

Разработка системы автоматизированного сбора, формирования и форматирования публикаций в новостных каналах в социальных сетях

Сидоров. И.О.¹, Шамис А.Р.¹, Тюфяков Н.Е.¹

Научный руководитель – ординарный доцент Болдырева Е.А.

¹Университет ИТМО

sidorov-2004@list.ru, alshamis2004@yandex.ru, post8888@list.ru

Введение

В эпоху цифровой коммуникации социальные сети стали ключевым каналом распространения информации, где оперативность и качество публикаций напрямую влияют на вовлечённость аудитории. Ручной сбор данных, подготовка и форматирование контента требуют значительных временных затрат. Особенно остро данная проблема проявляется при ведении нескольких блогов или при сборе информации из множества источников одновременно. В данной работе рассматривается разработка информационной системы автоматизированного сбора, обработки и публикации контента в социальных сетях.

Основная часть

Предложенное решение обеспечивает полный цикл работы с данными: получение информации из различных источников, форматирование текста и изображений и последующую отправку в целевые каналы. Система построена по принципу микросервисной архитектуры, для обеспечения масштабируемости и отказоустойчивости. Взаимодействие между сервисами осуществляется с помощью брокера сообщений RabbitMQ. Точкой входа в систему является транспортный слой, реализующий взаимодействие с API социальных сетей, для первоначальной настройки параметров пользователем, которые в последствии хранятся в базе данных PostgreSQL, и последующей доставки публикаций. Во время настройки пользователь указывает в параметрах источники, из которых в последствии сервис сбора информации собирает данные для публикации. Эти данные попадают в сервис обработки и форматирования, где с помощью llm преобразуют исходный текст к заданному в параметрах пользователя форматированию. Далее, если исходные данные содержали изображения, они преобразуются с помощью заданных пользователем шаблонов (или же выбранных по умолчанию) для поддержания единой стилистики блога. Готовая публикация с помощью транспортного слоя возвращается пользователю на проверку в той социальной сети, которую он выбрал для первоначальной настройки, и предлагается для публикации в подключенных социальных сетях.

Выводы

Разработанная система позволяет существенно сократить время подготовки поиска информации и подготовки публикаций, что существенно повысит актуальность информации. Автоматизация упростит процесс редакторам каналов, позволив редактировать пришедшие из системы публикации и позволит сосредоточиться на обсуждении новостей с вовлеченным сообществом.

Литература

1. Burns B. Designing Distributed Systems: Patterns and Paradigms for Scalable, Reliable Services. — Sebastopol: O'Reilly Media. 2018. 163 p.
2. Videla A., Williams J. RabbitMQ in Action: Distributed Messaging for Everyone. — Shelter Island: Manning Publications. 2012. 360 p.
3. Date C. J. Database Design and Relational Theory: Normal Forms and All That Jazz. — Sebastopol: O'Reilly Media. 2012. 278 p.
4. Akidau T., Chernyak S., Lax R. Streaming Systems: The What, Where, When, and How of Large-Scale Data Processing. — Sebastopol: O'Reilly Media. 2018. 330 p.