



Московский государственный технический университет имени Н.Э. Баумана

# Разработка метода тематического моделирования для новостей на русском языке

---

Автор:

студент группы ИУ7-81  
Маркин Кирилл Вадимович

Научный руководитель:

доцент, кандидат технических наук  
Клышинский Эдуард Станиславович

Консультант:

старший преподаватель  
Волкова Лилия Леонидовна

# Актуальность\\Введение

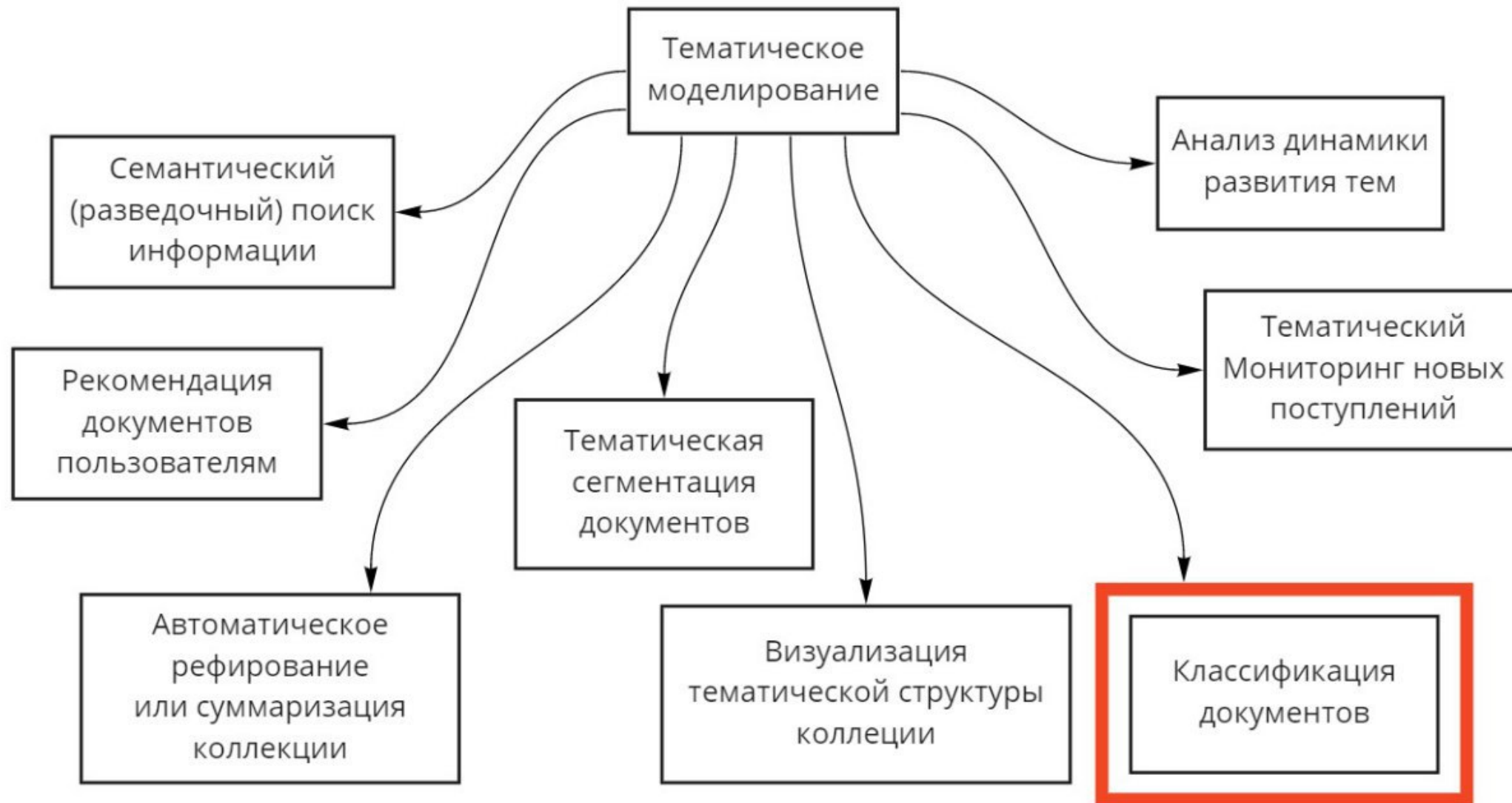
# Цели и задачи

Цель – целью работы является разработка метода тематического моделирования для новостей на русском языке.

## Задачи:

- анализ существующих решений и выбор базового алгоритма тематического моделирования для классификации новостей на русском языке;
- разработка программного продукта для сбора новостей на русском языке;
- разработка программного продукта для подготовки данных для последующего анализа;
- подбор методов улучшения алгоритма и значений его параметров;
- обучение модели;
- проведение параметризации метода;
- проведение апробации метода;
- составление рекомендаций о применимости предложенного метода.

# Задачи тематического моделирования



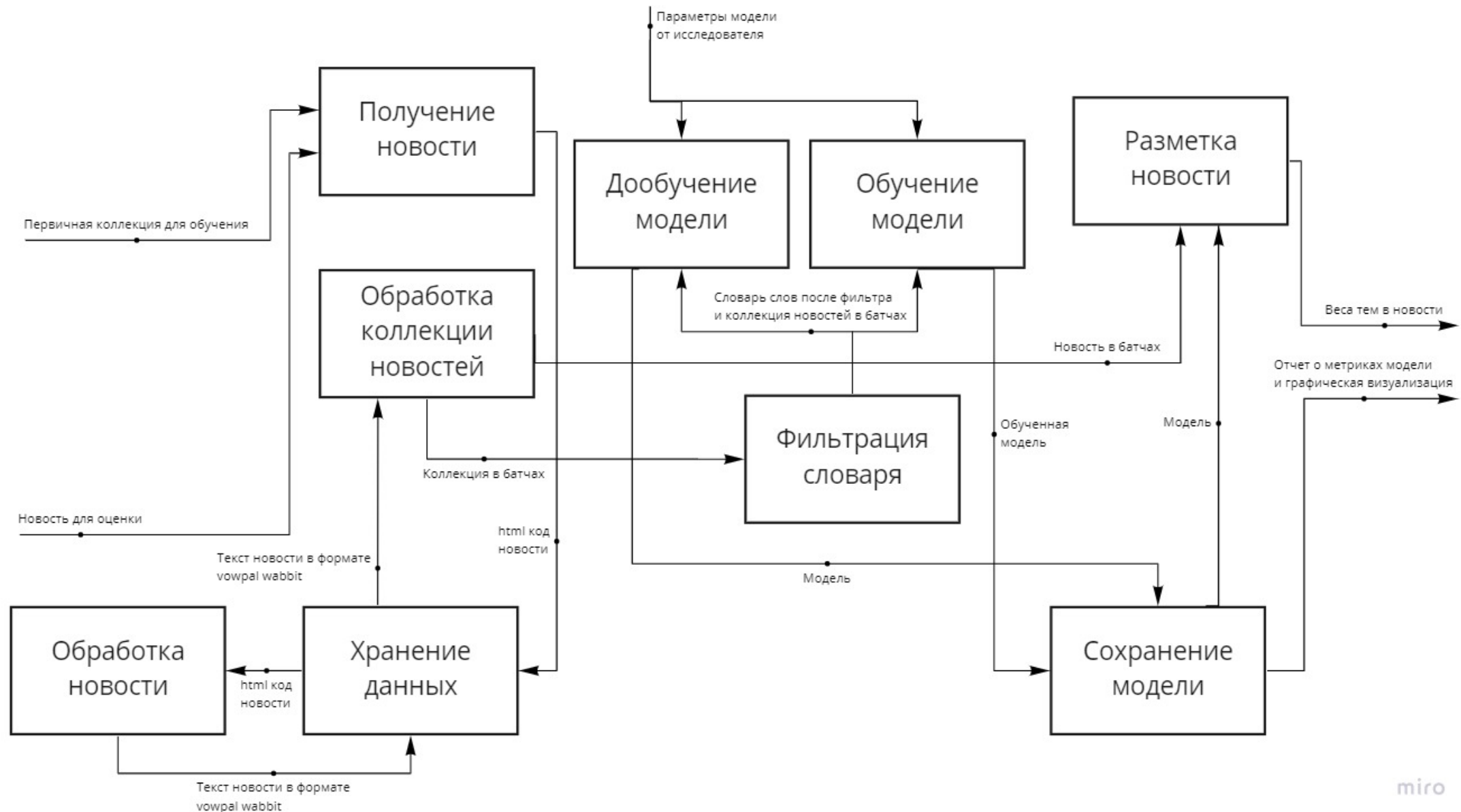
# Методы тематического моделирования



# Диаграмма метода



# Диаграмма метода



# Структура ПО // стек разработки



# Оценки

## Таблица 1. Результаты

- Можно исследовать сравнивая похожее с моим и посмотреть разницу
  - Можно исследовать собственный инструментарий на сложность, разные параметры и т д если решение совсем уж оригинальное
    - Можно в табличном виде - вот параметры, мой алгоритм лучше
    - Можно с помощью графиков

## Таблица 2. Результаты

# Заключение

В результате данной работы был разработан метод тематического моделирования новостей на русском языке.

## Были решены следующие задачи:

- проанализированы существующие решения и выбран базовый алгоритм тематического моделирования для классификации новостей на русском языке;
- разработан программный продукт для сбора новостей на русском языке и подготовки данных для последующего анализа;
- разработан программный продукт для подготовки данных для последующего анализа;
- подобраны методы улучшения алгоритма и значений его параметров;
- обучена модель;
- проведена параметризация метода;
- проведена апробации метода;
- составлены рекомендации о применимости предложенного метода.