

## О Т З Ы В

на выпускную квалификационную работу магистра  
по направлению 02.04.02 «Фундаментальная информатика и информационные технологии»  
основной образовательной программы №19/5503/1  
«Технологии баз данных»

**ТАРАСОВА Никиты Андреевича,**

выполненную на тему:

**«Суммаризация мультязычных дискуссий в социальных сетях»**

Выпускная квалификационная работа Тарасова Н. А. посвящена такому актуальному направлению в области анализа текстовых данных, как суммаризация текстов в контексте задачи анализа пользовательских дискуссий в социальных сетях. Данная область имеет широкое практическое применение для анализа социальных, политических событий, проведения маркетинговых исследований, которое основано на изучении сгенерированных пользователями социальных сетей больших данных.

Целью выпускной квалификационной работы Тарасова Н. А. являлись рассмотрение и анализ различных методов и подходов к суммаризации дискуссий в социальных сетях. Дополнительно, для методов, основанных на языковых моделях, рассматривалась возможность совместного их применения с методами анализа данных. Для достижения поставленных целей студентом был проведен обзор существующих подходов и архитектур суммаризации текстовых данных с использованием современной отечественной и зарубежной литературы.

В работе была представлена архитектура системы анализа корпусов текстовых сообщений, полученных из социальных сетей, включая в себя все основные этапы: сбор, обработку данных, применение методов суммаризации, тематического моделирования, анализа тональностей, визуализации итоговых результатов. Были описаны базовые подходы, основанные на методах тематического моделирования, проведено их сравнение с методом, основанным на кластеризации с предварительной кодировкой данных языковой предобученной моделью. Были проведены эксперименты, в которых использовались различные методы абстрактной суммаризации совместно с различными методами анализа тональности. Данные тесты показали эффективность методов в решении задач анализа комментариев к отдельным записям, полученных из различных сообществ сети Reddit и анализа временных рядов мультязычных записей и комментариев из сети Twitter.

Стоит отметить, что все результаты работы получены студентом лично. Автор достаточно хорошо изучил различные методы и инструменты анализа текстовых данных и овладел инструментами их практической реализации. Часть полученных результатов работы были доложены на нескольких международных конференциях и опубликованы в журналах, индексируемых в наукометрических базах Scopus и WoS. А на основе предложенных подходов автор разработал программный комплекс, по которому получил Свидетельство о государственной регистрации данной программы.

В ходе выполнения работы Тарасов Н. А. показал себя творческим и самостоятельным исследователем, способным анализировать результаты своей работы и предлагать варианты дальнейших улучшений реализованных подходов. Помимо этого, хочу отметить хорошую теоретическую подготовку студента, его высокий уровень самодисциплины и работоспособности.

На основании вышеизложенного считаю, что выпускная квалификационная

работа Тарасова Н. А. на тему «Суммаризация мультязычных дискуссий в социальных сетях» выполнена в соответствии с предъявляемыми требованиями и заслуживает оценку «отлично», а ее автор рекомендацию в аспирантуру.

**Научный руководитель**

Заведующий Кафедрой  
технологии программирования  
доцент, к.т.н.



Блеканов И.С.

4 июня 2021 г.