



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

Факультет компьютерных наук  
Департамент программной инженерии  
Выпускная квалификационная работа  
Программа подготовки протокола переговоров на  
основе распознавания речи

Выполнил студент группы БПИ132(2)  
образовательной программы  
09.03.04 «Программная инженерия»

Гнедин Павел Сергеевич

Научный руководитель:

Доцент департамента программной инженерии  
факультета компьютерных наук НИУ ВШЭ, к.т.н.

Ахметсафина Римма Закиевна

Более 5 000 000 организаций  
в Российской Федерации \*

(\* Согласно Росстату на 2016 год)

# ОСНОВНЫЕ ПОНЯТИЯ, ОПРЕДЕЛЕНИЯ, ТЕРМИНЫ

1. **API** – набор готовых классов, процедур, функций, структур и констант, предоставляемых приложением (библиотекой, сервисом) или операционной системой для использования во внешних программных продуктах;
2. **Бланк** – общая часть текста для документов одного типа;
3. **Распознавание речи** – перевод звука, записанного с помощью микрофона, в текстовую форму;
4. **Тональность текста** – эмоциональное отношение автора высказывания к некоторому объекту (объекту реального мира, событию, процессу или их свойствам/атрибутам), выраженное в тексте. Тональность всего текста в целом можно определить, как функцию (в простейшем случае сумму) лексических тональностей составляющих его единиц (предложений) и правил их сочетания;
5. **Трафарет** – основные сущности в тексте. Например: Персона, организация, дата и др.;
6. **Удобочитаемость** – свойство текстового материала, характеризующее лёгкость восприятия его человеком;
7. **Шаблон** – объединение бланка и набора трафаретов.

# ЦЕЛЬ И ЗАДАЧИ РАБОТЫ

## Цель работы

Целью данной работы является создание программы, с помощью которой можно записать речь, преобразовать ее «на лету» в текстовый вид, отредактировать текст и представить в удобном формате по заданным шаблонам.

## Задачи работы

1. Поиск и анализ существующих аналогов;
2. Обзор методов распознавания речи;
3. Сравнительный анализ и выбор API по распознаванию речи;
4. Изучение и выбор методов анализа текста;
5. Сравнительный анализ и выбор API анализа текста;
6. Разработка основных концепций;
7. Разработка схемы программы;
8. Выбор средств реализации;
9. Реализация серверной части программы;
10. Реализация WEB-оболочки программы;
11. Тестирование разработанной программы;
12. Подготовка технической документации согласно ГОСТ ЕСПД.

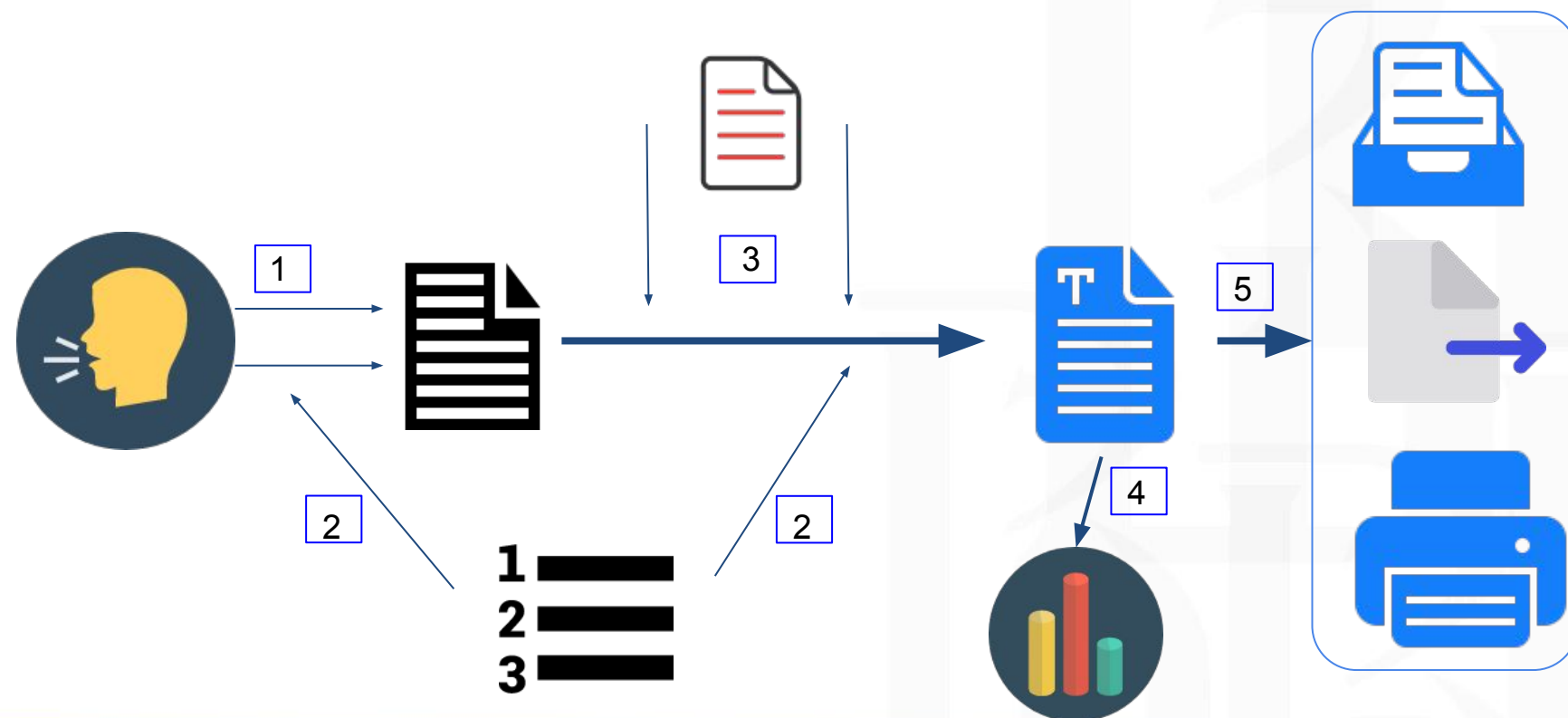
# АНАЛИЗ СУЩЕСТВУЮЩИХ РЕШЕНИЙ

Характ. Аналог	Web	Mobile	FREE	Реклама	Редакто р текста	Скачиван ие файла	Бланк	Анали з
Speech Logger	+	+	+	+	+	+	-	-
Dragon Dictatio n	-	+	+	-	+	-	-	-
Speech Notes	+	+	+	+	+	+	-	-
Speech Pad	+	-	+	+	+	+	-	-

Обзор данных аналогов показывает, что все они обладают ограниченной и недостаточной функциональностью для решения задачи подготовки протокола переговоров.

# Схема работы программы

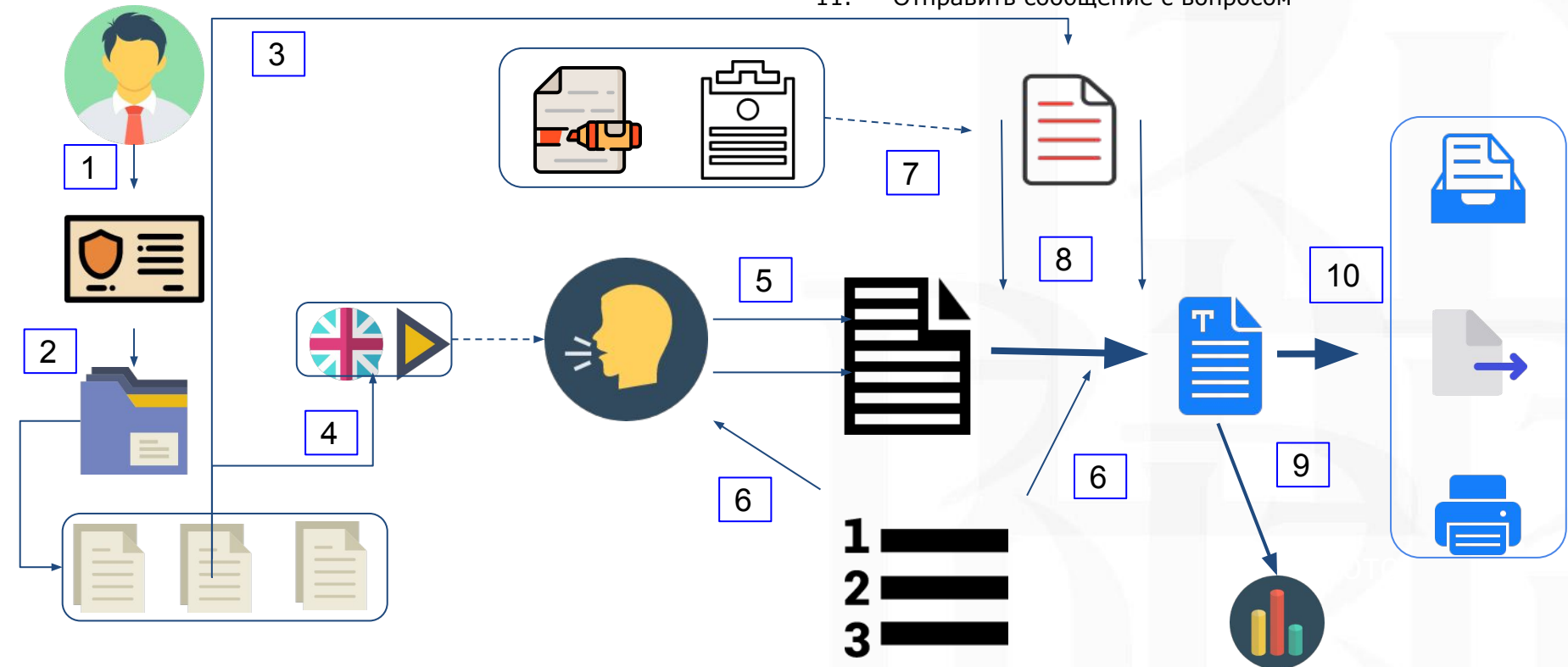
1. Распознать и перевести в текст речь;
2. В процессе записи распознанного текста редактировать его;
3. Применить шаблон к полученному тексту(бланк и трафареты);
4. При необходимости посмотреть статистику по полученному тексту;
5. Сохранить и/или распечатать/экспортировать итоговый документ.



# Схема работы программы (расширенная)

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

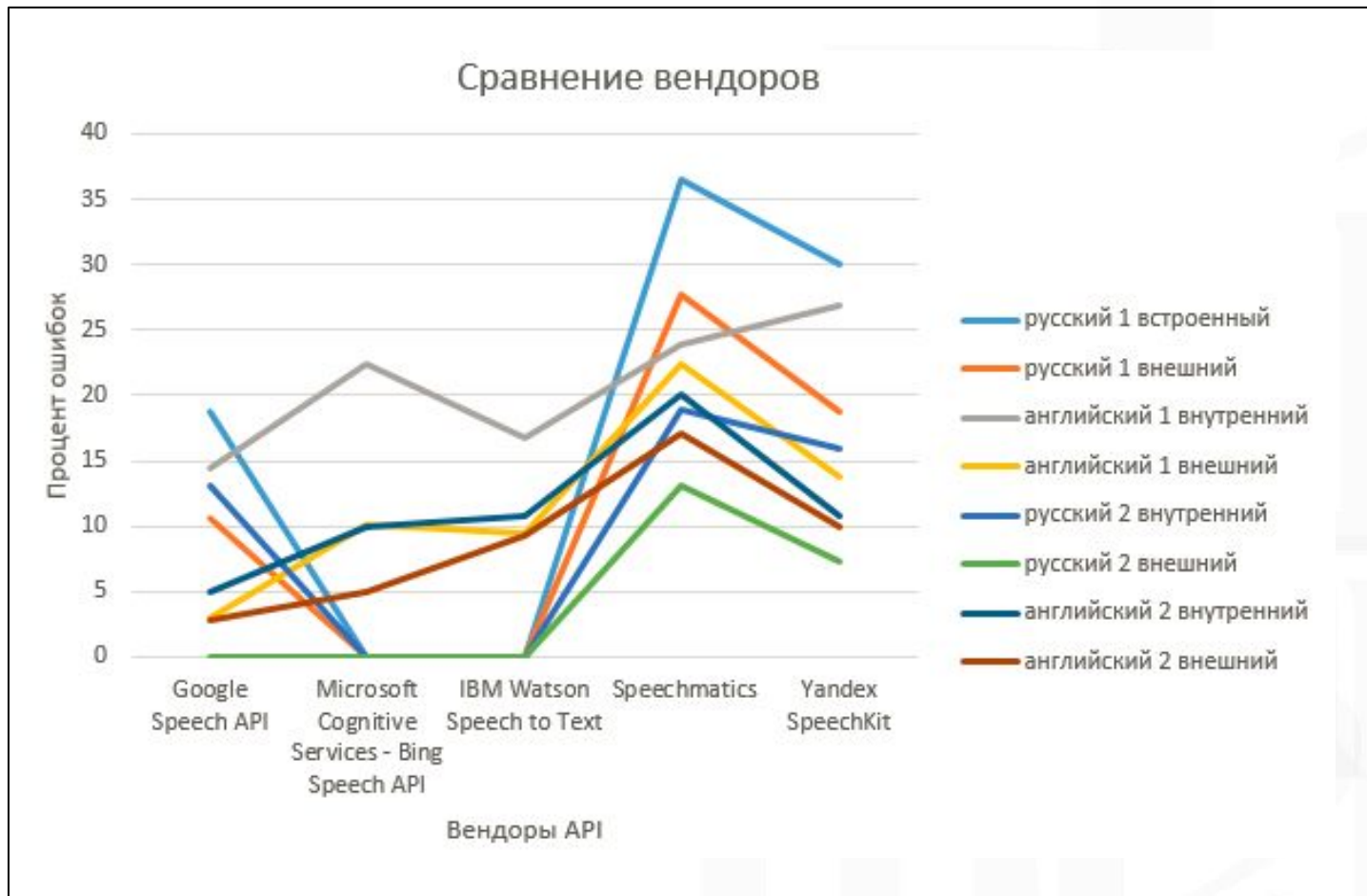
1. Авторизоваться;
2. Создать папку и необходимые файлы;
3. Выбрать шаблон, который применим к файлу или сменить его;
4. Выбрать язык и включить запись;
5. Распознать и перевести в текст речь;
6. В процессе записи распознанного текста редактировать его;
7. При необходимости отредактировать шаблон (бланк, имя, трафареты), либо создать новый;
8. Использовать шаблон (бланк и трафареты);
9. При необходимости посмотреть статистику по полученному тексту;
10. Сохранить и/или распечатать/экспортировать итоговый документ;
11. Отправить сообщение с вопросом



1. **Google Speech API** – API для распознавания речи, с помощью которого реализовано распознавание речи в данной программе;
2. **Django** – свободный фреймворк для веб-приложений на языке Python;
3. **Bootstrap 4 alpha 6** – свободный набор инструментов для создания сайтов и веб-приложений. Включает в себя HTML- и CSS-шаблоны оформления для веб-форм, кнопок, меток, блоков навигации и прочих компонентов веб-интерфейса, включая JavaScript-расширения;
4. **Anaconda** – дистрибутив языков программирования Python и R, включающий в себя набор библиотек;
5. **jQuery** – библиотека JavaScript, фокусирующаяся на взаимодействии JavaScript и HTML;
6. **uWSGI** и **Nginx** – это веб сервера. Nginx нужен для отдачи статики. uWSGI делает запросы питоновскому приложению и получает от него ответы. Взаимодействие между питоном и веб-сервером происходит по протоколу WSGI;
7. **Python** – основной язык для написания серверной части программы;
8. **html+css+javascript** – основные инструменты для реализации WEB-части программы;
9. **Aylien API** – библиотека для анализа англоязычного текста;
10. **Natasha** – библиотека для поиска и извлечения именованных сущностей (в тексте на русском языке);
11. **tinyMCE** – API редактора.

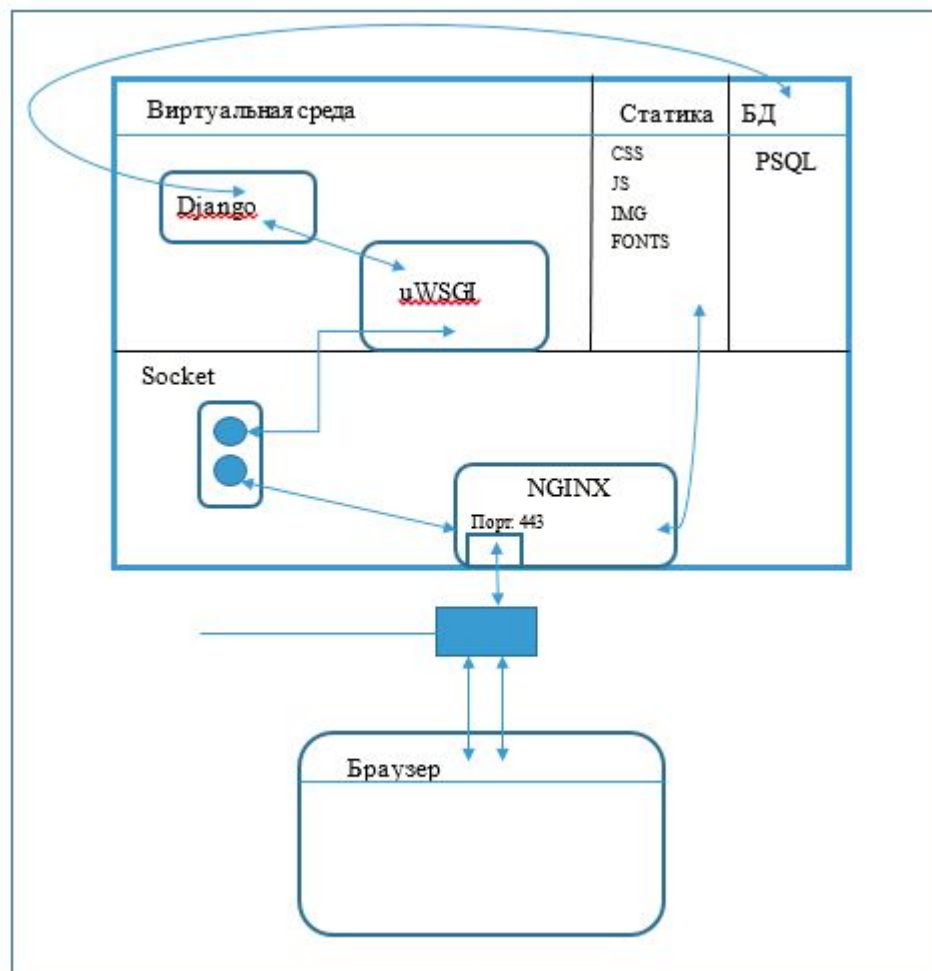


# АНАЛИЗ API ПО РАСПОЗНАВАНИЮ РЕЧИ



1. **Aylien API** – библиотека для анализа англоязычного текста (<http://aylien.com/>);
2. **Natasha** – библиотека для поиска и извлечения именованных сущностей в тексте на естественном языке (<http://natasha.readthedocs.io/ru/latest/>);
3. **TinyMCE** – API редактора (<https://www.tinymce.com/>).

# Схема программы



# Функциональность разработанного Web-приложения

В результате выполнения работы написано Web-приложение, обеспечивающее следующую функциональность:

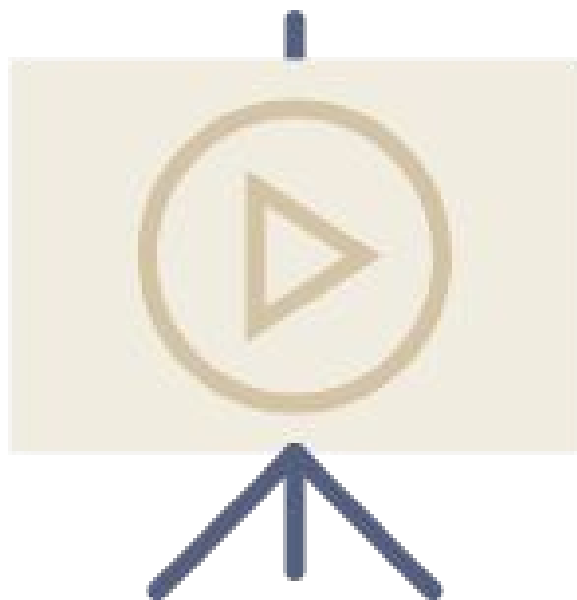
- Регистрация с помощью пользовательского логина и пароля;
- Авторизация с помощью пользовательского логина и пароля;
- Выход из аккаунта;
- Редактирование данные пользователя:
  - о Имя;
  - о Фамилия;
  - о Email;
  - о Компания;
  - о Должность.
- Распознавание английской и русской речи (перевод в текстовую форму);
- Сохранение результатов перевода;
- Создание, редактирование и использование файлов;
- Группировка файлов по папкам;
- Возможность экспортировать полученный текстовый файл;

# Функциональность разработанного Web-приложения

- Редактирование полученного текстового файла, а также редактирование его в режиме реального времени (в процессе записи речи):
  - о Выделение текста курсивом;
  - о Выделение текста жирным шрифтом;
  - о Изменение цвета текста;
  - о Изменение цвета фона текста;
  - о Изменение положения текста в документе.
- Создание, редактирование и применение бланка (верхняя часть текста, которая заранее записана пользователем и может многократно применяться к различным файлам);
  - Выделение основных сущностей из текста :
    - о Для русского языка:
      - § Имен;
      - § Географической локации (страны, города);
      - § Организации;
      - § Даты.
    - о Для английского языка:
      - § Имен;
      - § Географической локации (страны, города);
      - § Организаций;
      - § Даты.

# Функциональность разработанного Web-приложения

- Редактирование и использование шаблона. Шаблон представляет собой совокупность бланка и трафаретов;
- Определение общих характеристик текста:
  - Для русского языка:
    - § Определение удобочитаемости текста;
    - § Подсчет аналитики.
  - Для английского языка:
    - § Определение тональности текста;
    - § Подсчет аналитики;
    - § Выделение темы текста (классификация);
    - § Определение удобочитаемости текста;
    - § Подготовка аннотации.
- Выбора языка (русский или английский языки) при распознавании речи, а также при работе с текстом;
- Возможность работы с программой с правами более высокого уровня (наличие администратора), с выполнением следующих функций:
  - Добавление и изменение цветов, которые ставятся в соответствии трафаретам;
  - Добавление и изменение документов всех пользователей;
  - Добавление и изменение шаблонов всех пользователей;
  - Добавление и изменение трафаретов всех пользователей;
  - Добавление и изменение папок всех пользователей;
- Возможность работы с модулем помощи: возможность авторизованного пользователя отправить сообщение администратору.



# Пути дальнейшего развития

1. Интеграция с юристами/нотариусами;
2. Развитие мультязычности;
3. Развитие других платформ (мобильных);
4. Распознавание говорящего.



# СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- [1] An Exploratory Study on the Readability and Sentiment of Selected Philippine Senators' Microposts [Электронный ресурс] // <https://arxiv.org/pdf/1508.01321.pdf> (Дата обращения: 17.04.2017, режим доступа: свободный).
- [2] Anaconda [Электронный ресурс] //URL: <https://www.continuum.io/downloads/> (Дата обращения: 17.04.2017, режим доступа: свободный).
- [3] Automated readability index [Электронный ресурс] // <http://www.dtic.mil/dtic/tr/fulltext/u2/667273.pdf> (Дата обращения: 17.04.2017, режим доступа: свободный).
- [4] AYLIEN Text Analysis API Knowledge Center [Электронный ресурс] // <http://docs.aylien.com> (Дата обращения: 17.04.2017, режим доступа: свободный).
- [5] Bootstrap [Электронный ресурс] //URL: <http://getbootstrap.com/> (Дата обращения: 17.04.2017, режим доступа: свободный).
- [7] Django [Электронный ресурс] //URL: <https://www.djangoproject.com/> (Дата обращения: 17.04.2017, режим доступа: свободный).
- [8] Dragon Dictation [Электронный ресурс] // <https://itunes.apple.com/ru/app/dragon-dictation/id341446764?mt=8> (Дата обращения: 17.04.2017, режим доступа: свободный).
- [9] The uWSGI project [Электронный ресурс] //URL: <https://uwsgidocs.readthedocs.io/en/latest/> (Дата обращения: 17.04.2017, режим доступа: свободный).
- [10] tinymce [Электронный ресурс] <https://www.tinymce.com/> (Дата обращения: 17.04.2017, режим доступа: свободный).
- [11] What is jQuery? [Электронный ресурс] //URL: <https://jquery.com/> (Дата обращения: 17.04.2017, режим доступа: свободный).



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

# Спасибо за внимание!

Гнедин Павел Сергеевич  
[gnedinpavelrabota@gmail.com](mailto:gnedinpavelrabota@gmail.com)

Москва - 2017

# ДОПОЛНИТЕЛЬНЫЕ СЛАЙДЫ



Flesch-Kinkaid  
Dale-Chale readability formula  
Coleman-Liau index  
SMOG  
Automated Readability Index