

Анастасия Александровна Шibaева, студент  
Национальный исследовательский университет  
«Высшая школа экономики», Санкт-Петербург,  
Россия  
*э-почта:* shibaevanastasia@gmail.com

Anastasia Alexandrovna Shibaeva, Student  
National Research University Higher School  
of Economics, St. Petersburg, Russia  
*e-mail:* shibaevanastasia@gmail.com

Алина Андреевна Тельнова, студент  
Национальный исследовательский университет  
«Высшая школа экономики», Санкт-Петербург,  
Россия  
*э-почта:* lina.telnova.01@mail.ru

Alina Andreevna Telnova, Student  
National Research University Higher School  
of Economics, St. Petersburg, Russia  
*e-mail:* lina.telnova.01@mail.ru

Владислав Владимирович Лобанов, студент  
Национальный исследовательский университет  
«Высшая школа экономики», Санкт-Петербург,  
Россия  
*э-почта:* vlobanovpdrnv@gmail.com

Vladislav Vladimirovich Lobanov, Student  
National Research University Higher School  
of Economics, St. Petersburg, Russia  
*e-mail:* vlobanovpdrnv@gmail.com

Полина Сергеевна Соловьева, студент  
Национальный исследовательский университет  
«Высшая школа экономики», Санкт-Петербург,  
Россия  
*э-почта:* polinasolov2001@gmail.com

Polina Sergeevna Solovyova, Student  
National Research University Higher School  
of Economics, St. Petersburg, Russia  
*e-mail:* polinasolov2001@gmail.com

## **ДАВАЙ ПОКОРОЧЕ: СУММАРИЗАЦИЯ ТЕХНИЧЕСКИХ ТЕКСТОВ НА РУССКОМ ЯЗЫКЕ, СУЩЕСТВУЮЩИЕ РЕШЕНИЯ И ПЕРСПЕКТИВЫ**

В статье приводится обзор и анализ существующих методов суммаризации технических текстов с помощью LLM и примеров их применения на практике на материале технических стандартов организации, а также намечается дальнейший план исследования.

*Ключевые слова:* суммаризация; экстрактивная суммаризация; абстрактная суммаризация; технический стандарт; LLM; промпт-инжиниринг.

### **TOO LONG, DIDN'T READ: TECHNICAL TEXT SUMMARIZATION IN RUSSIAN, EXISTING SOLUTIONS AND PROSPECTS**

An overview of existing methods for summarizing technical texts using LLMs, along with some practical examples based on a company's technical standards. We also outline our further research plan.

*Key words:* summarization; extractive summarization; abstractive summarization; technical standard; LLM; prompt engineering.

На данный момент метод суммаризации текста применяется для работы с медицинскими, научными и юридическими документами, что позволяет оптимизировать поиск и усвоение релевантной информации. Интерес к этому виду обработки данных также возник в секторе бизнеса, что связано, в частности, с потребностью в суммаризации технических документов – в создании кратких содержаний, которые уменьшают время и усилия на изучение больших объемов текста.

Цель работы – определение наиболее эффективного инструмента для автоматической суммаризации стандартов организации на русском языке.

Задачи: 1) анализ и сравнение подходов к суммаризации (экстрактивный и абстрактный); 2) выбор наиболее эффективной большой языковой модели на основе метрик оценки качества суммаризации; 3) валидация результатов.

Эта работа посвящена суммаризации технической документации, для исследования было взято десять СТО (стандартов организации) компании АО «Северсталь», восемь в формате docx, два в формате doc длиной от 7 до 45 страниц. Всего в компании около 200 стандартов организации. Документы имеют четко определенную и неизменную структуру. На первой странице располагается титульный лист с идентификационной информацией об СТО, на второй – содержание документа.

Для достижения цели исследования первоначально был опробован экстрактивный метод с двумя моделями, предобученными для суммаризации текстов на русском языке. Результаты экстрактивной суммаризации были неудовлетворительными, так как большое количество релевантной информации было утеряно в процессе. Было принято решение сосредоточиться на методе абстрактивной суммаризации.

Для осуществления абстрактивной суммаризации были задействованы две LLM, запущенные локально: Vikhr-Llama3.1-8B-Instruct-R-21-09-24 [Hugging Face] и QuantFactory/saiga\_gemma2\_10b-GGUF [Hugging Face].

Обе модели справились с суммаризацией, однако результаты суммаризации также имели существенные отличия при экспертной проверке. Во-первых, модель Vikhr-Llama сохранила повторяющиеся именованные сущности, а также дату положений. Saiga-Gemma, напротив, опустила некоторые именованные сущности, а также суммаризировала текст с помощью более коротких, понятных предложений. Во-вторых, Saiga-Gemma, в отличие от Vikhr-Llama, сохранила информацию о том, каким образом стандарт вводится в действие. Однако Saiga-Gemma допустила логическую ошибку в суммаризованном тексте, что является серьезным недостатком.

Также был проведен анализ результатов суммаризации приведенных моделей с помощью метрик оценки качества текстов BERTScore. BERTScore включает в себя три ключевых показателя: Точность, Полноту, F1 (гармоническое среднее от точности и полноты).

По результатам анализа двух вариантов суммаризации с помощью BERTScore модели Saiga-Gemma и Vikhr-Llama были определены как предпочтительные для дальнейшей работы с суммаризацией с помощью промпт-инжиниринга. Однако для достижения оптимальных результатов необходимо продолжать исследование в этом направлении, включая эксперименты с различными техниками промптинга, разработкой шаблонов для саммари и поиском решений, уменьшающих галлюцинации моделей.

Проведенный библиометрический анализ позволил выявить ключевые направления и методы, применяемые в данной области, а также определить существующие пробелы в исследованиях.

Абстрактивный метод суммаризации был определен как предпочтительный. Технические стандарты были суммаризированы с помощью двух языковых моделей с использованием промпт-инжиниринга.

Далее в рамках работы планируется провести анализ галлюцинаций, разработать вызов функции (function calling) для снижения уровня галлюцинаций и повышения контролируемости выданных, а также создать шаблоны, в которые выбранные модели будут вставлять извлеченные из текста суммаризованные отрывки СТО.