

РАЗВЁРТЫВАНИЕ ИНФРАСТРУКТУРЫ ДЛЯ ИИ-АССИСТЕНТА ПО УЧЕБНЫМ ДИСЦИПЛИНАМ

Рахманов В. А. (ИТМО)

Научный руководитель – Доцент, заведующий лабораторией факультета прикладной информатики Хлопотов М. В. (ИТМО)

Введение. ИИ-ассистент по учебным дисциплинам – проект, который представляет собой чат-бота с подключённым RAG, который имеет 2 роли:

- Преподаватель
- Студент

Со стороны преподавателя доступен функционал загрузки и управления учебными материалами, общения с чат-ботом и просмотр собранной статистики по взаимодействию студентов с ботом в рамках учебной дисциплины, закрепленной за преподавателем. Со стороны студента: просмотр учебных материалов и общение с чат-ботом. Проект направлен на помощь обеим сторонам, чтобы студенты имели возможность задать вопрос чат-боту, который будет давать ответы на вопросы на основе учебных материалов.

Основная часть. Для создания инфраструктуры такого чат-бота будут использовано облако, которое имеет сервисы для работы с RAG (YandexCloud/ SelectelCloud) и инструменты подхода IaC:

Terraform – для развёртывания инфраструктуры (облачного провайдера, виртуальные машины для бота, базы данных, RAG и дополнительные ресурсы при необходимости) и управления состоянием.

Ansible – для автоматизации конфигурации ВМ и используемых ресурсов после их создания. Ansible-плейбуки позволяют автоматизировать установку пакетов, настройку служб и развёртывание ассистента.

Выводы. Такой подход обеспечит отказоустойчивость и предоставит возможность управления всеми аспектами инфраструктуры, что улучшает видимость и контроль над ресурсами.

Список использованных источников:

1. Официальная документация Terraform [URL: <https://developer.hashicorp.com/terraform/docs>]
2. Официальная документация Ansible [URL: <https://docs.ansible.com>]
3. Официальная документация Selectel [URL: <https://docs.selectel.ru/terraform/quickstart/>].