

УДК 007.5

РАЗРАБОТКА ГОЛОСОВОГО АССИСТЕНТА ДИСПЕТЧЕРСКОЙ СЛУЖБЫ ДЛЯ АВТОМАТИЗАЦИИ ОБРАБОТКИ ОБРАЩЕНИЙ ГРАЖДАН

Коряковцев С.Д., Федоров Д.А. (ИТМО)

Научный руководитель – кандидат технических наук, доцент Федоров Д.А.
(ИТМО)

Введение. Технологии искусственного интеллекта из-за активного развития, продолжающегося и по сей день, продолжают проникать во все большее количество сфер деятельности и способствовать решению множества рутинных или повторяющихся повседневных задач, требующих определенного количества человеческих ресурсов, одной из которых как раз является задача автоматизации сбора данных, получаемых в результате голосовых коммуникаций. На рынке голосовых ассистентов существует множество решений данной задачи, как платных и ограниченно-бесплатных, так и бесплатных и находящихся в открытом доступе. Однако большинство существующих решений предлагают либо излишне широкий функционал для решения поставленной задачи, либо, как в случае платных аналогов, предлагают решения, которые можно сделать более универсальными и уменьшить трату ресурсов затрачиваемых как на разработку решения, так и на дальнейшее использование. Чтобы сделать решение более универсальным и гибким в рамках поставленной задачи, требуется актуализировать существующую информацию по теме исследования и непосредственно применить полученные знания на практике, путем разработки предлагаемого решения, учитывающего поставленные требования.

Основная часть. В рамках исследования значимости и существующих разработок на смежные с настоящей темой путем обзора литературы и анализа архитектурных особенностей альтернативных решений с целевым реализованным функционалом были изучены:

- 1) аспекты интеграции со сторонними системами [1], применения ИИ для анализа данных [2], использования различных архитектур систем [3] и внедрения многоканальной интеграции [4] — во всех перечисленных темах была отмечена значимость текущего исследования,
- 2) для учета в процессе формирования требований, проектирования и реализации основного функционала характеристики некоторых существующих аналогов [5, 6], которые включают доступность, использованные технологии реализации, встроенные возможности и сложности в использовании.

По результатам исследования путем сбора информации из открытых источников [7–10] об оценке работы планирующихся для использования моделей ИИ (синтез речи, анализ текста LLM, распознавание речи) в качестве наиболее оптимальных аналогов для реализации и использования разрабатываемого решения в рамках поставленной задачи были выбраны:

- 1) Whisper (для быстрой интеграции) или Vosk 0.54 (для долгосрочной оптимизации) — для транскрибации речи,
- 2) SmolLM3–3B, phi4–mini–3.8B или Qwen3–4B — для анализа текста и извлечения намерений, где выбор зависит от имеющихся ресурсов для использования и субъективной оценки качества,
- 3) Piper или XTTS–v2 — для синтеза речи, в зависимости от требуемой языковой поддержки и баланса между качеством и скоростью работы.

Была проведена разработка решения, включающая:

- 1) этап проектирования, в который входит построение монолитной архитектуры как с точки зрения взаимодействия всех модулей пайплайна, так и диаграмм классов реализации и диаграмм последовательностей взаимодействия методов данных классов с учетом последующей возможности интуитивной интеграции решения с существующими системами напрямую или через смежные микросервисы, содержащие логику взаимодействия с

клиентами, используя результаты обработки разрабатываемого сервиса,

2) этап разработки, в рамках которого кроме непосредственной реализации были выбраны технологии для нее, включающие, за исключением ранее выбранных моделей ИИ, язык программирования Java 17, фреймворк для более быстрой, гибкой и масштабируемой реализации API Spring Boot 3, исходя из выбранного языка фреймворк тестирования JUnit 5, программный контейнеризатор Docker, а также был выявлен пайплайн для организации сотовой коммуникации с клиентами [11] и рассмотрен ряд решений для использования в качестве автоматизации приема звонков и передачи аудио между сервисом и клиентом,

3) этап тестирования, включающий модульное, по результатам которого было написано 35 тестов бизнес-логики реализованной системы, засчет которых было достигнуто 95 процентное покрытие тестами всего кода, и мануальное, которое показало работоспособность интеграции всех частей реализации и способность обрабатывать поставленные в задаче форматы данных и выводить ожидаемый результат в кратчайшие в соответствии с выделенными ресурсами на использование сроки — 1 минута 44 секунды для генерации и передачи аудио и 20 секунд на обработку ответа клиента и выдачи результата в подходящем формате.

Выводы. Для подтверждения применимости разработанной системы решается задача автоматизации обработки обращений граждан с целью передачи информации, содержащей цель запроса предоставления коммунальных услуг и данные о местоположении (квартиры, подъезд, участок) для их предоставления. Исследование формирует требования и предоставляет информацию о создании основ для разработки высокоэффективного, масштабируемого и интегрируемого сервиса сбора данных, способного соответствовать требованиям современных информационных систем и цифровых экосистем, а разработка инструмента на основе полученной базы знаний показывает и доказывает их применимость, а также предоставляет пример реализации с использованием передовых технологий для решения поставленной задачи с учетом установленных требований и проверки соответствия им.

Список использованных источников:

1. Microservice API Evolution in Practice: A Study on Strategies and Challenges / Lercher Alexander, Glock Johann, Macho Christian, Pinzger Martin. — Текст : непосредственный // Journal of Systems and Software. — 2024. — No 215.

2. Мокшанов М.В. Применение искусственного интеллекта в анализе данных: обзор текущего состояния и будущих направлений // Universum: технические науки : электрон. научн. журн. 2024. 5(122). URL: <https://7universum.com/ru/tech/archive/item/17513> (дата обращения: 16.01.2025).

3. A Systematic Mapping Study on Architectural Approaches to Software Performance Analysis / Zhao Yutong, Xiao Lu, Wei Chenhao [и др.]. — Текст : непосредственный // JOURNAL OF LATEX CLASS FILES. — 2015. — No 14. — С. 1-27.

4. Multichannel Digital Marketing Optimizations through BigData Analytics in the Tourism and Hospitality Industry / P. S. Damianos, P. R. Dimitrios, C. T. Marina, Vassilakis Costas. — Текст : непосредственный // Journal of 37 theoretical and applied electronic commerce research. — 2022. — No 17. — С. 1383-1408.

5. speech-assistant-openai-realtime-api-python — Текст : электронный // GitHub : [сайт]. — URL: <https://github.com/twilio-samples/speech-assistant-openai-realtime-api-python?tab=readme-ov-file> (дата обращения: 21.01.2025).

6. call-center-ai — Текст : электронный // GitHub : [сайт]. — URL: <https://github.com/microsoft/call-center-ai/blob/main/README.md> (дата обращения: 21.01.2025).

7. Открытые модели для распознавания русской речи 2025 // AlphaСephei URL: <https://alphacephei.com/nsh/2025/04/18/russian-models.html> (дата обращения: 16.01.2026).
8. Phi-4-Mini Technical Report: Compact yet Powerful Multimodal Language Models via Mixture-of-LoRAs // arXiv URL: <https://arxiv.org/pdf/2503.01743> (дата обращения: 18.01.2026).
9. alibayram/smollm3:latest // Ollama URL: <https://ollama.com/alibayram/smollm3:latest> (дата обращения: 18.01.2026).
10. Evaluation of Russian TTS models // AlphaСephei URL: <https://alphacephei.com/nsh/2024/07/12/russian-tts.html> (дата обращения: 19.01.2026).
11. Облачная АТС своими руками: нужно ли это вам // 1АТС URL: <https://1ats.ru/blog/oblachnaya-ats-svoimi-rukami/> (дата обращения: 20.02.2026).