



Цифровая платформа агрегации и анализа научно-технической информации

Руководство пользователя

Оглавление

1. Sciapp	3
2. Регистрация.....	4
3. Вход	5
4. Поиск заимствований.....	5
5. Семантический поиск.....	11
6. Поиск похожих документов.....	13
6.1. Поиск похожих.....	14
6.2. Ключевые слова.....	15
7. Помощник автора	15
8. Работа с результатами поиска	17
8.1. Анализ.....	18
8.2. Ключевые слова.....	22
9. Пользовательские подборки	23
9.1. Создание директории.....	23
9.2. Создание коллекции.....	24
9.3. Действия с подборкой.....	25
9.3.1. Общий доступ.....	25
9.3.2. Копировать.....	26
9.3.3. Переименовать.....	27
9.3.4. Переместить.....	27
9.3.5. Удалить	28
9.3.6. Добавить документы в коллекцию	28
9.3.7. Просмотр.....	29
9.4. Операции.....	29
9.4.1. Кластеризация.....	30
9.4.2. Расширение	30
9.4.3. Ключевые слова.....	30
9.4.4. Сравнение.....	31
9.4.5. Объединение	32
9.4.6. Пересечение	33
9.4.7. Разность.....	34
9.5. Аналитика	35
9.6. Искать.....	35
9.7. Просмотр документов коллекции.....	35
10. Работа с документом	36

1. Sciapp

Цифровая платформа агрегации и анализа научно-технической информации (SciApp) - программно-аппаратный комплекс интеллектуального поиска и анализа больших массивов текстов.

SciApp предназначен для решения следующих задач:

- 1) Индексация больших массивов текстовых документов из открытых источников интернета, баз данных, корпоративных хранилищ и т.п. с выделением сопутствующей метаинформации
- 2) Быстрое ознакомление с темой документов с помощью автоматически генерируемых по тексту ключевых словосочетаний
- 3) Эксплоративный поиск документов (тематически похожих на заданный эталонный документ)
- 4) Быстрое ознакомление с содержанием документов с помощью автоматически генерируемых резюме текстов
- 5) Поиск текстовой информации по запросу пользователя с поддержкой фасетного поиска по метаданным (как текстовым, так и нетекстовым) в режимах семантического, вопросно-ответного, концептного, а также поиска по отдельным ключевым словам
- 6) Формирование, сопоставление и тематический анализ пользовательских коллекций документов (тематическая кластеризация, построение ключевых тематических словосочетаний и т.п.)
- 7) Поиск смысловых текстовых заимствований
- 8) Тематический анализ коллекций документов

Сервис доступен по адресу sciapp.ru. При переходе на указанный адрес будет открыта начальная страница сайта(Рисунок 1).

Сервис предоставляет возможность входа в систему (см. Вход), выбора языка (вкладка "Язык" на начальной странице), а также поиска заимствований (см. Поиск заимствований), семантического поиска(см. Семантический поиск) и оценки качество публикации (см. Помощник автора). Регистрация (см. Регистрация) необходима для создания собственных подборок документов (см. Пользовательские подборки).



Рисунок 1 - Начальная страница

2. Регистрация

Для регистрации в системе нажмите вкладку «Войти» в правом верхнем углу начальной страницы (Рисунок 1), откроется форма авторизации (Рисунок 2). В данной форме выберите «Зарегистрироваться».

В форме регистрации (Рисунок 3) заполните необходимые поля и нажмите «Зарегистрироваться». После этого учетная запись создана и можно произвести вход в систему (см. Вход).

Вход в систему

Логин

Пароль

Войти

Зарегистрироваться

Восстановить доступ

Вернуться на главную страницу

Рисунок 2 - Форма авторизации

Регистрация

Адрес электронной почты

Пароль

Повторите пароль

Фамилия

Имя

Отчество

Номер телефона

Должность

Организация

Зарегистрироваться

[Вход в систему](#)

[Вернуться на главную страницу](#)

Рисунок 3 - Форма регистрации

3. Вход

Для авторизации в системе нажмите вкладку «Войти» в правом углу начальной страницы (Рисунок 1), откроется форма авторизации (Рисунок 2). В данной форме заполните необходимые поля и нажмите «Войти». В случае корректно введенных данных будет произведен вход в систему.

4. Поиск заимствований

Данный инструмент предназначен для анализа документа на наличие в нем заимствований из документов, содержащихся в коллекциях. Для доступа к соответствующему разделу интерфейса пользователя необходимо перейти по ссылке «Поиск заимствований» (Рисунок 1).

Для начала поиска заимствований необходимо загрузить для проверки документ в одном из поддерживаемых форматов или ввести текст в поле «Введите текст» (Рисунок 4).

SCI-APP Поиск текстовых заимствований

Пользовательские подборки GUEST@gmail.ru

Поиск заимствований

Семантический поиск

Поиск похожих

Помощник автора

Язык

Введите текст

или выберите файл

Файл не выбран

Выборить

Выберите коллекции

☐ Все коллекции

☐ Английская Википедия

☐ ФИПС. Изобретения

☐ ФИПС. Полезные модели

☐ Статьи энциклопедического характера

☐ ФИПС. Промышленные образцы

☐ Иностранные журналы

☐ Зарубежные статьи медицинской тематики

☐ Патенты USPTO

☐ Русскоязычная Википедия

☐ Российские журналы

☐ Авторефераты диссертаций

Годы: с 1992 по 2020

Год выхода документа:

2020

Автор:

Название:

Проверить на заимствования

☐ Использовать распознавание текста

Рисунок 4 - Поиск заимствований

Для загрузки файла нажмите «Выбрать». Откроется системное меню выбора файла (Рисунок 5). Выберите файл и нажмите кнопку «Открыть».

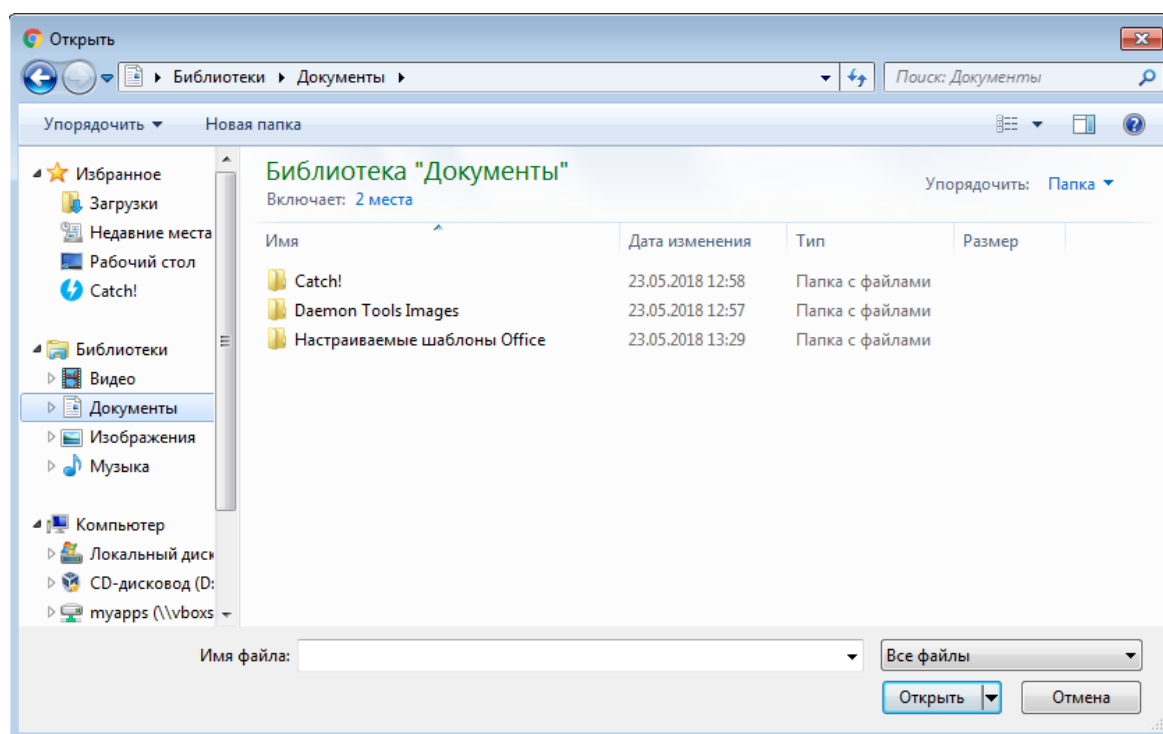


Рисунок 5 - Меню выбора файла

Выберите коллекции, по которым будет производиться поиск. Для этого отметьте галочками те коллекции, по которым вы хотите произвести поиск (Рисунок 6), или отметьте «Все коллекции» для поиска по всем коллекциям.

Выберите коллекции

<input type="checkbox"/> Все коллекции	<input type="checkbox"/> Статьи энциклопедического характера	<input type="checkbox"/> Зарубежные статьи медицинской тематики	<input checked="" type="checkbox"/> Российские журналы
<input type="checkbox"/> Английская Википедия	<input checked="" type="checkbox"/> ФИПС. Промышленные образцы	<input type="checkbox"/> Патенты USPTO	<input type="checkbox"/> Авторефераты диссертаций
<input checked="" type="checkbox"/> ФИПС. Изобретения	<input type="checkbox"/> Иностранные журналы	<input type="checkbox"/> Русскоязычная Википедия	
<input checked="" type="checkbox"/> ФИПС. Полезные модели			

Рисунок 6 - Выбор коллекций

Введите годы публикации документов, по которым вы хотите производить поиск, а также заполните поля «Год выхода документа», «Название» и «Авторы», если вам известны соответствующие характеристики загружаемого документа, выберите, необходимо ли использовать распознавание текста (Рисунок 7). Нажмите кнопку «Проверить заимствования».

Год выхода документа:

Автор:

Название:

☐ Использовать распознавание текста

Рисунок 7 - Метаполя документа

Результат проверки выглядит следующим образом (Рисунок 8). Выводится оценка оригинальности документа, проценты условно корректных и некорректных заимствований, диаграмма, показывающее соотношение заимствований и оригинальности текста, а также документы, из которых были произведены заимствования (информацию по работе с результатами поиска см. в разделе Работа с результатами поиска). Подробно ознакомиться с результатом проверки можно, нажав на кнопку «Просмотр заимствований в документе», откроется интерфейс просмотрщика документа (Рисунок 9).

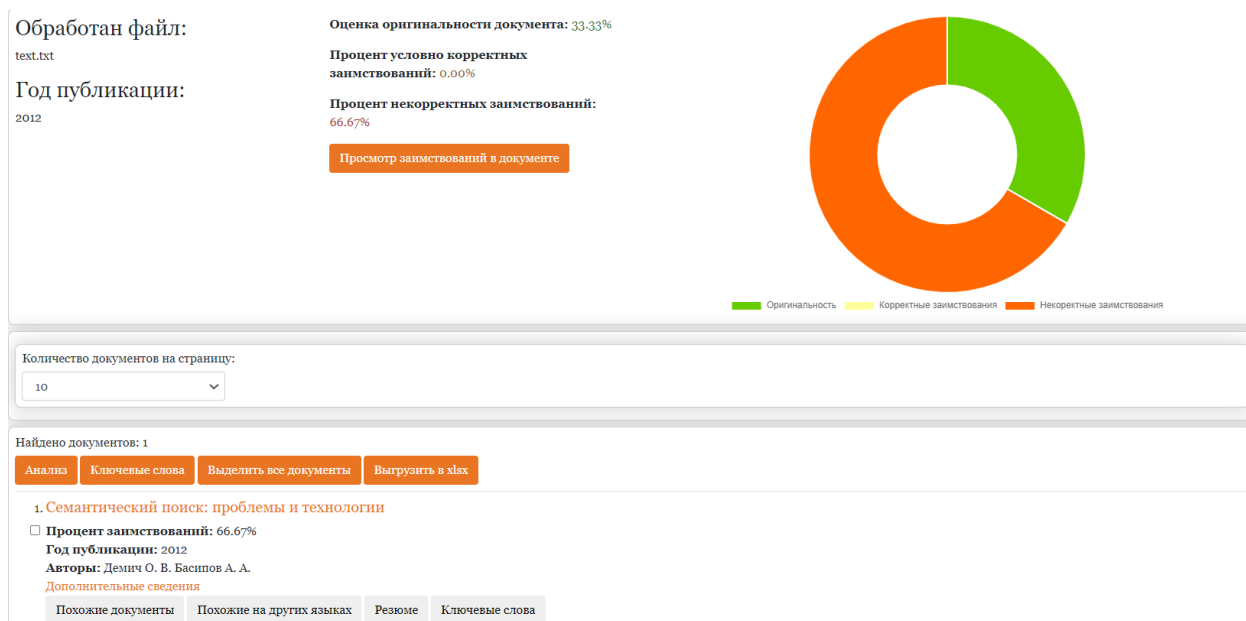


Рисунок 8 - Результат проверки на заимствования

Навигация по страницам

1

23.25% 0.0% 76.75%

Скачать PDF отчет Скачать PDF справку

Скачать расширенный отчет

Обозначить заимствования по источникам

Оригинальный текст

Непроверенный текст

Заимствования из нескольких источников

Показать все Скрыть все

Скрыть меньше 1%

1. Семантический поиск: ...
https://cyberleninka.ru/article/n/sema...
Басипов Андрей Алексеевич, Демич...
2012 76.75%

Семантическая паутина (Semantic Web) является расширением традиционного Интернета и нацелена на упрощение поиска и распределения информации. Данная технология основывается на элементах, построенных с использованием стандартных языков онтологий, таких как OWL. Обычные поисковые системы основываются на поиске ключевых терминов запроса в документе и не могут использовать его смысловое значение для получения результатов, поэтому сообщество исследователей семантической паутины предложило использовать семантические поисковые технологии, среди которых OntoSearch, Semantic Portals, Semantic Wikis, мультиагент R2R, семантические системы маршрутов (запросов), вопросно-ответные системы, использующие онтологии для хранения баз знаний [1]. Документ семантической паутины SWD (Semantic Web Document) можно рассматривать как набор данных контекстом которого является либо онтология либо обычный документ, размеченный определенными тегами, взятыми из онтологии предметной области. Такие Интернет-документы могут быть распределены по множеству различных категорий, относящихся к типам онтологий, используемых для разметки документа. Примерами таких категорий являются тяжеловесная или легковесная онтология. Рассмотрим существующие технологии семантической паутины в контексте следующих проблем: автоматическое создание формального запроса (онтологии запроса) и получение коллекций документов, структура которых не известна заранее (распределенные и семантически разнородные данные). Стандартный Интернет-поиск по ключевым словам базируется на поисковых технологиях, основой которых является обнаружение строкового (лексического) соответствия запрашиваемых терминов терминам, содержащимся в Интернет-документах. Обычно Интернет-поиск по ключу применяется для поиска неструктурированных Интернет-документов (текст без семантической разметки).

Рисунок 9 - Просмотрщик документа

Просмотрщик выполняет функции отображения загруженного документа, а также отображения найденных в нем совпадений с документами из коллекций. Просмотрщик содержит следующие блоки: текст исходного документа; список документов коллекций, в которых было найдено совпадение с текстом исходного документа; навигатор страниц; процентное соотношение оригинального текста к заимствованному; блок выгрузки отчетов.

Навигатор страниц. Расположен в верхней части интерфейса просмотрщика. Белым помечены номера страниц, на которых не обнаружено совпадений с документами коллекций. Частично залитые красным номера страниц помечают найденные совпадения,

причем, чем больше совпадений было найдено на странице, тем большая часть номера будет залита красным (Рисунок 10).

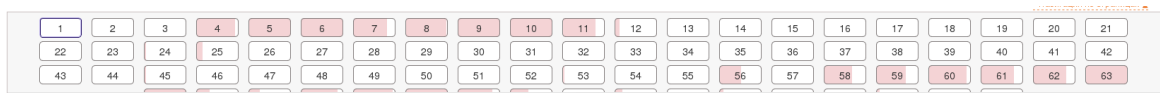


Рисунок 10 - Навигатор страниц

Процентное соотношение оригинального текста к заимствованному. Блок расположен в верхней левой части просмотрщика под навигатором страниц (Рисунок 11). Зеленым отмечен процент оригинального текста, желтым, заимствованного условно корректно, красным – некорректно заимствованного текста.

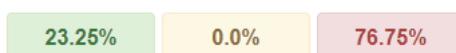


Рисунок 11 - Процентное соотношение

Список источников. Блок расположен в левой части просмотрщика (Рисунок 12). Для каждого источника указаны его название, автор и ссылка на документ.

В случае, если проверяющий не считает найденный документ источником его можно исключить из расчета, сняв галочку в левой верхней части блока с информацией о нем.

Примечание. Это действие приведет к полному исключению данного документа из расчета. Он не будет отображаться в списке источников в скачанных отчетах, а процентное соотношение оригинального и заимствованного текста будет пересчитано.

В случае если проверяющего не интересует список источников заимствования в документе, его отображение можно отключить, сняв галочку «Обозначить заимствования по источникам». Также присутствует возможность убрать источники, заимствования из которых составляет меньше 1%, для этого необходимо нажать соответствующую кнопку «Скрыть меньше 1%».

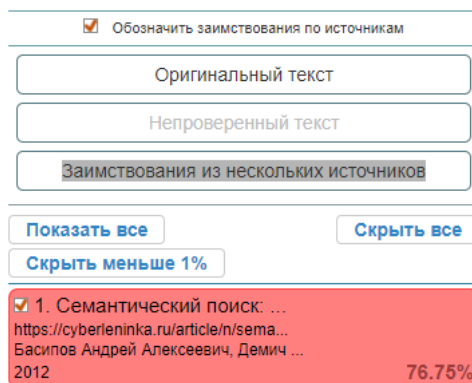


Рисунок 12 - Список источников

Текст исходного документа. Область отображения текста исходного документа расположена в центральной части просмотрщика (Рисунок 13).

Семантическая паутина (Semantic Web) является расширением традиционного Интернета и нацелена на упрощение поиска и распределения информации. Данная технология основывается на элементах, построенных с использованием стандартных языков онтологий, таких как OWL. Обычные поисковые системы основываются на поиске ключевых терминов запроса в документе и не могут использовать его смысловое значение для получения результата, поэтому сообщество исследователей семантической паутины предложило использовать семантические поисковые технологии, среди которых OntoSearch, Semantic Portals, Semantic Wikis, мультиагент P2P, семантические системы маршрутов (запросов), вопросно-ответные системы, использующие онтологии для хранения баз знаний [1]. Документ семантической паутины SWD (Semantic Web Document) можно рассматривать как набор данных, контентом которого является либо онтология, либо обычный документ, размеченный определенными тегами, взятыми из онтологии предметной области. Такие Интернет-документы могут быть распределены по множеству различных категорий, относящихся к типам онтологий, используемых для разметки документа. Примерами таких категорий являются тяжеловесная или легковесная онтологии. Рассмотрим существующие технологии семантической паутины в контексте следующих проблем: автоматическое создание формального запроса (онтологии запроса) и получение коллекций документов, структура которых не известна заранее (распределенные и семантически разнородные данные). Стандартный Интернет-поиск по ключевым словам базируется на поисковых технологиях, основой которых является обнаружение строкового (лексического) соответствия запрашиваемых терминов терминам, содержащимся в Интернет-документах. Обычно Интернет-поиск по ключу применяется для поиска неструктурированных Интернет-документов (текст без семантической разметки).

Рисунок 13 - Область отображения текста исходного документа

В этой области отображается текст исходного документа, с оригинальным форматированием.

В случае включенного флажка «Отображать заимствования по источникам» в тексте документа разными цветами помечены заимствования из различных документов. При нажатии на подсвеченный текст отображается всплывающее окно, в котором указано, из какого источника произведено заимствование, и приведен заимствованный фрагмент текста оригинального документа (Рисунок 14).

Рисунок 16 - Семантический поиск

Для выполнения поиска необходимо ввести запрос в поле «Поисковый запрос». Это может быть как обычная текстовая строка, так и строка форматированная в соответствии с языком запросов, подробную справку по которому можно получить, перейдя по ссылке «Справка по языку запросов» (Рисунок 17).

Рисунок 17 - Справка по языку запросов

Далее необходимо указать коллекции, по которым будет производиться поиск, а также при необходимости указать годы публикации искомых документов, их авторов и названия. Также можно указать правообладателя, классы МПК, код диссовета, место защиты и код ВАК, нажав «Дополнительные параметры» (Рисунок 18).

Рисунок 18 - Дополнительные параметры

Для выполнения поиска необходимо нажать кнопку «Поиск».

После выполнения поиска (это может занять некоторое время, в зависимости от сложности запроса) на странице появятся результаты поиска, отсортированные в порядке релевантности введенному запросу (Рисунок 19).

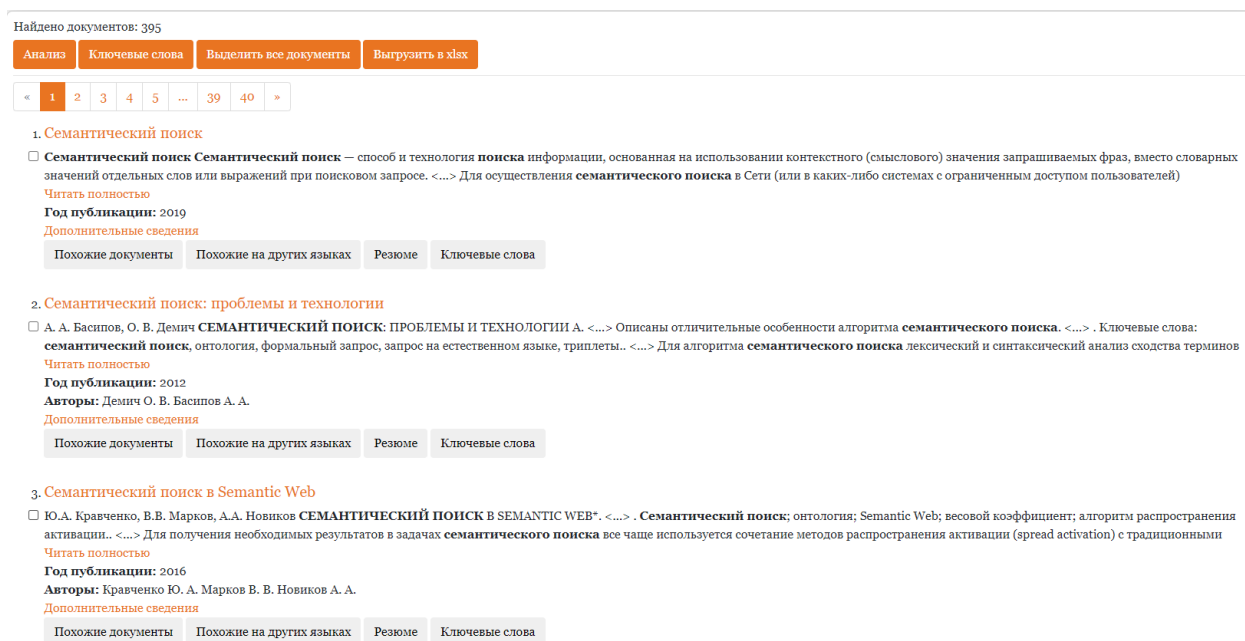


Рисунок 19 - Результаты поиска

Информацию по работе с результатами поиска см. в разделе Работа с результатами поиска.

6. Поиск похожих документов

Данный инструмент предназначен для поиска документов семантически близких загруженному документу или введенному тексту.

Для доступа к соответствующему разделу интерфейса пользователя необходимо перейти по ссылке «Поиск похожих». Откроется страница поиска похожих документов (Рисунок 20).

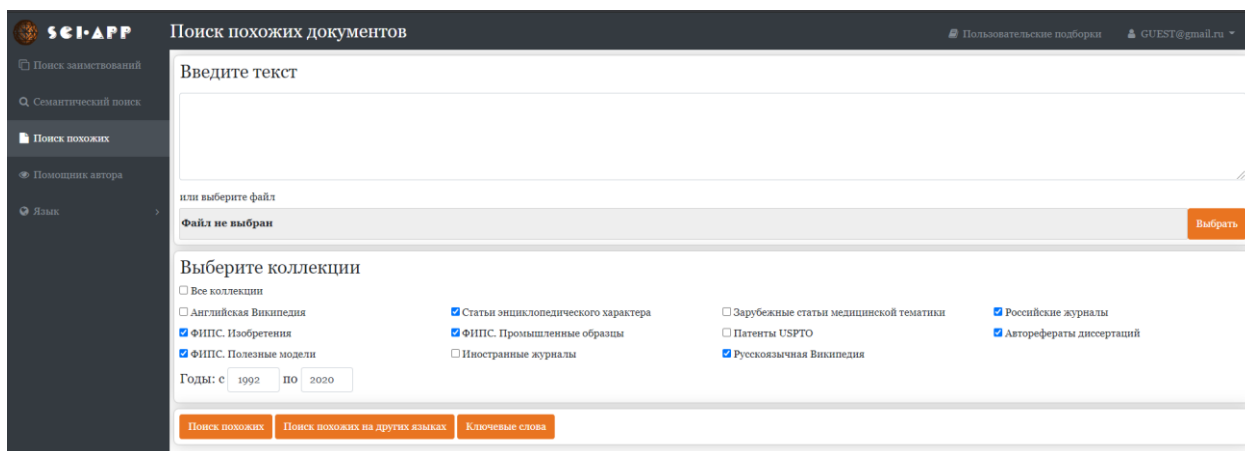


Рисунок 20 - Поиск похожих документов

Для выполнения поиска похожих документов необходимо загрузить проверяемый файл или ввести текст для проверки. Данный процесс полностью аналогичен описанному в разделе «Поиск заимствований».

После загрузки файла или введения текста выберите коллекции, по которым необходимо проводить поиск и годы публикации искомых документов. После этого выберите, какой тип действий вы хотите произвести.

6.1. Поиск похожих.

При нажатии на кнопку «Поиск похожих» будет выполнен поиск документов, семантически близких загруженному/введенному на том же языке на котором написан загруженный/введенный документ. Поиск может занять некоторое время, в зависимости от объема загруженного/введенного документа и количества документов в выбранных коллекциях. После выполнения поиска на странице отобразятся документы, семантически близкие загруженному/введенному (Рисунок 21).

По умолчанию установлен порог сходства – 0,2. В зависимости от необходимости, вы можете увеличить или уменьшить значение порога, воспользовавшись ползунком под надписью «Порог сходства» на странице результатов поиска похожих документов (Рисунок 21 - Результаты поиска похожих документов).

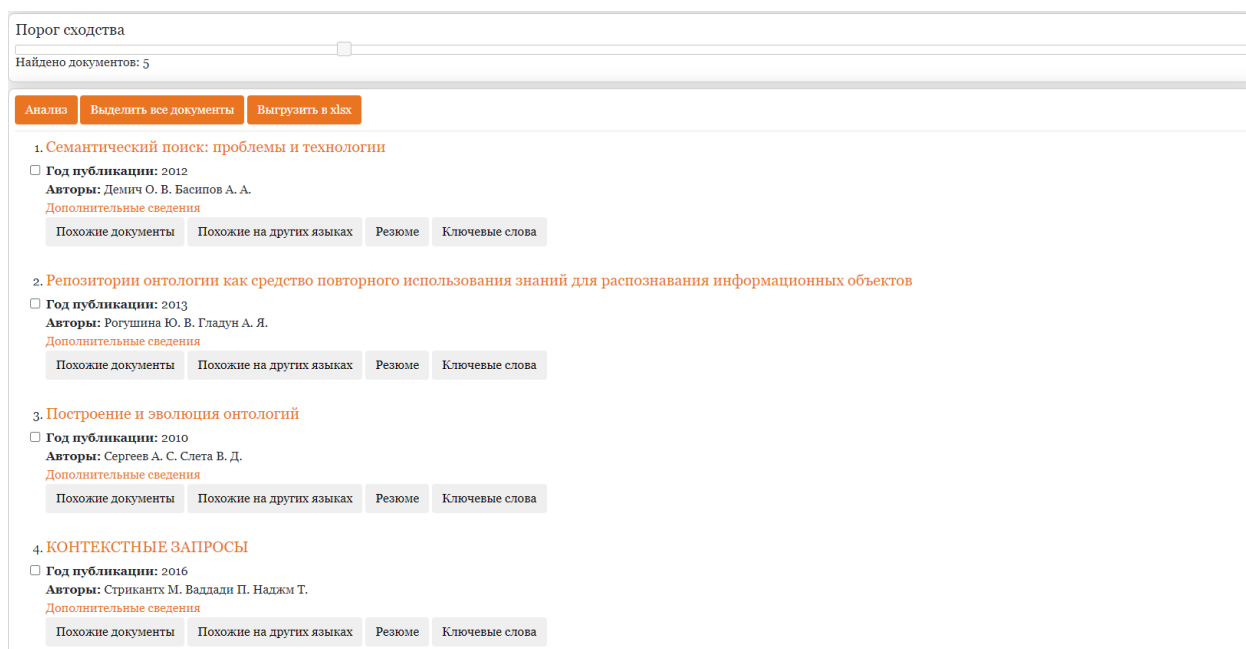


Рисунок 21 - Результаты поиска похожих документов

Информацию по работе с результатами поиска см. в разделе Работа с результатами поиска.

6.2. Ключевые слова.

При нажатии на кнопку «Ключевые слова» будет выполнен анализ документа с целью выделения из него ключевых слов, то есть слов, в той или иной мере отражающих смысл загруженного/введенного документа (Рисунок 22).

Для ключевых слов можно настроить минимальный вес, используя ползунок, расположенный под надписью «Ключевые слова документа». Также их можно отсортировать по алфавиту или весу, выбрав соответствующее в списке «Сортировка».

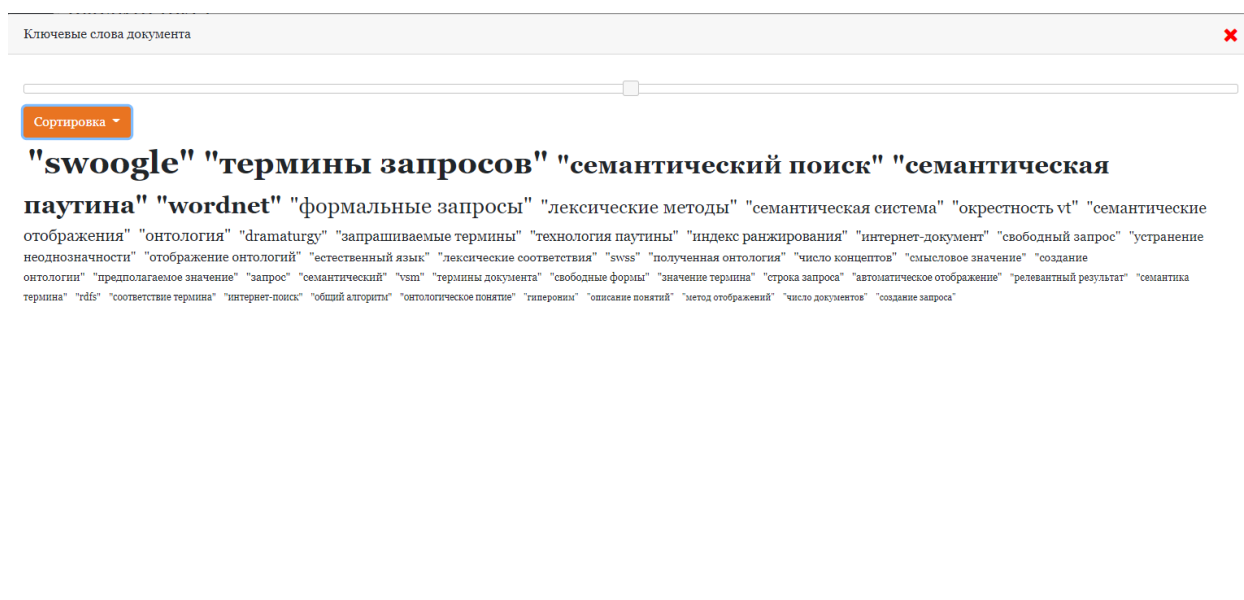


Рисунок 22 - Ключевые слова документа

7. Помощник автора

Данный инструмент предназначен для анализа степени научности загруженного/введенного документа по различным метрикам.

Для доступа к соответствующему разделу интерфейса пользователя необходимо перейти по ссылке «Помощник автора». Откроется страница анализа научных текстов (Рисунок 23).

Для начала анализа необходимо загрузить документ или ввести текст в поле «Введите текст». После этого для запуска анализа необходимо нажать кнопку «Анализировать».

Рисунок 23 - Помощник автора

После выполнения анализа (это может занять некоторое время, в зависимости от объема загруженного/введенного документа) на странице отобразятся следующие результаты анализа (Рисунок 24):

- Общая оценка научности документа;
- Оценка соответствия текста документа формальным требованиям;
- Количество речевых дефектов;
- Формулировки результатов;
- Формулировки определений.

Рисунок 24 - Результат анализа научного текста

Кроме того, отображается график распределения тематически похожих документов по журналам (Рисунок 25). Для просмотра документов соответствующего журнала нужно нажать на синий или оранжевый столбец напротив соответствующего журнала (Рисунок 26).

Информация о работе с результатами см. в разделе Работа с результатами поиска.



Рисунок 25 - Распределение тематически похожих документов по журналам

Аналитическая выборка Пользовательские подборки GUEST@gmail.ru

Количество документов на страницу: 10
Всего документов: 7

[Выделить все документы](#) [Выгрузить в xlsx](#)

- Проектирование лингвистических онтологий для информационных систем в широких предметных областях**
☐ Год публикации: 2015
 Авторы: Добров Б. В. Лукашевич Н. В.
[Дополнительные сведения](#)
[Похожие документы](#) [Похожие на других языках](#) [Резюме](#) [Ключевые слова](#)
- Технология генерации онтологии базы данных с помощью онтологического толкового словаря**
☐ Год публикации: 2012
 Авторы: Соболева Н. В. Медведев Д. В. Габдрахманов И. Н. Кучутанов В. Н. Мокроусов М. Н.
[Дополнительные сведения](#)
[Похожие документы](#) [Похожие на других языках](#) [Резюме](#) [Ключевые слова](#)
- Онтологии и персонализация профиля пользователя в рекомендующих системах третьего поколения**
☐ Год публикации: 2014
 Авторы: Тушканова О. Н. Городецкий В. И.
[Дополнительные сведения](#)
[Похожие документы](#) [Похожие на других языках](#) [Резюме](#) [Ключевые слова](#)
- Разработка онтологической модели информационной потребности пользователя при семантическом поиске**
☐ Год публикации: 2014
 Авторы: Рогущина Ю. В.
[Дополнительные сведения](#)
[Похожие документы](#) [Похожие на других языках](#) [Резюме](#) [Ключевые слова](#)

Рисунок 26 - Тематически схожие документы в журнале

8. Работа с результатами поиска

Случаи, в которых появляется возможность работы с результатами поиска описаны в разделах Поиск заимствований, Семантический поиск, Поиск похожих документов, Помощник автора.

Пример результата поиска представлен на Рисунок 8 - Результат проверки на заимствования. Кнопка «Выделить все документы» позволяет выделить найденные документы на странице для добавления в подборки (см. Пользовательские подборки). Также возможно выборочное добавление в пользовательские подборки, с помощью отметки необходимого документа флажком слева от названия документа (см. Пользовательские подборки). Также с помощью кнопки «Выгрузить xls» можно сохранить результаты запроса в формате Excel.

8.1. Анализ.

Функция позволяет построить различные графики найденных документов: кумулятивная масса документов (Рисунок 27), распределение документов по годам (Рисунок 28), распределение документов по авторам (Рисунок 29), распределение патентов по правообладателям (Рисунок 30), распределение документов по коллекциям (Рисунок 31), топ 5 государственной принадлежности правообладателей (Рисунок 32), распределение по классам МПК (Рисунок 33), Место защиты диссертации (Рисунок 34).

Все графики можно скачать, нажав кнопку «Скачать графики» (Рисунок 27).

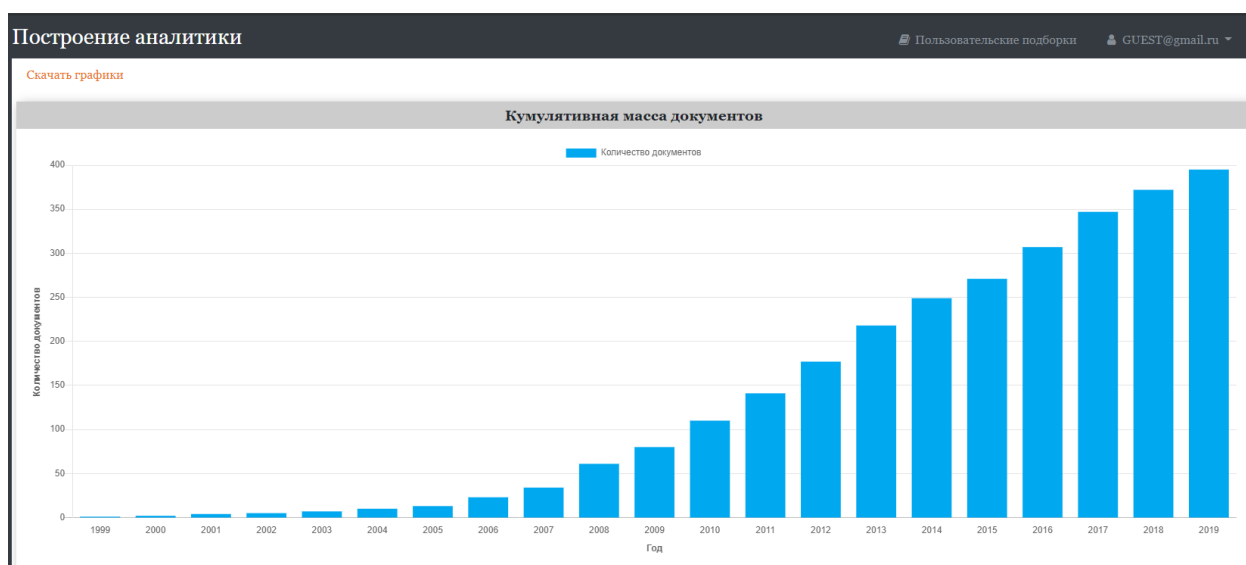


Рисунок 27 - Кумулятивная масса документов



Рисунок 28 - Распределение документов по годам

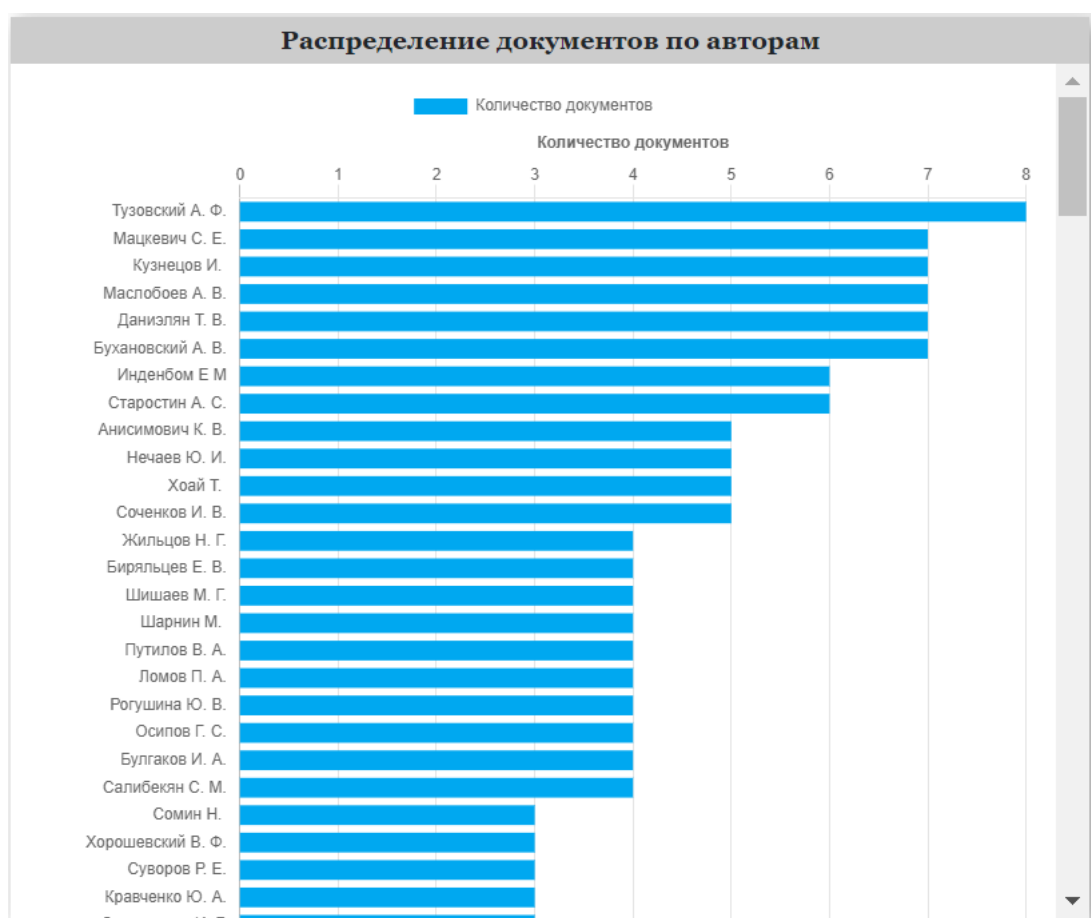


Рисунок 29 - Распределение документов по авторам



Рисунок 30 - Распределение патентов по правообладателям

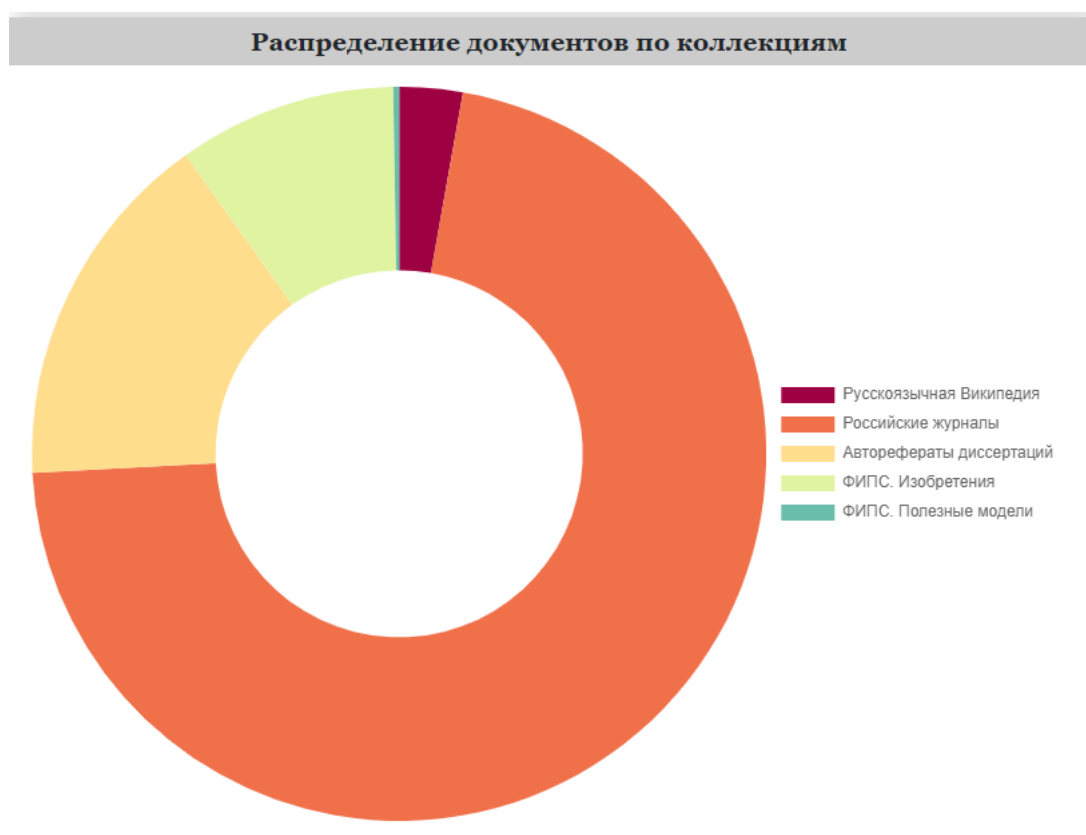


Рисунок 31 - Распределение документов по коллекциям

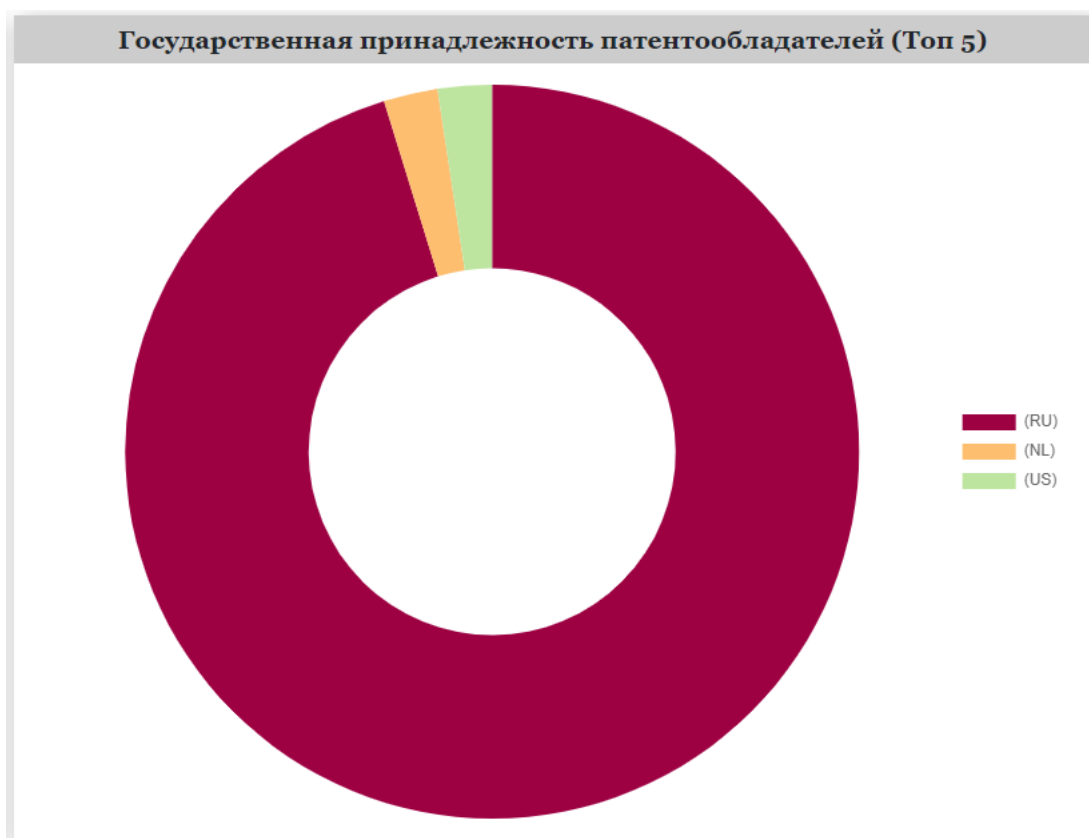


Рисунок 32 - Государственная принадлежность правообладателей (Топ 5)

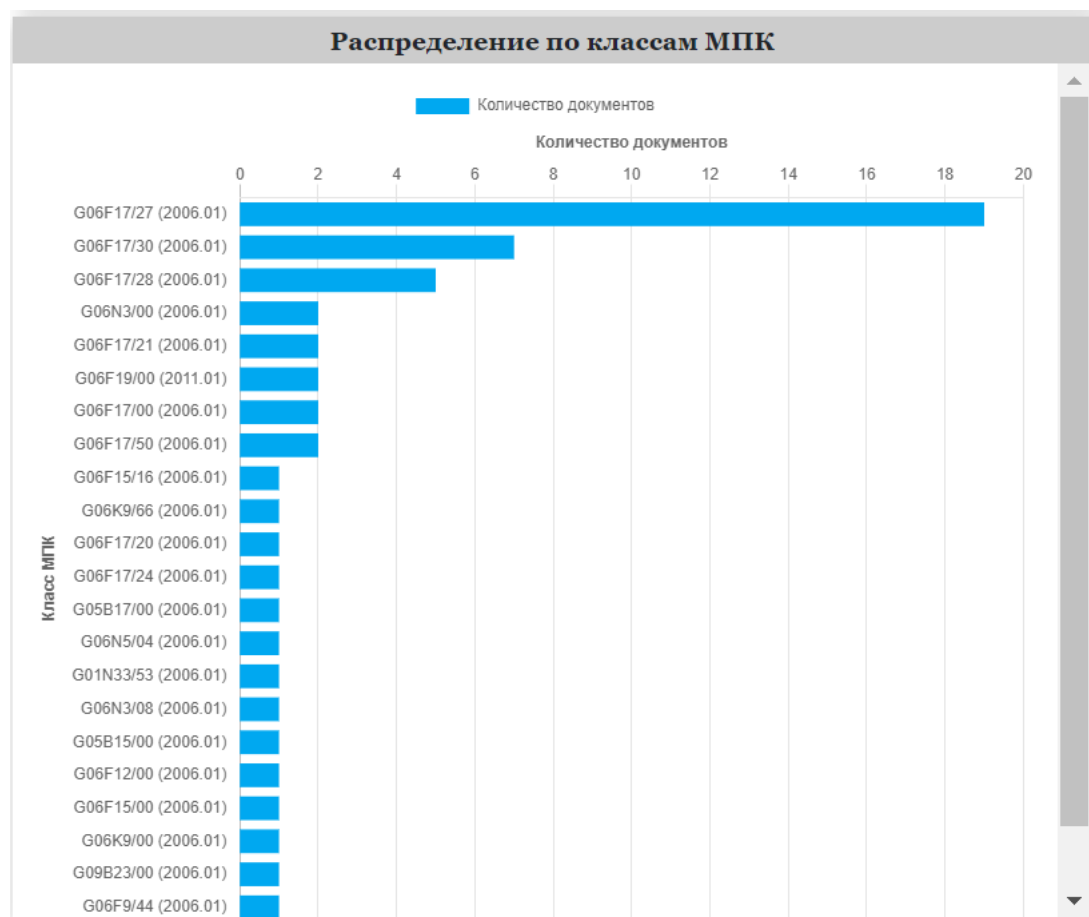


Рисунок 33 - Распределение по классам МПК

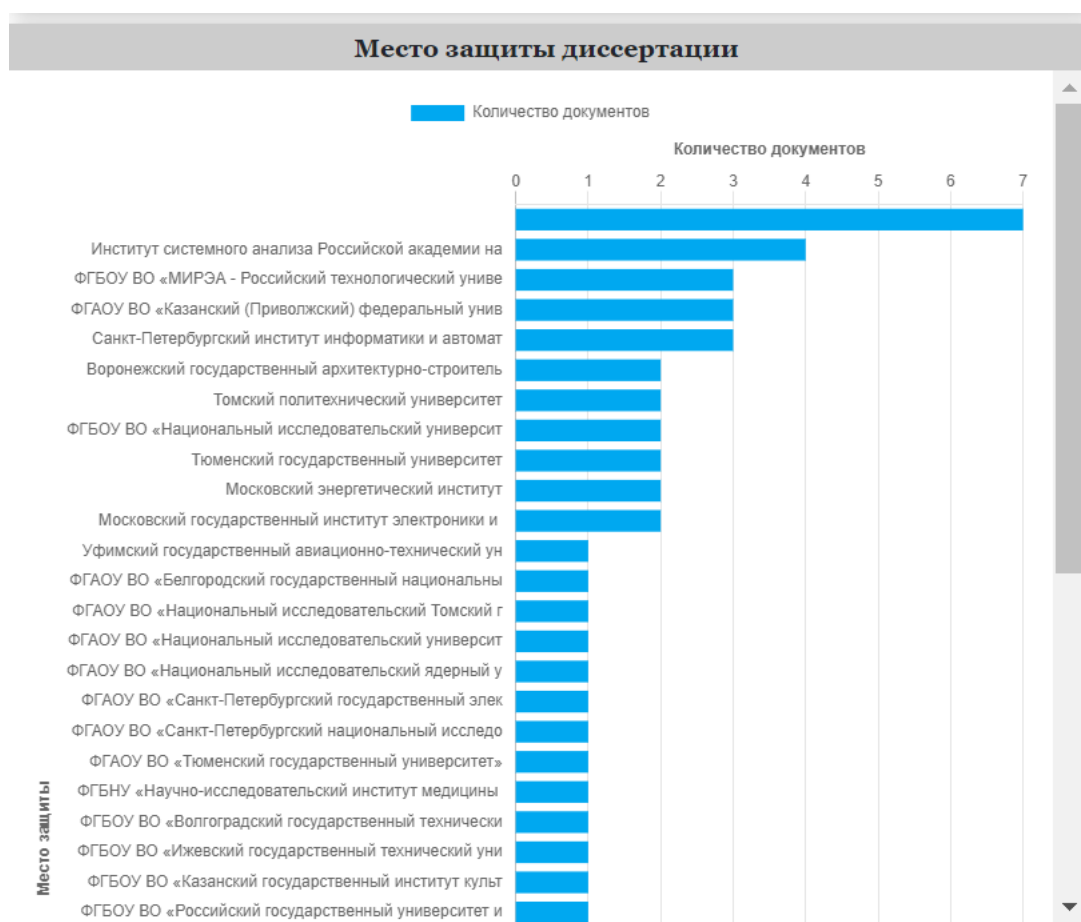


Рисунок 34 - Место защиты диссертации

8.2. Ключевые слова.

При нажатии на кнопку «Ключевые слова» будет выполнен анализ найденных документов с целью выделения из них общих ключевых слов, то есть слов, в той или иной мере отражающих смысл найденных документов (Рисунок 35).

Для ключевых слов можно настроить минимальный вес, используя ползунок, расположенный под надписью «Ключевые слова поисковой выдачи». Также их можно отсортировать по алфавиту или весу, выбрав соответствующее в списке «Сортировка».

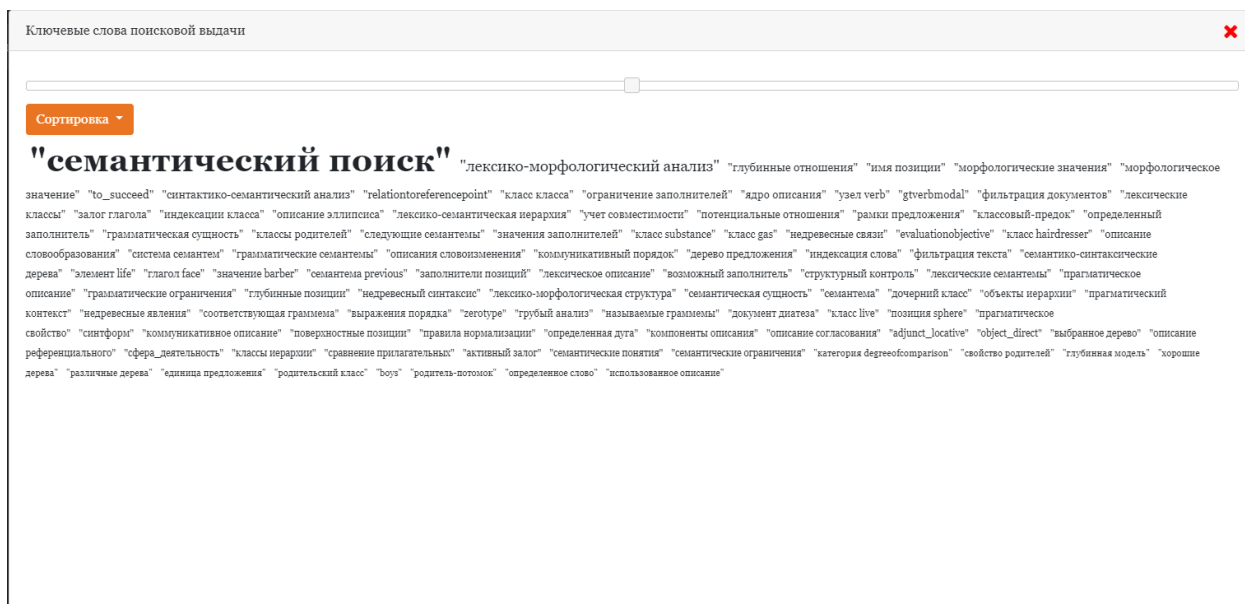


Рисунок 35 - Ключевые слова поисковой выдачи

9. Пользовательские подборки

Данный набор инструментов позволяет взаимодействовать с пользовательскими подборками документов, включая создание новых подборок, редактирование существующих подборок, добавление и удаление документов из подборок, сравнение существующих подборок. В подборки входят директории и коллекции.

9.1. Создание директории.

Для создания директории необходимо перейти в подраздел «Пользовательские подборки» в правом верхнем углу главной страницы (Рисунок 36).

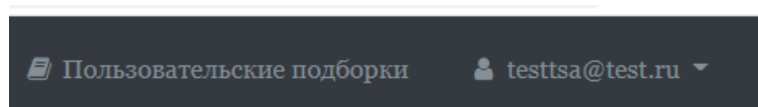


Рисунок 36 - Пользовательские подборки

Откроется окно просмотра ваших подборок (Рисунок 37). Перейдите в раздел "Мои подборки".

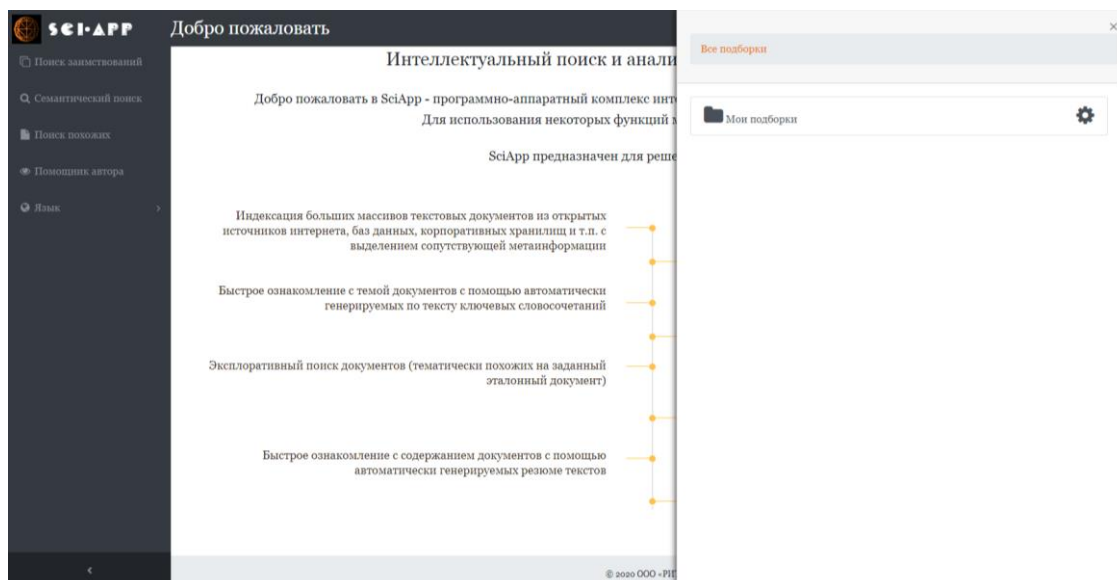


Рисунок 37 - Просмотр подборок

Далее необходимо нажать на кнопку «Добавить» и нажать «Директория». После этого в появившемся поле (Рисунок 38) необходимо ввести название создаваемой директории и нажать на кнопку «Подтвердить». После выполнения этих действий новая директория появится в списке подборок на странице просмотра подборок (Рисунок 39).

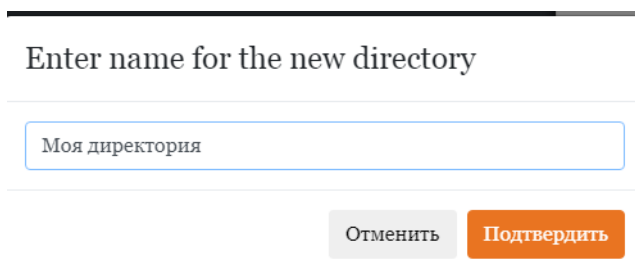


Рисунок 38 - Добавление директории

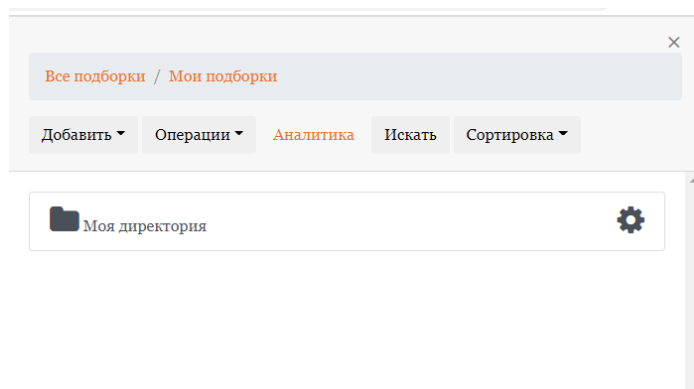


Рисунок 39 - Добавлена новая директория

9.2. Создание коллекции

Последовательность действий аналогична созданию директории, однако после нажатия кнопки "Добавить" необходимо выбрать "Коллекцию". Далее также вводим название и нажимаем "Подтвердить" (Рисунок 40).

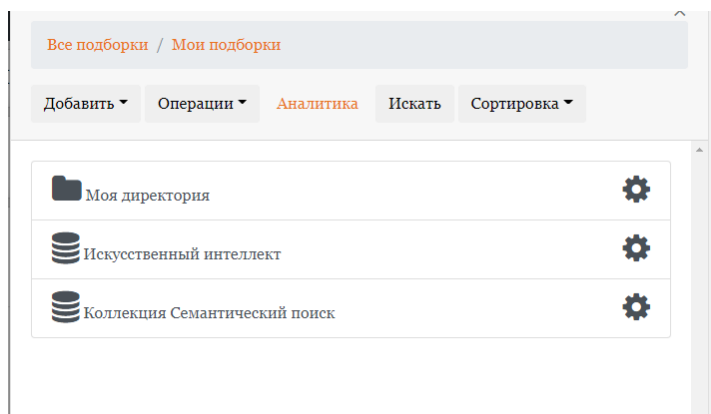


Рисунок 40 - Добавлены коллекции

9.3. Действия с подборкой

При нажатии на шестеренку показаны действия, доступные для выполнения с отдельной подборкой (Рисунок 41).

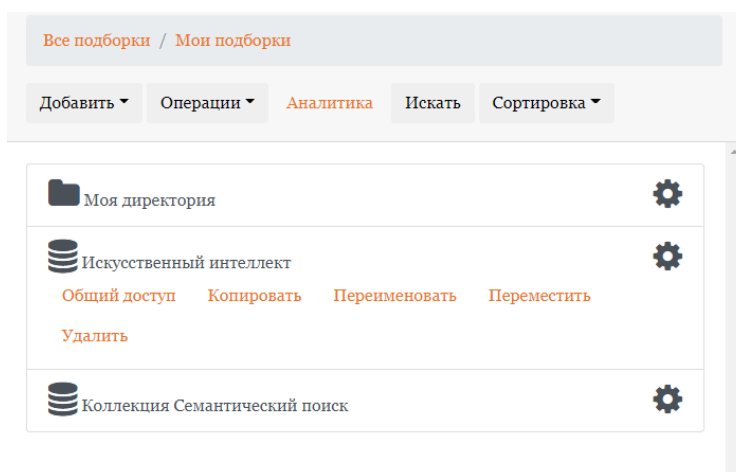


Рисунок 41 - Действия с подборкой

9.3.1. Общий доступ

Для предоставления доступа к вашей подборке нажмите "Общий доступ"(Рисунок 42), далее выберите "Добавить" и в появившемся окне введите email пользователя, которому вы хотите открыть доступ к подборке, нажмите "Подтвердить" (Рисунок 43).

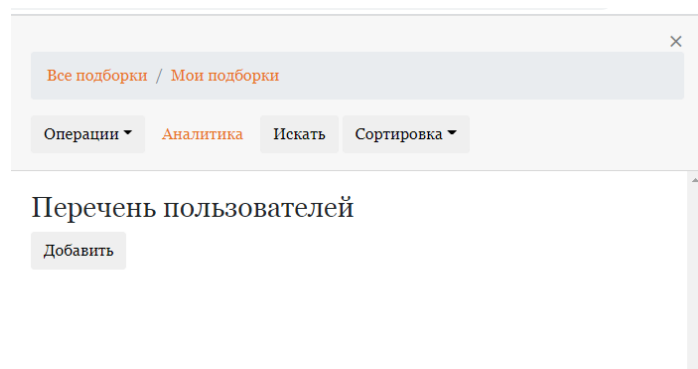


Рисунок 42 - Общий доступ

Enter user's email

Отменить Подтвердить

Рисунок 43 - Предоставление общего доступа

9.3.2. Копировать

При выборе "Копировать" в левой части странице появится окно, в котором необходимо выбрать место, в которую вы хотите поместить копию выбранной подборки (Рисунок 44) и имя копии. Нажмите "Подтвердить" и копия будет создана (Рисунок 45).

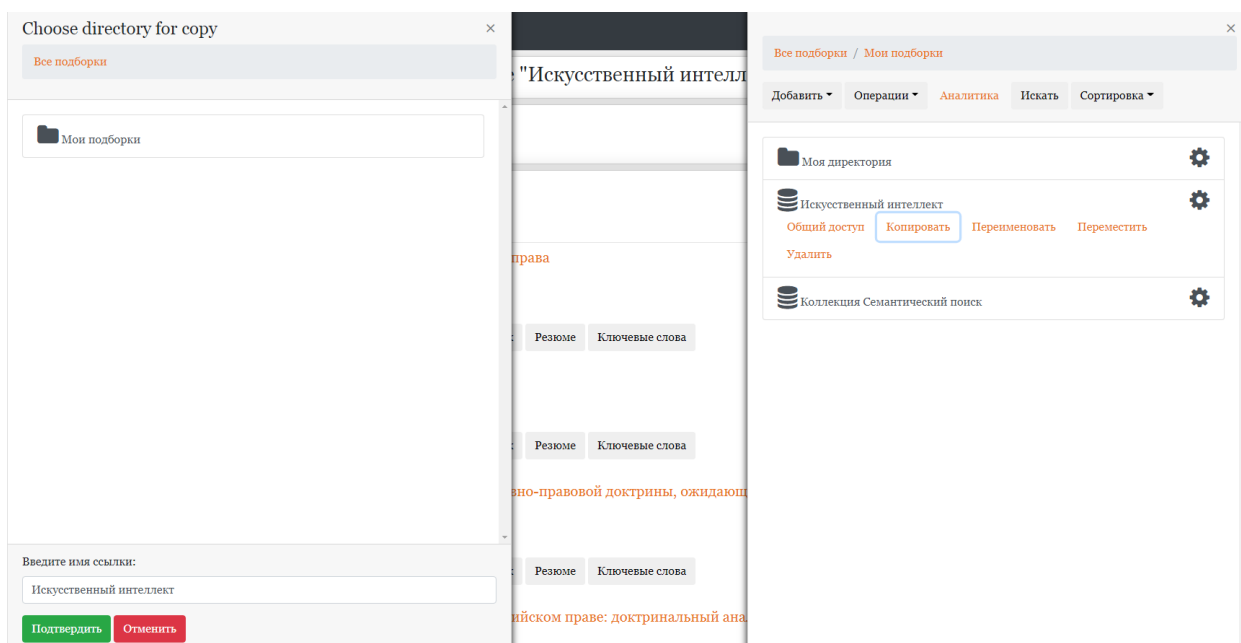


Рисунок 44 - Копирование подборки

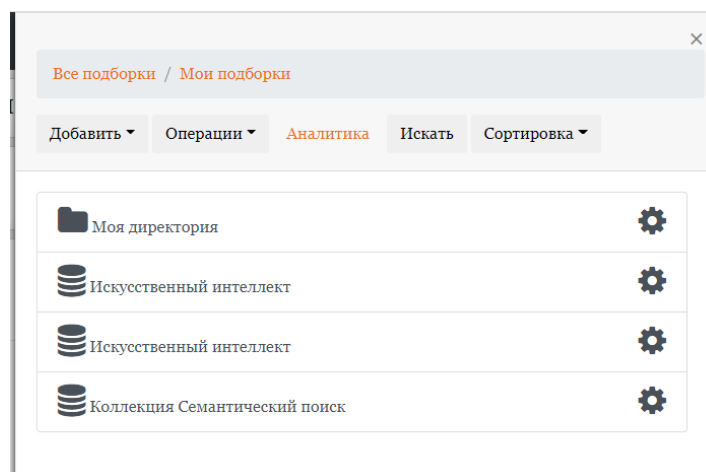


Рисунок 45 - Копия создана

9.3.3.Переименовать

Чтобы переименовать подборку нажмите "Переименовать" и в появившемся окне введите новое название подборки (Рисунок 46).

Enter new name:

Рисунок 46 - Переименование

9.3.4.Переместить

Для перемещения подборки нажмите "Переместить", после чего окно появится в левой части страницы, в котором необходимо указать место, куда вы хотите переместить подборку, далее нажмите "Подтвердить" (Рисунок 47).

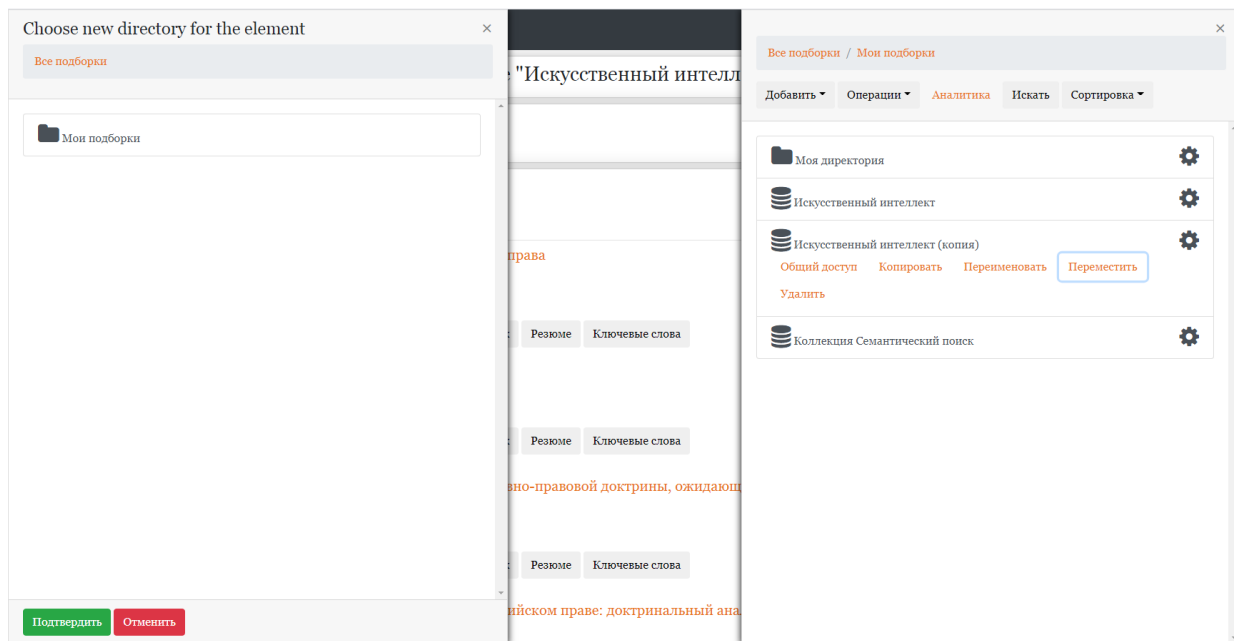


Рисунок 47 - Перемещение подборки

9.3.5. Удалить

Для удаления подборки выберите "Удалить" и в окне нажмите "Подтвердить" (Рисунок 48).

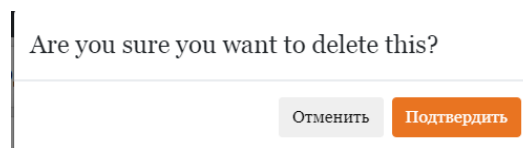


Рисунок 48 - Удаление подборки

9.3.6. Добавить документы в коллекцию

Для добавления документов в коллекцию выберите необходимые документы из найденных результатов (см. Работа с результатами поиска). Далее нажмите "Пользовательские подборки" (Рисунок 36), выберите необходимую коллекцию (Рисунок 49). "Добавить выбранное" позволяет добавить выбранные документы в коллекцию, "Добавить весь запрос" добавляет все найденные по запросу документы.

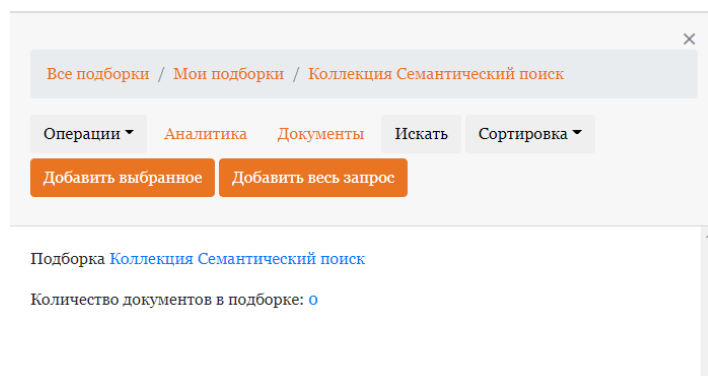


Рисунок 49 - Добавление документов в коллекцию
Документы добавлены в коллекцию (Рисунок 50).

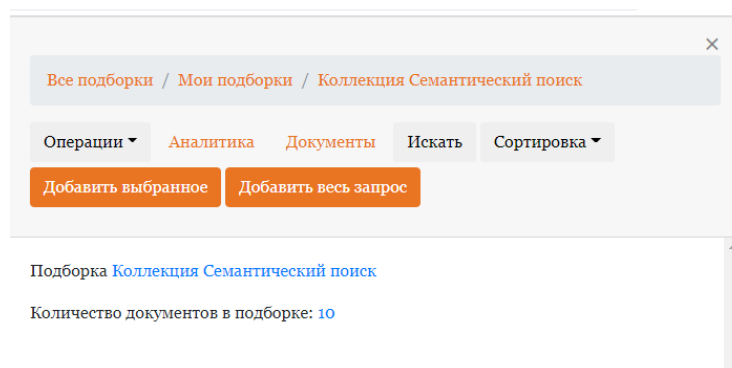


Рисунок 50 - Документы добавлены в коллекцию

9.3.7. Просмотр

Для просмотра содержимого директории или коллекции достаточно нажать на ее название.

9.4. Операции

Система позволяет выполнять различные операции с коллекциями документов. Для просмотра списка операций выберите "Операции" (Рисунок 51).

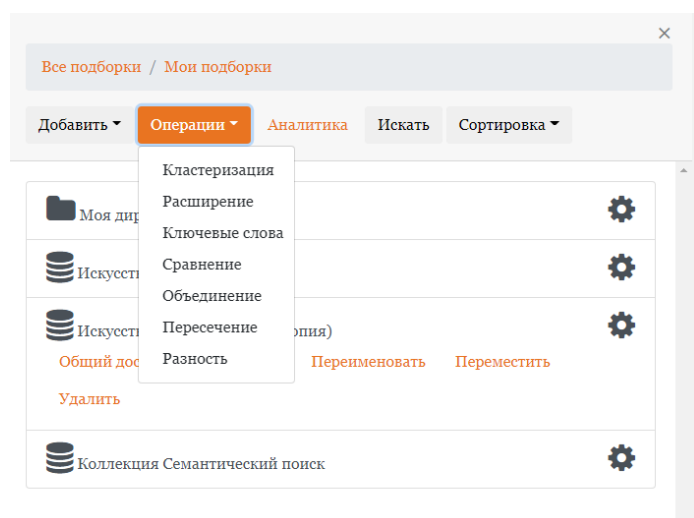


Рисунок 51 - Операции

9.4.1.Кластеризация

Кластеризация возможно как в рамках отдельной коллекции, так и в рамках директории. Для кластеризации выберите "Кластеризация" (Рисунок 52). В появившемся окне будет показано количество выделенных кластеров, количество документов в кластере, ключевые слова кластера. При нажатии "View documents" можно просмотреть документы кластера.

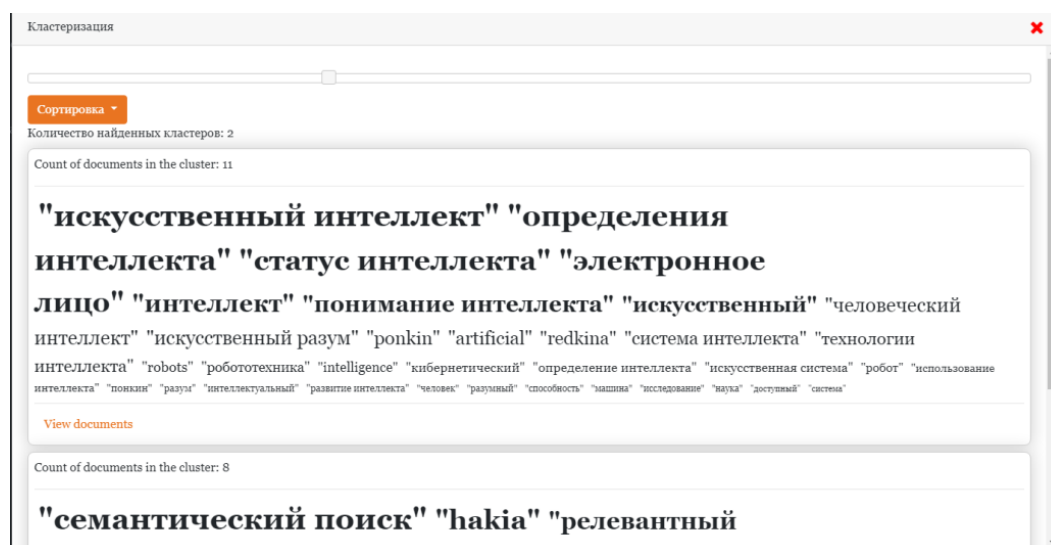


Рисунок 52 - Кластеризация

9.4.2.Расширение

Расширение подборки позволяет найти похожие документы для документов, находящихся в подборке (подробнее см. Поиск похожих документов).

9.4.3.Ключевые слова

Для выделения ключевых слов нажмите "Ключевые слова"(Рисунок 53).

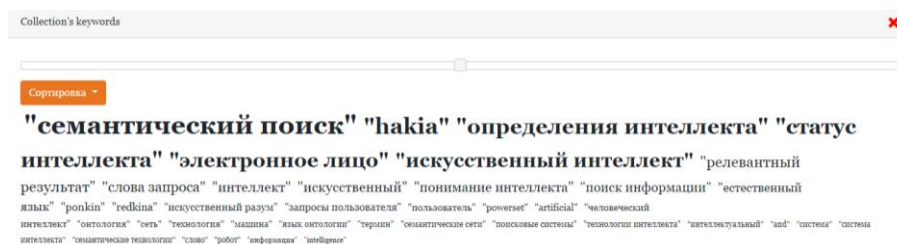


Рисунок 53 - Ключевые слова подборки

9.4.4. Сравнение

Операция "Сравнение" позволяет сравнивать подборки документов. Для этого выберите необходимую коллекцию, в списке "Операции" нажмите "Сравнение". В появившемся слева окне выберите подборку, с которой хотите провести сравнение и нажмите "Подтвердить" (Рисунок 54).

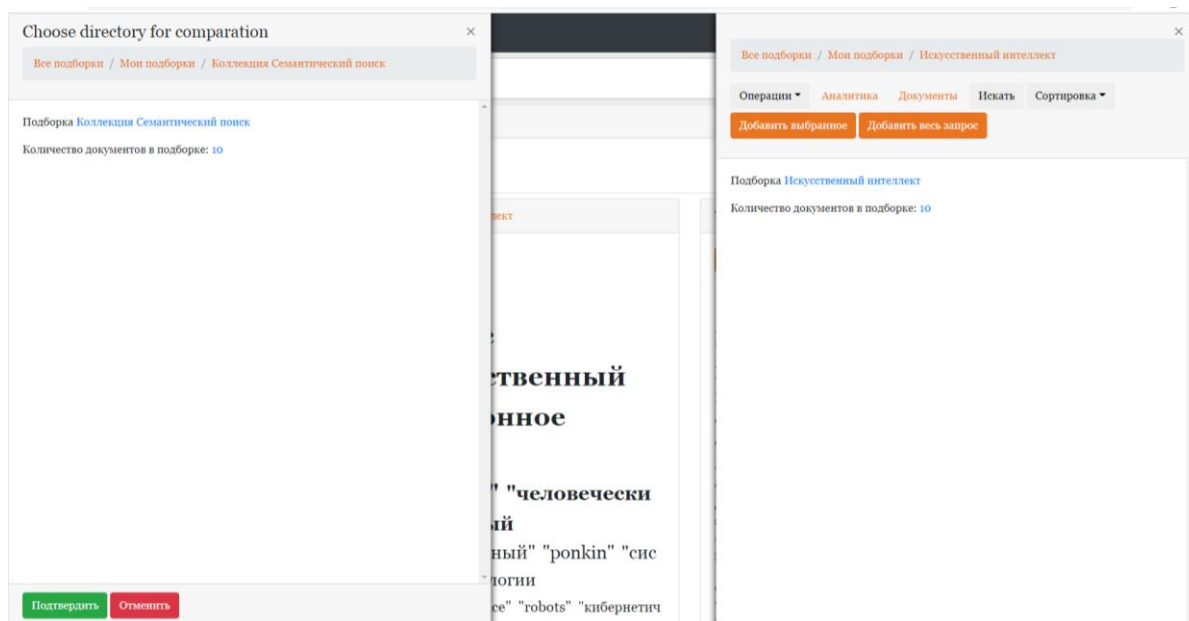


Рисунок 54 - Сравнение подборок

Для выбранных подборок будет определено тематическое сходство в процентах, выделены общие ключевые слова, а также уникальные для коллекции слова (Рисунок 55).

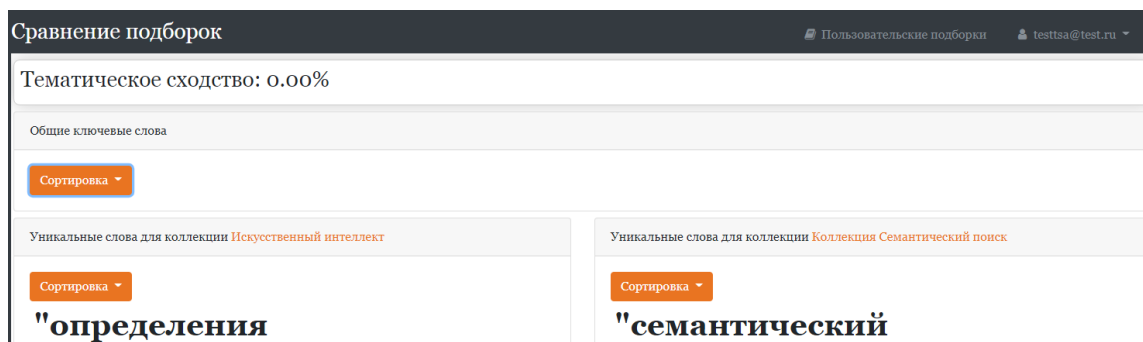


Рисунок 55 - Результат сравнения

9.4.5. Объединение

Операция "Объединение" позволяет вывести объединение документов двух коллекций. Для этого выберите необходимую коллекцию, в списке "Операции" нажмите "Объединение". В появившемся слева окне выберите подборку, с которой хотите провести объединение и нажмите "Подтвердить" (Рисунок 56). В итоге будет выведен список документов, содержащихся или в первой коллекции, или во второй (Рисунок 57). Эти документы можно добавить в новую коллекцию или выгрузить в формате Excel (см. Работа с результатами поиска).

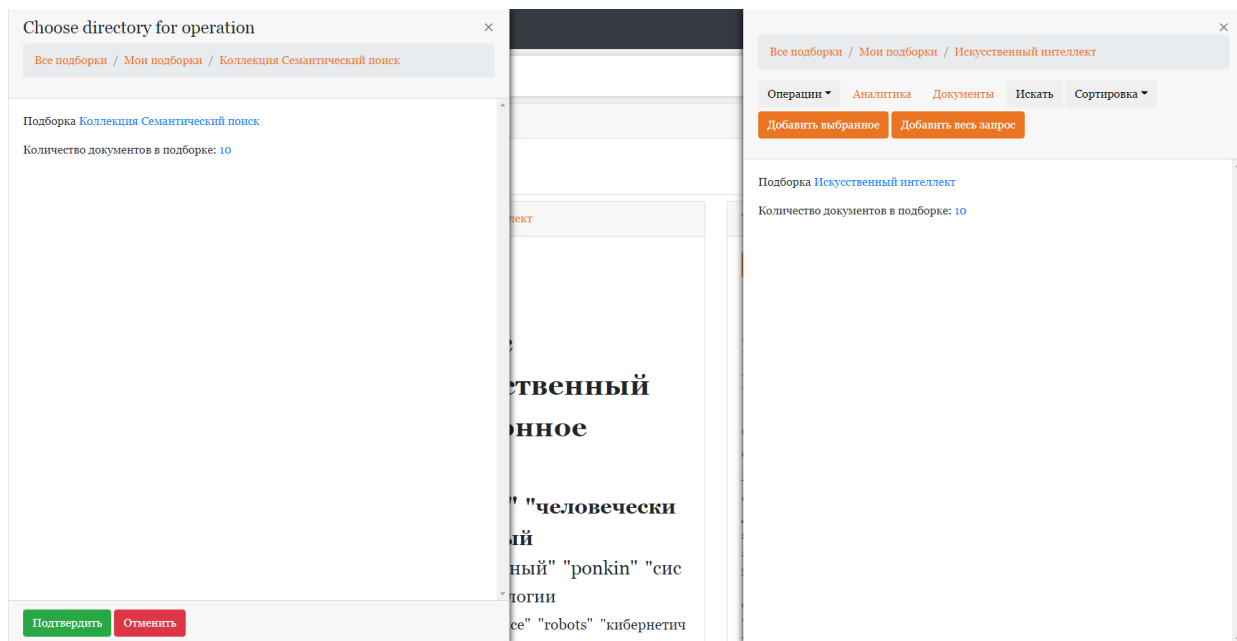


Рисунок 56 - Объединение

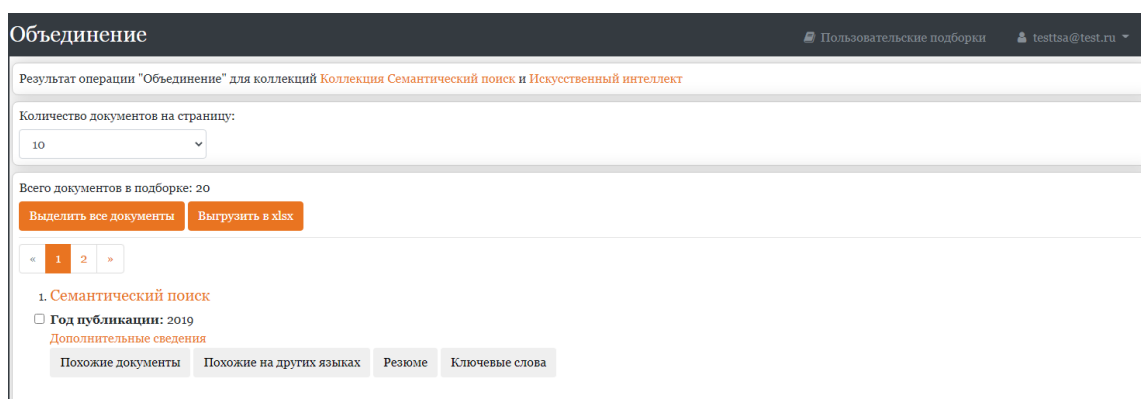


Рисунок 57 - Результат объединения

9.4.6. Пересечение

Операция "Пересечение" позволяет вывести пересечение двух коллекций, то есть документы общие для выбранных двух коллекций. Для этого выберите необходимую коллекцию, в списке "Операции" нажмите "Пересечение". В появившемся слева окне выберите подборку, с которой хотите провести пересечение и нажмите "Подтвердить" (Рисунок 58). В итоге будет выведен список документов, содержащихся в первой коллекции и во второй (Рисунок 59). Эти документы можно добавить в новую коллекцию или выгрузить в формате Excel (см. Работа с результатами поиска).

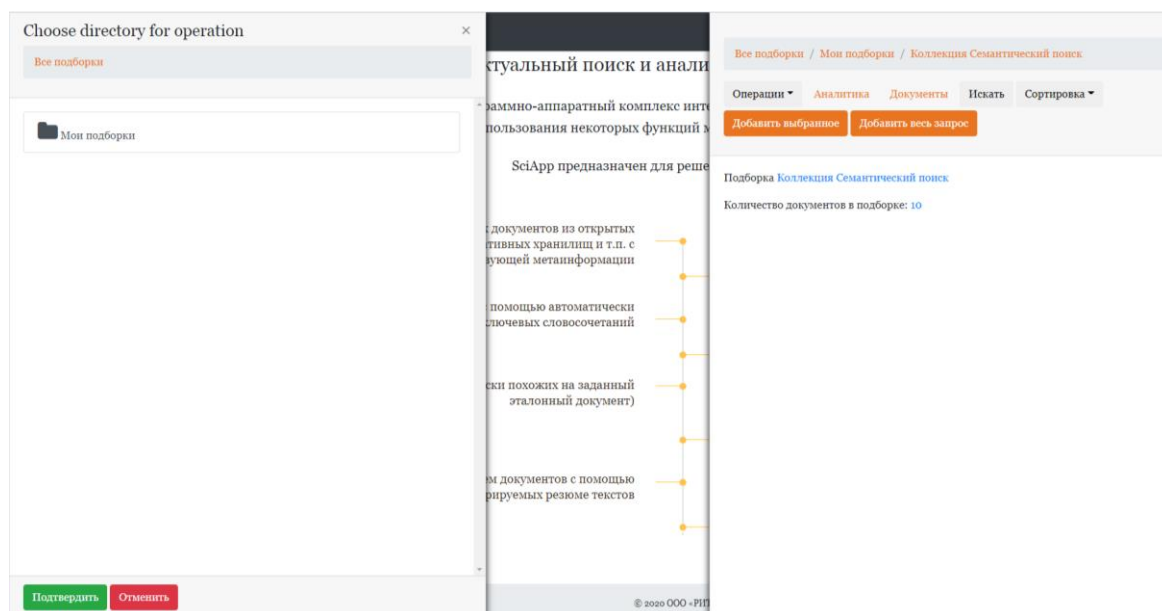


Рисунок 58 - Пересечение

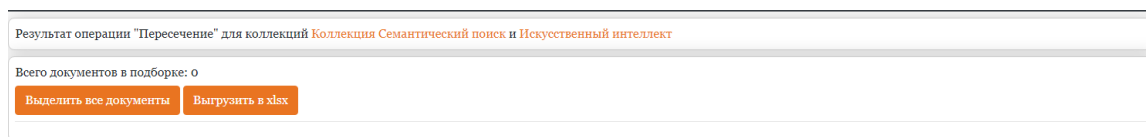


Рисунок 59 - Результат пересечения

9.4.7.Разность

Операция "Разность" позволяет вывести разность двух коллекций, то есть документы, которые содержатся в первой коллекции, но не содержатся во второй. Для этого выберите необходимую коллекцию, в списке "Операции" нажмите "Разность". В появившемся слева окне выберите подборку, с которой хотите провести разность и нажмите "Подтвердить" (Рисунок 60). В итоге будет выведен список документов, содержащихся в первой коллекции, но не содержащихся и во второй (Рисунок 61). Эти документы можно добавить в новую коллекцию или выгрузить в формате Excel (см. Работа с результатами поиска).

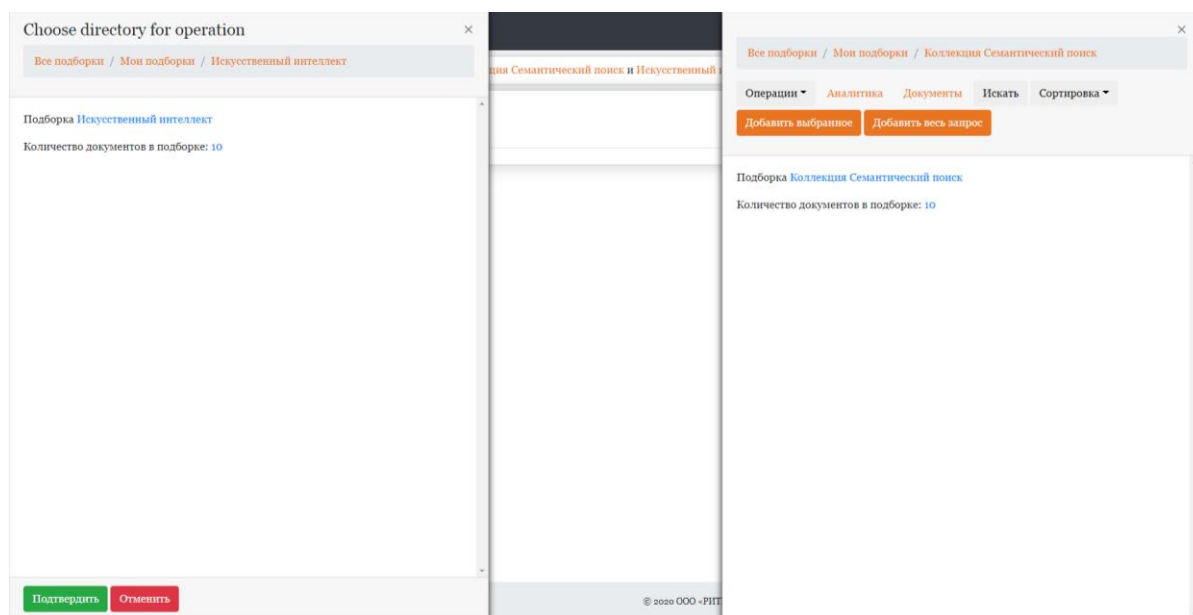


Рисунок 60 - Разность

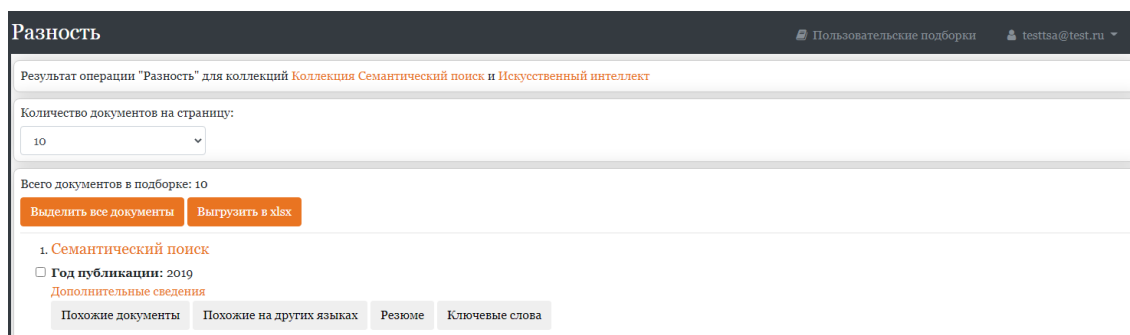


Рисунок 61 - Результат разности

9.5. Аналитика

Для пользовательских подборок возможно проведение аналитики, то есть построения различных графиков, рассмотренных в Работа с результатами поиска.

9.6. Искать

Эта функция позволяет осуществлять семантический поиск в рамках пользовательских коллекций (см. Семантический поиск). Результатов выполнения будет список документов, работа с которым рассмотрена в разделе Работа с результатами поиска.

9.7. Просмотр документов коллекции

Для просмотра документов, содержащихся в коллекции, перейдите в коллекцию и выберите "Документы" (Рисунок 62). Результатом будет список документов, содержащихся в коллекции (Рисунок 63). Помимо стандартных возможностей, описанных в разделе Работа с результатами поиска, присутствует возможность удаления документа из подборки. Для этого отметьте документы галочкой слева от названия или нажмите "выбрать все документы" для выделения всех на странице, появится кнопка "Удалить выбранные документ" (Рисунок 64). После нажатия на него появится окно (Рисунок 65), в котором для удаления необходимо выбрать "Подтвердить".

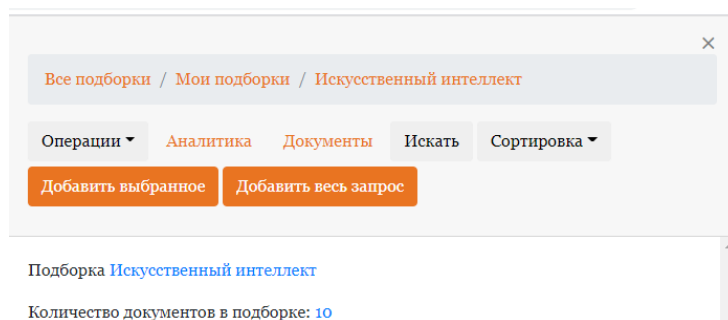


Рисунок 62 - Информация о коллекции

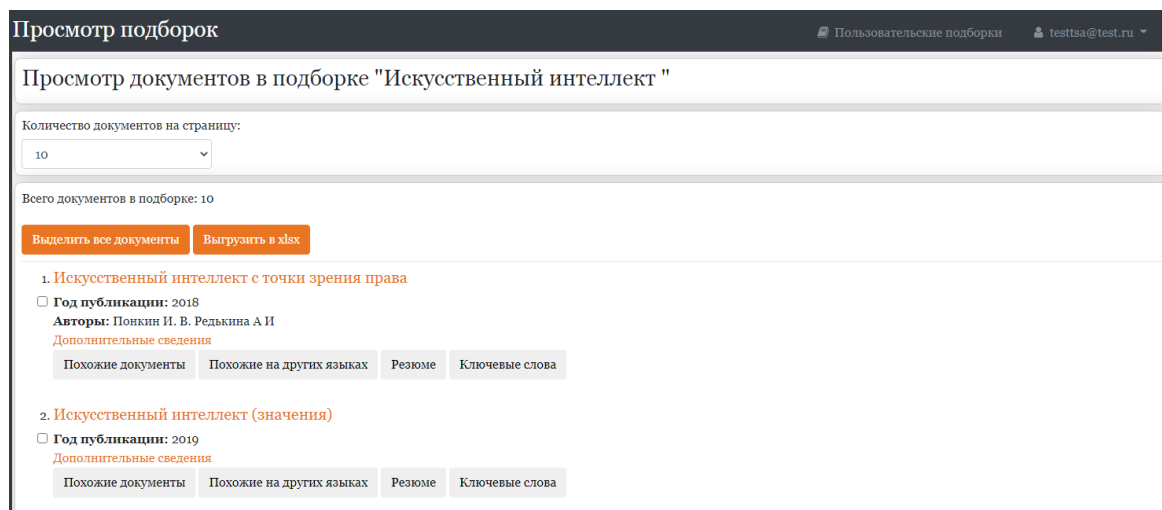


Рисунок 63 - Список документов

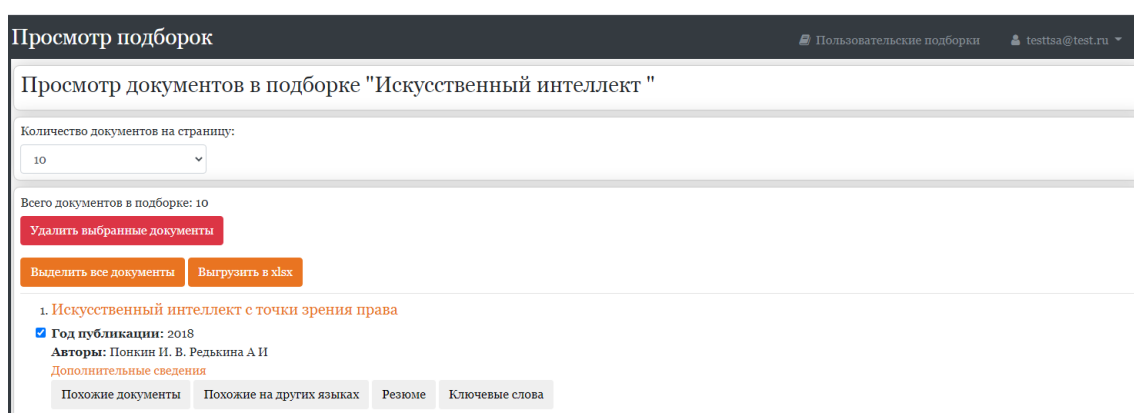


Рисунок 64 - Удаление документа из подборки

Уверены, что хотите удалить?

Отменить

Подтвердить

Рисунок 65 - Окно удаления

10. Работа с документом

Данный раздел посвящен функциональным возможностям для работы с отдельным документом из аналитической выборки или пользовательской подборки. Для каждого найденного документа (Рисунок 66) указано его название, год публикации, авторы, а также сниппет документа, содержащий слова введенного запроса.

Также для каждого документа возможно просмотреть его текст, нажав на название документа, воспользоваться поиском похожих на него документов (см. Поиск похожих документов), похожих документов на других языках (см. Поиск похожих документов),

просмотреть автоматически сгенерированное резюме документа (Рисунок 67), а также просмотреть ключевые слова документа (Рисунок 68).

1. Семантический поиск

☐ **Семантический поиск** Семантический поиск — способ и технология поиска информации, основанная на использовании контекстного (смыслового) значения запрашиваемых фраз, вместо словарных значений отдельных слов или выражений при поисковом запросе. <...> Для осуществления **семантического поиска** в Сети (или в каких-либо системах с ограниченным доступом пользователей) [Читать полностью](#)
Год публикации: 2019
[Дополнительные сведения](#)

[Похожие документы](#) [Похожие на других языках](#) [Резюме](#) [Ключевые слова](#)

Рисунок 66 - Найденный документ

Резюме документа

Семантический поиск Семантический поиск — способ и технология **поиска** информации, основанная на использовании контекстного (смыслового) **значения** запрашиваемых **фраз**, вместо словарных значений отдельных слов или выражений при поисковом запросе. <...> Улучшение **результатов поиска** при обработке запросов достигается за счет более точной интерпретации поисковых намерений пользователя. <...> При семантическом **поиске** учитывается информационный контекст, местонахождение и **цель поиска** пользователя, словесные **вариации**, **синонимы**, обобщенные и специализированные запросы, язык запроса, а также другие особенности, позволяющие получить **соответствующий результат**. <...> Ряд крупных поисковых систем, таких как Google и Bing, используют некоторые **элементы** семантического **поиска**, не являясь таковыми в чистом виде. <...> **Цель** семантического **поиска** - определять особенности пользователя и предоставлять ему наиболее **релевантные результаты**. <...> **Семантический поиск** возник из **семантической** сети, которая строится на онтологиях. <...> Почти все, что связано с запросами или сайтами, может считаться частью семантической области, имеющей отношение к **результатам поиска**. <...> **Семантический поиск** зависит как от **семантической** разметки веб-сайтов, так и от огромного количества **семантической** информации, которое она за собой влечет. <...> **Умные** поисковые **системы** учитывают целый ряд факторов для того, чтобы получить наиболее соответствующие и подходящие поисковые запросы, включая: 1. <...> Текущие **тенденции** Если выборы президента только что закончились, и кто-то ищет информацию: «Кто стал новым президентом?», семантический поиск должен понять **запрос** и дать соответствующие результаты, основанные на текущих **тенденциях** и новостях. <...> Местонахождение пользователя Если человек ввел **запрос** «Какая сейчас температура?», **семантическая** поисковая **система** должна предоставить результаты, основанные на месте его нахождения в момент **запроса**. <...> Например, для ростовского жителя результатом **запроса** будут данные о температуре воздуха в городе Ростове-на-Дону. <...> **Цель поиска** Семантический **поиск** основывается на выдаче подходящих результатов, основанных на **цели поиска** пользователя, а не определенных слов, использованных при обращении к поисковой системе. <...> **Вариации слов** в семантическом поиске Семантический поиск должен учитывать лингвистические особенности запроса (падеж, число и время) и предлагать подходящие результаты для всех семантических **вариаций слов**, то есть видеть различия между близкими по написанию **словами**. <...> **Синонимы** Семантические поисковые **системы** должны понимать **синонимы** и давать более или менее похожие результаты на любые синонимичные слова, запрашиваемые пользователем. <...> Например, при **запросах** «наивысший пик» или «наибольший пик» ответ должен быть одинаковым. <...> **Общие** и специализированные <...>

Рисунок 67 - Резюме документа

Ключевые слова документа

Сортировка ▾

"семантический поиск" "семантическая система" "поисковые системы" "цель поиска" "семантический" "вариации слова" "поисковый" "общие запросы" "релевантный результат" "умная система" "соответствующие результаты" "семантический веб" "тематические базы" "запрос" "значение фраз" "результаты поиска" "элементы поиска" "концепция запроса" "обработка синонимов" "тенденция запроса" "семантическая оптимизация" "hakiа" "обработка вариаций" "количество метаданных" "текущие тенденции" "атрибуты поиска" "freebase" "технология поиска" "осуществление поиска" "проблема движения" "простой язык" "подходящие результаты" "плат задач" "поисковые запросы" "powerset" "компания афк" "морфологические вариации" "подходящий результат" "эволюция поиска" "семантическая аннотация" "значение запросов" "черта поисков" "поиск google" "вычислительный путь" "число решений" "статистические частоты" "суть поиска" "консорциум паутины" "местонахождение пользователя" "различные сектора" "тематический запрос" "работа поиска" "интерпретации намерений" "сверхсложная задача" "особенности запросов" "строительство эстакады" "поиск" "текстовый поиск" "вычисление задачи" "частота запросов" "google" "пользователь" "слова запроса" "семантические технологии" "шiroвая точка" "нужный ответ" "информационный контекст" "найденные результаты" "дорожное движение" "mq1" "указ дороги" "аггрегированная информация" "категории задач" "веб" "байты данных" "контекстная информация" "особенности пользователя" "новые чертания" "точность поиска" "теория поиска" "упоминание слова"

Рисунок 68 - Ключевые слова документа