Бюджетное учреждение высшего образования Ханты-Мансийского автономного округа – Югры «Сургутский государственный университет»

Политехнический институт

Кафедра информатики и вычислительной техники

**Реферат**

«Платформа для сбора и краткого аннотирования новостей в области IT»

Работу выполнил студент группы 607-01

Кузнецов Кирилл Владимирович

Проверил работу преподаватель:

Столбов Дмитрий Александрович

Сургут

2023 г.

СОДЕРЖАНИЕ

[СЛОВАРЬ ТЕРМИНОВ 3](#_Toc154177531)

[ВВЕДЕНИЕ 5](#_Toc154177532)

[АКТУАЛЬНОСТЬ 14](#_Toc154177533)

[ПОСТАНОВКА ЗАДАЧИ 15](#_Toc154177534)

[ДОПОЛНИТЕЛЬНАЯ СПЕЦИФИКАЦИЯ 17](#_Toc154177535)

[КОНЦЕПТ ПРОЕКТА 21](#_Toc154177536)

[Прецедент: Просмотр новостей на веб-ресурсе 23](#_Toc154177537)

[Прецедент: Поиск новостей на веб-ресурсе 25](#_Toc154177538)

[Прецедент: Отметить Понравившиеся Статьи 26](#_Toc154177539)

[Прецедент: Добавить Статьи в Закладки 27](#_Toc154177540)

[Использование веб-платформы 28](#_Toc154177541)

[1. Читатель (Пользователь) 38](#_Toc154177542)

[2. Интерфейс Пользователя (UI) 39](#_Toc154177543)

[3. Сервер Flask API (BackEnd) 39](#_Toc154177544)

[4. Искусственный Интеллект (AI) 39](#_Toc154177545)

[5. Парсер 39](#_Toc154177546)

[ЗАКЛЮЧЕНИЕ 41](#_Toc154177547)

# СЛОВАРЬ ТЕРМИНОВ

|  |  |  |
| --- | --- | --- |
| **Термин** | **Определение** | **Синонимы** |
| API | Интерфейс программирования приложений, набор методов и инструментов для взаимодействия между программами | Application Programming Interface (API), прикладной программный интерфейс |
| REST | Архитектурный стиль разработки веб-приложений, основанный на принципах передачи состояния и взаимодействия между клиентом и сервером | Representational State Transfer (REST) |
| Агрегаторы | Сервисы или приложения, собирающие и объединяющие информацию из разных источников | Aggregators |
| Искусственный интеллект | Область компьютерных наук, изучающая создание умных программ и устройств, способных имитировать человеческое мышление | Artificial Intelligence (AI) |
| Категории | Группы или разделы, в которые объединяются объекты по общим признакам | Categories, классификации |
| Парсер | Программа или скрипт, выполняющий сбор и анализ данных | Parser, анализатор, синтаксический анализатор |
| Парсинг | Автоматический анализ и извлечение данных из веб-страниц или текстовых документов | Parsing |
| Персональные рекомендации | Индивидуальные рекомендации, основанные на предпочтениях пользователя | Personalized Recommendations |
| Платформа | Система, на базе которой строится программное обеспечение | Platform, платформа разработки, программная среда |
| Сервисы | Веб-приложения или онлайн-ресурсы, предоставляющие определенные услуги | Services |
| Фильтры | Критерии для отбора или сортировки данных | Filters, фильтрация, отбор данных |

# ВВЕДЕНИЕ

В современном мире информационные технологии играют ключевую роль. Информационные технологии (IT) являются одной из самых быстроразвивающихся отраслей в мире современных технологий. С каждым днем в этой области происходят значительные изменения и инновации. С увеличением числа компьютеров и смартфонов растет и интерес пользователей к новостям из мира IT. В условиях постоянно меняющейся технологической среды доступ к актуальной информации становится важным для широкой аудитории.

Целью проекта, создание информационной системы в данном контексте заключается в обеспечении пользователей доступом к надежным, актуальным и интересным новостям из области информационных технологий. Эта система призвана удовлетворить потребности как профессионалов IT-сферы, так и пользователей, интересующихся последними трендами в мире технологий. Информационная система будет охватывать разнообразные аспекты IT, включая технологические инновации, криптовалюты, вопросы конфиденциальности и меры обеспечения безопасности информации. Она будет предоставлять удобный и интуитивно понятный интерфейс для поиска и чтения новостей, а также возможность настройки персонализированных новостной ленты в соответствии с интересами каждого пользователя, добавление заинтересовавших статей в закладки для дальнейшего прочтения их в свободное время. С учетом активного развития информационных технологий и растущего интереса к этой области, создание такой информационной системы обещает быть востребованным и полезным как для пользователей, стремящихся быть в курсе последних событий в мире IT, так и для людей, занимающихся этим профессионально.

Система будет предоставлять следующие возможности:

* Агрегирование новостей из различных источников в одном месте.
* Категоризация новостей для удобной навигации.
* Поиск новостей, по ключевым словам, и фразам.
* Просмотр полных статей и источников новостей.
* Возможность пользователей отмечать понравившиеся статьи и добавлять их в список "Избранное" для дальнейшего чтения.

Исходя из описанных выше проблем, целью выпускной квалификационной работы будет разработка информационного портала с названием «IT News». Поставим ключевые технические задачи проекта:

* Изучить предметную области для получения полного понимания системы и выявления существующих проблем;
* Подготовить технические задание;
* Провести анализ существующих аналогов.
* Разработка концептуальных решений для информационного портала «IT News», включая выбор дизайна и функциональности.
* Выбор программно-инструментальных средств, необходимых для разработки системы.
* Изучение программного инструментария, который будет использоваться при разработке портала.
* Проектирование и разработка веб-приложения с учетом всех ранее определенных требований и концептуальных решений.
* Тестирование разработанной системы с целью выявления и устранения ошибок.
* Отладка системы для обеспечения ее стабильной и надежной работы.
* Подготовка подробной пояснительной записки, в которой будет представлена вся информация о разработанной системе, ее особенностях и принципах работы.

В рамках данной работы рассматривается разработка главной страницы веб-ресурса, страницы новостных статей, реализация модуля пользователя, который включает в себя важные функции, такие как процесс регистрации и авторизации, создание и управление личными кабинетами пользователей, просмотр страницы закладок статей.

Возможности

Definition Statements

* Автоматический сбор новостей с разных источников.
* Аннотирование статей и выделение главных мыслей.
* Фильтрация новостей по интересам пользователя.
* Уведомления о новых статьях в интересующих категориях.
* Быстрый доступ к актуальным новостям.
* Персональные рекомендации на основе пользовательской истории.

System Features

* Автоматический поиск и сбор новостей.
* Аннотирование новостных статей для пользователя + ссылка на источник.
* Классификация новостей по категориям и ключевым словам.
* Пользовательские настройки фильтров.
* Хранение истории просмотренных новостей пользователя.

Рынок

SWOT-анализ

|  |  |  |
| --- | --- | --- |
|  | **Возможности** | **Угрозы** |
| Внешние | - Растущий интерес к IT-новостям и трендам - Увеличение числа онлайн-пользователей и активных читателей IT-новостей - Повышенный спрос на персонализированные новостные сервисы для всех заинтересованных IT-сферой - Возможность расширения на мировой рынок и привлечение аудитории из разных стран | - Сильная конкуренция на рынке новостных агрегаторов - Быстро меняющиеся технологии и требования пользователей - Закрытость крупных медиа-компаний и площадок к сотрудничеству - Потенциальные правовые ограничения в разных странах |
| Внутренние | - Уникальная система аннотирования и выделения главных мыслей статей, что улучшает опыт чтения пользователей - Гибкость в настройке рекомендаций, позволяющая удовлетворить потребности разных категорий пользователей - Возможность внедрения дополнительных функций, таких как аналитика трендов и статистика по интересам | - Ограниченные ресурсы для маркетинга и продвижения - Риск технических сбоев и потери данных - Низкая лояльность пользователей к новым платформам - Риск долгого времени ожидания окупаемости проекта - Отсутствие контроля над качеством контента |

Экономические предпосылки:

* Растущий интерес к IT-новостям и трендам.
* Потребность в быстром доступе к актуальной информации.
* Возможности монетизации через рекламу и подписку.

Тенденции:

* Увеличение интереса к технологическим инновациям.
* Рост количества IT-новостных источников.

Заинтересованные лица

Основные задачи высокого уровня

|  |  |  |  |
| --- | --- | --- | --- |
| **Заинтересованное лицо** | **Цель высокого уровня** | **Проблемы, возможности и замечания** | **Текущие решения** |
| Эксперт | Получить актуальные новости и тренды в различных сферах IT. | Потеря времени на поиск информации. | Поиск информации на большом количестве ресурсов. |
| Журналист | Получить материал для статей и обзоров. | Затраты времени на прочтение и структурирование. | Ознакомление с информацией на различных новостных источниках. |
| IT-компания | Изучить конкурентов и рынок. | Трудности в анализе данных.  Финансовые затраты на мониторинг рыка. | Анализ информации в ручном режиме. |

Позиционирование.

Платформа для всех, кто интересуется сферой IT, с быстрым и удобным доступом к актуальной информации.

Продвижение

* Социальные сети.
* IT-сообщества.

Расчет рынка

Описание бизнес-модели

* Монетизация через рекламу и подписку.

Объем рынка

Потенциальная аудитория: миллионы людей, интересующихся IT-сферой.

Сегментирование и выявление своих клиентов

* Профессиональный уровень (начинающие, опытные, эксперты)
* Область интересов (разработка, кибербезопасность, и т. д.)
* География (региональная, мировая)

Ценообразование

* Расчет цены в зависимости от:
* Вероятности продажи.
* Стремления к максимальной прибыли.
* Затрат на продажу и эксплуатацию.

Конкуренты

|  |  |  |
| --- | --- | --- |
| **Конкурент** | **Описание** | **Отличия платформ** |
| news.google.com ("Google News") | Глобальная новостная платформа, предоставляющая широкий выбор новостей и агрегирующая информацию из разных источников. | - Агрегирует новости из различных источников по всему миру.  - Имеет персонализированные рекомендации для пользователей. |
| yandex.ru ("Yandex News") | Популярная новостная платформа от Яндекса, предлагающая новости и события из различных областей. | - Фокусируется на событиях в России и странах СНГ.  - Имеет интегрированный поиск и другие сервисы Яндекса. |
| news.yahoo.com ("Yahoo News") | Крупная новостная платформа Yahoo, предоставляющая новости, статьи и финансовую информацию. | - Основное внимание уделяется новостям и финансовой информации.  - Имеет разделы с разнообразным контентом, включая статьи и видео. |

Преимущества IT News платформы

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Параметр** | **IT News** | **Yahoo News** | **Google News** | **Yandex News** |
| **Региональная специализация** | Глобальный охват IT-новостей. | Фокус на новостях в основном в США. | Мировые и региональные новости. | Основное внимание региональным событиям. |
| **Персонализированный контент** | Выделение тематических категорий, источников и ключевых слов для персональных рекомендаций. | Общие категории, менее персонализированный контент. | Персонализированные рекомендации, но с ограниченными возможностями настройки. | Ограниченные возможности настройки персональных рекомендаций. |
| **Гибкость в агрегации** | Агрегация новостей с различных источников, включая технические блоги, форумы и официальные источники компаний. | Агрегация основных новостных источников. | Агрегация новостей из известных источников. | Агрегация новостей с упором на популярные источники. |
| **Источники** | Выбор надежных источников с актуальными новостями и качественным контентом. | Широкий выбор источников, но разная степень достоверности. | Надежные источники, но охватывают не все аспекты IT. | Разнообразие источников, но могут быть менее достоверными. |

Перспективы решения

Прогноз:

* Ожидается рост числа пользователей и прибыли.
* Расширение функционала для удобной работы с платформой.

Исходя из сравнения в таблице, можно сделать вывод, что каждый из перечисленных порталов имеет свои уникальные особенности и преимущества в предоставлении информации о событиях в IT-отрасли.

* «Yahoo News» фокусируется на новостях в США, предоставляя общие категории с разной надежностью источников.
* «Google News» предлагает персонализированные рекомендации с ограниченными настройками, охватывая мировые и региональные новости, но не все аспекты IT.
* «Yandex News» сосредотачивается на региональных событиях, но источники могут быть менее достоверными.

# АКТУАЛЬНОСТЬ

В рамках разработки информационной системы «IT News», актуальность данного проекта проявляется в контексте современных информационных технологий и возрастающей потребности в актуальной и качественной информации в области IT.

Современный мир информационных технологий характеризуется постоянными изменениями и инновациями. В этой динамичной сфере, происходит множество событий, разработок и изменений каждый день. Люди, работающие в IT-сфере, журналисты, эксперты и представители компаний нуждаются в оперативном доступе к актуальным новостям и аналитическим материалам для принятия решений. В этом контексте, разработка информационной системы, которая позволит пользователям быстро получать доступ к IT-новостям, настраивать свою ленту новостей и иметь доступ к разнообразным источникам информации, становится актуальной задачей. Пользователи смогут находить новости по интересующим их темам, будь то технологии, криптовалюты, конфиденциальность или безопасность. Актуальность данного проекта также подкрепляется тем, что в современном мире доступ к информации является ключевым фактором конкурентоспособности. IT-компании, журналисты и эксперты должны быть в курсе последних событий и трендов, именно поэтому создание платформы, обеспечивающей доступ к актуальным IT-новостям, становится неотъемлемой частью информационной среды.

Таким образом, проект IT News обладает актуальностью как в научном, так и в практическом контексте, удовлетворяя потребности пользователей в оперативной и структурированной информации в сфере информационных технологий.

# ПОСТАНОВКА ЗАДАЧИ

Информационная система предоставляет пользователям доступа к актуальным и интересным новостям в сфере информационных технологий. Информационная система охватывает разнообразные аспекты IT, включая технологические инновации, криптовалюты, вопросы конфиденциальности и меры обеспечения безопасности информации.

Система предоставляет удобный и интуитивно понятный интерфейс для пользователей, позволяя осуществлять поиск и чтение новостей, а также настраивать персонализированные новостные ленты в соответствии с интересами каждого пользователя. Пользователи могут добавлять понравившиеся статьи в закладки для дальнейшего прочтения в свободное время. Учитывая активное развитие информационных технологий и повышенный интерес к данной области, создание такой информационной системы обещает быть востребованным как для тех, кто стремится быть в курсе последних событий в мире IT, так и для профессионалов этой области.

Система предоставляет уникальные функциональности, которые делают процесс получения информации более удобным и эффективным. Пользователям предоставляется возможность получить доступ к новостям из различных источников через централизованный портал. Одной из ключевых функций, является возможность осуществлять поиск, по ключевым словам, и фразам, что позволяет пользователям быстро находить интересующие их новости. Система также обладает удобной функцией фильтрации новостей по категориям, что значительно упрощает навигацию и позволяет сосредотачиваться на интересующих темах.

Помимо этого, пользователи могут читать полные статьи и источники новостей непосредственно на платформе, что делает процесс чтения более удобным. Функционал "Избранное" позволяет сохранить понравившиеся статьи для последующего прочтения, обеспечивая возможность возвращения к ним в удобное время.

В рамках данной работы рассматривается разработка главной страницы веб-ресурса, страницы новостных статей, реализация модуля пользователя, который включает в себя важные функции, такие как процесс регистрации и авторизации, создание и управление личными кабинетами пользователей, просмотр страницы закладок статей.

# ДОПОЛНИТЕЛЬНАЯ СПЕЦИФИКАЦИЯ

Проект "IT News" представляет собой инновационную платформу, которая парсит новостные статьи с различных агрегаторов, аннотирует их и предоставляет пользователям возможность читать статьи в сокращенном виде, а также просматривать оригиналы в источниках. Цель проекта - упростить доступ к актуальным новостям в области информационных технологий (IT) и обогатить пользовательский опыт чтения новостей.

Функциональность

Регистрация событий и обработка ошибок

* Реализация системы регистрации событий для непрерывного отслеживания действий пользователей и оперативной обработки возможных ошибок.
* Внедрение механизмов анализа журналов событий с целью улучшения функционала и обеспечения стабильности системы.

Удобство использования

Человеческие факторы

* Анализ поведения пользователей для оптимизации интерфейса и улучшения удобства использования веб-платформы.
* Проведение тестирований с участием пользователей для оценки уровня удобства и выявления потенциальных проблем.

Надежность

Возможность восстановления информации

* Регулярное создание резервных копий данных для обеспечения возможности восстановления информации в случае сбоев или потери данных.
* Автоматизированный процесс восстановления данных для минимизации времени простоя системы.

Производительность

* Оптимизация алгоритмов парсинга и аннотирования для обеспечения высокой производительности при обработке больших объемов данных.
* Мониторинг производительности и оптимизация системы в реальном времени для предотвращения задержек и сбоев.

Возможности поддержки

Адаптация системы

* Возможность быстрой адаптации системы под изменяющиеся требования рынка IT-новостей.
* Поддержка автоматических обновлений и модульная архитектура для удобства добавления новых функций.

Конфигурирование

* Предоставление пользовательских настроек для конфигурирования интерфейса и отображения контента.
* Максимальная гибкость в настройке интерфейса, включая возможность выбора тем оформления, языка и типа контента, отображаемого на главной странице.

Ограничения

* Ограничение доступа к некоторым функциям в зависимости от уровня пользователя (например, администратор или обычный пользователь).

Приобретаемые компоненты

* Оплата использования API веб-платформы согласно тарифному плану или подписке.

Бесплатные компоненты на основе открытого кода

* Использование открытых библиотек и фреймворков для оптимизации процессов парсинга и аннотирования.
* Активное участие в сообществе разработчиков для обмена опытом и получения поддержки по использованию открытого кода.

Интерфейсы

Важные интерфейсы и аппаратные средства

* Оптимизация интерфейса под различные устройства, включая компьютеры, планшеты и смартфоны.

Программные интерфейсы

* Создание API для интеграции с другими информационными системами и сервисами.
* Поддержка публичных API для разработчиков сторонних приложений и сервисов.

Бизнес-правила

|  |  |  |  |
| --- | --- | --- | --- |
| **Имя** | **Правило** | **Возможность изменения** | **Источник** |
| Мониторинг производительности | Система должна непрерывно мониторить производительность и автоматически оптимизировать работу при необходимости. | Да | - |
| Интеграция с другими системами | Система должна предоставлять API для интеграции с другими информационными системами. | Да | Технические возможности системы |

Вопросы законодательства

* Веб-платформа должна соответствовать законам о защите данных и обеспечивать конфиденциальность пользовательской информации.
* Система должна соблюдать авторские права при парсинге и отображении новостных статей.

Информация из предметной области

* Непрерывное обновление и расширение базы данных новостей для обеспечения актуальности информации.
* Анализ трендов в области IT для определения наиболее востребованных новостей и тем.

# КОНЦЕПТ ПРОЕКТА

Разрабатываемый проект — веб-платформа для обработки новостной информации. Этот проект объединяет в себе технологии в области анализа текста с интеграцией искусственного интеллекта и информационных технологий с целью автоматизации сбора, анализа и представления актуальных новостей. Ниже представлена концептуальная модель данного проекта, которая визуализирует ключевые компоненты и взаимосвязи между ними.

Новостные статьи (News Articles):

* + Заголовок (Заголовок): Элемент содержит информацию о заголовке новостной статьи.
  + Автор (Автор): указывает на автора статьи.
  + Дата публикации (Дата публикации): содержит информацию о времени публикации статьи.
  + Ссылка на картинку (Ссылка на картинку): указывает на изображение, связанное со статьей.
  + Ссылка на источник (Ссылка на источник): содержит ссылку на оригинальный источник новости.
  + Текст (Текст): представляет собой основной текст новости.

Категории (Categories):

* + Категории группируют новостные статьи по общим темам или тематикам.

Аннотации (Annotations):

* + Этот элемент предназначен для краткого изложения содержания новостной статьи.

Парсер (Parser):

* + Статья (Статья): Этот элемент представляет функцию Собрать(), которая отвечает за сбор данных, необходимых для анализа новостных ресурсов.
  + Новостные ресурсы (News Resources):
  + Название (Название): Содержит название новостного ресурса.
  + Ссылка (Ссылка): Указывает на URL новостного ресурса.

Теги (Tags):

* + Теги используются для категоризации новостных статей и обеспечивают более точное их описание.

Нейронная сеть (Neural Network):

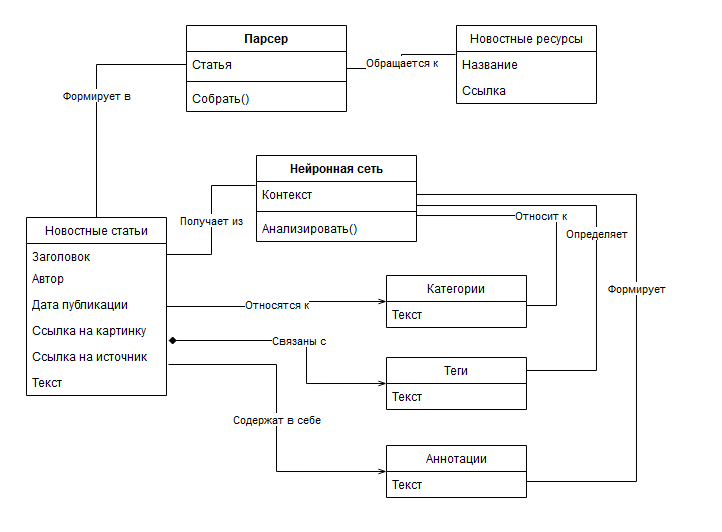
* + Контекст (Контекст): Представляет функцию Анализировать(), которая выполняет анализ новостей, формирует контекст и выводит результаты.

Контекст (Context):

* + Контекст содержит результаты анализа новостей, представленные текстовой информацией.

Взаимосвязи между элементами модели определяют поток данных и влияние. Например:

* Парсер собирает информацию из новостных ресурсов и формирует статьи.
* Нейронная сеть анализирует статьи и формирует контекст.
* Новостные статьи могут быть категоризированы, а их теги использованы для точной классификации.



*Рис. 1. Концептуальная-схема веб-платформы.*

Рассмотрим, как в рамках разрабатываемой системы происходит взаимодействует с пользователями. Приведена схема сценария использования, чтобы лучше понять, как различные люди будут взаимодействовать с веб-платформой и какие процессы будут происходить. Рассмотрим некоторые прецедент:

### Прецедент: Просмотр новостей на веб-ресурсе

**Описание:** Пользователь заходит на веб-ресурс информационной системы для просмотра актуальных новостей в области IT.

**Шаги:**

1. **Открыть Веб-Ресурса:**
   * Пользователь вводит адрес веб-ресурса в адресной строке браузера.
   * Система загружает главную страницу веб-ресурса.
2. **Навигация по Сайту:**
   * Система отображает список категорий новостей (например, "Технологии", "Криптовалюты", "Конфиденциальность", "Безопасность").
   * Пользователь изучает доступные разделы Новостей.
3. **Выбрать Категории:**
   * Пользователь выбирает определенную категорию новостей, например, "Криптовалюты".
   * Система загружает страницу с новостями в выбранной категории.
4. **Просмотр Новостей:**
   * Пользователь просматривает заголовки новостей, выбирает интересующую статью для прочтения.
   * Система открывает полный текст статьи внутри веб-ресурса.
5. **Чтение и Взаимодействие:**
   * Пользователь читает статью, может прокручивать вниз для прочтения текста.
   * Пользователь может открыть источник статьи, для прочтения полной версии текста.
6. **Возврат к Навигации:**
   * После прочтения статьи пользователь может вернуться к списку новостей в категории или вернуться на главную страницу для выбора других разделов.

**Результат:** Пользователь получает информацию о последних новостях в области IT, взаимодействует с контентом и может отмечать понравившиеся статьи.

### Прецедент: Поиск новостей на веб-ресурсе

**Описание:** Пользователь ищет конкретные новости или темы в области IT на веб-ресурсе информационной системы.

**Шаги:**

1. **Открыть Поиск:**
   * Пользователь находится на главной странице веб-ресурса.
   * Пользователь видит поле поиска, где он может ввести ключевые слова или фразы.
2. **Ввести Ключевые Слова:**
   * Пользователь вводит ключевые слова или фразы, связанные с интересующей его темой (например, "искусственный интеллект", "блокчейн технологии").
   * Пользователь нажимает кнопку "Поиск".
3. **Поиск Новостей:**
   * Система выполняет поиск по введенным ключевым словам в сохраненных новостях и отображает результаты поиска.
4. **Просмотр Результатов:**
   * Пользователь видит список новостей, соответствующих его запросу.
   * Каждая новость включает в себя заголовок, краткое описание, картинку и источник.
5. **Выбрать Новости:**
   * Пользователь выбирает конкретную новость из результатов поиска, которая его заинтересовала.
   * Пользователь нажимает на заголовок или краткое описание новости, чтобы прочитать текст.
6. **Чтение и Взаимодействие:**
   * Пользователь читает выбранную новость и взаимодействует с контентом (отмечать понравившиеся статьи).
   * Пользователь может вернуться к результатам поиска и выбрать другую новость для прочтения.
7. **Навигация в Результатах:**
   * После прочтения статьи пользователь может вернуться к списку результатов поиска или вернуться на главную страницу для новых запросов.

**Результат:** Пользователь находит и читает новости, соответствующие его запросу, и взаимодействует с содержимым. Поиск позволяет пользователю быстро и эффективно находить информацию в области IT, которая соответствует его интересам и потребностям.

### Прецедент: Отметить Понравившиеся Статьи

**Описание:** Пользователь отмечает статьи, которые ему понравились, после чтения, чтобы сохранить их для дальнейшего возврата.

**Шаги:**

1. **Чтение Статьи:**
   * Пользователь читает статью на веб-ресурсе.
   * После чтения, пользователь решает, что статья ему понравилась и он хочет сохранить ее.
2. **Отметить Статьи:**
   * Под статьей отображается опция "Понравилось" или "Добавить в Избранное".
   * Пользователь кликает на эту опцию, отмечая статью как понравившуюся.
3. **Просмотр Понравившихся Статей:**
   * Пользователь может вернуться в раздел "Избранное" или "Понравившиеся статьи" на веб-ресурсе.
   * Система показывает список всех статей, которые пользователь отметил, как понравившиеся.
4. **Управление Отметками:**
   * Пользователь может удалить статью из списка понравившихся, если его мнение изменится.
   * Пользователь может поделиться отмеченными статьями в социальных сетях или отправить ссылку на статью друзьям.
5. **Уведомления:**
   * Пользователь может настроить уведомления, чтобы получать оповещения о новых статьях от тех же авторов или на те же темы, что и статьи, отмеченные им как понравившиеся.

**Результат:** Пользователь может эффективно управлять статьями, которые ему интересны и которые он хочет сохранить для дальнейшего прочтения. Этот прецедент позволяет пользователям сохранять контент, который вызвал у них интерес, для последующего возврата и более детального изучения.

### Прецедент: Добавить Статьи в Закладки

**Описание:** Пользователь добавляет статью в закладки для дальнейшего ознакомления в случае, если статья его заинтересовала, но в данный момент у него нет возможности прочитать ее.

**Шаги:**

1. **Интерес к Статье:**
   * Пользователь видит статью на веб-ресурсе, которая его заинтересовала, но у него нет времени или возможности прочитать ее в данный момент.
2. **Добавить в Закладки:**
   * Под статьей отображается опция "Добавить в Закладки".
   * Пользователь кликает на эту опцию, добавляя статью в свои закладки.
3. **Управление Закладками:**
   * Пользователь может перейти в раздел "Закладки" на веб-ресурсе.
   * В этом разделе пользователь видит список всех статей, которые он добавил в закладки.
4. **Чтение в Удобное Время:**
   * Когда у пользователя появляется свободное время, он может вернуться в раздел "Закладки" и выбрать статью для прочтения.
5. **Удалить из Закладок:**
   * После прочтения статьи, если она больше не представляет интерес или отметить статью как понравившуюся, пользователь может удалить ее из закладок.

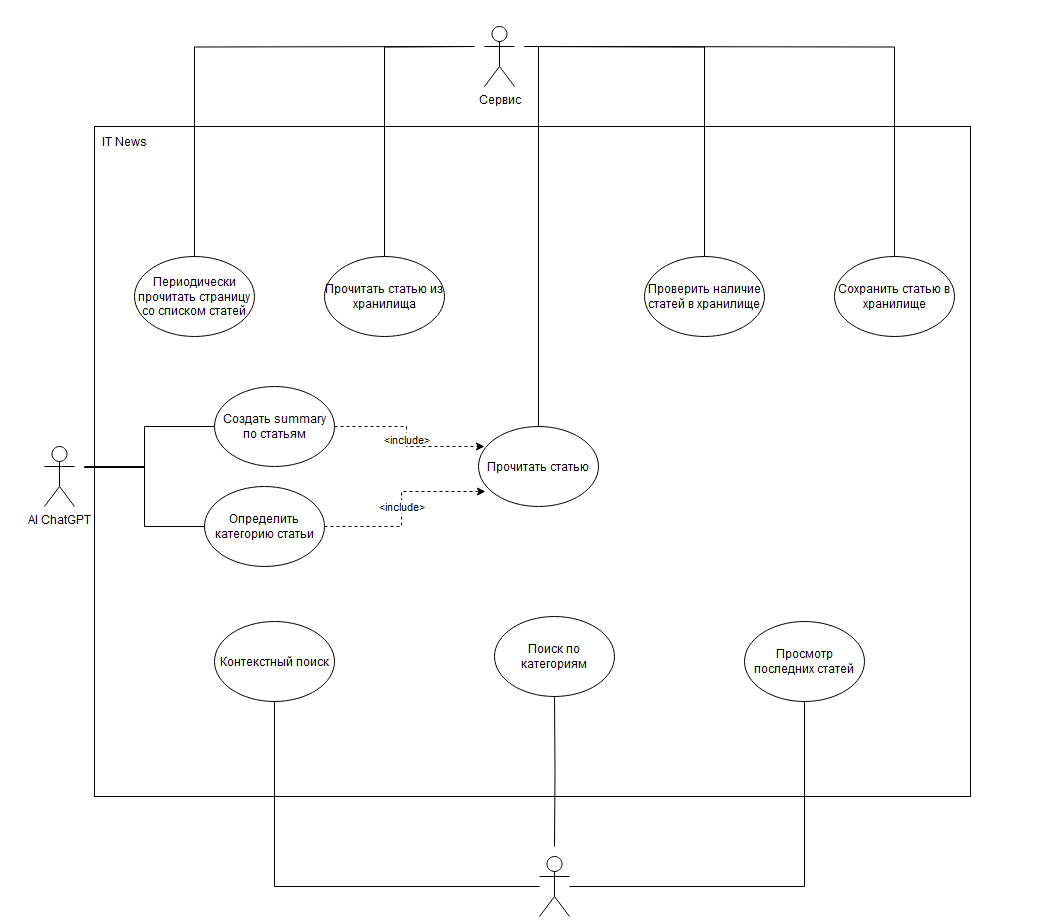
**Результат:** Пользователь может сохранить интересные статьи для отложенного чтения в разделе "Закладки", если у него нет времени изучить заинтересовавшую его статью сразу. Этот прецедент позволяет пользователям эффективно управлять своим временем и изучать интересующий контент в более удобное время.

### Использование веб-платформы

* Просмотр последних статей
  + Описание: Пользователь может просматривать список последних опубликованных статей на веб-сайте.
  + Действие: Пользователь выбирает опцию "Просмотр последних статей" и видит список свежих новостей.
* Поиск по категориям
  + Описание: Пользователь имеет возможность искать статьи по определенным категориям или темам.
  + Действие: Пользователь выбирает определенную категорию (например, "Искусственный Интеллект") и получает список статей, связанных с этой темой.
* Контекстный поиск
  + Описание: Пользователь может вводить ключевые слова или фразы для поиска конкретной информации.
  + Действие: Пользователь вводит ключевые слова (например, "Интернет вещей") и получает результаты, соответствующие запросу.

Система "IT News"

* Сервисы Обработки данных
  + Описание: Система включает в себя сервисы, обрабатывающие данные и обеспечивающие функциональность веб-приложения.
  + Детали:
    - Сохранение статьи в хранилище: обеспечивает сохранение новых статей в базе данных для последующего отображение на ресурсе.
    - Чтение статьи из хранилища: позволяет веб-платформе получать содержимое сохраненных статей.
    - Создание аннотации по статьям: автоматически генерирует краткое описание статей для облегчения понимания содержания.
    - Определение категории статьи: автоматически определяет тематическую принадлежность статьи (например, "Технологии" или "Искусственный Интеллект").
* Дополнительные Сервисы
  + Описание: Система включает в себя сервис, отвечающий за мониторинг новых статей и обновлений в хранилище данных.
  + Детали:
    - Периодическое чтение страницы со списком статей: Система регулярно сканирует веб-сайты для выявления новых статей.
    - Проверка наличия статей в хранилище: предотвращает дублирование статей в базе данных.
    - Чтение конкретной статьи: получает полный текст статьи для сохранения и отображения пользователям.
* Взаимодействие с Искусственным Интеллектом (AI ChatGPT)
  + Описание: Система взаимодействует с искусственным интеллектом для повышения уровня сервиса для пользователей.
  + Детали:
    - Создание аннотации по статьям: Искусственный интеллект автоматически формирует краткие обзоры содержания статей для быстрого ознакомления.
    - Определение категории статьи: Искусственный интеллект определяет тематическую принадлежность статей, облегчая пользователям поиск по интересующим темам.



*Рис. 2. Сценарий использования.*

Основываясь на информации выше, спроектируем структуру проекта для выполнение поставленных задач. Система будет предоставлять следующие возможности, а именно:

* Агрегирование новостей из различных источников в одном месте.
* Категоризация новостей для удобной навигации.
* Поиск новостей, по ключевым словам, и фразам.
* Просмотр полных статей и источников новостей.
* Возможность пользователей отмечать понравившиеся статьи и добавлять их в список "Избранное" для дальнейшего чтения.

Определим основные компоненты системы. Разберем, как каждый из них взаимодействует, обеспечивая работоспособность и функциональность платформы.

Основные Компоненты:

1. Клиент (Browser):

* Клиент взаимодействует с платформой через браузер, заходя на веб-сайт по предоставленной ссылке.

1. Сервер (Docker Container):

* Docker контейнер является центральным элементом системы, предоставляя доступ к основным сервисам.
* Nginx имеет внешний порт 80:80, через который клиент может обращаться к веб-странице платформы.

1. Back-end (Flask):

* Отвечает за обработку бизнес-логики и взаимодействие с базой данных.
* Запущен на локальном порту 5566 внутри Docker контейнера.

1. База данных (PostgreSQL):

* Хранит данные, необходимые для функционирования платформы.
* Запущена на локальном порту 5432 внутри Docker контейнера.

1. Front-end (Vue JS):

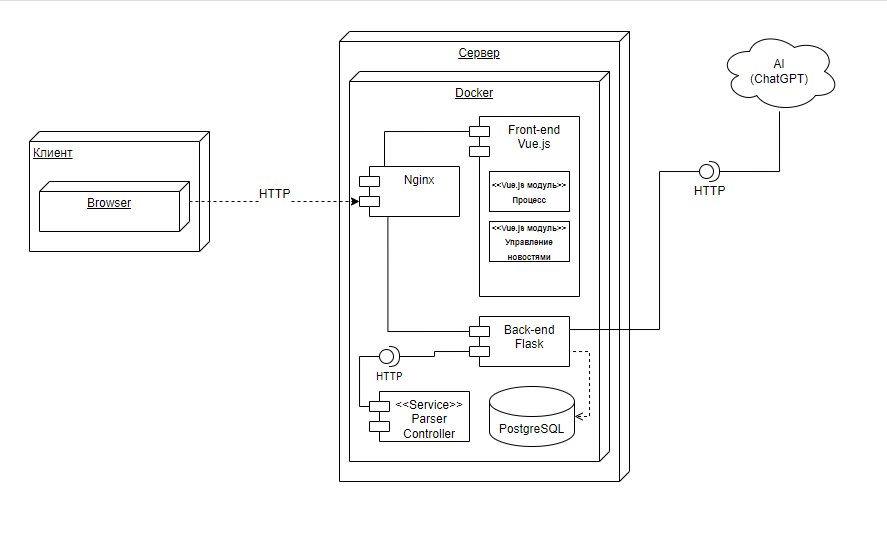
* Обеспечивает пользовательский интерфейс для клиента.
* Взаимодействует с бэкэндом для получения и отображения данных.
* Доступен по внешнему порту 80:80.

Взаимодействие Компонентов:

* Клиент, используя браузер, отправляет запрос на веб-сайт платформы через внешний порт 80:80 Docker контейнера.
* Nginx работает на порту 80.
* Front-end (Vue JS) принимает запрос от клиента и взаимодействует с бэкэндом (Flask) для получения необходимых данных.
* Back-end (Flask) в свою очередь обращается к базе данных (PostgreSQL) для извлечения и обработки информации.
* Результаты обработки передаются обратно во Front-end, который формирует и отображает данные на веб-странице для клиента.

Подробности Docker Контейнера:

* Внутри Docker контейнера размещены два локально запущенных сервиса: Back-end (Flask) и база данных (PostgreSQL).
* Back-end слушает запросы на порту 5566.
* База данных PostgreSQL слушает порт 5432.
* Docker контейнер имеет внешний порт 80:80, Nginx направляет трафик на веб-страницу Front-end (Vue JS).



*Рис. 3. Диаграмма компонентов.*

В таком случае программная архитектура проекта будет иметь следующий вид, где на схеме представлена структура модели данных, описывающая веб-платформу с различными связанными аспектами, такими как категории, теги и аннотации.

Article (Статья):

* id: long: Уникальный идентификатор статьи.
* title: String: Название статьи.
* author: String: Автор статьи.
* publicationDate: date: Дата публикации статьи.
* pictureLink: String: Ссылка на изображение, связанное со статьей.
* articleLink: String: Ссылка на саму статью.
* body: String: Текст статьи.

Category (Категория):

* id: long: Уникальный идентификатор категории.
* name: String: Название категории.

Tag (Тег):

* id: long: Уникальный идентификатор тега.
* name: String: Название тега.

Annotations (Аннотации):

* id: long: Уникальный идентификатор аннотации.
* body: String: Текст аннотации.

AI (Искусственный интеллект):

* model: String: Модель искусственного интеллекта.
* prompt: String: Подсказка для модели.
* getResponse(): String: Метод для получения ответа от модели.
* setMessage(String): type: Метод для установки сообщения модели.
* setModel(String): type: Метод для установки модели.

Parser (Парсер):

* targetSource: <Source>: Целевой источник данных.
* sendData(): bool: Метод для отправки данных.

Source (Источник):

* name: String: Название источника.
* url: List<String>: Список URL-ов, связанных с источником.

Связи между элементами:

* Статья связана с категорией, тегами и аннотациями.
* Искусственный интеллект использует данные, полученные при помощи парсера с различных ресурсов, для обработки данных.
* Парсер получает данные от источника.

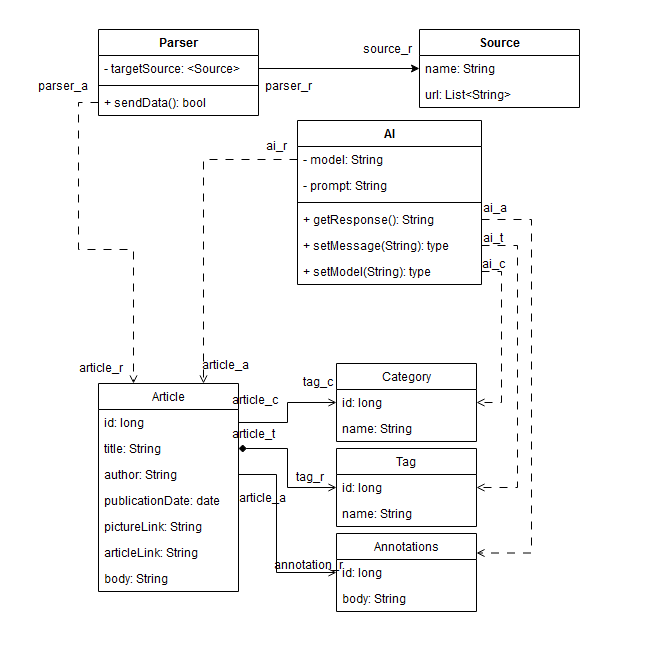
Такая структура позволяет описать систему, в которой статьи сгруппированы по категориям и тегам, имеют аннотации, а также взаимодействуют с искусственным интеллектом для обработки данных из определенного источника.

*Таблица 2 – Наименование элементов диаграммы классов.*

|  |  |
| --- | --- |
| **Сущность** | **Наименование** |
| parser | Парсер |
| source | Ресурс |
| ai | Нейронная сеть |
| article | Тип данных |
| category | Тип информации |
| tag | Тип информации |
| annotations | Тип информации |

*Таблица 3 – Связи между элементами диаграммы классов.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Сущность** | **Ключ** | **Атрибуты** | **Обязательность** | **Тип** | **Отношение** | **Описание** |
| parser |  | targetSource | + | Source |  | Тип данных, содержащий название ресурса и доступные и ссылки на списки статей. |
| parser |  | source\_r | + |  | 1..1 |  |
| source | + | name | + | String |  | Идентификатор ресурса. |
| source | + | url | + | List[String] |  | Содержит ссылки на страницы с списком статьями. |
| article | + | id | + | Long |  | Содержит идентификатор новостной статьи. |
| article |  | title | + | String |  | Информация с названием новостной статьи. |
| article |  | author | + | String |  | Содержит информацию о авторе новостной статьи. |
| article |  | publicationDate | + | date |  | Информация с датой публикации. |
| article |  | pictureLink | + | String |  | Содержит ссылку на картинку статьи. |
| article |  | sourceLink | + | String |  | Содержит ссылку на страницу новостной статьи на новостном ресурсе. |
| article |  | body | + | String |  | Содержит основную информацию новостной статьи. |
| article |  | article\_c | + | category | 1..1 |  |
| article |  | article\_t | + | tag | 1..\* |  |
| article |  | article\_a | + | annotation | 1..1 |  |
| category | + | id | + | Long |  | Содержит идентификатор новостной статьи. |
| category |  | name | + | String |  | Содержит название категории статьи. |
| tag | + | id | + | Long |  | Содержит идентификатор новостной статьи. |
| tag |  | name | + | String |  | Содержит один из тегов статьи. |
| annotations | + | id | + | Long |  | Содержит идентификатор новостной статьи. |
| annotations |  | body | + | String |  | Содержит результат работы нейронной сети по статье. |
| ai |  | prompt | + | String |  | Содержит запрос для нейронной сети. |



*Рис. 4. Диаграмма классов.*

Диаграмма пояснительной модели IT News наглядно представляет взаимодействие между ключевыми компонентами системы.

### **1. Читатель (Пользователь)**

* **Действие**: Пользователь начинает взаимодействие с системой, запрашивая статьи.
* **Роль**: Инициирующий актор, требующий контента.

### **2. Интерфейс Пользователя (UI)**

* **Действие**: UI принимает запрос от пользователя и направляет его на сервер Flask API для обработки.
* **Роль**: Активный посредник между пользователем и системой.

### **3. Сервер Flask API (BackEnd)**

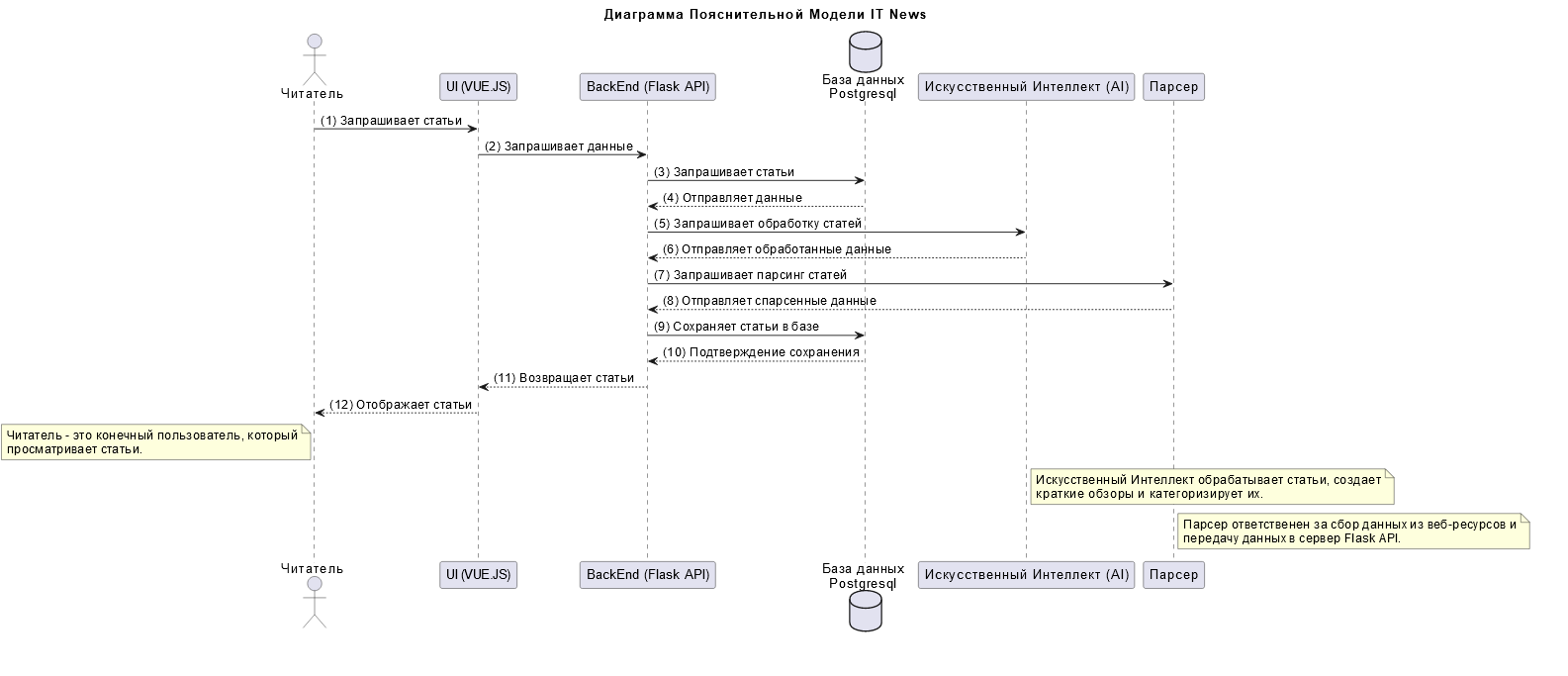
* **Действия**:
  + Запрашивает данные из базы данных PostgreSQL.
  + Обрабатывает данные, используя искусственный интеллект (AI) для создания кратких обзоров и категоризации.
  + Обращается к парсеру для получения дополнительных данных.
  + Сохраняет статьи в базе данных и возвращает их UI для отображения.
* **Роли**:
  + **Контролирующий**: Координирует все этапы обработки статей.
  + **Агент обработки данных**: Производит обработку и категоризацию статей.

### **4. Искусственный Интеллект (AI)**

* **Действие**: Обрабатывает статьи, создает краткие обзоры и категоризирует контент.
* **Роль**: Интеллектуальный агент, обогащающий статьи и делая их более удобными для чтения.

### **5. Парсер**

* **Действие**: Отвечает за сбор данных из веб-ресурсов.
* **Роль**: Поставщик данных, обеспечивающий сервер Flask API актуальной информацией.



*Рис. 5. Диаграмма последовательностей.*

# ЗАКЛЮЧЕНИЕ

В условиях постоянного развития информационных технологий и увеличения потока информации существует критическая потребность в эффективных платформах для сбора и краткого аннотирования новостей из области информационных технологий.

Платформа для тех, кто стремится получать не только самые актуальные новости, но и обширный обзор мнений, аналитики и сообщества в области информационных технологий. В дальнейшем развитие платформы может потребоваться уделять больше внимания персонализации для более точного соответствия потребностям широкого спектра пользователей.