

СРАВНЕНИЕ РЕЛЯЦИОННЫХ И ДОКУМЕНТНЫХ БАЗ ДАННЫХ ДЛЯ ХРАНЕНИЯ ЦИФРОВОГО СЛЕДА В ОБРАЗОВАНИИ

В данной статье представлен комплексный сравнительный анализ реляционных и документных баз данных в контексте управления цифровыми следами в образовании. Исследование углубляется в отличительные характеристики реляционных и документных баз данных, имеющими схожие характеристики и принципиальные отличия в использовании, целеполагании применения. В работе анализируются различные критерии обеих архитектур баз данных с учетом тонкостей данных, включая формат хранения цифрового следа в образовательной среде, источники сбора цифрового следа, целостность данных, масштабируемость и гибкость разработки. Каждый из параметров систематически оценивается и подлежит сравнению, чтобы стало возможным сильные и слабые стороны каждого типа базы данных в образовательном контексте. В статье также подчеркиваются практические выводы, полученные в результате анализа, проливающие свет на сценарии, в которых одна архитектура баз данных может превосходить другую, и определяющие потенциальные варианты использования гибридного подхода.

База данных, цифровой след, образование, анализ преподавательской деятельности.

Kopteva Anna Valerievna., Shadrina Valentina V.

СРАВНЕНИЕ РЕЛЯЦИОННЫХ И ДОКУМЕНТНЫХ БАЗ ДАННЫХ ДЛЯ ХРАНЕНИЯ ЦИФРОВОГО СЛЕДА В ОБРАЗОВАНИИ

This article presents a comprehensive comparative analysis of relational and document databases in the context of digital footprint management in education. The study delves into the distinctive characteristics of relational and document databases,

which have similar characteristics and fundamental differences in use and purpose of application. The paper analyzes the various criteria of both database architectures, taking into account the intricacies of the data, including the format for storing the digital footprint in the educational environment, sources of digital footprint collection, data integrity, scalability and development flexibility. Each of the parameters is systematically assessed and compared so that the strengths and weaknesses of each type of database in an educational context become possible. The article also highlights the practical implications of the analysis, shedding light on scenarios in which one database architecture may be superior to another and identifying potential use cases for a hybrid approach.

Database, digital footprint, education, analysis of teaching activities.

Основная часть

Стандарт сбора цифрового следа в образовании разработали и утвердили в Университете 20.35. «Университет 20.35» – финансируемая государством онлайн-платформа, аккумулирующая образовательный контент на темы, связанные со сквозными технологиями цифровой экономики. По словам разработчиков данной платформы, цифровой след собирается через аудио- и видеозаписи Zoom-конференций, чат-логи, данные об активности в рамках проектной работы, реакции учащихся, взаимные оценки и пр.

При формировании цифрового следа профессиональной и образовательной деятельности человека учитываются данные диагностики, образовательного содержания, процесса, индивидуального образовательного опыта, достижения, словом, тот багаж знаний, который человек собирает на протяжении всей жизни.

Цифровой след в образовательной среде – это коллективный след студентов и преподавателей, который состоит из целенаправленно собранных записей определенных действий в сети, взаимодействия учащихся или преподавателей, а также их вклада в коллективное развитие и в научную деятельность. Он охватывает данные и информацию, генерируемые и оставленные в результате взаимодействия людей с цифровыми платформами, образовательными инструментами, корпоративными сетями и другими онлайн-ресурсами в контексте образования [1].

Ключевые компоненты цифрового следа в образовательной среде включают в себя 4 категории.

Первая – онлайн-взаимодействие. Деятельность на образовательных веб-сайтах, в системах управления обучением, дискуссионных форумах и

платформах для корпоративного сотрудничества. Дополнительно к такого рода взаимодействию можно отнести участие в обсуждениях в аудитории офлайн и онлайн, отправку заданий и использование учебных материалов онлайн [2].

Вторая – использование цифровых инструментов: интеграция технологий в образовательный процесс предполагает использование различных цифровых инструментов, приложений и программного обеспечения. Каждое взаимодействие с этими инструментами, будь то в классе или во время дистанционного обучения, увеличивает цифровой след (BPC, MS Teams, MS Outlook и т.д.).

Третья – профессиональные (когнитивные) профили: создание когнитивного портрета каждого из педагогов может помочь в формировании коллективного цифрового следа. Эти профили демонстрируют квалификацию преподавателя, его опыт работы и научной деятельности, а также вклад в эту область.

Четвертая – проектирование базы данных для хранения данных о цифровых следах в сфере образования требует тщательного рассмотрения типов задействованных данных, вопросов конфиденциальности и безопасности, масштабируемости и конкретных потребностей образовательного учреждения [3].

Ниже приведены основные требования к базе данных для хранения данных о цифровых следах в образовательных учреждениях:

- 1) схема и структура данных: необходимо определить четко и хорошо организованную схему базы данных, отражающую форматы данных цифрового следа, которые будут храниться для дальнейшей интерпретации;
- 2) масштабируемость: необходимо учесть потенциальное увеличение объема информации, числа пользователей, дисциплин и т.д.;
- 3) нормализация: важно нормализовать базу данных, чтобы минимизировать избыточность данных и улучшить согласованность данных;
- 4) меры безопасности: критично важно внедрить надежные меры безопасности для защиты конфиденциальной информации;
- 5) резервное копирование и восстановление: нужно установить регулярные процедуры резервного копирования для предотвращения потери данных;
- 6) протоколы аудита: внедрение журналов аудита для регистрации изменений в базе данных;

7) управление параллелизмом: важно внедрить механизмы управления параллелизмом для одновременного доступа и изменения данных несколькими пользователями;

8) интеграция с другими системами: важно иметь возможность интегрировать базу данных по сбору и анализу цифрового следа с другими образовательными системами;

9) аутентификация и авторизация пользователя: требуется внедрить строгие механизмы аутентификации и авторизации пользователей для контроля доступа к базе данных, а также иерархию прав – кто имеет право на просмотр, кто на редактирование [4].

Данные требования к структуре базы данных для хранения артефактов цифрового следа в образовании критически важны при разработке надежной и безопасной системы хранения данных.

Реляционные базы данных можно эффективно использовать для хранения и обработки цифрового следа в образовании. Цифровые следы в образовании охватывают широкий спектр данных, генерируемых студентами, преподавателями и образовательными системами посредством взаимодействия в онлайн-среде и не только.

Данный тип баз данных предоставляет структурированный и организованный способ хранения, управления и извлечения разнообразных форматов данных [5].

Реляционные базы данных хорошо подходят для хранения и обработки цифровых следов в сфере образования благодаря нескольким ключевым характеристикам и преимуществам:

1) модель структурированных данных: РБД используют структурированную модель данных, организовывая информацию в таблицы с заранее определенными связями;

2) целостность данных: реляционные РБД обеспечивают целостность данных посредством ограничений, гарантируя, что данные соответствуют предопределенным правилам;

3) отношения и объединения: реляционная модель облегчает установление отношений между различными объектами;

4) гибкие запросы: Реляционные базы данных предоставляют стандартизированный язык запросов (SQL), который позволяет пользователям гибко извлекать, фильтровать и агрегировать данные;

5) масштабируемость: РБД могут масштабироваться для обработки широкого спектра объемов данных. Поскольку цифровой след в образова-

нии может включать в себя значительный объем данных, масштабируемость реляционных баз данных гарантирует, что они смогут удовлетворить растущие информационные потребности образовательных учреждений;

6) меры безопасности: РБД предлагают надежные механизмы безопасности, включая аутентификацию, авторизацию и шифрование пользователей;

7) нормализация данных: РБД поддерживают нормализацию данных, уменьшая избыточность и повышая эффективность;

8) согласованное резервное копирование и восстановление: РБД предоставляют надежные механизмы регулярного резервного копирования и восстановления;

9) интеграция с другими системами: РБД могут быть легко интегрированы с другими системами управления в ВУЗе, обеспечивая единую и связную среду данных;

10) историческая справка и анализ: РБД позволяют хранить исторические данные, позволяя образовательным учреждениям отслеживать изменения с течением времени. Это ценно для анализа тенденций, оценки эффективности образовательных инициатив и получения комплексного представления о цифровом следе [6].

Суммируя выше сказанное, можно сделать промежуточный вывод, что структурированная реляционная природа реляционных баз данных делает их подходящими для обработки разнообразных и взаимосвязанных данных, генерируемых в рамках цифрового следа в образовании. Их способность обеспечивать целостность данных, поддерживать гибкие запросы и обеспечивать надежные меры безопасности, делает их надежным выбором для образовательных учреждений, управляющих и анализирующих цифровые следы, но по-прежнему необходимо учесть все минусы данного вида БД.

Документные базы данных могут использоваться для хранения и обработки цифрового следа в образовании, особенно при работе с полуструктурированными или неструктурированными данными.

Документные базы данных относятся к категории баз данных NoSQL (не только SQL) и предназначены для обработки широкого спектра форматов данных, включая JSON, XML и другие форматы, ориентированные на документы [7].

Основные преимущества ДБД, отличающие их от баз данных с реляционной структурой:

1) гибкость с полуструктурированными данными: цифровые следы часто включают в себя разнообразные и развивающиеся типы данных, такие как взаимодействия в корпоративных сетях, выполнение внутренних заданий, отчеты в форматах .excel и .word, а также – мультимедийный контент;

2) поддержка изменчивости в структуре данных: ДБД могут адаптироваться к различным форматам документов в одной структуре;

3) масштабируемость: базы данных документов разработаны с возможностью горизонтального масштабирования, что упрощает обработку больших объемов данных;

4) производительность операций чтения и записи: ДБД могут обеспечивать отличную производительность как для операций чтения, так и для операций записи, что делает их подходящими для сценариев, где важен быстрый поиск и обновление данных [8];

5) расширенные запросы и индексирование: ДБД предоставляют мощные возможности запросов, позволяющие извлекать конкретную информацию из сложных документов;

6) интеграция с облачными сервисами: ДБД часто проектируются для беспрепятственной работы с облачными сервисами;

7) динамические обновления и эволюция моделей данных: ДБД могут легко поддерживать динамические обновления структур данных, не требуя заранее определенной схематизации;

8) простота разработки: ДБД удобны для разработки и предлагают более интуитивный и гибкий подход к моделированию данных [9].

Хотя базы данных документов предполагают преимущества для определенных аспектов управления цифровым следом, выбор между базой данных документов и реляционной базой данных зависит от конкретных требований образовательного учреждения, характера данных и общей архитектуры предполагаемой интеллектуальной системы управления данными внутри структуры ВУЗа. Чтобы определить основные отличия между ДБД и РБД, важно наглядно представить их отличия.

В табл. 1 показаны ключевые сравнительные характеристики реляционных баз данных и баз данных документов:

Таблица 1

Сравнение Реляционной базы данных и документной с целью хранения и обработки цифровых следов в образовании

| Критерии | Реляционная БД | Документная БД |
|---------------------------|--|--|
| Модель данных | Таблицы со структурированными отношениями | Сборка документов разных форматов |
| Схема | Фиксированная схема, predetermined структура | Динамическая схема, адаптируемая к изменениям |
| Целостность данных | Свойства ACID (Атомичность, Консистенция, Изоляция, Долговечность) | Может пожертвовать некоторыми свойствами ACID для гибкости (например, возможной последовательности) |
| Язык Запросов | SQL (Структурированный Язык Запросов) | Языки запросов, специфичные для базы данных (например, язык запросов MongoDB для MongoDB) |
| Сложные Отношения | Хорошо подходит для сложных отношений | Может обрабатывать более простые отношения, но не высокая оптимизация для обширных соединений |
| Масштабируемость | Вертикальная масштабируемость (добавление большего количества ресурсов на один сервер) | Горизонтальная масштабируемость (добавление большего количества серверов для распределения нагрузки) |
| Типы данных | В первую очередь подходит для структурированных типов данных | Подходит для структурированных и неструктурированных типов данных |
| Изменения в схеме | Изменения схемы мо- | Изменения схемы могут |

| | | |
|---------------------------------------|--|---|
| | гут быть сложными и требуют тщательного планирования | быть более плавными и приниматься в моменте |
| Скорость разработки | Может потребоваться предварительный дизайн и планирование | Гибкая и более быстрая разработка и прототипирование |
| Требования к согласованности | Сильная последовательность по умолчанию | Может обеспечить опосредованную согласованность, в зависимости от конфигурации |
| Интеграция с другими системами | Хорошо зарекомендовавшая себя интеграция и зрелая экосистема | Интеграции могут варьироваться; растущая экосистема, часто используемая в современных облачных архитектурах |
| Сценарии использования | Хорошо подходит для приложений с четко определенными отношениями и там, где согласованность данных имеет решающее значение | Подходит для приложений с развивающимися или разнообразными структурами данных, высокими потребностями в масштабируемости и быстрыми циклами разработки |
| Среды реализации | MySQL, PostgreSQL, Oracle | MongoDB, CouchDB, Firebase Firestore |

Выводы

Из проанализированных данных следует, что выбор базы данных складывается из следующих параметров: изначальная неструктурированность данных, разнообразие форматов и рассредоточение данных на разных устройствах и в цифровых ресурсах.

Исходя из этих критически важных параметров, при отсутствии четко структурированных данных, для хранения и исследования целесообразно использование документного типа баз данных.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Мартишин С.А., Симонов В.Л., Храпченко М.В.* Проектирование и реализация баз данных в СУБД MySQL с использованием MySQL Workbench. Учебное пособие. – М.: Форум, 2018. – 160 с.
2. *Мартишин С.А., Симонов В.Л., Храпченко М.В.* Базы данных. Практическое применение СУБД SQL- и NoSQL-типа для применения проектирования информационных систем. – М.: Форум, 2018. – 368 с.
3. *Илющечкин В.М.* Основы использования и проектирования баз данных. Учебник. – М.: Юрайт, 2017. – 214 с.
4. *Парфенов Ю.П.* Постреляционные хранилища данных. Учебное пособие. – М.: Юрайт, Издательство Уральского Университета, 2017. – 122 с.
5. *Попова Е.П., Решетникова К.В.* Основы использования и проектирования баз данных. Учебник. – М.: Юрайт, 2017. – 214 с.
6. *Мартишин С.А., Симонов В.Л., Храпченко М.В.* Базы данных. Практическое применение СУБД SQL и NoSQL. Учебное пособие. – М.: Форум, Инфра-М, 2016. – 368 с.
7. *Чаллавава Ш., Лакхатария Д., Мехта Ч., Патель К.* MySQL 8 для больших данных. – М.: ДМК Пресс, 2018. – 226 с.
8. *Шустова Л.И., Тараканов О.В.* Базы данных. Учебник. – М.: Инфра-М, 2018. – 304 с.
9. *Нестеров С.* Базы данных. Учебник и практикум. – М.: Юрайт, 2017. – 230 с.