРУ" (RU)**

(72) Изобретатель:  
**Балюра Денис Олегович,  
Маглеванный Александр Юрьевич  
(RU)**

(74) Представитель:  
**Луцковский М.Ю., Корниец Р.А. (RU)**

(56) US-A1-20220293089  
US-B2-10498898  
CN-A-111696576  
CN-A-111858884

---

(57) Изобретение относится к способу построения цифрового ассистента в голосовых диалоговых системах. Способ включает создание нового робота, построение общей логики диалога, формирование отдельных фрагментов диалога и объединение этих данных с набором шаблонизированных логик диалогов и с набором диалогов, полученных с помощью генеративного искусственного интеллекта. Технический результат заключается в возможности создания и конфигурирования голосового робота под конкретные бизнес-задачи с минимальным временем задержки для подготовки ответа.

---

**B1**

**047172**

**047172**

**B1**

Изобретение относится к области вычислительной техники, в частности к способу построения голосового интерфейса в диалоговых системах [H04M3/493, G06F40/30, G06Q30/00, G10L25/00].

Из уровня техники известна ИНТЕЛЛЕКТУАЛЬНАЯ ГОЛОСОВАЯ РОБОТИЗИРОВАННАЯ СИСТЕМА [CN 112333341 A, опубл. 2021-02-05], содержащая модуль управления услугами, интеллектуальный бизнес-модуль и базовый модуль выполнения задач, модуль управления услугами содержит модуль службы уровня управления развитием бизнеса и модуль службы мониторинга службы. Интеллектуальный бизнес-модуль содержит блок управления диалогами бизнес-процессов и блок механизма семантического синтаксического анализа. Базовый встроенный исполнительный модуль содержит блок ACD центра обработки вызовов, блок мультимедийных услуг MS и блок обработки голоса. При этом отображается путь связи с клиентом, и точно оценивается намерение клиента, поддерживается прерывание вызова в любое время, улучшается качество обслуживания клиентов и повышается качество взаимодействия. Все содержимое звонка записывается и точно преобразуется в символы, чтобы содержание звонка было удобно проверять. Недостатком аналога является невозможность создания или конфигурирования голосового робота под конкретные бизнес-задачи.

Также из уровня техники ИНТЕЛЛЕКТУАЛЬНАЯ ГОЛОСОВАЯ СИСТЕМА НА ОСНОВЕ WECHAT [CN 113691686 A, опубл. 2021-11-23], содержащая ПК-терминал Enterprise WeChat, который устанавливается на компьютер или сервер и используется для установления голосовой связи с терминалом WeChat клиента. Контроллер RPA, который используется для формулирования стратегий исходящих вызовов в соответствии с потребностями бизнес-сценариев, а также для управления вызовом. Диалоговый робот с ИИ, который используется для формирования плана голосового вещания в соответствии с потребностями бизнес-сценария, а также получает и анализирует обратную связь аудиопотока с терминала WeChat клиента, выполняет план голосового вещания на основе полученного аудиопотока и подсчитывает результаты звонка. Первая виртуальная звуковая карта, которая используется для передачи аудиопотока, полученного корпоративным ПК WeChat, диалоговому роботу с ИИ. Вторая виртуальная звуковая карта используется для передачи аудиопотока, сгенерированного диалоговым роботом с искусственным интеллектом, на корпоративный ПК WeChat.

Недостатком аналога является невозможность создания или конфигурирования голосового робота под конкретные бизнес-задачи.

Наиболее близким по технической сущности заявленному изобретению является СПОСОБ ФОРМИРОВАНИЯ ГОЛОСОВОЙ ИНФОРМАЦИИ РОБОТА [CN 114179083 A, опубл. 2022-03-15], включающий следующие этапы: ведущий робот обнаруживает, что посетителю требуется помощь ведущего, получение голосового сигнала посетителя; определяется тип языка голосового сигнала; на основе типа языка определяется режим генерации ответного речевого сигнала ведущего робота, причем режим генерации включает в себя локальный режим и серверный режим; и формирование ответного речевого сигнала, соответствующего типу языка посетителя, на основе режима генерации, и воспроизведение ответного речевого сигнала на ведущем роботе. В соответствии с различными типами языка ответные голосовые сигналы для ответа посетителям генерируются на ведущем роботе или сервере так, что ведущий робот может отвечать большему количеству посетителей на разных языках.

Недостатком прототипа является невозможность создания или конфигурирования голосового робота под конкретные бизнес-задачи. Кроме того, к недостаткам прототипа можно отнести наличие ощутимой задержки при формировании ответов робота.

Задачей настоящего изобретения является устранение недостатков прототипа.

Технический результат изобретения заключается в возможности создания и обучения голосового робота под конкретные бизнес-задачи, при этом ввиду того, что голосовой робот создается для выполнения конкретных задач, то одновременно каждый такой робот может поддерживать диалог с практически неограниченным количеством людей и с минимальным временем задержки для подготовки ответа клиенту.

Технический результат изобретения достигается за счет того, что способ обучения голосового робота характеризуется тем, что с помощью модуля веб-приложения:

- a) создается новый голосовой робот,
  - b) для нового голосового робота задается общая логика диалога, при этом общая логика диалога задается посредством построения древовидной структуры диалога,
  - c) формируются отдельные фрагменты диалога, не входящие в общую логику диалога,
  - d) данные, полученные на этапах a)-c), сохраняются в модуле данных веб-приложения;
- после чего в модуль веб-приложения передается:

набор шаблонизированных логик диалогов из модуля единых шаблонизированных логик диалогов, набор диалогов, полученных с помощью генеративного искусственного интеллекта;

в модуле веб-приложения осуществляется объединение набора шаблонизированных логик диалогов с данными, сохраненными на этапе d), и с набором диалогов, полученных с помощью генеративного искусственного интеллекта.

В частности, модуль веб-приложения содержит модуль пользовательского интерфейса, посредством которого осуществляется ввод данных на этапах a)-c).

В частности, на этапах b) и c) общая логика диалога и отдельные фрагменты диалога формируются с помощью блоков.

В частности, построение диалога происходит через последовательное построение блоков с фразами робота и блоков с ответами человека.

В частности, блок с ответами человека может включать в себя ожидаемый ответ, включающий общее намерение человека, точную фразу или молчание, а также может быть настроен на прием от человека данных.

В частности, набор шаблонизированных логик диалогов содержит логики диалогов под разные сочетания отраслей и бизнес-задач для робота, а также содержит облака синонимов.

В частности, модуль веб-приложения содержит модуль серверной части веб-приложения.

В частности, набор шаблонизированных логик диалогов и набор диалогов, полученных с помощью генеративного искусственного интеллекта, передаются в модуль серверной части веб-приложения.

В частности, набор диалогов, полученных с помощью генеративного искусственного интеллекта, создается с помощью внешнего чат-бота с искусственным интеллектом.

В частности, объединение набора шаблонизированных логик диалогов с данными, сохраненными на этапе d), и с набором диалогов, полученных с помощью генеративного искусственного интеллекта, осуществляется в модуле серверной части веб-приложения.

#### **Краткое описание чертежей**

На фиг. 1 показана блок-схема платформы для создания голосовых роботов.

На фиг. 2 показана блок-схема способа обучения голосового робота.

На фиг. 3 приведен фрагмент пользовательского интерфейса.

На фиг. 1 обозначено: 1 - пользователь; 2 - модуль внешних API; 3 - модуль веб-приложения; 4 - модуль пользовательского интерфейса; 5 - модуль серверной части веб-приложения; 6 - модуль данных веб-приложения; 7 - модуль микросервисов; 8 - модуль единых шаблонизированных логик диалогов; 9 - модуль планировщика заданий; 10 - модуль поиска; 11 - модуль планировщика звонков; 12 - модуль SIP прокси; 13 - модуль VOIP; 14 - файловое хранилище; 15 - модуль внутреннего генеративного искусственного интеллекта; 16 - модуль внешнего генеративного искусственного интеллекта.

На фиг. 2 обозначено: 100 - этап создания нового голосового робота; 101 - этап построения общей логики диалога посредством древовидной структуры диалога для голосового робота; 102 - этап формирования отдельных фрагментов диалога, не входящих в общую логику; 103 - этап передачи набора шаблонизированных логик диалогов; 104 - этап передачи набора диалогов, полученных с помощью генеративного искусственного интеллекта; 105 - этап объединения набора шаблонизированных логик диалогов с данными, полученными на этапах 100-102, и с набором диалогов, полученных с помощью генеративного искусственного интеллекта; 106 - этап получения обученного голосового робота.

На фиг. 3 обозначено: 200 - пример построения общей логики диалога; 201 - пример построения отдельных фрагментов диалога.

#### **Раскрытие сущности изобретения**

Для начала необходимо отметить, что под пользователем 1 понимается пользователь веб-приложения (он же аналитик или "учитель" робота), который осуществляет ввод данных для настройки и конфигурирования голосового робота. А под человеком (клиентом) понимается получатель услуги голосового робота, то есть человек, который общается с помощью голоса с голосовым роботом.

Настоящий способ обучения голосового робота выполняется с помощью платформы для создания голосовых роботов, которая раскрыта на фиг. 1. Основными частями данной платформы являются следующие модули.

Модуль веб-приложения 3, который в свою очередь состоит из модуля пользовательского интерфейса 4, модуля серверной части веб-приложения 5 и модуля данных веб-приложения 6. При этом модуль пользовательского интерфейса 4 может быть создан с использованием фреймворка Angular, модуль серверной части веб-приложения 5 может быть создан на платформе Node.js, модуль данных веб-приложения 6 может быть создан с использованием документоориентированной системы управления базами данных MongoDB. Модуль серверной части веб-приложения 5 может быть развернут на сервере с помощью Docker контейнеризации.

Модуль единых шаблонизированных логик диалогов 8, который может быть создан на языке Rust и развернут на сервере с помощью Docker контейнеризации, а набор единых шаблонизированных логик диалогов может храниться с помощью системы управления базами данных PostgreSQL. При этом модуль единых шаблонизированных логик диалогов 8 соединен с модулем серверной части веб-приложения 5 по протоколу HTTP (HTTPS). Модуль внутреннего генеративного искусственного интеллекта 15 - это модуль искусственного интеллекта, способный генерировать текст. Генеративный искусственный интеллект использует генеративные модели, такие как большие языковые модели, для статистической выборки новых данных на основе набора обучающих данных, который использовался для их создания. Данный модуль 15 может быть создан на языке Python и развернут на сервере с помощью Docker контейнеризации. При этом модуль внутреннего генеративного искусственного интеллекта 15 соединен с модулем серверной части веб-приложения 5 по протоколу HTTP (HTTPS).

Дополнительно модуль внутреннего генеративного искусственного интеллекта 15 подключен к модулю внешнего генеративного искусственного интеллекта 16, в качестве которого может выступать, например, чат-боты ChatGPT-3, ChatGPT-4 от компании OpenAI или чат-бот Bard от компании Google. Подключение модуля внутреннего генеративного искусственного интеллекта 15 к модулю внешнего генеративного искусственного интеллекта 16 осуществляется посредством соответствующего API.

Кроме того, платформа может содержать модуль внешних API 2, модуль микросервисов 7, модуль планировщика заданий 9 и модуль поиска 10.

Также платформа содержит модули, позволяющие роботу осуществлять голосовые вызовы. К таким модулям относятся модуль планировщика звонков 11, модуль SIP прокси 12 и модуль VOIP 13, который также содержит модули Text-to-speech и Speech-to-text (на фиг. не показаны).

Настоящий способ обучения голосового робота выполняется преимущественно с помощью модуля веб-приложения 3, модуля пользовательского интерфейса 4, модуля серверной части веб-приложения 5, модуля данных веб-приложения 6, модуля единых шаблонизированных логик диалогов 8, а также с помощью модуля внутреннего генеративного искусственного интеллекта 15 и модуля внешнего генеративного искусственного интеллекта 16.

Сначала в модуле пользовательского интерфейса 4 пользователем 1 инициируется создание нового голосового робота (этап 100).

Затем для нового голосового робота в модуле пользовательского интерфейса 4 пользователем 1 задается общая логика диалога (этап 101). Общая логика диалога задается посредством построения древовидной структуры диалога (200).

Далее для фрагментов диалога (201), не входящих в общую логику диалога (древовидную структуру), в модуле пользовательского интерфейса 4 пользователем 1 формируются отдельные фрагменты диалога (этап 102).

При выполнении этапов 100, 101 и 102 пользователь 1 выступает в роли аналитика или "учителя" робота. Он строит целенаправленный диалог, обучает робота понимать смысл сказанного и строит логику ответов на каждое сообщение, а также формирует дальнейшую логику диалога.

Общая логика диалога (этап 101) и отдельные фрагменты диалога (этап 102) формируются с помощью блоков. Построение диалога происходит через последовательное построение блоков с фразами робота и блоков с ответами человека. При этом блок с ответами человека может включать в себя ожидаемый ответ, включающий общее намерение человека, точную фразу или молчание, а также может быть настроен на прием от человека данных.

То есть на каждый блок с фразой робота ("робот говорит") может приходиться один или несколько блоков с ответами человека ("клиент говорит").

Кроме того, блок с ответами человека помимо группы ожидаемых ответов имеет возможность получения от клиента переменных/данных. Каждый такой блок может принять фразу человека целиком и передать ее дальше по диалогу. Также блок с ответами человека может содержать какую-либо переменную, полученную от клиента, например, озвученный клиентом адрес.

После построения всей логики работы голосового робота, конфигурационные данные, полученные на этапах 100, 101 и 102, сохраняются в модуле данных веб-приложения 6.

Затем из модуля единых шаблонизированных логик диалогов 8 в модуль серверной части веб-приложения 5 передается набор шаблонизированных логик диалогов (этап 103). Этот набор содержит логики диалогов под разные сочетания отраслей и бизнес-задач для робота. Кроме того, набор шаблонизированных логик диалогов содержит облака синонимов, с помощью которых система тренируется понимать реакции человека на сказанное роботом. Например, пользователь 1 в одном месте ставит ожидаемый ответ "Да" и платформа автоматически добавляет более двухсот не повторяющихся формулировок утвердительного ответа/согласия.

Одновременно с этим из модуля внутреннего генеративного искусственного интеллекта 15 в модуль серверной части веб-приложения 5 передается набор диалогов, полученных с помощью генеративного искусственного интеллекта (этап 104). Этот набор диалогов формируется в модуле внутреннего генеративного искусственного интеллекта 15 посредством промптов, которые позволяют сгенерировать необходимые ответы от модуля внешнего генеративного искусственного интеллекта 16, в роли которого может выступать чат-бот, такой как ChatGPT-3, ChatGPT-4 или Bard.

Генерирование набора диалогов с помощью генеративного искусственного интеллекта позволяет на порядок увеличить разнообразие используемых ответов голосового робота. При этом этот набор диалогов может включать в себя как узкоспециализированные диалоги, так и диалоги общего характера.

Далее на этапе 105 в модуле серверной части веб-приложения 5 происходит объединение данных, полученных на этапах 100, 101 и 102, набора шаблонизированных логик диалогов и набор диалогов, полученных с помощью генеративного искусственного интеллекта. В итоге получается предобученный голосовой робот, сформированный под конкретные бизнес-задачи. При этом объединение узкоспециализированных сценариев диалогов со сценариями диалогов под разные сочетания отраслей и бизнес-задач позволяет обучить голосового робота таким образом, что с одной стороны он будет неотличим от живого человека по вокабуляру (объему используемой лексики) и по пониманию клиента, а с другой стороны

среднее время задержки ответа будет заметно ниже имеющихся решений.

### ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ обучения голосового робота, характеризующийся тем, что с помощью модуля веб-приложения:

- a) создается новый голосовой робот,
- b) для нового голосового робота задается общая логика диалога, при этом общая логика диалога задается посредством построения древовидной структуры диалога,
- c) формируются отдельные фрагменты диалога, не входящие в общую логику диалога,
- d) данные, полученные на этапах а)-с), сохраняются в модуле данных веб-приложения;

после чего в модуль веб-приложения передается:

набор шаблонизированных логик диалогов из модуля единых шаблонизированных логик диалогов, набор диалогов, полученных с помощью генеративного искусственного интеллекта;

в модуле веб-приложения осуществляется объединение набора шаблонизированных логик диалогов с данными, сохраненными на этапе d), и с набором диалогов, полученных с помощью генеративного искусственного интеллекта;

модуль веб-приложения содержит модуль пользовательского интерфейса, посредством которого осуществляется ввод данных на этапах а)-с);

набор шаблонизированных логик диалогов содержит логики диалогов под разные сочетания отраслей и бизнес-задач для робота, а также содержит облака синонимов;

набор диалогов, полученных с помощью генеративного искусственного интеллекта, создается с помощью внешнего чат-бота с искусственным интеллектом.

2. Способ по п.1, отличающийся тем, что на этапах b) и c) общая логика диалога и отдельные фрагменты диалога формируются с помощью блоков.

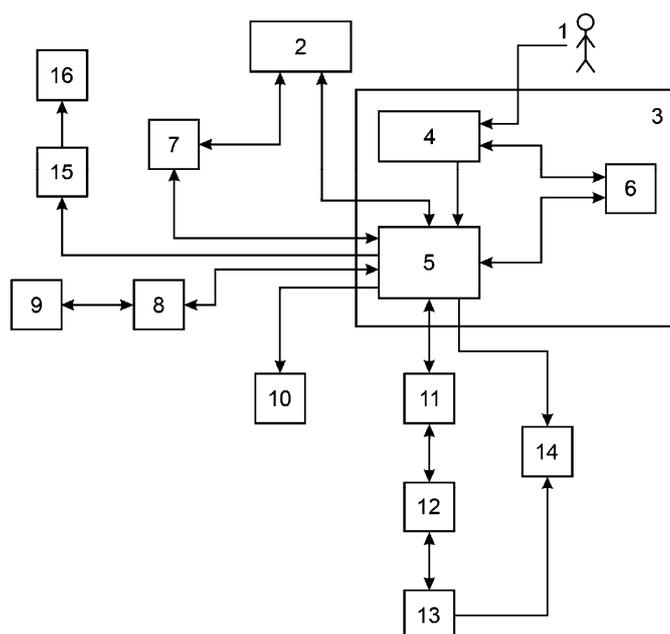
3. Способ по п.2, отличающийся тем, что построение диалога происходит через последовательное построение блоков с фразами робота и блоков с ответами человека.

4. Способ по п.3, отличающийся тем, что блок с ответами человека может включать в себя ожидаемый ответ, включающий общее намерение человека, точную фразу или молчание, а также может быть настроен на прием от человека данных.

5. Способ по п.1, отличающийся тем, что модуль веб-приложения содержит модуль серверной части веб-приложения.

6. Способ по п.5, отличающийся тем, что набор шаблонизированных логик диалогов и набор диалогов, полученных с помощью генеративного искусственного интеллекта, передаются в модуль серверной части веб-приложения.

7. Способ по п.5, отличающийся тем, что объединение набора шаблонизированных логик диалогов с данными, сохраненными на этапе d), и с набором диалогов, полученных с помощью генеративного искусственного интеллекта, осуществляется в модуле серверной части веб-приложения.



Фиг. 1

