

Цифровизация ЦНТБ: разработка базы данных для тематического структурирования депозитарного фонда

Курган Елена Геннадьевна,
заместитель директора
по вопросам интеллектуальной собственности,
канд.экон.наук, доцент,
Петрова Ксения Сергеевна, инженер-программист,
Собчишен Андрей Сергеевич, инженер-программист 1 категории,
Институт научно- технической информации,
Донецкая Народная Республика

Отдел «Центральная научно-техническая библиотека» ФГБНУ «Институт научно-технической информации» является старейшей и единственной специализированной библиотекой Донбасса.

Библиотека содержит более 700000 ед. хранения научно-технической литературы. Вся литература связана с отраслевой спецификой Донбасса. Самое старое издание датируется 1826 годом.

Библиотека была создана в 1948 г. и расцвет ее деятельности (с точки зрения финансирования, пополнения фондов, посещаемости) совпал с периодом послевоенного восстановления и индустриализации Донбасса. В этот же период в ЦНТБ был сформирован депозитарный фонд, пополнение которого осуществлялось с 1975 г. по 1991 г.

Основу депозитарного фонда составила научно-техническая литература по угольной промышленности, но есть издания, касающиеся других отраслей, таких как машиностроение, энергетика и пр. Депозитарный фонд ЦНТБ имеет большое количество изданий, которые можно отнести к категории ценных и редких книг. Это и периодика, и отраслевые издания, и книги с автографами и пр.

ЦНТБ в составе ЦБНТИ прекратила свою деятельность в 2014 г., с началом военных действий, и открылась уже в период Донецкой Народной Республики, в составе ГБУ «Институт научно-технической информации» в 2016 г.

В период 2016-2022 г. здание библиотеки не отапливалось, дважды получало повреждения в результате артобстрелов, текла крыша, фонды библиотеки перемещались по

зданию без надлежащего учета. В этот период практически не было финансирования, штат библиотеки составлял 3 штатные единицы.

В таких условиях велась работа по обслуживанию читателей и приему обязательного экземпляра неопубликованных материалов периодических изданий ДНР, а также по написанию контрольных талонов на документы, находящиеся на абонементе. Сверка талонов с учетными документами не проводилась. В целом работы по инвентаризации фондов выполнять было невозможно.

В 2022 г. Донецкая Народная Республика вошла в состав Российской Федерации в качестве субъекта. Наше учреждение перешло в федеральное подчинение, и библиотека смогла возобновить свою деятельность. Штатная численность на 01.08.2024 г. составляет 17 штатных единиц, получено финансирование на выполнение государственного задания, в том числе связанного с вопросами изучения структуры депозитарного фонда ЦНТБ по тематическому направлению «Угольная промышленность».

Сведения, полученные в ходе исследования, оцифровки, структурирования и систематизации депозитарного фонда, упорядоченные в виде тематических коллекций, путеводителей, обзоров могут быть представлены в качестве электронной базы данных.

В современных условиях информационные технологии оказывают значительное влияние на эффективность работы организаций. Для ЦНТБ оцифровка депозитарного и патентного фонда является не только шагом в будущее, но и насущной необходимостью для обеспечения устойчивости и гибкости в управлении информацией. Это требует внедрения современной системы управления данными, и поэтому выбор подходящей СУБД стал ключевым элементом в этой стратегической трансформации.

Для этой цели разработана база данных на платформе PostgreSQL, обеспечивающая надежное хранение и гибкое управление большими объемами информации. Были тщательно проанализированы различные СУБД, учитывая их потенциал в обеспечении производительности, надежности, безопасности, а также интеграции с существующими системами ЦНТБ. Создан пользовательский интерфейс, упрощающий работу с базой данных, даже для пользователей без специальной подготовки в области SQL.

Такой подход позволяет модернизировать техническую инфраструктуру и значительно облегчить взаимодействие пользователей с системой. Данная статья представляет результаты оценки различных СУБД, демонстрирует преимущества выбранного решения, также приводится детальное сравнение различных СУБД, их характеристик и возможностей (табл. 1), что позволило определить оптимальную платформу для реализации амбициозных целей по цифровой трансформации ЦНТБ.

Таблица 1 – Анализ СУБД

Показатели	PostgreSQL	Microsoft SQL Server	MySQL	Oracle Database	MongoDB
Лицензия	Open Source. PostgreSQL. Бесплатная.	Proprietary. Microsoft. Платная.	Open Source. GPL. Бесплатная.	Proprietary. Oracle. Платная.	Open Source. SSPL/AGPL. Бесплатная.
Поддержка JSON формата	Глубокая поддержка JSON	Ограниченная поддержка	Основная поддержка	Поддержка через плагин	Основная функция
Степень соблюдения стандартов	Высокая	Высокая	Средняя	Очень высокая	Нет (NoSQL)
Масштабируемость	Высокая	Высокая	Ограниченная (федерация)	Очень высокая	Массивная (за счёт кластера)

Продолжение таблицы 1 – Анализ СУБД

Показатели	PostgreSQL	Microsoft SQL Server	MySQL	Oracle Database	MongoDB
Расширяемость	Высокая (модули, плагины)	Средняя	Низкая	Низкая	Высокая
Сообщество	Большое сообщество	Профессиональная	Большое сообщество	Профессиональная	Активное сообщество
Поддержка сложных запросов	Отличная	Отличная	Ограниченная	Отличная	Ограниченная
Наличие локальной версии	Есть (Postgres Pro)	Нет	Нет	Нет	Нет

Основываясь на проведённом анализе, выбрана СУБД PostgreSQL, так как она собрала все необходимые преимущества для проекта, а именно:

1. Лицензия: PostgreSQL — это полностью бесплатная open-source СУБД. Это дает пользователям свободу использования и изменения без финансовых ограничений;
2. Поддержка JSON: глубокая поддержка JSON делает PostgreSQL идеальной для работы с гибридными данными;
3. Соответствие стандартам: высокая степень соблюдения SQL-стандартов гарантирует надежность и портативность кода;
4. Масштабируемость: PostgreSQL легко справляется с ростом данных и увеличением нагрузки;
5. Расширяемость: экосистема модулей и плагинов позволяет адаптировать СУБД под нужды проекта;

6. Сообщество: активное сообщество поддерживает разработку и решение проблем, что обеспечивает регулярные обновления;

7. Поддержка сложных запросов: отличная оптимизация сложных запросов обеспечивает эффективность работы с данными;

8. Локальная версия: возможность развертывания локальной версии предоставляет дополнительную гибкость.

Далее рассмотрим основные элементы и особенности структуры разработанной базы данных, схема которой изображена на рисунке 1. В базе данных таблица «Книга» играет ключевую роль, так как она хранит все свойства книг и служит связующим элементом для других таблиц. В числе других таблиц базы данных также имеются справочные таблицы, такие как коллекции, подколлекции, языки, авторы и другие, которые используются для поддержки и организации данных. Структура базы данных была тщательно разработана, чтобы оптимально отразить депозитарный фонд ЦНТБ, обеспечивая гибкую классификацию и организацию книг в соответствии со всеми требованиями и ожиданиями.

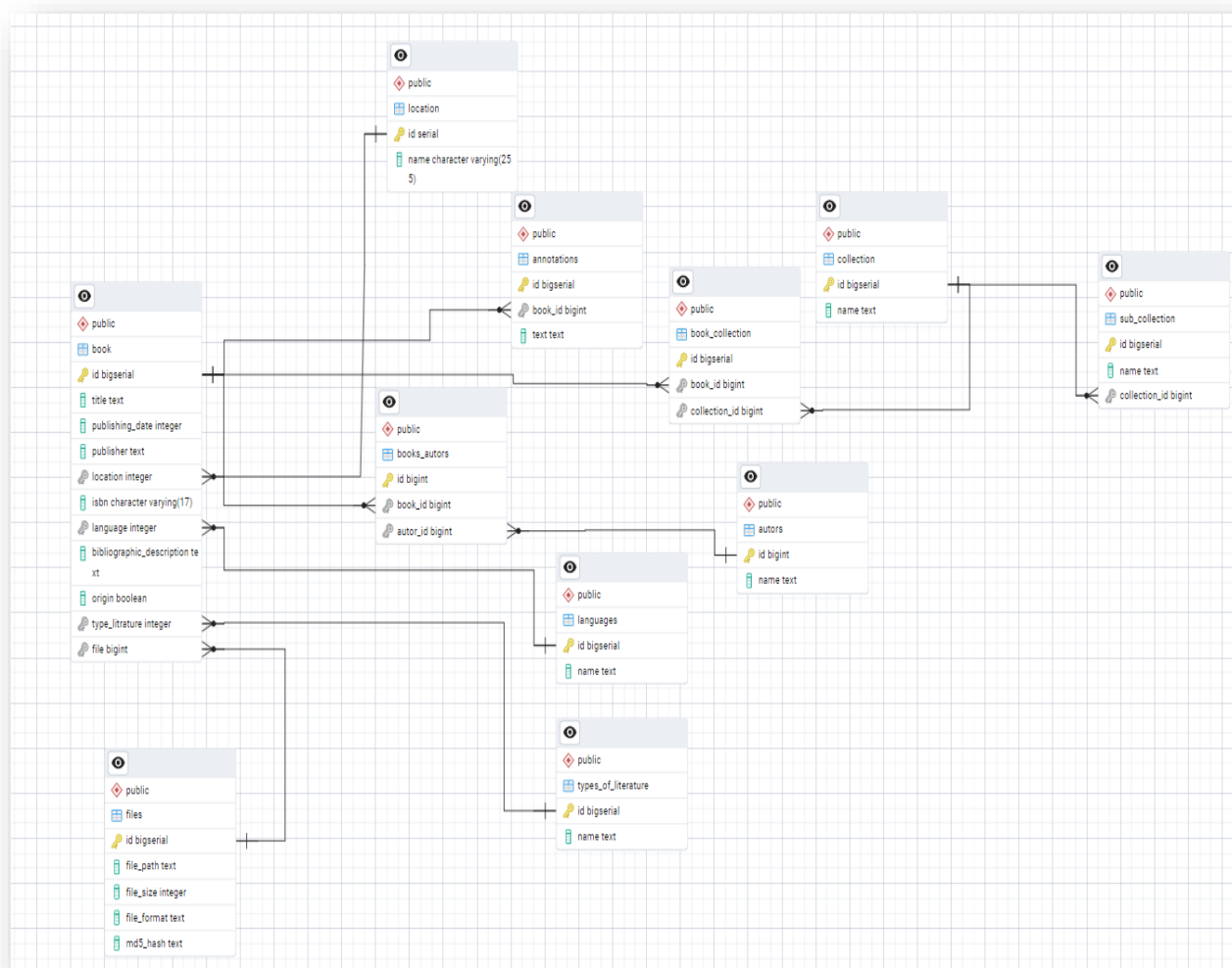


Рисунок 1 – Схема базы данных

При разработке базы данных особое внимание было уделено созданию удобного пользовательского интерфейса (рис.2). Он спроектирован таким образом, чтобы пользователи могли легко добавлять, редактировать и управлять данными без необходимости обладать глубокими техническими знаниями. Такой интерфейс значительно упрощает взаимодействие с системой, делая её доступной и интуитивно понятной, что является ключевым элементом для успешной реализации проекта.

The screenshot shows a web interface for adding a new book. At the top, there are two buttons: "Просмотр всех документов" (View all documents) and "Добавление справочной информации" (Add reference information). The main form is titled "Добавление новой книги" (Add new book). It contains the following fields and options:

- Библиографическое описание:** A text area containing a detailed bibliographic entry.
- Название книги:** A text input field with the value "Проектирование бурения скважин на твердые полезные ископаемые".
- Оригинал:** Radio buttons for "Да" (selected) and "Нет".
- Добавить перевод/оригинал (если нужно):** A text input field with the value "Ря".
- Разрушение горных пород:** A text input field.
- Аннотация:** A text area with the value "Учебно-методический комплекс дисциплины предназначен для студентов по направлению подготовки «130208 Технологии геологической разведки»".
- Автор:** A text input field.
- Бюки Б.И. X:** A text input field.
- Год издания:** A text input field with the value "2013".
- Издательство:** A text input field with the value "Сибирский федеральный университет, Красноярск".
- Место хранения:** A dropdown menu with the value "Абонемент".
- Язык:** A dropdown menu with the value "Русский".
- Вид литературы:** A dropdown menu with the value "Справочная".
- ISBN:** A text input field with the value "978-5-8038-0568-7".
- Коллекция:** A dropdown menu with the value "Справочная, учебная и научная литература по горному делу".
- Подколлекция:** A dropdown menu with the value "Энциклопедии и справочники".
- Прикрепить файл (PDF или Word):** A section with a "Выбрать файл" button and the text "Файл не выбран".
- Добавить книгу:** A button at the bottom of the form.

Рисунок 2 – Интерфейс пользователя для добавления новой книги

Форма позволяет прописать следующие свойства о книге: библиографическое описание, название, является ли книга оригиналом, аннотация, авторы, издательство, год издания, язык, вид литературы, ISBN, месторасположение книги в библиотеке, выбрать коллекцию и подколлекцию к которой принадлежит книга, а также прикрепить оцифрованный файл. Форма позволяет пользователю не только внести всю информацию о книге, но и связать её с уже существующим переводом или оригинальной версией в базе данных. Это значительно

упрощает процесс управления литературными произведениями, так как пользователи могут легко отслеживать различные версии книги, их переводы и оригинальные тексты, сохраняя целостность данных. Благодаря такой функциональности, система обеспечивает более точную каталогизацию, облегчает поиск информации и поддерживает высокую степень интеграции данных, что особенно важно для библиотек и архивов, работающих с большими объемами материалов.

Сверху страницы есть небольшое меню – две кнопки, которые обеспечивают весь необходимый функционал. Первая кнопка имеет название «Просмотр всех документов» и обеспечивает переход на вторую страницу, на которой можно просмотреть все книги в БД. Вторая кнопка позволяет добавить данные в справочные таблицы с помощью всплывающей модальной формы.

На рисунке 3 изображены модальные всплывающие формы, позволяющие добавлять авторов, коллекции, подколлекции и типы литературы в справочные таблицы.

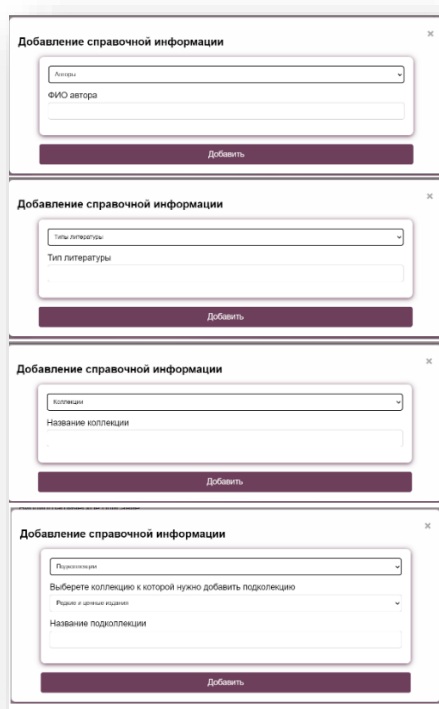


Рисунок 3 – Модальные формы добавления справочной информации

Далее более подробно рассмотрим вторую страницу разрабатываемой формы (рис. 4). Ключевым элементом этой страницы является таблица, содержащая основную информацию о книге – библиографическое описание, месторасположение книги и отметку о наличии оцифрованного файла.

Добавление нового документа Добавление справочной информации

Просмотр всех документов

Библиографическое описание	Место хранения	Файл
Нескоромных, Вячеслав Васильевич. Направленное бурение и основы кернотриемии : учебное пособие / В. В. Нескоромных, Д. В. Лысков ; ред.: П. С. Лушмин, П. Ф. Еловых ; Сибирский федеральный университет. Институт цветных металлов. – Электрон. текстовые дан. (6,5 Мб). – Красноярск : СФУ, 2022 (2023-10-06). – 164 с., 10,3 усл. тек. л., ил., проф., цв. ил., табл. – Загл. с титул. экрана. – Электрон. версия лек. публикации. – Библиогр. с. 154. – 100 экз. – ISBN 978-5-7638-4770-3. Б. ц. – Изд. № 2022-17622. – Текст : электронный.	Абонемент	
Проектирование бурения скважин на твердые полезные ископаемые : учебно-методический комплекс [для студентов напр. подготовки «130200 Технология геологической разведки»] / Сиб. федер. ун-т. Ин-т горн. дела, геологии и геотехнологий ; сост. В. В. Нескоромных. – Электрон. текстовые данные (самораспространяющийся архив, 43,11 Мб). – Красноярск : СФУ, 2013. – [Электронная библиотека СФУ. Учебно-методические комплексы дисциплин в авторской редакции]. – Загл. с титул. экрана. – ISBN 978-5-8038-0566-7. – Текст : электронный.	Читальный зал	
Разрушение горных пород : учебно-методический комплекс [для студентов напр. 131000.62 «Нефтегазовое дело», профиль 131000.62.01 «Бурение нефтяных и газовых скважин»] / Сиб. федер. ун-т. Ин-т горн. дела, геологии и геотехнологий ; сост. В. В. Нескоромных. – Электрон. текстовые данные (самораспространяющийся архив, 58,5 Мб). – Красноярск : СФУ, 2013. – [Электронная библиотека СФУ. Учебно-методические комплексы дисциплин в авторской редакции]. – Загл. с титул. экрана. – ISBN 978-5-7638-2810-6. – Текст : электронный.	Депозитарное хранение (дх)	

Рисунок 4 – Страница просмотра всех книг

Также информацию о книге можно отредактировать с помощью иконки карандаша, расположенной в первом столбце каждой строки. Всплывающая модальная форма редактирования информации изображена на рисунке 5. Всплывающая форма идентична форме добавления, только уже имеет заполненные поля, которые можно отредактировать.

Редактирование книги

Библиографическое описание:
 Нескоромных, Вячеслав Васильевич. Направленное бурение и основы кернотриемии : учебное пособие / В. В. Нескоромных, Д. В. Лысков ; ред.: П. С. Лушмин, П. Ф. Еловых ; Сибирский федеральный университет. Институт цветных металлов. – Электрон. текстовые дан. (6,5 Мб). – Красноярск : СФУ, 2022 (2023-10-06). – 164 с., 10,3 усл. тек. л., ил., проф., цв. ил., табл. – Загл. с титул. экрана. – Электрон. версия лек. публикации. – Библиогр. с. 154. – 100 экз. – ISBN 978-5-7638-4770-3. Б. ц. – Изд. № 2022-17622. – Текст : электронный.

Название книги:
 Направленное бурение и основы кернотриемии

Оригинал:
 Да Нет

Добавить перевод/оригинал (если нужно):
 Пр

Проектирование бурения скважин на твердые полезные ископаемые

Аннотация:
 «авторская и основанная термины, предназначенно для студентов вузов, обучающихся по специальности 21.05.03 «Технология геологической разведки», специализация 21.05.03.03 «Технология и техника разведки месторождений полезных ископаемых» (ФГОС-2020 (3+3+)).»

Автор:
 Бокий Б.И. ✕ Зворыкин А.А. ✕

Год издания:
 2023

Издательство:
 Сибирский федеральный университет, Красноярск

Место хранения:
 Абонемент

Язык:
 Русский

Вид литературы:
 Учебная

ISBN:
 978-5-7638-4770-3

Коллекция:
 Редкие и ценные издания

Подколлекция:
 Документы с автографами авторов

Прикрепить файл (PDF или Word):
 Файл не выбран Выбрать файл

Сохранить изменения

Рисунок 5 – Форма редактирования книги

Таким образом, реализация проекта на основе PostgreSQL предоставляет ЦНТБ не только инструмент для текущего управления фондами, но и потенциал для дальнейшего роста и развития в условиях быстро меняющихся технологий и требований к хранению данных.

Разработанный удобный пользовательский интерфейс, включая форму для добавления и редактирования данных, значительно упрощает работу с системой, обеспечивая лёгкость управления информацией. Это делает процесс оцифровки и систематизации данных более эффективным и доступным.

Патентный и депозитарный фонд ЦНТБ обладает обширной коллекцией литературы. Чтобы эффективно управлять таким объемом информации, необходимо оптимизировать как структуру базы данных, так и запросы к ней. В PostgreSQL предоставляются мощные инструменты для такой оптимизации:

1. Индексирование: поддержка различных типов индексов, таких как B-tree (по умолчанию), Hash, GIN и GiST, позволяет значительно ускорить выполнение запросов.

2. Партиционирование таблиц: позволяет распределять данные на более управляемые части, что упрощает обработку больших объемов информации и улучшает производительность.

3. Планировщик запросов (Query Planner): анализирует запросы и выбирает наилучший способ их выполнения. Команда EXPLAIN предоставляет информацию о планах выполнения запросов, что позволяет выявить узкие места.

4. Настройка конфигурации: параметры, такие как `shared_buffers`, `work_mem`, и `maintenance_work_mem`, можно настроить для достижения оптимальной производительности в зависимости от конкретных условий.

5. Кэширование: активное использование кэширования позволяет хранить часто запрашиваемые данные в памяти, ускоряя доступ к ним.

6. Анализ производительности: инструменты, такие как `pg_stat_statements`, собирают статистику о выполнении запросов и помогают выявить потенциальные проблемы в производительности.

7. Оконные функции и агрегация: помогают эффективно обрабатывать большие объемы данных, минимизируя объем передаваемой информации.

Использование вышеперечисленных инструментов и методов позволяет обеспечивать высокую производительность и масштабируемость базы данных ЦНТБ, удовлетворяя требования к управлению огромным массивом литературы.

В перспективе база данных будет интегрирована в создаваемую цифровую платформу, предоставляющую читателям прямой доступ к фондам ЦНТБ.

