

РАЗРАБОТКА СИСТЕМЫ УПРАВЛЕНИЯ КОНФИГУРАЦИЯМИ ПРИ ОБРАБОТКЕ ПОТОКОВЫХ ДАННЫХ

Долматова К. А.¹

Научный руководитель – канд. техн. наук, доцент Забашта А. С.¹

¹Университет ИТМО

fantastic.drogon@gmail.com

Введение

В настоящее время в одном из крупных банков России все события, связанные с активностью пользователей в его мобильных приложениях, поступают и сохраняются на автоматизированных системах (АС). Такие системы выступают в роли источников, обрабатывающих до 20000 уникальных типов событий. Для выполнения бизнес задач командам разработки нужно подключаться к АС, чтобы получить доступ к потоковым данным. Сейчас процесс подключения происходит так: сначала разработчик заполняет заявку с конфигурацией, потом она перейдёт в отдел безопасности, где будет проверена на предмет возможности предоставления запрошенных данных конкретному разработчику или нет, и только после этого будет выдан доступ. Разработчику нужно найти нужное ему событие из большого числа данных. Из-за этого очень просто допустить ошибку в составлении конфигурации запроса, что приведёт к отказу в доступе и последующему созданию нового запроса. Так же неправильные заявки создают дополнительную нагрузку на отдел безопасности. Для уменьшения количества неправильных заявок и снижение нагрузки на отдел безопасности, данный процесс необходимо автоматизировать.

Основная часть

В качестве решения разработана система управления конфигурациями с использованием архитектуры RAG (Retrieval-Augmented Generation - генерация, дополненная поиском). Принцип работы: сначала пользователь пишет запрос в произвольной форме, где указывает описание события или его название, потом отправляет этот запрос в систему. По этому запросу система извлекает необходимые данные и создаёт их векторное представление для поиска релевантных событий во внешней векторной базе (retriever). Найденные события показываются разработчику для выбора именно тех событий, которые ему нужны. После этого они передаются в LLM (Large Language Model - большая языковая модель) для генерации точной конфигурации (generator).

Применение RAG-системы позволило решить следующие задачи:

- Автоматизация процесса подключения команд разработки к данным. Теперь разработчик может написать описание данных, к которым нужен доступ в произвольной форме, а система создаст ему необходимую конфигурацию.
- Минимизация ошибок, связанная с человеческим фактором. Так как пользователю необходимо выбрать события из предложенных, ему не нужно вводить названия, значит не будет ошибок в написании.
- Повышение удобства пользовательского взаимодействия. Так как минимизировался ручной ввод в интуитивно понятной системе.

Выводы

В результате работы разработана система, с помощью которой удалось автоматизировать процесс создания и проверки конфигурации. Внедрение этого решения позволило минимизировать временные затраты на получение доступа к необходимым данным и исключить ошибки, связанные с некорректным заполнением заявок.

Литература

1. Уэске Ф., Калаври В. Поточковая обработка данных с Apache Flink. Москва : ДМК Пресс, 2021. 298 с.
2. Пселтис Э. Поточковая обработка данных. Конвейер реального времени. Москва : ДМК Пресс, 2018. 218 с.
3. Душкин Р.В. RAG-системы от теории к практике. Москва : ДМК Пресс, 2025. 286 с.