

Введение. Для оценивания результативности научной деятельности все чаще используют наукометрические показатели. Эти показатели основаны на количестве публикаций автора и на количестве ссылок на его работы. Возросший интерес к наукометрическим показателям вызван в первую очередь возможностью автоматизации процесса оценивания.

Информационная система (ИС) — это совокупность взаимосвязанных компонентов, которые работают вместе для сбора, обработки, хранения и передачи информации. ИС необходимы для повышения эффективности организации.

Создание информационной системы для генерации отчетов о публикациях профессорско-преподавательского состава позволит обеспечить хранение и обработку информации о публикациях сотрудников кафедры, автоматизацию расчета наукометрических показателей, наглядное отображение данных для дальнейшего анализа и выгрузку данных в формате Excel для удобства составления отчетов о публикационной активности преподавателя.

Целью магистерской работы является разработка информационной системы для генерации отчетов о публикациях ППС, которая позволит в удобной форме хранить данные о публикационной активности преподавателей и проводить анализ этих данных.

Для достижения указанной цели были поставлены следующие задачи:

- Обзор наукометрических показателей публикационной активности.
- Обзор программных средств для разработки информационной системы.
- Проектирование UML-модели системы с помощью UML-диаграмм.
- Проектирование структуры данных базы данных.
- Получение данных с портала Google Scholar.
- Реализация информационной системы для генерации отчетов о публикациях ППС средствами Python и Django.
- Анализ наукометрических показателей по результатам реализации информационной системы.
- Генерация отчетов в формате Excel.

Наукометрические показатели публикационной активности. Для оценивания результативности научной деятельности используют наукометрические показатели. Эти показатели основаны на количестве публикаций

автора и количестве ссылок на его работы. Основными наукометрическими показателями являются индекс цитирования и индекс Хирша.

Индекс цитирования — это суммарное количество ссылок в научных публикациях на работы автора.

Индекс Хирша (h-индекс) — это максимальное целое число h , указывающее, что автор опубликовал h статей, каждая из которых процитирована хотя бы h раз. Эти h статей составляют ядро Хирша.

Проектирование информационной системы. Информационная система хранит информацию о публикациях сотрудников кафедры и дает возможность загружать и обновлять данные в автоматическом и ручном режимах.

Диаграмма прецедентов позволяет визуализировать различные типы ролей в системе и то, как эти роли взаимодействуют с системой. В соответствии с рисунком 1 представлена диаграмма прецедентов.

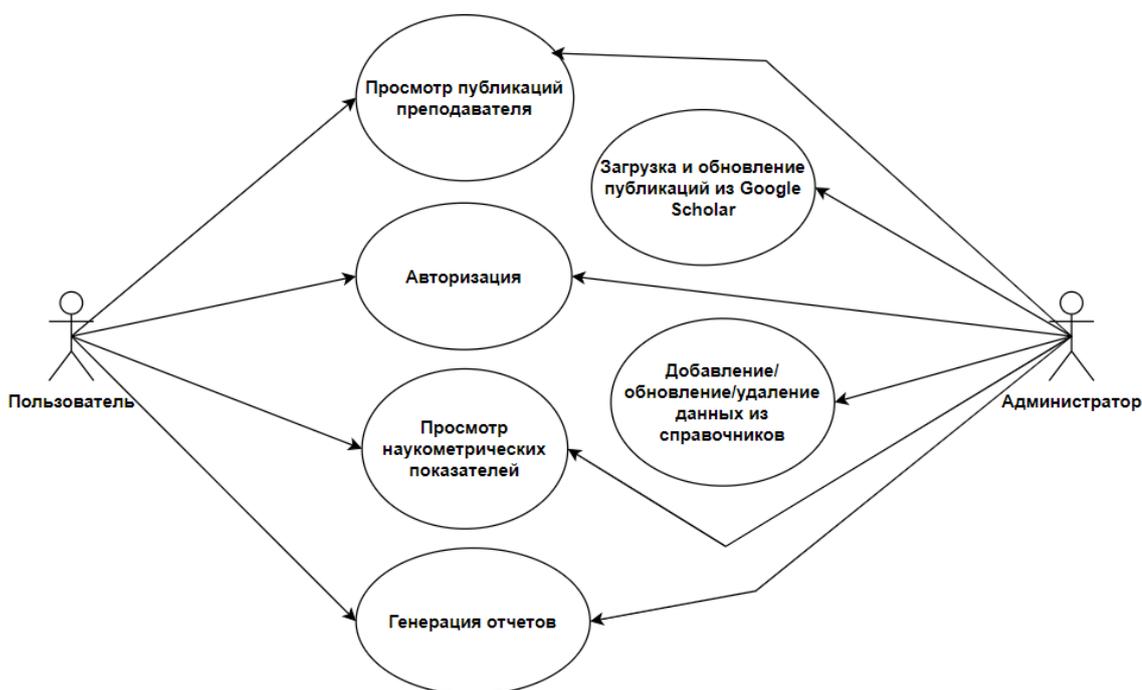


Рисунок 1 — Диаграмма прецедентов

Была построена диаграмма последовательности для сценария «Загрузка и обновление публикаций». Сценарий «Загрузка и обновление публикаций» предполагает следующие шаги:

- Пользователь выбирает необходимого преподавателя на странице веб-сайта и переходит в профиль.
- Сервер отправляет запрос в БД.
- БД в случае успешного взаимодействия отправляет необходимые данные на сервер.
- На странице преподавателя отображается список публикаций и рассчитываются наукометрические показатели.
- При нажатии кнопки «Обновить публикации» происходит обновление данных в БД из Google Scholar.
- При нажатии на кнопку «Генерация отчета» генерируется отчет в виде книги Excel.

Была построена диаграмма последовательности для сценария «Обновление данных в справочниках». Сценарий «Обновление данных в справочниках» предполагает следующие шаги:

- Пользователь переходит на страницу добавления или обновления данных.
- Пользователь вводит данные в форму.
- При корректном заполнении сервер формирует запрос на добавление или обновление данных в БД.

Была построена диаграмма последовательности для сценария «Удаление данных справочников». Сценарий «Удаление данных справочников» предполагает следующие шаги:

- Пользователь переходит на страницу обновления данных.
- Пользователь нажимает на кнопку «Удалить».
- Сервер формирует запрос на удаление данных из БД.

Так же была построена ER диаграмма информационной системы.

База данных. База данных хранит необходимые данные для генерации отчетов о публикационной активности. Информация о публикациях хранится в системе управления базами данных PostgreSQL.

БД имеет следующие таблицы:

- Publication — таблица публикаций.
- Author — таблица авторов.
- Journal — таблица журналов.

- Book — таблица книг.
- Conference — таблица конференций.
- Departament — таблица кафедр.
- AuthorsOfPublications — осуществляет связь «многие ко многим» для авторов и публикаций.
- TypeOfPublications — осуществляет связь публикаций и типов публикаций

Модель данных. Модель данных состоит из семи классов:

- Класс «Journal» — сущность журнала с атрибутами идентификатора журнала, названия журнала и издателя.
- Класс «Book» — сущность книги с атрибутами идентификатора книги и названия книги.
- Класс «Conference» — сущность конференции с атрибутами идентификатора конференции и названия конференции.
- Класс «Departament» — сущность кафедры с атрибутами идентификатора кафедры и названия кафедры.
- Класс «TypeOfPublication» — сущность типа публикации с атрибутами идентификатора и названия.
- Класс «Publication» — сущность публикации с атрибутами идентификатора публикации, заголовка публикации, года, номера, тома, страниц, цитирований публикации, автора, журнала, книги, конференции и типа публикации.
- Класс «Author» — сущность преподавателя с атрибутами идентификатора, ФИО, места работы, кафедры и ссылки на Google Scholar.

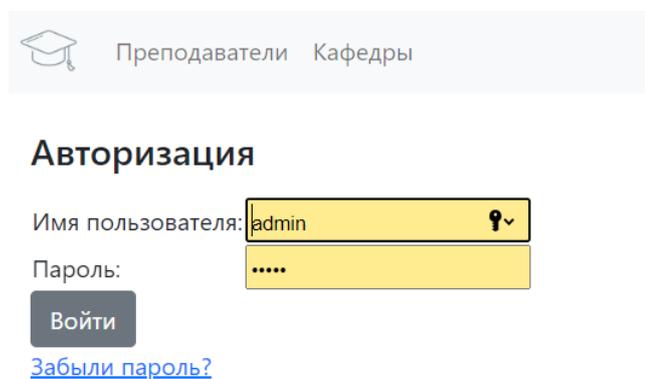
Импорт данных. Данные в базу данных загружаются из Google Scholar. Для обработки данных с портала Google Scholar используется библиотека Scrapy.

Создан отдельный проект с помощью библиотеки Scrapy. Описан класс «PublParseSpiderSpider» для считывания данных с портала Google Scholar. В класс передается ссылка на страницу преподавателя, после чего начальной страницей устанавливается ссылка на страницу введенного преподавателя. Для перехода по страницам при считывании данных описано правило перехода: считывается список ссылок на страницы публикаций со страницы

профиля автора, затем происходит переход по каждой ссылке. Основной метод класса «parse_item» считывает информацию в словарь, затем данные записываются в файл формата JSON.

Авторизация. Авторизация и аутентификация пользователя реализована на основе фреймворка работы с сессиями. Прописаны настройки в параметрах проекта для использования авторизации Django.

Для разграничения доступа описана функция «check_is_personal», которая проверяет является ли пользователь персоналом. Функции, к которым не должно быть полномочий у обычных пользователей используют специальный валидатор. В соответствии с рисунком 2 представлена страница авторизации.



Преподаватели Кафедры

Авторизация

Имя пользователя: admin

Пароль:

Войти

[Забыли пароль?](#)

Рисунок 2 — Страница авторизации

Описание реализации информационной системы. Приложение реализовано на языке Python с помощью фреймворка Django.

Главная страница представлена в соответствии с рисунком 3. На главной странице отображается навигационная панель с ссылками на главную страницу, страницу кафедр, а так же на вход или выход.

Также на главной странице располагается таблица с ссылками на страницы преподавателей и другими данными, такими как место работы, кафедра и количество публикаций. Для наглядности на страницу добавлены диаграммы общего количества публикаций по преподавателей и общего количества публикаций по годам.

При входе с аккаунта администратора на главной странице добавляется кнопка «Добавить преподавателя». При нажатии на данную кнопку осуществляется переход на страницу с формой для добавления преподавателя.

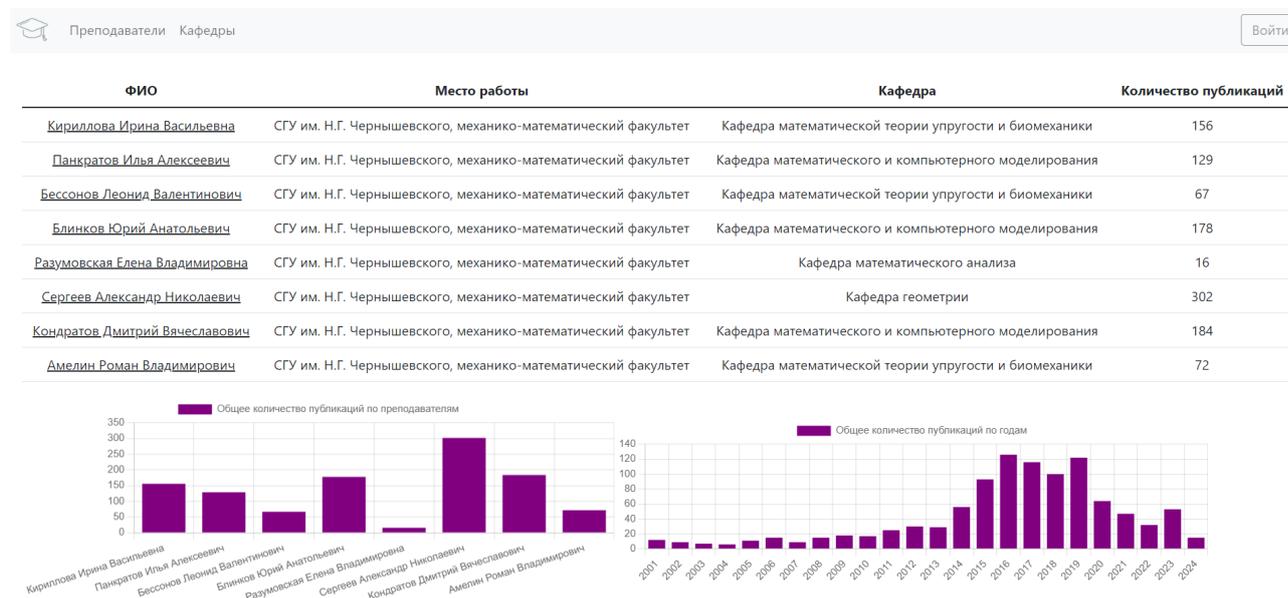


Рисунок 3 — Главная страница

Страница «Кафедры» представлена в соответствии с рисунком 4.

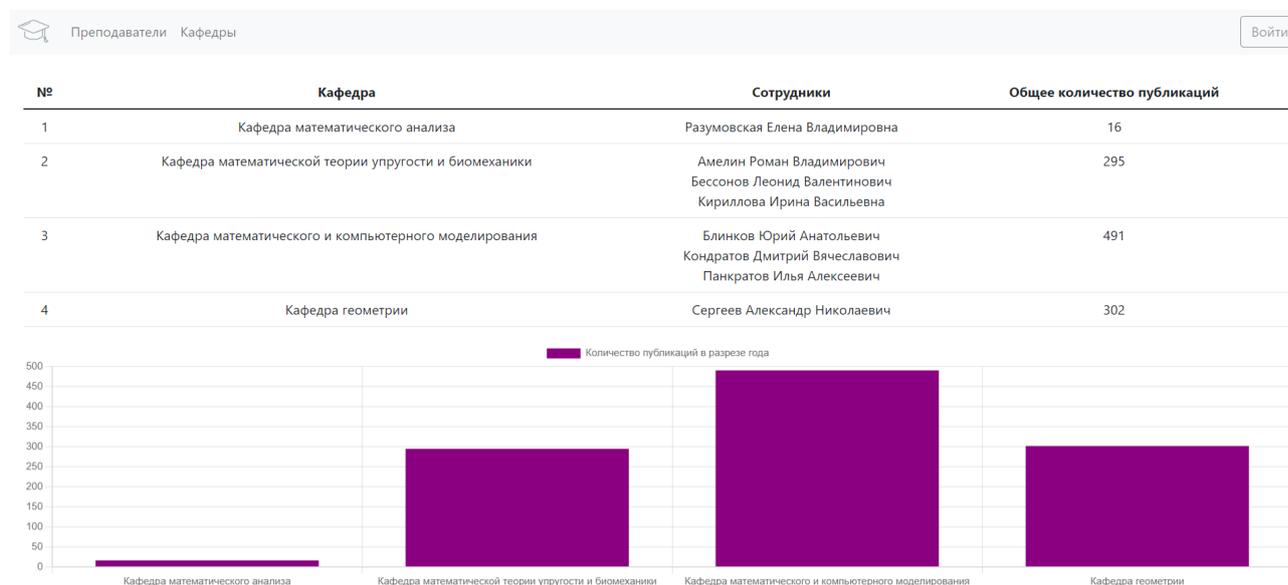


Рисунок 4 — Страница «Кафедры»

На странице «Кафедры» отображается таблица с перечнем кафедр, их сотрудников и общим количеством публикаций по кафедре. Также на данной

странице располагается диаграмма общего количества публикаций различных кафедр. При входе с аккаунта администратора на странице добавляется кнопка «Добавить кафедру». При нажатии на значок карандаша открывается страница для редактирования названия или удаления кафедры. При нажатии на кнопку «Добавить кафедру» осуществляется переход на страницу для добавления кафедры.

Страница преподавателя представлена в соответствии с рисунком 5. Сверху на странице преподавателя отображается информация о преподавателе, кнопка «Генерация отчета» и данные анализа. При нажатии кнопки «Генерация отчета» формируется Excel-файл с публикациями преподавателя и показателями. Так же на странице преподавателя отображаются диаграммы: первая — количество публикаций преподавателя в разрезе года, вторая — количество цитирований преподавателя в разрезе года.

При входе на страницу с правами администратора предоставляется возможность редактирования данных преподавателя. При нажатии на кнопку «Сохранить изменения» данные в БД обновляются. При нажатии на кнопку «Удалить преподавателя» преподаватель удаляется из БД, кроме того, удаляются все его публикации.

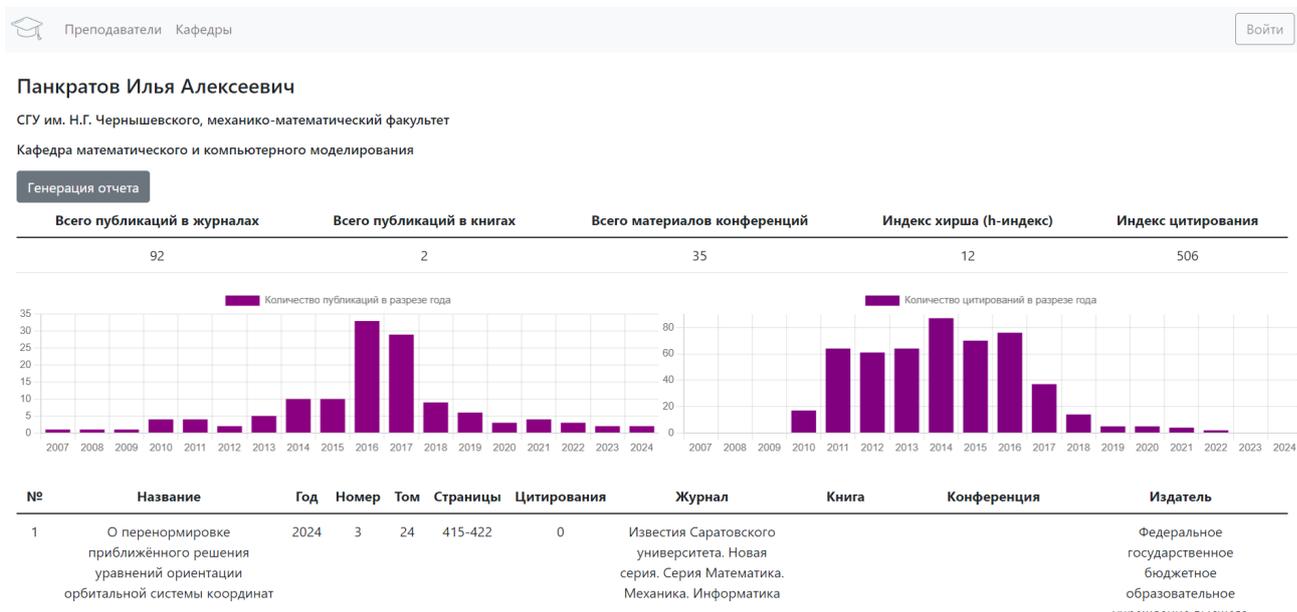


Рисунок 5 — Страница преподавателя

Также реализована возможность вручную добавлять публикации. При нажатии на кнопку «Добавить публикацию» осуществляется переход на страни-

цу добавления публикации. После внесения данных и нажатия кнопки «Добавить публикацию» в БД заносятся данные.

При нажатии на кнопку «Обновить публикации» запускается обновление данных из Google Scholar.

Таблица публикаций при входе за администратора имеет значок карандаша рядом с каждой публикацией, при нажатии на который открывается страница обновления публикации. Так же на данной странице содержится кнопка «Удалить публикацию», которая удаляет публикацию из БД.

Анализ данных о публикациях ППС. С помощью JavaScript были созданы диаграммы для анализа данных о публикациях ППС. Для отрисовки диаграмм в шаблоны передаются массивы данных в необходимом для диаграмм формате, описана область Canvas для интерактивного отображения диаграмм и скрипт с JS-функцией отрисовки. Для визуализации данных и проведения анализа построены диаграммы и приведена аналитика в текстовом формате. Это позволяет более детально изучить и оценить различные аспекты активности преподавателей и научные результаты.

Для детального анализа данных добавлены следующие показатели:

- Всего публикаций в журналах преподавателя.
- Всего публикаций в книгах преподавателя.
- Всего материалов конференций преподавателя.
- Индекс хирша (h-индекс) преподавателя.
- Индекс цитирования преподавателя.
- Диаграмма количества публикаций в разрезе года.
- Диаграмма количества цитирований в разрезе года.
- Общее количество публикаций по преподавателям.
- Общее количество публикаций по годам.
- Общее количество публикаций по кафедрам.
- Диаграмма общего количества публикаций в разрезе преподавателей.
- Диаграмма общего количества публикаций в разрезе кафедр.

Для демонстрации работы системы и анализа данных были взяты преподаватели из нескольких кафедр механико-математического факультета и загружены их публикации. После загрузки публикаций выбранных препода-

вателей в базе данных содержится более тысячи записей о научных публикациях.

В соответствии с рисунком 6 представлена диаграмма, которая отображает общее количество публикаций Панкратова И. А. в разрезе года. Такая диаграмма позволяет наглядно увидеть динамику публикационной активности по годам. Наиболее активный год по количеству публикаций Панкратова И. А. — 2016. В 2016 году было сделано 33 публикации. Наименее активные годы — 2007 — 2009, в них по одной.

В соответствии с рисунком 7 представлена диаграмма, которая отображает общее количество цитирований Панкратова И. А. в разрезе года. Данная диаграмма позволяет рассмотреть динамику цитирований научных публикаций по годам. Наиболее активный год по цитированию публикаций Панкратова И. А. — 2014. В этом году было сделано 87 цитирований различных публикаций преподавателя. Наименее активный год — 2022, в нем было сделано 2 цитирования.



Рисунок 6 — Диаграмма количества публикаций



Рисунок 7 — Диаграмма количества цитирований

В соответствии с рисунком 8 представлены другие показатели публикационной активности преподавателя. Для более детального анализа публикационной активности преподавателя ИС рассчитывает индекс Хирша. Индекс

Хирша Панкратова И. А. равен 12, что является достаточно высоким показателем публикационной активности. Так же на странице преподавателя отображается индекс цитирования. В данном случае индекс цитирования равен 506, что так же является высоким показателем.

Отображение результатов по типам публикаций позволяет обратить внимание преподавателя на преобладающие виды своей публикационной активности и сделать выводы относительно своей дальнейшей научной работы. В публикациях Панкратова И. А. преобладают публикации в журналах — 92 публикации.



Рисунок 8 — Показатели публикационной активности

В соответствии с рисунком 9 представлена диаграмма общего количества публикаций по кафедрам в рамках загруженных данных. Данная диаграмма позволяет проводить анализ публикационной активности в рамках кафедр, а в дальнейшем и факультетов. На диаграмме видно, что в рамках загруженных данных, кафедра математического и компьютерного моделирования имеет наибольшее количество публикаций.

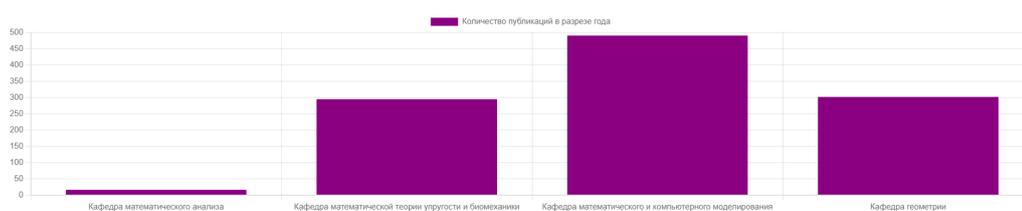


Рисунок 9 — Диаграмма активности по кафедрам

В соответствии с рисунком 10 представлена диаграмма общего количества публикаций преподавателей факультета в разрезе года. Такая диаграмма позволяет проводить анализ по данным прошлых годов и более эффективно планировать научную деятельность сотрудников в будущем. В рамках загруженных данных наиболее активным годом является 2016 год.



Рисунок 10 — Активность факультета по годам

Заключение. В рамках магистерской работы были изучены основные программные средства для реализации информационной системы для генерации отчетов о публикациях ППС, разработана БД и реализована ИС, позволяющая легко и быстро анализировать публикационную активность преподавателя.

Все поставленные задачи были выполнены:

- Проведен обзор наукометрических показателей публикационной активности (индекс Хирша, индекс цитируемости публикаций).
- Проведен обзор программных средств для разработки ИС (PostgreSQL, Python, Django, Scrapy).
- Спроектирована структура данных ИС, построены UML-диаграммы.
- Получены данные с портала Google Scholar.
- Реализована информационная система для генерации отчетов о публикациях ППС средствами Python и Django.
- Проведен анализ наукометрических показателей.
- Реализована выгрузка отчетов в формате Excel.