

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»**

Кафедра Математического и компьютерного моделирования

Разработка информационной системы поддержки учета работы

диссертационного совета

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 411 группы

направление 01.03.02 — Прикладная математика и информатика

механико-математического факультета

Яфарова Шамиля Валериевича

Научный руководитель
проф., д.ф.-м.н., доцент

Д.В. Кондратов

Зав. кафедрой
зав. каф., д.ф.-м.н., доцент

Ю.А. Блинков

Саратов 2023

ВВЕДЕНИЕ. Современные информационные технологии стремительно проникают во все сферы человеческой деятельности, в том числе и в образование. Степень развития информационных ресурсов и характер их использования оказывают существенное влияние на благосостояние общества и эффективность выполнения определенных видов профессиональной деятельности.

Информация является одним из основных понятий электронной обработки данных. В повседневной жизни мы используем информацию, хранящуюся в нашей памяти. Мы получаем эту информацию из различных источников, таких как книги, газеты, телевидение, радио и т.п. Работа, которую мы выполняем, может включать большие объемы информации, отдельные "части" которой похожи по структуре, но отличаются по содержанию. При работе с большими объемами информации скорость поиска информации (или доступ к ней) имеет решающее значение. В некоторых случаях поиск адреса в адресной книге, отсортированной по алфавиту, может быть быстрее, чем такой же поиск в ящике с несортированными карточками. Увеличение объема обрабатываемых данных выдвигает на первый план вопрос об эффективности средств организации и доступа к обрабатываемым данным. К ним относятся и базы данных.

Лидирующее положение компании определяется, прежде всего, ее способностью использовать новейшие информационные технологии. Особую роль играет образовательная сфера, где формируется интеллектуальный потенциал страны по потреблению и созданию своих информационных ресурсов.

Диссертационные советы можно рассматривать как предприятия, хотя и с определенной спецификой, поэтому проблема массового информирования актуальна для диссертационных советов, как и для любого предприятия.

Внедрение новых информационных технологий в организацию работы диссертационных советов может улучшить работу председателя по учету информации членов диссертационных советов и соискателей ученых степеней. Понятно, что в работе любой организации существует множество рутинных задач. Хранение документов в электронном виде на диске, безусловно, более целесообразно, чем хранение их на бумаге.

Проблемы по поиску нужной информации занимают много времени и сил, что может снизить быстроту и качество производимой работы. Именно поэтому было создано это приложение.

Цель бакалаврской работы состоит в разработке информационной системы поддержки учета работы диссертационного совета, которая позволяет хранить и обрабатывать данные, осуществлять быстрый поиск и отбор по введенным условиям. Для достижения поставленной цели необходимо решить следующие задачи:

- проанализировать деятельность диссертационного совета и его организационную структуру;
- исследовать особенности ведения информационной системы поддержки учета работы совета;
- осуществить сбор данных;
- спроектировать и разработать информационную систему для работы с данными.

Объектом исследования является диссертационный совет.

Предметом исследования является информационная система.

Структура работы соответствует логике исследования и включает в себя: введение, четыре главы, заключение, список использованной литературы и приложение.

В первом разделе рассматривается аналитическая часть.

В первом подразделе описывается общая характеристика деятельности диссертационного совета.

Во втором подразделе описывается организационная структура диссертационного совета.

В третьем подразделе были определены цели и задачи проектирования информационной системы.

В четвертом подразделе выполняется обзор существующего программного обеспечения.

В пятом подразделе показываются особенности ведения информационной системы.

Во втором разделе рассматривается разработка и реализация проектных решений.

В первом подразделе описывается сбор данных.

Во втором подразделе осуществляется анализ источников по преобразованию данных.

В третьем подразделе описывается обоснование проектных решений.

В третьем разделе рассматривается проектирование и реализация информационной системы.

В первом подразделе выполняется реализация сбора данных.

Во втором подразделе описываются созданные классы и используемые библиотеки.

В третьем подразделе описывается Установка SQLite в C#.

В четвертом подразделе выполняется подключение к базе данных SQLite.

В пятом подразделе описывается выбор системы управления базами данных.

В шестом подразделе описывается интерфейс модуля.

В четвертом разделе рассматривается тестирование программного продукта.

Основная часть.

Диссертационный совет (также диссовет, или совет по защите докторских и кандидатских диссертаций, или иногда специализированный учёный совет) — орган, создающийся при некоторых научных организациях и вузах.

Деятельность диссертационного совета направлена на координацию научных исследований в определенной области знаний, а также на оценку и защиту диссертационных работ соискателей научной степени.

Диссертационный совет создается в учебно-научных институтах высшего образования, научно-исследовательских учреждениях, вузах и академиях. Он является законодательным органом, который устанавливает правила, процедуры и требования к подготовке и защите диссертационных работ.

Основная функция совета - это оценка соответствия диссертационных работ академическим стандартам, научной актуальности, оригинальности и значимости для данной отрасли. Деятельность диссертационного совета включает:

- определение тем диссертационных работ, которые могут быть защищены в данном совете;

- выявляет сильные и слабые стороны работы;
- проведение открытых научных заседаний по обсуждению диссертаций;
- проверка соответствия диссертационной работы требованиям научной этики и научной морали;
- принятие решения об ученой степени, которую может получить автор диссертации;
- оформление удостоверения о присвоении ученой степени.

Также экспертная комиссия несет ответственность за следующее:

- качество и объективность экспертизы диссертаций;
- обоснованность принимаемых решений совета;
- составление заключения по научным трудам, которые планируется защищать в других диссертационных советах.

Помимо прочего, диссовет способствует созданию максимально благоприятных условий для проведения защиты диссертации оказанию всяческой помощи соискателю.

Соискателем научной степени считается специалист с высшим профессиональным образованием, который хочет получить учёную степень и (или) звание кандидата или доктора наук. Работа над диссертацией не предполагает курирования со стороны преподавателя. И в течение подготовки он прикрепляется к научному руководителю или научно-исследовательскому учреждению. Форма обучения подходит следующим категориям лиц:

- форма получения учёного звания удобна для исследовательских учреждений, которые могут подготовить научные кадры без обязательного зачисления в ВУЗ;
- сотрудники исследовательских организаций не проводят вступительных экзаменов. Аспирант, которые является соискателем степени, не сдаёт вступительных мероприятий, не проходит конкурсы аттестатов и дипломов;
- соискатель научной степени не ходит на пары, не посещает ВУЗ для изучения вопросов касаясь своей темы научной работы;
- соискатель может консультироваться с научным педагогическим составом. На кафедре соискатель числится, как внештатный специалист, который пишет диссертацию;

- кафедра может включить в штат соискателя на основании положительно характеристики. Она составляется после собеседования с ректором по научной работе. При этом преподавательская деятельность считается необязательной для него;
- как только соискатель становится прикрепленным к научному руководителю, будучи зачисленным на кафедру, он готовится к сдаче кандидатского минимума.

Таким образом, деятельность диссертационного совета важна для поддержки научных исследований и продвижения научных знаний в обществе.

Организационная структура диссертационного совета включает:

- председатель диссертационного совета. Обычно это деятель науки или профессор высшей категории, который имеет соответствующие знания и опыт в соответствующей области знания;
- члены диссертационного совета. Они являются профессорами, докторами наук или высококвалифицированными специалистами в соответствующей области знания, представленными университетом или другими организациями;
- секретарь диссертационного совета. Он управляет рабочим процессом диссертационного совета, отвечает за организацию работы совета и подготовку необходимых документов.

Все члены совета работают исключительно на добровольных началах, без материального поощрения. Эта должность считается престижной, поэтому претендентов на «пост» члена диссовета, как правило, много.

Организационная структура диссертационного совета приведена в соответствии с рисунком 1»



Рисунок 1 — Организационная структура диссертационного совета

Организационная структура диссертационного совета может варьироваться в зависимости от университета или другой организации, но она обычно включает те же основные элементы. Информация о соискателе ученой степени включает ввод основных сведений о соискателе, диссертации, публикацию объявления о защите и загрузку полного комплекта документов аттестационного дела для отправки в Министерство науки и высшего образования РФ.

В соответствии с поставленной целью в рамках бакалаврской работы требуется решить следующие задачи:

- проанализировать деятельность диссертационного совета и его организационную структуру;
- исследовать особенности ведения информационной системы поддержки учета работы совета;
- осуществить сбор данных;
- спроектировать и разработать информационную систему для работы с данными.

Поскольку основным результатом бакалаврской работы должна стать информационная система для поддержки работы диссертационного совета, необходимо сформулировать основные требования к функциональности разрабатываемого программного обеспечения.

База данных должна предоставлять пользователям информацию, необходимую для быстрого поиска записей о статьях за последние пять лет и информации о документах соискателей ученой степени.

В настоящее время автоматизировать работу диссертационных советов можно либо путем разработки собственных информационных систем и информационных технологий, полностью отвечающих всем требованиям работы компании, либо путем приобретения готовых программ с целью автоматизации существующих процессов в отдельных отделах и в организации в целом. В последние несколько лет наблюдается тенденция, когда руководители организаций предпочитают покупать готовые продукты автоматизации. Это связано с тем, что им не нужно тратить много времени и трудовых ресурсов на их создание. Кроме того, они могут добавить необходимые элементы и функциональные возможности к готовому продукту. В ответ на эту тенденцию производители готовых продуктов автоматизации также несколько

меняют свои методы продаж. В настоящее время производители таких продуктов создают новую систему, предлагают своим клиентам базовую версию, а затем настраивают ее под нужды отдельных организаций с учетом специфики их бизнеса и управления. При этом разработчик предоставляет заказчику все необходимые инструкции для того, чтобы внедрить новую систему в деятельность компании, использовать ее более эффективно и минимизировать время, необходимое для повышения квалификации существующих сотрудников.

Разработка информационной системы - это процесс создания комплексной информационной системы, включающий в себя анализ, проектирование, реализацию и внедрение программных продуктов или комплексных решений на основе современных информационных технологий. Внедрение информационной системы будет способствовать:

- уменьшению рутинной работы председателя и ученого секретаря;
- уменьшению времени на обработку информации, фильтры, поиск, аналитика данных способствуют облегчению работы с данными;
- обеспечению достоверности информации (уменьшение возможности допустить ошибки);
- повышению уровню защищенности информации от несанкционированного доступа, потери или искажения;
- автоматизация процесса подачи заявлений на участие в диссертационных советах и контроля за их выполнением;
- создание единой базы данных диссертационных работ и кандидатов для облегчения поиска и анализа информации;
- увеличение прозрачности и общественной открытости деятельности диссертационных советов;
- обеспечение соблюдения правил и требований, установленных законодательством, при прохождении диссертационных процедур;
- замене бумажных носителей данных на электронные, что приведет к более рациональной организации переработки информации на компьютере и снижению объемов документов на бумаге.

Функциональным назначением программы является предоставление пользователю актуальных данных о членах диссертационного совета, заносить данные в базу данных для предоставления отчетности.

Парсер - это программа или функция, которая анализирует текстовые данные, разбирает их на составляющие элементы и позволяет обрабатывать эту информацию в соответствии с определенными правилами или форматами. Парсеры используются в различных областях программирования, как для обработки текстовых файлов, так и для работы с различными видами данных, такими как XML, JSON, HTML и других. Они позволяют извлекать нужную информацию из больших объемов данных и переводить ее в удобный для использования формат, как правило, выкладываются в базу данных, представляются в виде файла

Парсер представляет собой программный компонент, который принимает входные данные (часто текст) и строит структуру данных — часто своего рода дерево разбора, абстрактного синтаксического дерева или другой иерархической структуры, что дает структурное представление ввода при проверке правильности синтаксиса

Парсер для информационной системы диссертационного совета нужен для автоматического сбора данных из различных источников (например, из реестра ВАК, из базы данных научных журналов и т.д.) и их обработки для последующего использования в системе. Это позволяет ускорить процесс сбора и обработки информации, а также снизить вероятность ошибок при ручном вводе данных. Кроме того, парсер может использоваться для регулярного обновления данных в системе, что также повышает ее эффективность и точность.

Перед разработкой универсального парсера файлов требуется обосновать технологию парсинга, как эффективную технологию для преобразования данных.

Главный источник информации для диссертационного совета является Российская электронная библиотека. eLibrary.ru - это Российская электронная библиотека, на которой представлены научные, учебные и художественные книги, журналы, статьи и другие публикации на русском языке.

eLibrary.ru предоставляет возможность использования книг и статей в личных и научных целях, а также в учебном процессе. Некоторые работы защищены авторским правом и могут использоваться только с разрешения правообладателя.

Python является одним из самых популярных языков программирования для парсинга данных с веб-сайтов. Вот несколько причин, почему:

- широкая доступность библиотек и пакетов: Python имеет множество библиотек и пакетов, специализированных для парсинга данных с сайтов. Некоторые из них включают в себя BeautifulSoup, Scrapy, Selenium и Requests;
- простота использования: Python является легким для обучения языком программирования. Это делает его подходящим выбором для начинающих и опытных программистов;
- экономичный процесс разработки: Построение парсера данных на Python может быть более экономичным по сравнению с другими языками программирования, отчасти потому, что нужно меньше строк кода;
- универсальность: Python легко интегрируется с другими языками программирования и инструментами. Это позволяет использовать его в любых условиях, что значительно упрощает процесс разработки приложений;
- оптимальность: благодаря концепции «быстрого прототипирования», возможности Python позволяют быстрее получить результат, чем в других языках.

Большинство из перечисленных выше преимуществ Python делают подходящим выбором для парсинга данных и программирования в целом.

Учитывая возможности, а также учитывая требования к программному обеспечению, необходимо создать программный продукт. Программный продукт должен устранить существующие недостатки в процессе работы членов диссертационного совета, а также повысить уровень информационной безопасности. Поэтому для разработки интерфейса предоставления информации была выбрана среда разработки Microsoft Visual Studio 2022 и язык программирования C#. Microsoft Visual Studio — полнофункциональная интегриро-

ванная среда разработки (IDE) с поддержкой популярных языков программирования, среди которых C, C++, VB.NET, C#, F#, JavaScript, Python.

Программа сбора данных предназначена для преобразования информации с сайта в файлы других форматов. Программное обеспечение разработано для диссертационного совета для парсинга электронной библиотеки, которая хранит данные об авторах и их статьях.

Цель и назначение парсера:

- оптимизация процесса обработки текстовой информации;
- обеспечение качества отчетности;
- повышение качества работы сотрудников.

Система должна отвечать следующим критериям:

- скорость обработки данных;
- надежность и безопасность;
- простота эксплуатации системы.

Пример входных данных текстового файла, диссертационного совета приведен в соответствии с рисунком 2.

```
Председатель - authorid=2932
Зам.председателя - authorid=63805
Учёный секретарь - authorid=42238
Член дис. совета - authorid=13326
Член дис. совета - authorid=175520
Член дис. совета - authorid=107800
Член дис. совета - authorid=174006
Член дис. совета - authorid=10937
Член дис. совета - authorid=109940
Член дис. совета - authorid=126940
Член дис. совета - authorid=3409
Член дис. совета - authorid=73470
Член дис. совета - authorid=10353
Член дис. совета - authorid=17323
Член дис. совета - authorid=150252
Член дис. совета - authorid=683327
Член дис. совета - authorid=16120
Член дис. совета - authorid=16116
Член дис. совета - authorid=3428
Член дис. совета - authorid=14208
```

Рисунок 2 — Пример входных данных текстового файла

Назначение программного обеспечение чтение из текстового файла и парсинг информации из eLibrary.ru в другие форматы. Поэтому важными моду-

лями данного программного обеспечения являются использование сессий для парсинга сайтов.

Библиотека `requests` является стандартным инструментом для составления HTTP — запросов в Python. Любой, кто когда-либо пытался парсить сайты на python начинал с простого запроса `get` библиотеки `requests`. Запрос `get` выгружает html код страницы, который можно обрабатывать под свои нужды.

HTML — это HyperText Markup Language («язык гипертекстовой разметки»), который работает за счет распространения элементов документа со специальными тегами. В HTML есть много разнообразных тегов, но стандартный шаблон включает три основных: `html`, `head` и `body`. Они организуют весь документ.

Но иногда данные доступны только после авторизации на ресурсе, для ее прохождения нужно иметь логин-пароль и библиотеку `requests`.

Использование сессий дает преимущества в скорости парсинга данных и исключает блокировку учетной записи. Если приходится выгружать данные — страницу сайта за страницей, при каждом новом запросе будет создаваться новый запрос с новым подключением к сайту, при использовании сессии она создается один раз и используется на всем протяжении работы.

Для получения содержимого страницы с сайта используется библиотека `Beautiful Soup`. `Beautiful Soup` — это библиотека на языке Python, которая используется для парсинга HTML- и XML-документов. Она позволяет разбирать и анализировать данные, полученные из интернета, и извлекать интересующую информацию.

В результате выполнения программы, полученные данные лучше всего записать в базу данных. Для этого была использована библиотека `SQLite` и библиотека `Pandas`.

`SQLite` — это библиотека Python, которая позволяет использовать базу данных `SQLite` в Python приложениях.

Библиотека `pandas` позволяет загружать данные из разных форматов файлов, таких как `CSV`, `Excel`, `JSON`, `SQL` и т. д., а также проводить различные операции с данными, включая сортировку, фильтрацию, агрегацию, группировку и многое другое.

Pandas широко используется в области анализа данных, науки о данных, финансов и экономики для обработки и анализа больших объемов данных.

На сегодняшний день SQLite является одной из самых популярных решений такого типа. Она используется практически повсеместно: в embedded приложениях, мобильных телефонах, различных платформах и т.д. Например браузер Google Chrome использует SQLite для хранения Cookies. SQLite обладает феноменальной надежностью (зачастую в ущерб производительности, если её неправильно “готовить”), поэтому, несмотря на то, что она является открытой, разработчики практически не принимают сообщения об ошибках.

Для того, чтобы иметь возможность работать с БД SQLite в C# потребуется пакет System.Data.SQLite. Чтобы его установить, достаточно воспользоваться менеджером пакетов NuGet.

Выбор между базами данных SQL и базами данных, отличными от SQL, обычно сводится к различиям в структуре. Однако, когда рассматривается несколько решений SQL, критерии гораздо более искажены. Рассмотрим аспекты более точно и проанализируем базовый функционал. Выполним сравнение СУБД: MySQL, Postgresql и SQL server.

MySQL является одной из самых популярных баз данных. Это определенный лидер среди SQL-решений, используемых Google, LinkedIn, Amazon, Netflix, Twitter и другими. MySQL популярность сильно растет, потому что команды все чаще предпочитают решения с открытым исходным кодом вместо коммерческих.

PostgreSQL - реляционная база данных, которая известна поддержкой большого количества типов данных, интуитивно понятным хранением бесструктурных данных и богатой функциональностью. Некоторые разработчики заходят даже так далеко, что утверждают, что это самая передовая база данных с открытым исходным кодом на рынке. Не нужно поддерживать или опровергать, но это, безусловно, очень универсальное решение.

В отличие от Postgresql против MySQL, SQL Server является коммерческим решением. Его предпочитают компании, которые регулярно сталкиваются с большими нагрузками трафика. Она также считается одной из наиболее совместимых систем со службами Windows.

Инфраструктура SQL Server включает в себя множество дополнительных средств, таких как службы отчетов, системы интеграции и аналитика. Для компаний, которые управляют несколькими командами, эти инструменты имеют большое значение в повседневной работе.

Интерфейс программы реализован следующим образом на входе пользователя встречает главная форма, на которой имеется загруженная база данных и десять кнопок: «|<», «<», «>», «>|», «Delete», «Add New», «Save», «Refresh Data», «Exit<», «Aspirant».

При нажатии на кнопку «Aspirant» появляется форма, предназначенная для работы с документами соискателей ученой степени, содержащая таблицу, содержащую документы, поле графика и 6 кнопок: «Upload», «Insert», «Update», «Delete», «Back», «Plot a graph». При нажатии на кнопку:

Тестирование – процесс выполнения прикладных программ с целью выявления ошибок.

Тестирование программного продукта проводилось несколькими способами. Было проведено полное тестирование правильного выполнения всех транзакций программы.

Следующим этапом тестирования являлось тестирование переходов между окнами диалогов. На этом этапе была проверена правильность связей частей программы между собой. Были исследованы все возможные переходы из одних диалоговых окон в другие.

Отдельно проводилось тестирование всего программного продукта после завершения программирования. Для этих целей в качестве пользователя привлекали несколько человек, не имеющих отношения к программированию. На этом этапе проверялась понятность программы для неопытного пользователя.