

УДК 004.41

Богданович А.О.

учащийся УО «Национальный Детский Технопарк»,
г. Минск, РБ

Нистюк Д.А.

учащийся УО «Национальный детский технопарк»,
г. Минск, РБ

Андрейчук О.Н.

преподаватель высшей категории предметов общепрофессионального
и специального циклов УО «БГУИР» филиал «МРК», м.т.н.,
г. Минск, РБ

Середа И.А.

старший преподаватель кафедры ИИТ УО «БГУИР», м.т.н.,
г. Минск, РБ

Сницко В.А.

старший преподаватель кафедры ИИТ УО «БГУИР», м.т.н.,
г. Минск, РБ

АНАЛИЗ АНАЛОГОВ И РАЗРАБОТКА ВЕБ - ПОРТАЛА НАУЧНЫХ КОНФЕРЕНЦИЙ ПО ИНТЕРЕСАМ ПОЛЬЗОВАТЕЛЯ

Аннотация

В статье представлен процесс разработки и реализации веб - портала научных конференций, ориентированного на мероприятия в Республике Беларусь. Описаны методы

сбора данных из разнородных источников, включая устаревшие сайты вузов. Особое внимание уделено решению ключевой проблемы автоматизации — извлечению структурированной информации (научное направление, тип мероприятия) из заголовков с применением большой языковой модели Gemini 2.0 Flash. Приводятся технические аспекты реализации парсинга, дизайна адаптивного веб - интерфейса и развертывания работающего ресурса.

Ключевые слова

Научные конференции, конференции, искусственный интеллект, Gemini 2.0 Flash, парсинг.

Введение. В современном мире человек сталкивается с огромным объёмом информации в Интернете, однако на её поиск и обработку требуется много времени. Цель данной работы – разработать ориентированный на пользователя веб - портал, который станет агрегатором информации о большинстве конференций, проходящих на территории Республики Беларусь. Разработка такого веб - портала позволит ученикам и преподавателям эффективнее находить информацию о предстоящих научных конференциях на территории Республики Беларусь, не затрачивая время на поиск информации на различных веб - сайтах и ручной отбор по нужным критериям.

Основная часть. В процессе поиска и анализа действующих сайтов–агрегаторов, содержащих информацию о проводимых научных конференциях в различных высших учебных заведениях, особое внимание было уделено ресурсам «Конференции.RU» и «Научные конференции России».

Следует отметить, что оба упомянутых ресурса обладают технически устаревшим интерфейсом, что может затруднять эффективный поиск и восприятие информации для современных пользователей.

Помимо двух ранее упомянутых сайтов–агрегаторов научных конференций, были также выявлены веб–сайты отдельных высших учебных заведений, содержащие информацию о конференциях, организуемых непосредственно этими учреждениями (например, Белорусский государственный университет, Белорусско - Российский университет, Национальная академия наук Беларуси и другие). Данные интернет–ресурсы будут использованы в дальнейшем для парсинга информации в единую базу данных.

Также был обнаружен документ в формате Word, опубликованный Министерством образования Республики Беларусь и содержащий таблицу с большим количеством научных конференций, запланированных к проведению в 2025 году в различных белорусских учебных заведениях. Несмотря на значительный объём записей, данный документ не охватывает большинство существующих конференций. Тем не менее, он будет использован для парсинга и последующего наполнения базы данных.

В процессе парсинга собираемая информация записывается в базу данных в следующем формате: уникальный идентификатор каждого мероприятия, полное название проводимого мероприятия, список научных направлений, присутствующих на конференции, название учебного заведения (организатора), которое проводит мероприятие, форма проведения мероприятия: очная, заочная или очно - заочная, название города, в котором будет проводиться мероприятие, даты проведения мероприятия (некоторые мероприятия проходят в один конкретный день, а некоторые растягиваются на несколько, поэтому важно хранить именно диапазон дат вместо одной конкретной), номера контактных телефонов, адреса электронной почты и факс, тип мероприятия (например, «Конкурс» или «Конференция»).

Однако в рассматриваемой таблице отсутствует ряд значимых сведений, в частности — научное направление и тип мероприятия. Эти данные, как правило, могут быть интуитивно

распознаны человеком на основе анализа названия конференции, однако автоматическое извлечение данной информации с использованием традиционных алгоритмов парсинга практически невозможно. С учётом стремительного развития технологий искусственного интеллекта, данную задачу целесообразно решать с применением больших языковых моделей (LLM) - нейронных лингвистических сетей, обученных на огромных корпусах данных для понимания и обработки текста.

Использование искусственного интеллекта позволяет реализовать довольно точное автоматическое извлечение информации о научной направленности и типе мероприятия непосредственно из заголовков (например, конференция, семинар, симпозиум и т.д.). В данном проекте использована большая языковая модель Gemini 2.0 Flash, разработанная компанией Google. Модель предоставляет условно бесплатный API, а также обладает гибкими возможностями формирования запросов, что делает её подходящей для решения поставленной задачи.

После получения ответа от языковой модели парсер объединяет извлечённую информацию с ранее подготовленными данными и сохраняет готовые мероприятия в базу данных. Описанный процесс парсинга — от извлечения наименования мероприятия до генерации информации о научной направленности, типе мероприятия и последующей записи данных в базу — выполняется для каждого из собранных открытых источников каждую ночь с целью поддержания актуальности базы данных.

Веб - интерфейс разработан с использованием архитектуры Single Page Application. Использование данной архитектуры позволяет избежать долгой загрузки при навигации по веб - portalу.

Для создания комплексных элементов пользовательского интерфейса, таких как слайдер источников и карточки конференций, предварительно разрабатывались макеты в графическом редакторе Figma. Это позволило на ранних этапах продумать логику взаимодействия пользователя с интерфейсом, а также определить визуальный стиль и компоновку элементов до непосредственной реализации в коде.

Особое внимание при разработке дизайна было уделено адаптивности интерфейса. Все элементы страницы корректно отображаются на экранах различных размеров — от мобильных устройств до широкоформатных мониторов. Это реализовано с использованием медиа - запросов, изменяющих стили в зависимости от размера экрана, flexbox и разбивки веб - страницы на компоненты, что обеспечивает удобство использования ресурса вне зависимости от устройства (рис. 3).

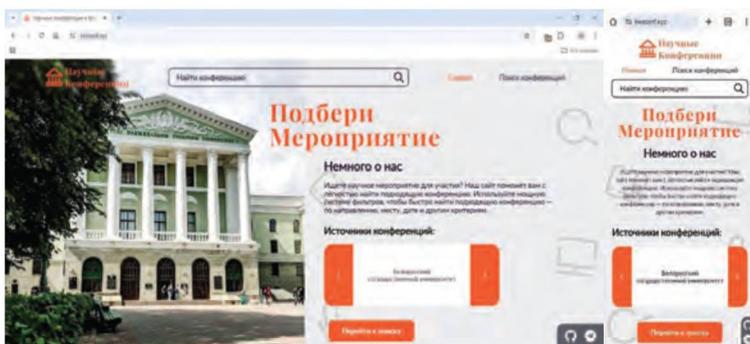


Рис. 3. Внешний вид веб - сайта на широкоформатном экране и на экране смартфона

Веб - портал может быть размещён в открытом доступе. Для этого необходимо арендовать сервер у одного из хостинг - провайдеров. Оптимальным решением для рассматриваемого проекта является использование виртуального выделенного сервера. Виртуальный выделенный сервер (VDS) — это услуга виртуального хостинга, при которой клиенту выделяется виртуальный сервер целиком с полными административными правами, которые дают возможность установить на сервер любое программное обеспечение. Функционально виртуальный сервер ничем не отличается от физического.

Помимо аренды сервера, для обеспечения удобного доступа пользователей к веб - portalу необходимо зарегистрировать доменное имя. Доменные имена – ключевая составляющая инфраструктуры Интернета. Они предоставляют человеку - читаемые адреса веб - серверов, доступных в Интернете. В рамках данного проекта доменное имя было привязано к арендованному серверу с использованием сервиса Cloudflare, который предоставляет расширенные возможности управления DNS - записями, а также обеспечивает дополнительную защиту сайта от распределённых атак типа DDOS.

Для открытия сайта с использованием протокола HTTPS и не определялся браузером как небезопасный, необходимо настроить SSL - сертификат. SSL - сертификат — это цифровой сертификат, удостоверяющий подлинность веб - сайта и позволяющий использовать зашифрованное соединение. Использование HTTPS не только повышает доверие пользователей, но и положительно влияет на ранжирование сайта в поисковых системах. В рамках данного проекта сертификат был получен бесплатно от центра сертификации Let's Encrypt с использованием утилиты Certbot, которая автоматизирует процесс получения, установки и обновления сертификатов. Это позволило быстро и надёжно настроить защищённое соединение.

Разработанный веб - ресурс был опубликован по адресу <https://belconf.xyz/>, что дало возможность всем желающим получить к нему доступ без необходимости установки дополнительного программного обеспечения. Для оценки удобства и стабильности работы ресурса на различных устройствах было проведено тестирование на настольных компьютерах, ноутбуках, планшетах и смартфонах с экранами разных размеров и разрешений.

Заключение. Разработанный проект обладает большим потенциалом для дальнейшего развития. Ниже приведены перспективные направления, в которых можно улучшать и расширять функциональность веб - ресурса:

- увеличение количества обрабатываемых открытых источников. Это возможно за счёт систематического мониторинга образовательных и научных веб - ресурсов, публикующих сведения о предстоящих мероприятиях. Такое расширение обеспечит более высокую полноту и актуальность собираемых данных;

- выход за пределы Республики Беларусь. Разработанный проект можно масштабировать за счёт добавления конференций, проводимых в соседних странах, например, в Российской Федерации. Можно начать с таких городов, как Москва или Санкт - Петербург;

Источник: разработано авторами

© Богданович А.О., Нистюк Д.А., Андрейчук О.Н., Серета И.А., Сицко В.А., 2025