

ПОВЫШЕНИЕ ПЕРТИНЕНТНОСТИ РЕЗУЛЬТАТОВ ПОИСКА ПО ИСТОРИИ СООБЩЕНИЙ ВНУТРИ КОРПОРАТИВНОГО МЕССЕНДЖЕРА

Бойко В.А. (ИТМО)

Научный руководитель – Письмак А.Е.

(ИТМО)

Введение. На сегодняшний день важнейшими средствами корпоративной коммуникации стали различные мессенджеры. Простота и удобство их использования сделали эти инструменты очень популярными в последнее время и за последние годы их использования, в базе сообщений мессенджеров в корпоративной среде накопилось довольно много данных, по которым часто приходится осуществлять поиск необходимой информации. Полнотекстовый поиск не удовлетворяет запросам пользователей из-за специфики хранения информации в формате коротких сообщений. В подобных окружениях необходим более интеллектуальный подход к решению задачи, который позволит осуществлять поиск с учетом смысловой нагрузки текстов сообщений.

Основная часть. В рамках рабочего процесса была сформирована следующая задача: внедрить алгоритмы семантического поиска [1] в мессенджер “Rocket.Chat”. Для её решения необходимо было произвести следующие модификации:

- подготовить базу данных, хранящую сообщения, внедрив индексацию, описанную в спецификации алгоритма семантического поиска;
- модифицировать модуль приложения, отвечающий за выборку из базы данных сообщений, чтобы он генерировал запрос в соответствии со спецификацией алгоритма семантического поиска и правильно возвращал пользователю полученные результаты;
- модифицировать интерфейс введения запросов пользователем, чтобы добавить возможность уточнять запрос, выбирая синонимичные конструкции из предлагаемого списка;
- модифицировать формат запроса от клиентской части к серверной на поиск по истории сообщений, чтобы принять во внимание модификацию, описанную в предыдущем пункте;
- добавить к серверной части приложения зависимое программное обеспечение, регламентированное алгоритмом поиска;
- изменения в семантической сети для проведения экспериментов в заданной области [2].

Выводы. В рамках работы был модифицирован ряд модулей приложения «Rocket.Chat», отвечающих за поиск по истории сообщений. Теперь вместо обычного поиска по подстроке приложение предлагает пользователю уточнить запрос выбрав из списка синонимичных токенов те, что лучше отражают вложенный в запрос смысл. Основываясь на этой дополнительной информации и данных, содержащихся в семантической сети, построенной заранее по нескольким базам знаний, система формирует более широкий запрос к базе данных, хранящей сообщения, и предлагает пользователю результаты, более полно охватывающие смысловую нагрузку запроса.

Список использованных источников:

1. Pismak A., Klimenkov S., Tsopa E., Yarkhev A., Nikolaev V., Gavrilov A. Method of Semantic Refinement for Enterprise Search // IC3K 2020 - Proceedings of the 12th

International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management - 2020, Vol. 2, pp. 307-312

2. Клименков С.В., Николаев В.В., Харитонов А.Е., Гаврилов А.В., Письмак А.Е., Покид А.В. Применение семантической сети для хранения слабоструктурированных данных // Инженерный вестник Дона [электронный журнал] – 2020. – № 2(62). – С. 27.