

<https://doi.org/10.52676/1729-7885-2024-3-12-20>

УДК 004.4: 004.6

РАЗРАБОТКА БАЗЫ ДАННЫХ ДЛЯ ХРАНЕНИЯ И АНАЛИЗА НАУЧНЫХ ПУБЛИКАЦИЙ

И. А. Бачурин*, Н. В. Ларионова, А. В. Топорова

Филиал «Институт радиационной безопасности и экологии» РГП НЯЦ РК, Курчатов, Казахстан

* E-mail для контактов: bachurina@nnc.kz

В статье рассмотрены ключевые функциональные требования к базе данных (БД) для эффективного хранения, организации и анализа научных публикаций. На основе анализа материалов за 30-летний период научной деятельности Института радиационной безопасности и экологии, а также учитывая интересы пользователей, были определены требования к функциональности и структуре БД. В результате была спроектирована и разработана БД с использованием средств и технологий, обеспечивающих взаимодействие пользователя и сервера. Созданный пользовательский интерфейс предоставляет удобные средства для работы с данными. Определено, что БД является эффективным инструментом для анализа данных, поскольку обладает спектром функций, позволяющих проводить разнообразные виды анализа: возможность выборки данных в соответствии с заданными критериями, их сортировка по различным параметрам и извлечение для последующего анализа.

Ключевые слова: база данных, публикации, статьи, клиент-серверное приложение, пользовательский интерфейс.

ВВЕДЕНИЕ

Обмен научной информацией, идеями, полученными экспериментальными данными, является важной составляющей науки и играет ключевую роль в ее развитии. Различные формы публикаций, такие как отчеты, монографии, тезисы, сборники трудов служат главными средствами распространения результатов исследований. Обязательным условием для ведения научной деятельности является опубликование статей, цель которых заключается в оперативном информировании общественности о полученных результатах.

Институт радиационной безопасности и экологии РГП НЯЦ РК активно внедряется в научные процессы, принимая участие в государственных программах и проводя исследования в области радиационной безопасности и экологии. За 30 лет своего существования Институт накопил обширный объем научных материалов, содержащих важные данные и результаты исследований. Филиал и сегодня задействован во многих научных программах, одной из которых является программа BR21882086 «Разработка устойчивого управления земельными ресурсами и водными объектами на территории бывшего Семипалатинского испытательного полигона». В рамках данной программы запланировано проведение семинаров и тренингов по популяризации мер радиационной и экологической безопасности и внедрению научных результатов программы.

Для популяризации результатов научной деятельности Института встал вопрос о создании инструмента, который бы обеспечивал управление, хранение и доступ к научным материалам. Данный ресурс должен обеспечивать многопользовательский доступ, поиск и анализ информации, что позволило бы более качественно использовать научный потенциал Института и способствовать дальнейшему развитию об-

ласти радиационной безопасности и экологии. Таким средством могут выступать базы данных – структурированные коллекции данных, организованные и доступные для использования любым пользователем [1], важным аспектом которых является возможность поиска нужной информации.

Существует большое количество БД, содержащих информацию по разным предметным областям, для всех них поиск информации осуществляется с помощью:

- использования метаданных для классификации и систематизации научного материала. Это могут быть названия публикаций, ФИО авторов, ключевые слова, аннотации и т.д.
- поисковых алгоритмов, учитывающих не только прямое соответствие данных запросу, но и семантические аспекты запроса. Поисковые алгоритмы также позволяют осуществлять поиск по всему тексту статьи.
- элементов машинного обучения с возможностью предоставления рекомендаций по похожим публикациям и исследованиям. Алгоритмы анализируют поведение пользователей и их предпочтения для предложения наиболее подходящих материалов.
- фильтрации и сортировки данных по различным параметрам. Чаще всего это дата опубликования, журнал, предметная область, организация и т.п.
- интеграции с другими БД. Например, возможность просмотра публикаций на платформе ORCID через профиль в БД Scopus, используя ORCID-идентификатор (уникальный числовой идентификатор исследователя).

Рассмотрим некоторые известные БД. PubMed – крупнейшая база медицинских научных публикаций, включающая статьи из различных областей медицины и биомедицинских исследований (<https://pubmed.ncbi.nlm.nih.gov/>) [2]. PubMed использует

расширенные функции поиска, такие как Medical Subject Headings. MeSH – это словарь, используемый для индексирования статей, MeSH термины позволяют пользователям более точно находить релевантные статьи. БД IEEE Xplore – содержит более 5 000 000 публикаций, отчетов и стандартов в области электротехники, компьютерных наук и связи (<https://ieeexplore.ieee.org/Xplore/home.jsp>) [3]. IEEE Xplore предоставляет возможности для поиска по конкретным областям инженерии и технологий, с фильтрацией по типу публикации, автору, дате опубликования и другим критериям. БД Scopus – мультидисциплинарная база данных, предоставляющая информацию о научных публикациях, статьях, книгах и конференциях в различных областях знаний, (<https://www.scopus.com/search/form.uri?display=basic#basic>) [4]. Отличительной чертой Scopus является поддержка анализа цитирования, а также использование индекса Хирша, как оценки «научной продуктивности» ученых.

МАТЕРИАЛЫ И СРЕДСТВА

Объектом исследования работы является научный материал, включающий статьи, тезисы (материалы конференций), монографии, авторефераты, сборники трудов, буклеты и отчеты о научно-исследовательской работе, созданные в рамках деятельности Института радиационной безопасности и экологии за весь период его существования. Этот материал содержит результаты исследований в области радиационной безопасности и экологии. На данный момент проанализировано и занесено в БД более 860 источников.

Создание БД включает в себя несколько этапов:

- определение основных функций и возможностей инструмента разработки;
- проектирование и разработка БД (создание структуры БД, определение таблиц, полей и связей между ними, а также разработка пользовательского интерфейса для ввода, редактирования и поиска данных);
- наполнение БД научным материалом и проведение анализа с помощью инструментов, имеющихся в разработанной БД.

В качестве основных требований к БД определены следующие:

- *Хранение данных*: обеспечение возможности сохранения различных типов научных публикаций, таких как статьи, тезисы, монографии и др.
- *Организация и структурирование данных*: создание системы организации данных, позволяющей классифицировать информацию по различным критериям, таким как автор, название, ключевые слова, дата публикации и другие параметры.
- *Поиск и фильтрация данных*: разработка возможностей для эффективного поиска и фильтрации информации среди хранимых данных с целью быстрого доступа к необходимым публикациям.

- *Управление доступом*: обеспечение безопасного доступа к данным и возможность управления правами доступа для различных пользователей в соответствии с их ролями и функциями.

- *Анализ данных*: предоставление инструментов для анализа собранных данных с целью выявления закономерностей, тенденций и вклада различных факторов в область радиационной безопасности и экологии (сортировка, выборка).

- *Интерфейс пользователя*: разработка интуитивно понятного пользовательского интерфейса, который будет удобен для использования и позволит пользователям легко взаимодействовать с инструментом.

На основе анализа научного материала в виде статей, материалов конференций, материалов издательской деятельности и отчетов, семантическая модель разрабатываемой БД включает следующие основные сущности:

- *Публикации (publics)*: информация о научных публикациях, включая заголовок, тип документа, название журнала, дату создания, библиографические записи и связи с категориями и авторами.

- *Авторы (authors)*: данные об авторах публикаций, включая их полное имя, место работы и индекс Хирша.

- *Категории (categories)*: описание категорий научных публикаций.

- *Пользователи (users)*: информация о пользователях системы, включая администраторов и обычных пользователей.

- *Связи между публикациями и авторами (public_authors)*: связи между публикациями и их авторами для учета нескольких авторов публикации.

Определены функции для работы с этими данными через пользовательский интерфейс. Их можно разделить на три основные категории.

1) Функции доступа к данным:

- *selectAll(\$table, \$params = [])*: возвращает все записи из указанной таблицы с возможностью задания дополнительных параметров.

- *selectOne(\$table, \$params = [])*: возвращает одну конкретную запись из таблицы с учетом заданных параметров.

- *selectWithPaginationAndSorting (\$table, \$params, \$perPage, \$offset, \$sortBy, \$sortOrder)*: позволяет получать данные с учетом пагинации и сортировки.

2) Функции модификации данных:

- *insert(\$table, \$params)*: добавляет новые записи в указанную таблицу базы данных.

- *update(\$table, \$id, \$params)*: редактирует существующие записи в таблице на основе указанного идентификатора.

- *delete(\$table, \$params)*: удаляет записи из таблицы в соответствии с указанными параметрами.

– *deleteAuthors(\$stable, \$id)*: удаляет записи авторов из таблицы на основе указанного идентификатора.

3) Функции отладки и проверки:

– *checkDatabaseError(\$query)*: проверяет наличие ошибок при выполнении запроса к базе данных.

– *testView(\$value)*: отображает содержимое переменной в процессе отладки и анализа данных.

– *logError(\$message)*: регистрирует сообщения об ошибках и событиях для последующего анализа.

В качестве инструментов анализа, имеющихся в БД, выступают функции, позволяющие выполнить выборку по категориям публикаций, авторам, типам материала, годам: *getAllCategories()* – получение списка всех категорий, которые затем могут быть использованы для классификации научных публикаций; *getPublicationsByCategory(\$categoryId)* – позволяет получить все публикации в заданной категории; *getPublicationsByType(\$type)* – возвращает все публикации определенного типа; *selectTopTypes()* – позволяет выявить наиболее часто встречающиеся типы материалов; *getAllCategoriesWithPublications Count()* – возвращает список всех категорий с указанием количества публикаций в каждой из них; *getPublicationsByAuthor(\$authorId)* – возвращает все публикации по указанному идентификатору автора и т.д. Данные функции осуществляют сортировку и выборку данных, на основе которых можно выполнить анализ данных и выявить тренды и выполнить оценку влияния различных факторов на радиационную безопасность и экологию.

При проектировании и разработке БД для хранения и систематизации научных публикаций были использованы следующие технологии:

– *MySQL*: MySQL был выбран в качестве системы управления реляционными базами данных (СУБД) благодаря своей надежности, производительности и распространенности. Он предоставляет широкие возможности для создания, изменения и запросов данных, а также обладает хорошей поддержкой и активным сообществом разработчиков.

– *PHP (Hypertext Preprocessor)*: PHP использовался для создания динамических веб-приложений и взаимодействия с базой данных MySQL. PHP позволяет выполнять различные операции с данными, обрабатывать запросы пользователя и генерировать динамический контент на основе полученных данных из базы данных.

– *HTML (HyperText Markup Language) и CSS (Cascading Style Sheets)*: HTML и CSS применялись для создания пользовательского интерфейса и визуального оформления веб-приложения. HTML определяет структуру контента страницы, а CSS – ее внешний вид, что позволяет создавать привлекательные и удобные для использования интерфейсы.

– *JavaScript*: JavaScript использовался для добавления интерактивности и динамического поведения на веб-страницах, что улучшает пользовательский опыт. Он позволяет обрабатывать события пользователя, выполнять асинхронные запросы к серверу, создавать анимацию и многое другое.

Для разработки БД использовались такие программные средства, как AMPPS Version 4.3, Visual Studio Code Version 1.78.2, Google Chrome. Разработанная БД – это клиент-серверное приложение. AMPPS обеспечивал локальную серверную среду, включая Apache, MySQL, PHP, Perl и Python, позволяя создавать и тестировать веб-приложения локально. Visual Studio Code предоставлял интуитивно понятную среду разработки с множеством интегрированных инструментов и возможностей для написания и отладки кода.

AMPPS (Apache, MySQL, PHP, Perl, Python) – это мощный и удобный инструмент для разработки веб-приложений, который позволяет создавать локальные серверные среды на персональных компьютерах. Он включает в себя все необходимые компоненты для запуска веб-сервера и базы данных на компьютере, в том числе Apache, MySQL, PHP, Perl и Python. Это означает, что возможно создавать и тестировать веб-приложения локально, прежде чем загружать их на удаленный сервер. [5].

Visual Studio Code (VS Code) – это инновационный текстовый редактор, разработанный Microsoft, который предоставляет разработчикам широкий спектр возможностей для создания высококачественного программного обеспечения. VS Code предлагает удобную и интуитивно понятную среду разработки. Редактор обладает множеством интегрированных инструментов, таких как подсветка синтаксиса, автоматическое завершение кода, отладчик, управление Git и многое другое. Одной из особенностей VS Code является его высокая степень настраиваемости и расширяемости благодаря большому количеству доступных расширений из магазина расширений. [6].

Google Chrome – веб-браузер, который предоставляет разработчикам инструменты для анализа, отладки и оптимизации веб-приложений. Включенные в него инструменты разработчика, такие как DevTools, позволяют анализировать HTML, CSS и JavaScript в реальном времени, отслеживать сетевые запросы, отлаживать JavaScript-код и аудировать веб-страницы на производительность, доступность и SEO-оптимизацию. Кроме того, Google Chrome предоставляет возможность эмулировать различные устройства и разрешения экрана, что позволяет создавать адаптивные и отзывчивые веб-страницы.

В таблице 1 представлены основные характеристики средств, используемых для разработки БД научных публикаций.

Таблица 1. Основные характеристики средств разработки

| Программный продукт | AMPPS | Visual Studio Code | Google Chrome |
|----------------------|--|--|--|
| Тип | Локальная серверная среда | Интегрированная среда разработки (IDE) | Веб-браузер |
| Поддерживаемые языки | PHP, Perl, Python, MySQL и другие | JavaScript, TypeScript, HTML, CSS и многие другие | HTML, CSS, JavaScript и другие веб-технологии |
| Операционные системы | Windows, macOS, Linux | Windows, macOS, Linux | Windows, macOS, Linux, Android, iOS |
| Отладка | Поддержка отладки PHP-приложений | Встроенные инструменты отладки для различных языков программирования | Возможность отладки веб-страниц и веб-приложений |
| Расширяемость | Возможность установки дополнительных модулей и компонентов | Огромное количество плагинов и расширений для настройки среды разработки | Магазин Chrome Web Store для установки дополнений и расширений |
| Управление версиями | Отсутствует | Интеграция с системами контроля версий, такими как Git | — |

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В результате проведенных работ, определены следующие требования к БД:

- возможность аутентификации и авторизации пользователя (для пользователя-администратора);
- быстрый поиск материалов по автору и/или названию (ключевым словам) публикации;
- отображение статистического материала: разбивка по категориям, типам материала, авторам;
- отображение последних публикаций в порядке убывания даты их публикации;
- выборка публикаций по типу научного материала за определенный год или за все годы;
- возможность просмотра списка всех публикаций для отдельного автора;
- возможность извлечения библиографической записи для каждой публикации, позволяющая получить стандартизированную информацию о документе;
- для администратора возможность редактирования и удаления записей (новые категории, авторы, публикации).

Для эффективного хранения больших объемов данных была разработана корректная структуры БД и определены ее функции и возможности. Таким образом, логическая модель БД выглядит следующим образом.

Таблица **publics**: в этой таблице хранится информация о научных публикациях. Она содержит уникальный идентификатор *Id*, который является первичным ключом (PRIMARY KEY). Также в таблице есть внешний ключ *Id_category*, который связывает публикации с категориями из таблицы **categories**.

Id (INT, PK): уникальный идентификатор публикации.

Title (VARCHAR): заголовок публикации.

DOC (VARCHAR): имя файла документа.

Type (VARCHAR): тип документа (статья, тезисы, монография и т.д.).

Journal_title (VARCHAR): название журнала.

Created_date (DATE): дата опубликования / создания научного материала.

Bibl_rec_2017 (TEXT): библиографическая запись в формате ГОСТ от 2017 года.

Bibl_rec_2003 (TEXT): библиографическая запись в формате ГОСТ от 2003 года.

Id_category (INT, FK): внешний ключ, связывающий с таблицей *categories*, указывает на категорию публикации.

Id_user (INT, FK): внешний ключ, связывающий с таблицей *users*, указывает на пользователя, который создал публикацию.

Таблица **authors**: здесь хранится информация об авторах научных публикаций. Таблица имеет уникальный идентификатор *Id* в качестве первичного ключа. Кроме того, здесь используется индекс PRIMARY KEY, который обеспечивает уникальность значений поля *Id*.

Id (INT, PK): уникальный идентификатор автора.

Fullname (VARCHAR): полное имя автора.

Job_position (VARCHAR): организация, в которой работает сотрудник.

HIRSH (TINYINT): индекс Хирша автора.

Таблица **categories**: содержит информацию о категориях научных публикаций. Каждая категория имеет уникальный идентификатор *Id*, который является первичным ключом. Также в таблице есть индекс *Name*, который обеспечивает быстрый доступ к категориям по их названиям.

Id (INT, PK): уникальный идентификатор категории.

Name (VARCHAR): название категории.

Description (TEXT): описание категории.

Таблица **users**: хранится информация о пользователях системы. В таблице есть уникальный идентификатор *Id* в качестве первичного ключа. Также в таблице есть индекс *Email*, который гарантирует уникальность адресов электронной почты пользователей.

Id (INT, PK): уникальный идентификатор пользователя.

Username (VARCHAR): имя пользователя.

Email (VARCHAR): адрес электронной почты пользователя (уникальное поле).

Password (VARCHAR): пароль пользователя.

Admin (TINYINT): флаг, определяющий администратора (1) или обычного пользователя (0).

Createdate (TIMESTAMP): дата создания записи о пользователе.

Таблица **public_authors**: представляет собой связь между научными публикациями и их авторами. Она содержит уникальный идентификатор *id* в качестве первичного ключа. Кроме того, таблица содержит два внешних ключа: *Public_id*, который связан с идентификаторами публикаций из таблицы **publics**, и *Author_id*, который связан с идентификаторами авторов из таблицы **authors**. Эти внешние ключи обеспечивают целостность данных и поддерживают связь между различными сущностями в БД.

Id (INT, PK): уникальный идентификатор связи между публикацией и автором.

Public_id (INT, FK): внешний ключ, связывающий с таблицей **publics**, указывает на публикацию.

Author_id (INT, FK): внешний ключ, связывающий с таблицей **authors**, указывает на автора.

БД использует кодировку utf8mb4 для поддержки Unicode. Используется движок InnoDB для всех таблиц, что обеспечивает поддержку транзакций и внешних ключей. Индексы созданы для обеспечения быстрого доступа к данным по ключевым полям и поддержания целостности данных.

На рисунке 1 представлены основные таблицы БД.

Главная страница БД представлена на рисунке 2.

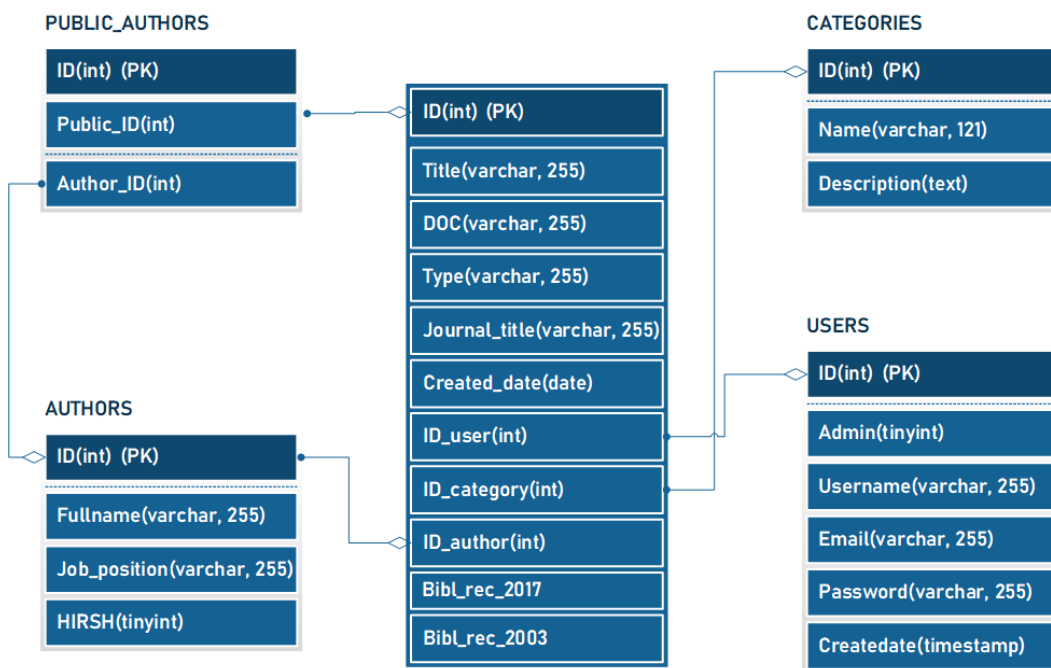


Рисунок 1. Структура БД, содержащая информацию о таблицах



Рисунок 2. Пользовательский интерфейс БД для хранения и систематизации научных публикаций

Условно, главную страницу можно разделить на 3 части: поиск, статистический материал, последние добавленные публикации. БД позволяет осуществлять поиск по автору и/или названию публикации/ключевым словам. Регистр не имеет значения, поиск осуществляется корректно в соответствии с запросом. Статистический материал наглядно показывает информацию в виде следующих категорий: Journal citations reports/Scopus; РИНЦ/Международные базы с ненулевым импакт-фактором; КОКНВО, материалы конференций; издательскую деятельность; прочие статьи и отчеты.

Рассмотрим каждую категорию отдельно.

– *Journal citations reports/Scopus*: научные работы в международных рецензируемых научных журналах, входящих в 1, 2 и 3 квартиль по данным Journal Citation Reports компании Clarivate Analytics, или имеющих в базе данных Scopus показатель процента по CiteScore хотя бы по одной из научных областей в соответствующих журналах, индексируемых в базе данных Web of Science Core (Arts and Humanities Citation Index, Science Citation Index Expanded, Social Sciences Citation Index), компании Clarivate Analytics. Такие материалы можно найти в базах Scopus и Web of Science.

– *РИНЦ/Международные базы с ненулевым импакт-фактором*: материалы, опубликованные в научных журналах, индексируемых РИНЦ и других международных базах с ненулевым импакт-фактором. Материалы чаще всего представлены на сайте e-libru.ru – крупнейшем российском информационно-аналитическом портале в области науки, технологии, медицины и образования, содержащем рефераты и полные тексты более 38 млн научных публикаций и патентов, в том числе их электронные версии [7].

– *КОКНВО*: материалы в научных изданиях, рекомендованных Комитетом по обеспечению качества в сфере образования Министерства просвещения Республики Казахстан.

– *Материалы конференций*: научные работы в материалах конференций, форумов, съездов, симпозиумов, конгрессов.

– *Издательская деятельность*: издательская деятельность включает в себя монографии, сборники трудов, буклеты и др.

– *Прочие статьи*: статьи, не попавшие ни в одну из перечисленных выше категорий.

– *Отчеты*: отчеты о научно-исследовательской деятельности.

Также, помимо категорий доступен вывод статистической информации о топ-авторах. Это те авторы, которые наиболее часто встречаются в публикациях, как «первый автор», так и «соавтор». Следующая статистика охватывает типы материалов, которые разделены на статьи, тезисы, монографии, авторефераты, сборники трудов.

Последние добавленные публикации представлены в виде карточек с отображением краткой информации о научном материале, об авторе, соавторах, названии и позволяющие скачать публикацию в формате pdf (рисунок 3).

Абсолютно всем пользователям доступна информация об авторах. При нажатии на любого автора выводится не только его личная информация, но и список всего научного материала, где он был автором или соавтором. Это относится и к категориям, и к типам материалов. При нажатии на целочисленное значение выводится список публикаций, доступных для прочтения.

Вся эта информация доступна для просмотра как обычному пользователю, так и администратору. Администратор имеет больше прав доступа и ему доступна админ-панель для занесения информации о публикациях, авторах, категориях, пользователях, редактирования и удаления этой информации. Таким образом, реализован многопользовательский режим, позволяющий регистрироваться новым пользователям и авторизовываться уже всем зарегистрировавшимся. При регистрации пользователей пароль в БД отображается уже как хешированный.

$\$pass = password_hash(\$passFirst, PASSWORD_DEFAULT)$

После выполнения этой строки переменная \$pass будет содержать хешированный пароль, который можно сохранить в БД или использовать для проверки при аутентификации пользователей. Своего рода, это один из методов защиты данных в приложениях, особенно в контексте безопасности паролей пользователей. Хеширование пароля делает его нечитаемым для посторонних лиц, даже если база данных с хешами паролей будет скомпрометирована. На рисунке 4 показана панель управления, доступная для администратора.

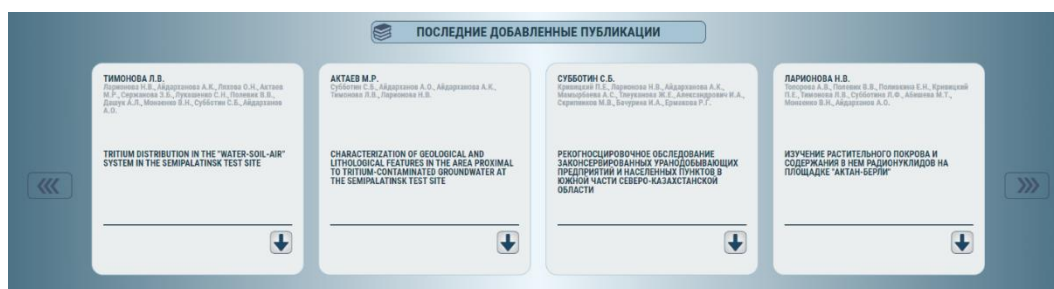


Рисунок 3. Последние добавленные публикации

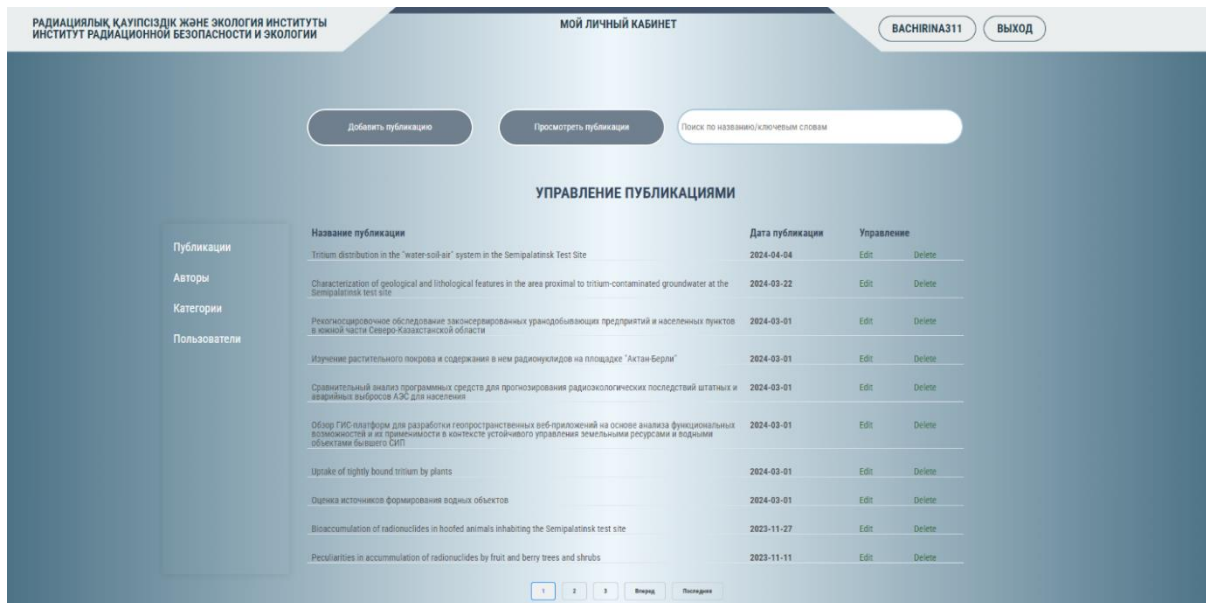


Рисунок 4. Панель администратора для управления публикациями

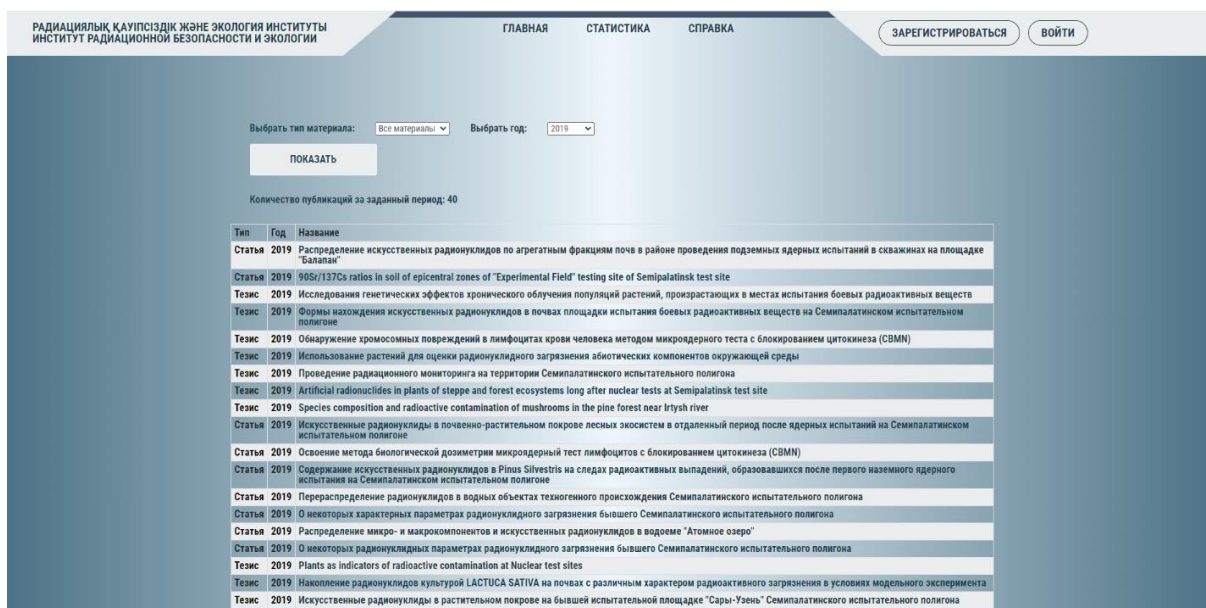


Рисунок 5. Научный материал Института за 2019 год

Отдельно имеется возможность предоставления статистических данных по типам материала за определенный промежуток времени, начиная с 1993 года по настоящее время, а также «все года». Это сделано для отображения количественных показателей, как результата деятельности за определенный период. Например, ниже представлен вывод всего научного материала Института радиационной безопасности и экологии за 2019 год (рисунок 5).

БД является эффективным инструментом для анализа данных, поскольку обладает широким спектром

функций. Так, например, с помощью БД можно определить количество научного материала Института за последние 5 лет, начиная с 2020 по 2024 годы (рисунок 6). Из полученных данных видно, что пик публикаций приходится на 2021 год, что, вероятно, связано с мировыми событиями, в частности, с периодом после коронавируса. Увеличение числа тезисов, в свою очередь, может быть связано с участием в различных конференциях и семинарах сотрудников Института.

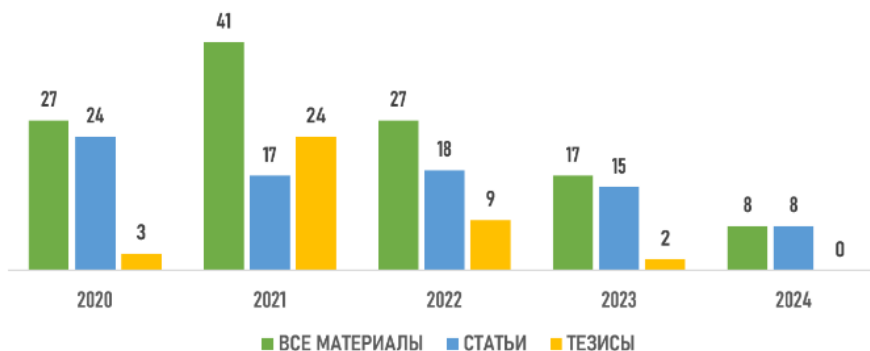


Рисунок 6. Статистика основных публикаций за последние 5 лет

Также с помощью инструментов БД можно оценить распределение публикаций по категориям. Для анализа были выбраны следующие типы научного материала: статьи; материалы конференций; издательская деятельность. Статьи в свою очередь подразделяются на следующие категории: Journal citations reports/Scopus; РИНЦ/Международные базы с ненулевым импакт-фактором; КОКНВО (визуализация представлена на рисунке 7). Из полученных статистических данных видно, что чаще всего научный материал публикуется в виде статей в КОКНВО – научных изданиях, рекомендованных Комитетом по обеспечению качества в сфере образования Министерства просвещения Республики Казахстан.

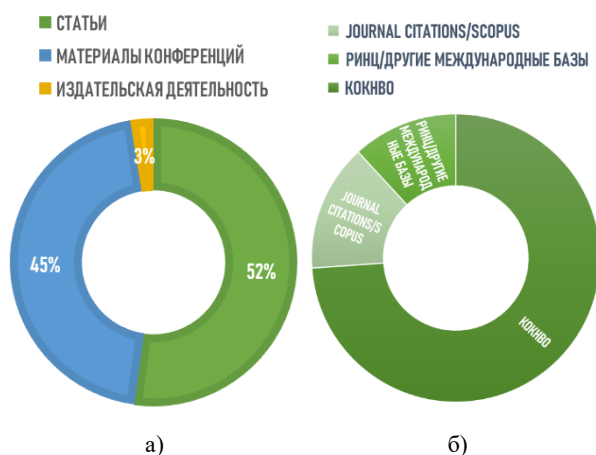


Рисунок 7. Распределение научного материала: публикаций по типам материала (а); статей по категориям (б)

ЗАКЛЮЧЕНИЕ

По результатам анализа научного материала Института радиационной безопасности и экологии, собранного за 30-летний период его деятельности, были определены ключевые функциональные требования к БД, направленные на хранение, организацию и анализ данных для обеспечения эффективного использования научной информации. В соответствии с требованиями к функциональности была организована структура БД, где определены таблицы, их поля и связи. Благодаря средствам (Ampps, Visual Studio Code, Google Chrome) и технологиям (HTML, CSS,

JS, PHP), спроектирован пользовательский интерфейс, который обеспечивает корректное и эффективное взаимодействие пользователя и сервера. Разработанный пользовательский интерфейс предоставляет удобные средства для ввода, редактирования и поиска данных, а также обеспечивает наглядное отображение результатов. Определено, что БД является эффективным инструментом для анализа данных, поскольку обладает широким спектром функций, позволяющих проводить разнообразные виды анализа. Среди ключевых функций БД можно выделить возможность выборки данных в соответствии с заданными критериями, их сортировки по различным параметрам и извлечения для последующего анализа. Благодаря этим функциям пользователи могут эффективно проводить аналитическую работу, выявлять закономерности, тенденции и взаимосвязи в данных, что делает БД важным инструментом в исследованиях и принятии решений.

Работа выполнена при финансовой поддержке Министерства образования и науки Республики Казахстан в рамках научной программы BR21882086 «Разработка устойчивого управления земельными ресурсами и водными объектами на территории бывшего Семипалатинского испытательного полигона».

ЛИТЕРАТУРА / REFERENCES

1. Бачурина И.А. База данных, как электронный архив научной деятельности Института радиационной безопасности и экологии / Бачурина И.А., Топорова А.В., Ларионова Н.В. // Тезисы. X Международная конференция «Семипалатинский испытательный полигон: наследие и перспективы развития научно-технического потенциала», 12-14 сентября 2023, стр. 110. [Bachurina I.A. Baza dannykh, kak elektronnyy arkhiv nauchnoy deyatelnosti Instituta radiatsionnoy bezopasnosti i ekologii / Bachurina I.A., Toporova A.V., Larionova N.V. // Abstracts. X International Conference “Semipalatinsk Test Site: Legacy and Prospects for Scientific-Technical Potential Development”, September 12–14, 2023, p. 110. (In Rus.)]

2. National Library of Medicine PubMed [Электронный ресурс]. – 2024. – URL: <https://pubmed.ncbi.nlm.nih.gov/> (дата обращения 09.04.2024).
3. IEEE Xplore Advancing technology for humanity [Электронный ресурс]. – 2024. – URL: <https://www.ieee.org/about/vision-mission.html> (дата обращения 09.04.2024).
4. Scopus: Comprehensive, multidisciplinary, trusted abstract and citation database [Электронный ресурс]. – 2024. – URL: https://www.elsevier.com/products/scopus?dgcid=RN_AGCM_Sourced_300005030 (дата обращения 09.04.2024).
5. Ampps. Docs [Электронный ресурс]. – 2024. – URL: <https://ampps.com/docs/> (дата обращения 12.04.2024).
6. Visual Studio Code. Learn to code with Visual Studio Code [Электронный ресурс]. – 2024. – URL: <https://code.visualstudio.com/learn> (дата обращения 12.04.2024).
7. Научная электронная библиотека e-library.ru. – 2024. – URL: <https://www.elibrary.ru/defaultx.asp> (дата обращения 12.04.2024) [Nauchnaya elektronnaya biblioteka e-library.ru. – 2024. – URL: <https://www.elibrary.ru/defaultx.asp> (date of access 12.04.2024) (In Rus.)].

ҒЫЛЫМИ ЖАРИЯЛАНЫМДАРДЫ САҚТАУ ЖӘНЕ ТАЛДАУ ҮШІН ДЕРЕКТЕР БАЗАСЫН ӘЗІРЛЕУ

И. А. Бачурина*, Н. В. Ларионова, А. В. Топорова

ҚР ҰАО РМК «Радиациялық қауіпсіздік және экология институты» филиалы, Курчатов, Қазақстан

**Байланыс үшін E-mail: bachurina@nnc.kz*

Мақалада ғылыми жарияланымдарды тиімді сақтау, ұйымдастыру және талдау үшін деректер базасына (ДБ) қойылатын негізгі функционалдық талаптар қарастырылады. Радиациялық қауіпсіздік және экология институтының ғылыми қызметінің 30 жылдық кезеңіндегі материалдарды талдау негізінде, сондай-ақ пайдаланушылардың мүдделерін ескере отырып, ДБ функционалдығы мен құрылымына қойылатын талаптар айқындалды. Нәтижесінде пайдаланушы мен сервердің сенімді өзара әрекеттесуін қамтамасыз ететін құралдар мен технологияларды пайдалана отырып, ДБ жобаланды және әзірленді. Жасалған пайдаланушы интерфейсі деректермен жұмыс істеудің ыңғайлы құралдарын ұсынады. Деректер базасы деректерді талдаудың тиімді құралы болып табылады, өйткені ол талдаудың әртүрлі түрлерін жүргізуге мүмкіндік беретін көптеген функцияларға ие: берілген критерийлерге сәйкес деректерді іріктеу мүмкіндігі, оларды әртүрлі параметрлер бойынша сұрыптау және кейінгі талдау үшін алу. Осы мүмкіндіктердің арқасында пайдаланушылар аналитикалық жұмысты тиімді жүргізе алады, мәліметтердегі заңдылықтарды, тенденциялар мен қатынастарды анықтай алады, бұл деректер базасын зерттеу мен шешім қабылдауда қажетті құралға айналдырады.

Түйін сөздер: деректер базасы, жарияланымдар, мақалалар, клиент-сервер қосымшасы, пайдаланушы интерфейсі.

DEVELOPMENT OF A DATABASE TO STORE AND ANALYZE SCIENTIFIC PUBLICATIONS

I. A. Bachurina*, N. V. Larionova, A. V. Toporova

RSE NNC RK Branch “Institute of Radiation Safety and Ecology”, Kurchatov, Kazakhstan

** E-mail for contacts: bachurina@nnc.kz*

The article addresses key functional database (DB) requirements for the effective storage, arrangement and analysis of scientific publications. Based upon the materials analyzed over the 30-year period of scientific activities of the Institute of Radiation Safety and Ecology and considering users' interests, requirements for the functionality and structure of a database have been defined. As a result, a database was designed and developed using tools and technologies that ensured a secure interaction between a user and a server. The user interface created provides convenient tools for working with data. DB was found to be an efficient tool to analyze data since it has a wide range of functions enabling to carry out a variety of analyses: a possibility to retrieve data pursuant to specified criteria, sort it out by different parameters and select followed by the analysis. Thanks to these functions, users are able to carry out analytical activities efficiently, reveal regularities, tendencies and interrelations in data, which makes DB an indispensable tool in research and decision-making.

Keywords: database, publications, articles, client-server application, user interface.