



**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ**  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ  
(национальный исследовательский университет)»

Кафедра 319 «Системы интеллектуального мониторинга»

## Классификация задач компьютерной лингвистики для создания инструмента по управлению потоками задач

Выполнил:  
студент группы М3О-121М-22  
Чересов П.А.

**Москва, 2023**

# Направления компьютерной ЛИНГВИСТИКИ

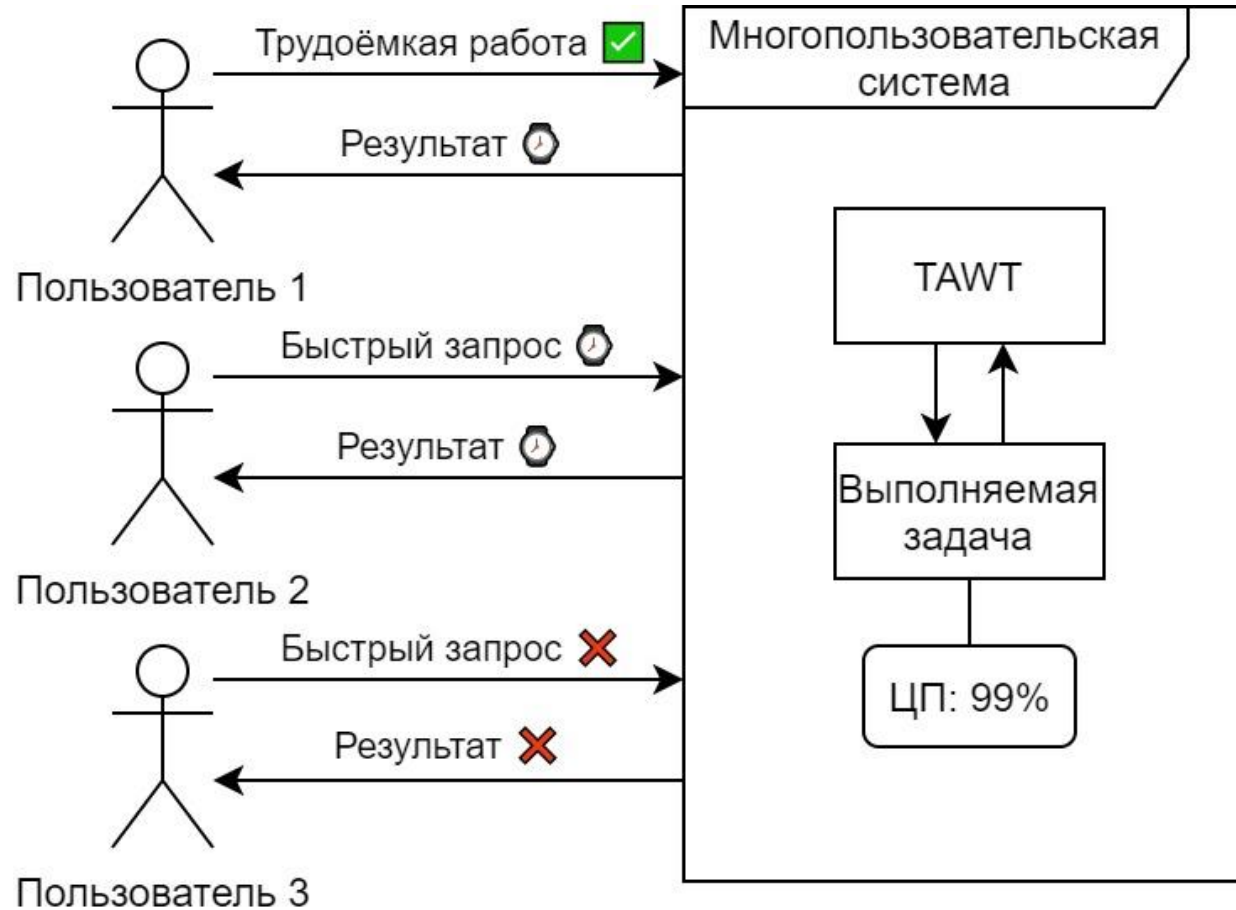
- Генерация текстов.
- Машинный перевод.
- Поиск заимствований.
- Поиск информации.
- Реферирование текстов.
- Аннотирование текстов.
- Извлечение ключевых слов.
- ...

# Проблемы многопользовательских систем

Большое количество разнотипных задач, которые приводят к:

- высокой нагрузке на оборудование;
- невозможности начать или завершить выполнение задачи;
- получению некорректных результатов.

# Проблемы многопользовательских систем



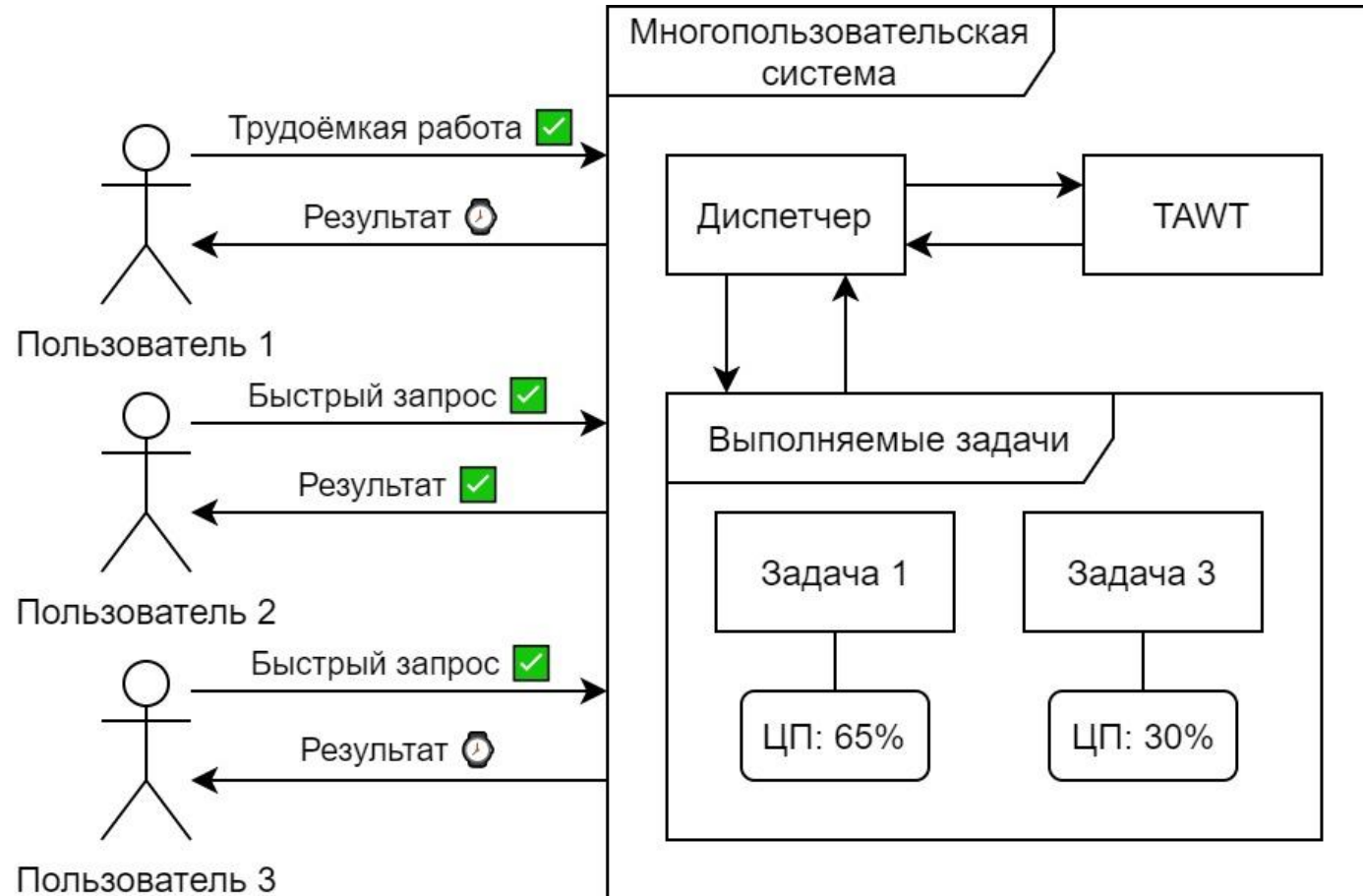
TAWT (Tools for Automated Work with Text) – набор инструментов для работы с текстом на русском языке.

# Предлагаемое решение

Создание инструмента по управлению потоками задач, ориентированного на автоматическую обработку текстов на естественном языке.

- Проанализировать существующие подходы управления потоками задач.
- Проанализировать и классифицировать задачи компьютерной лингвистики.
- Провести анализ сложности методов из библиотеки TAWT.
- Разработать требуемый инструмент.
- Протестировать его.

# Архитектура предлагаемого решения



# Классы задач

Базовая задача – какая-либо специфическая задача, решаемая на этапе анализа текста. Это может быть как выделение отдельных слов для графематического анализа, так и разрешение морфологической омонимии для морфологического анализа.

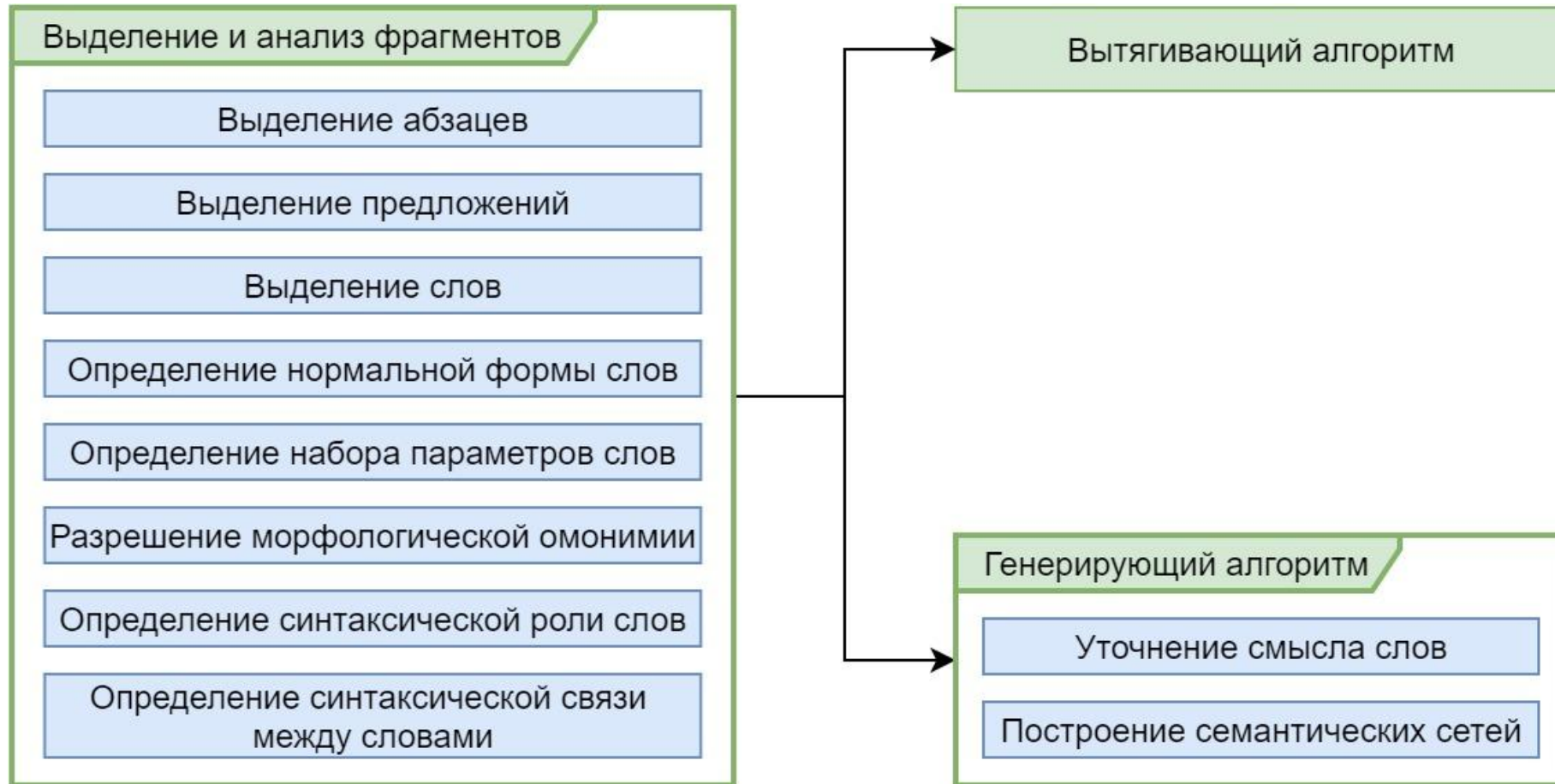
Прикладная задача – набор отдельных базовых задач или некоторый специфический алгоритм, работающий с текстовыми данными, полученными в результате обработки на предыдущих этапах, с целью получения практически ценного результата. Они также могут включать в себя другие прикладные задачи.

# Базовые задачи



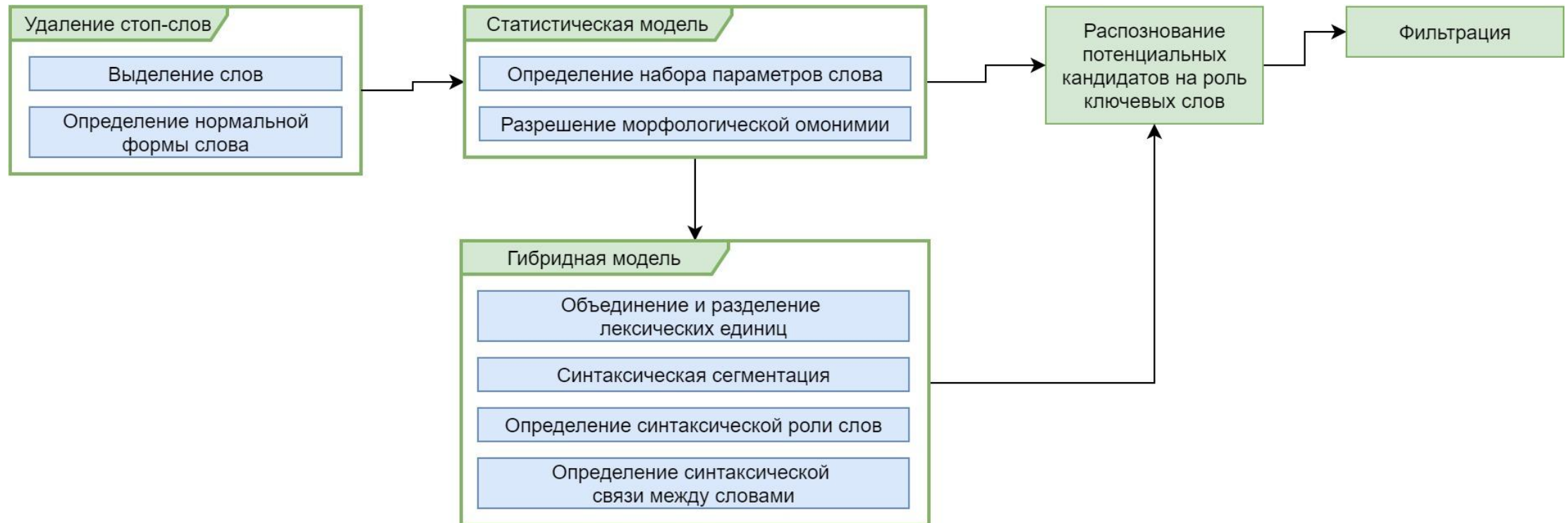


# Автоматическое аннотирование текстов



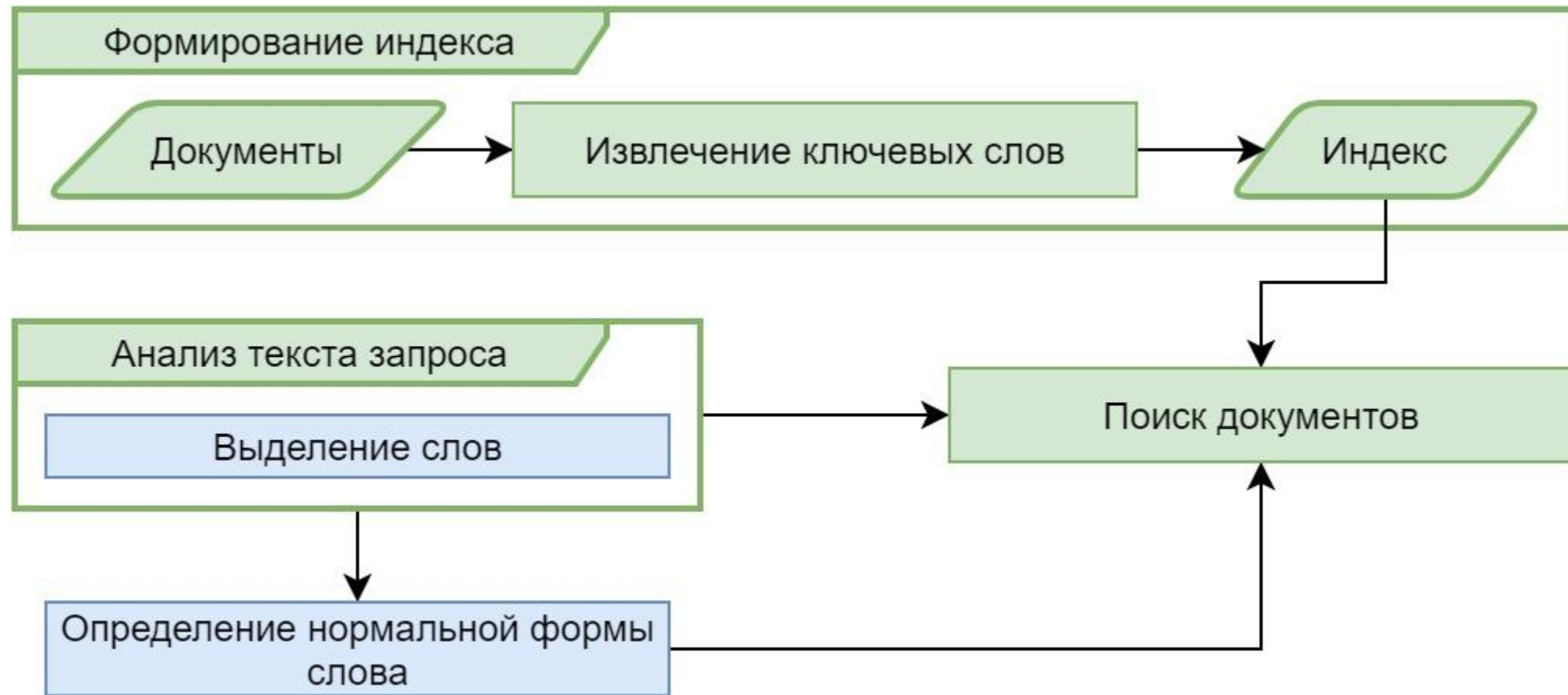
- – базовые задачи
- – прикладные задачи

# Автоматическое извлечение ключевых слов



- – базовые задачи
- – прикладные задачи

# Информационный поиск



- – базовые задачи
- – прикладные задачи

# Заключение

- Выделены основные задачи всех этапов анализа текста, которые являются базовыми при решении прикладных задач КЛ.
- Произведены анализ и классификация задач КЛ.
- На основании предложенной классификации при разработке инструмента будут определены весовые коэффициенты и приоритеты поступающих задач.

# Спасибо за внимание!

Классификация задач компьютерной лингвистики для создания  
инструмента по управлению потоками задач

Выполнил:  
студент группы М3О-121М-22  
Чересов П.А.