

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«ПЕНЗЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

Факультет
вычислительной техники

Кафедра
«Вычислительная техника»

Направление подготовки: 09.03.01 «Информатика и вычислительная техника»

Профиль подготовки: «Вычислительные машины, комплексы, системы и сети»

БАКАЛАВРСКАЯ РАБОТА

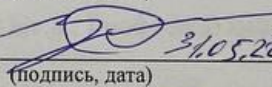
на тему

**НОВОСТНОЙ IT-ПОРТАЛ С
АВТОМАТИЧЕСКИМ ПОИСКОМ И
ПУБЛИКАЦИЕЙ АКТУАЛЬНОГО МАТЕРИАЛА**

Студент

 31.05.22 Воеков И.С.
(подпись, дата)

Руководитель

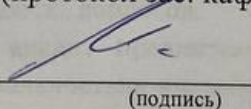
 31.05.22 Винкин С.А.
(подпись, дата)

Нормоконтролёр

 31.05.22 А.С. Бычков
(подпись, дата)

Работа допущена к защите (протокол заес. кафедры от 02.06.22 г. № 13)

Заведующий кафедрой


(подпись)

М.А. Митрохин

Работа защищена с отметкой хорошо (протокол заседания ГЭК от
20 июня № 15)

Секретарь ГЭК


(подпись)

Е.И. Гурин

Пенза 2022

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«ПЕНЗЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

Факультет

вычислительной техники

Кафедра

«Вычислительная техника»

«Утверждаю»

Заведующий кафедрой

М.А. Митрохин

«26» 04 2022 г.

**ЗАДАНИЕ
НА БАКАЛАВРСКУЮ РАБОТУ**

1. Студент Воеков Иван Сергеевич гр. 18вв1 факультета вычислительной техники.
2. Тема работы «Новостной IT-портал с автоматическим поиском и публикацией актуального материала.».

Тема утверждена приказом ПГУ № 357/о от "26" апреля 2022 г.

3. Руководитель бакалаврской работы д.т.н., профессор каф. ВТ ПГУ Зинкин С.А.

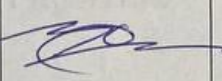
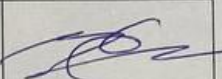
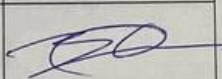
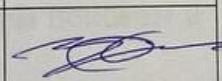
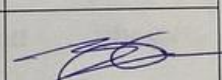
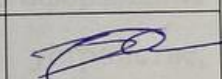
4. Задание на работу (назначение разработки, исходные данные и т.п.).

Назначение разработки – Новостной веб-портал, предназначенный для регулярной автоматической публикации новостей и статей из других источников. Веб-ресурс должен обладать удобным интерфейсом с правильной структурой страниц на всех видах устройств.

5. Перечень подлежащих разработке вопросов.

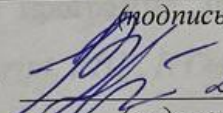
Требуется произвести анализ предметной области: изучить аналоги новостных порталов со схожей функциональностью, выделить достоинства и недостатки. Определить архитектуру приложения и средства разработки. Выбрать язык программирования и доступные варианты реализации проекта.

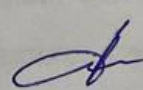
6. Календарный график выполнения работы

№ п/п	Наименование этапов работы	Объем работы	Срок выполнения	Подпись руководителя
1	Изучение предметной области	32	26.04.2022 – 28.05.2022	
2	Анализ задач	32	29.04.2022 – 30.05.2022	
3	Создание макета работы проекта	30	1.04.2022 – 3.05.2022	
4	Выбор инструментов, применяемых при разработке	40	4.05.2022 – 8.05.2022	
5	Разработка веб-ресурса	60	9.05.2022 – 15.05.2022	
6	Разработка и настройка всех модулей веб-ресурса	76	16.05.2022 – 23.05.2022	
7	Оформление пояснительной записки	54	24.05.2022 – 31.05.2022	

Дата выдачи задания "26" апреля 2022 г.

Руководитель бакалаврской работы  26.04.22 Зинкин С.А.
(подпись, дата)

Задание к исполнению принял студент  26.04.22 Воеков И.С.
(подпись, дата)

Работу к защите допустить
Декан факультета  17.06.2022 Фионова Л.Р.
(подпись, дата)

Реферат

Работа содержит 52 листа, 18 рисунков, 3 приложения.

САЙТ-АГРЕГАТОР, ВЕБ-ПОРТАЛ, ПАРСЕР, КОНТЕНТ, СКРИПТ, ХОСТИНГ, *CRON*, *SEO*, *PHP*, *JAVASCRIPT*, *HTML*, *API*, *FRONTEND*, *BACKEND*, *SSL*, *UML*, *DNS*, *FTP*, *DKIM*.

В рамках данной выпускной квалификационной работы был создан новостной веб портал, публикующий новости из других источников автоматически.

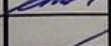
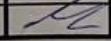
Цель работы – создать новостной веб-портал с удобным интерфейсом, статьи и новости на котором будут публиковаться автоматически.

Объектом разработки является веб-портал с автоматическим поиском и публикацией новостных статей по тематическим категориям.

В результате проведенной работы реализован проект, удовлетворяющий поставленной задаче, а именно: качественное оформление как на персональном компьютере, так и на мобильном устройстве, автоматизированная публикация новостных статей по категориям, правильное оформление и структура текстов, соблюдение правильной расстановки тэгов и заголовков.

Разработка проводилась в программе *Sublime Text*. Использованные языки программирования: *JavaScript*, *PHP*, *HTML*.

Созданный веб-сайт может быть использован для ознакомления с ним, в качестве новостного источника, а также образования большой ссылочной массы, вытекающий в сбор аудитории для какой-либо информационной компании.

					ПГУ 1.090301.06.001 ПЗ				
Изм.	Лист	№ докум.	Подпись	Дата	Новостной IT портал с автоматическим поиском и публикацией актуального материала Пояснительная записка	Лит.	Лист	Листов	
Разраб.		Воеков И.С.							
Проверил		Зинкин С.А.					2	54	
Н. Контр.		Бычков А.С.				ФВТ, каф. ВТ Гр. 18ВВ1			
Утверд.		Митрохин М.А.							

Содержание

Перечень сокращений и обозначений	5
Введение	6
1 Новостной IT-портал с автоматическим поиском и публикацией актуального материала	7
1.1 Анализ предметной области	7
1.2 Постановка задачи	9
1.3 Архитектура приложения	10
1.3.1 Архитектура веб ресурса	10
1.3.2 Шаблоны	11
1.3.3 <i>Search Automatic Plugin</i>	11
1.4 Инструменты разработки	12
1.4.1 <i>HTML</i>	12
1.4.2 <i>JavaScript</i>	13
1.4.3 <i>PHP</i>	13
1.4.4 <i>CSS</i>	14
1.4.5 Хостинг <i>ISPmanager</i>	14
1.4.6 <i>Sublime text</i>	15
1.4.7 <i>Cron</i>	16
1.5 Реализация	17
1.5.1 Создание сайта	18
1.5.2 Создание и использование плагина автонаполнения	20
1.6 Описание работы сайта	28
2 Охрана труда	34
2.1 Требования к организации рабочего места с ПК	34
2.2 Требования к освещению помещений и рабочих мест	35

2.3 Режим труда и отдыха при работе с компьютером	36
Заключение	38
Список использованных источников	39
Приложение А - <i>UML</i> - диаграммы.....	40
Приложение Б - листинг программы	45
Приложение В - презентация.....	48

Перечень сокращений и обозначений

API (Application Programming Interface) - интерфейс прикладного программирования, который представляет собой набор определений и протоколов для создания и интеграции прикладного программного обеспечения

CRON – планировщик задач

DKIM (Domain Keys Identified Mail) - цифровая подпись, доказывающая подлинность отправителя и гарантирует целостность доставленного письма до адресата. Подпись интегрируется к служебным заголовкам письма и невидима для пользователя

DNS (Domain Name System «система доменных имён») - является одной из основополагающих технологий современной мировой паутины и представляет собой распределенную систему хранения и обработки информации о доменных зонах

FTP - сервер, который работает по протоколу *File Transfer Protocol* и предназначен для отправки и приёма файлов через интернет или локальную сеть компьютера

SEO (Search Engine Optimization - оптимизация под поисковые системы) - это набор методов и мер, которые призваны повысить значимость сайта в глазах поисковых систем

SSL (secure sockets layer — уровень защищённых сокетов) - криптографический протокол для безопасной связи, его цель - обеспечить безопасную передачу данных. Для аутентификации используются асимметричные алгоритмы шифрования

					ПГУ 1.090301.06.001 ПЗ	Лист
						5
Изм.	Лист	№ докум.	Подпись	Дата		

Введение

С середины 20 века во всем мире наступила цифровая эпоха. Сдвиг всех индустрий пал на развитие себя в информационной сфере. Конкурентная борьба в разработке веб приложений, сайтов, технических инструментов для сбора статистических данных, дали еще больший толчок для развития сферы информационных технологий.

Современный среднестатистический человек получает 80% новой информации из интернета: уведомление о дне рождения друга в социальной сети, рецепт поминутно правильной варки яиц вкрутую, новый музыкальный клип от вашего любимого исполнителя, а также новости о ситуации в мире и не только. Именно с последним пунктом будет тесно связана моя дипломная работа.

Существует огромное количество различных новостных сайтов, на которые мы натываемся ежедневно, читая какую-либо новость в интернете. Мы можем разделить такие веб-ресурсы на два типа: сайты первоисточники, заполняющиеся вручную людьми, а также новостные сайты-агрегаторы, размещающие контент автоматически.

Цель этой работы – разработать новостной IT портал с автоматическим поиском и публикацией актуального материала, отвечающий всем требованиям функциональности, дизайна и структуры современного веб-ресурса.

В задачи подобных агрегаторов разработчик ставит получение различного рода аудитории на свой основной ресурс за счет большой ссылочной массы и, как следствия, высокого ранжирования в поисковых сетях.

					ПГУ 1.090301.06.001 ПЗ	Лист
						6
Изм.	Лист	№ докум.	Подпись	Дата		

1 Новостной IT-портал с автоматическим поиском и публикацией актуального материала

1.1 Анализ предметной области

Идея создать сайт-агрегатор на IT тематику мне пришла еще давно, так как эта сфера мне интересна, но знакомиться с новостями из этой области неудобно из-за малого наличия таких веб-порталов, а те, что существуют, имеют несовременный дизайн и устаревшую структуру, а самое главное, некачественно проработанную вёрстку для ознакомления с их контентом с мобильного устройства, она потеряла актуальность и просто имеет неудобные вид и структуру для среднестатистического пользователя.

Сфера IT и технологий сейчас невероятно широка и распространена во всех слоях общества, но полагаясь на свои предпочтения, новостной форум по душе я так и не нашёл. Именно эта причина и те, что указаны выше, натолкнули меня на создание собственного продукта в этой области.

Задачей стало создать качественный новостной веб портал агрегатор, который будет приятен глазу и удобен в использовании обычному пользователю. Также наполнение и контент, которым будет пополняться сайт не будет выглядеть как спам новости, содержание будет интересным и без лишней информации. Ещё на данном веб-ресурсе будут размещаться поучительные статьи в той же IT тематике, написанные лично мной, обучающие юзера вести себя в мировой сети интернет правильно и безопасно, а также пользователю в отдельной вкладке будет представлен список полезных программных утилит для личного пользования, который тот, в свою очередь, сможет скачать и установить.

На сегодняшний день существуют не только агрегаторы сми, но и многие другие. Агрегаторы товаров, объявлений, недвижимости, новостей.

Из сайтов-агрегаторов по IT тематике аналогичной направленности могу предложить ознакомиться с такими как “3D News”, “С News” и

					ПГУ 1.090301.06.001 ПЗ	Лист
						7
Изм.	Лист	№ докум.	Подпись	Дата		

“Ixbt.com”, их я привёл в пример не для того, чтобы мне выгодно выставить их недостатки для себя, а потому что когда я искал современные новостные форумы и веб-порталы по тематике инновационных технологий, компьютеров и так далее, то в самом верху списка находились три данных ресурса.

Этот список сайтов имеет такие плюсы, как:

- регулярное обновление и пополнение новостной ленты;
- большой доменный возраст;
- большое количество узких рубрик, а также буквально море информации, скопившейся за такой продолжительный промежуток времени в каждой из них.

Также были обнаружены недостатки:

- не отвечающий современным требованиям интерфейс и дизайн;
- данные сайты в избытке заполнены рекламой и просмотр новостей из-за этого становится неудобным;
- неудобная адаптация сайта на мобильных устройствах;
- приведённые выше веб ресурсы часто не указывают первоисточник статьи, что потенциально плохо влияет на её автора и деморализует его на дальнейший творческий путь.

На большинстве веб-порталах семантическое ядро сайта настроено некорректно. За такое длительное существование (10-20 лет), поиск статей по ключевым словам попросту не выдаёт пользователю их ресурс в поисковых выдачах. Решение данной проблемы заключается в правильной настройке семантического ядра сайта, то есть правильно сформированных слов и словосочетаний для описания тематики сайта.

Также нужно продемонстрировать удобный, функциональный дизайн, адаптирующийся под все виды устройств и удовлетворяющий всем современным требованиям.

					ПГУ 1.090301.06.001 ПЗ	Лист
						8
Изм.	Лист	№ докум.	Подпись	Дата		

1.2 Постановка задачи

Необходимо разработать веб ресурс, регулярно публикующий актуальные новости из других источников.

Данный веб ресурс должен удовлетворять следующим требованиям:

- качественное оформление как на персональном компьютере, так и на мобильном устройстве;

- правильное оформление и структура текстов, соблюдение правильной расстановки тэгов и заголовков;

- *SEO*-оптимизация текстов;

- правильное отображение страниц сайта в поисковых системах;

- отсутствие дубликатов страниц;

- запрет на индексацию ненужных страниц;

- размещение ссылок под новостями на релевантный источник;

- кроссбраузерность (видимость во всех браузерах);

- восстановление при возможных угрозах (резервное копирование и хранение на локальном компьютере или жестком диске даёт возможность сохранить информацию сайта, упростит его восстановление в случае хакерской атаки);

- в проекте будет реализована отдельная категория с полезным программным обеспечением для ежедневного использования;

- интерфейс нужно реализовать в виде стандартного новостного сайта, в котором нужно реализовать:

- 1) корректное отображение сайта на всех видах устройств;

- 2) разбиение новостей и статей на тематические области и свободное перемещение между ними;

- 3) возможность посмотреть публикации, набравшие наибольший охват среди пользователей на отдельной области сайта;

- 4) дизайн должен быть модернизированным и удобным для пользователя с любым опытом работы в интернете.

1.3 Архитектура приложения

1.3.1 Архитектура веб ресурса

В этом проекте прилагается вариант новостного веб портала с использованием автозаполнения. Для начала нужно определиться с правильным обозначением таких сайтов. На самом деле, термина “Автонаполняемый сайт” нет. Правильно их называть агрегатор. Это корректное техническое обозначение для подобных веб ресурсов.

Чтобы предоставить больше конкретики было взято основное обозначение из Википедии. Новостной агрегатор, или, менее точно, агрегатор СМИ — компьютерная программа или информационный интернет-ресурс (сайт или страница сайта), отображающий структурированную новостную ленту, формируемую им подборку из новостей различных СМИ, как правило по заданным критериям автоматически.

Во-первых, сайт должен удовлетворять такому требованию как, эффективность системы. Прежде всего, веб ресурс, конечно же, должен решать задачи и качественно выполнять свои функции при различных условиях. Сюда относятся такие характеристики, как надежность, безопасность, производительность, способность справляться с увеличением нагрузки (масштабируемость) и другие. Во-вторых, обладать гибкостью. Каждое приложение должно меняться со временем, так как требования в современном мире ежедневно меняются, добавляются новые. Чем быстрее и удобнее я смогу внести изменения в существующую функциональность. Чем меньше проблем и ошибок вызовет этот процесс, тем более гибкой и конкурентоспособной будет является вся система. Изменение одного фрагмента системы не должно влиять на другие. В-третьих, система должна быть легко расширяемой. Возможность добавления новых объектов и функций в систему без нарушения ее основной структуры. Архитектура должна позволять легко создавать дополнительные функциональные возможности, если это необходимо. Таким образом, выполнение наиболее

					ПГУ 1.090301.06.001 ПЗ	Лист
						10
Изм.	Лист	№ докум.	Подпись	Дата		

вероятных изменений требует наименьших усилий. В-четвёртых, легко масштабируемой в процессе разработки. Возможность сократить период развития путем добавления новых людей в проект. Архитектура должна позволять параллелизировать процесс разработки, чтобы многие люди могли работать над программой одновременно. Важным пунктом является наличие возможности повторного использования. Нужно, чтобы система была спроектирована так, чтобы ее фрагменты могли быть повторно использованы в других проектах, а также хорошо структурированный, читаемый и понятный код. Как правило, многие люди работают над программой — некоторые уходят, приходят новые. После написания сопровождать программу также, как правило, приходится новым людям, которые не участвовали в ее создании. Поэтому хорошая архитектура должна позволить новым людям легко и быстро понять, новую для них, систему. Проект должен быть хорошо структурирован, без дублирования, иметь хорошо сформированный код и документацию. А также в проекте применяются стандартные, общепринятые решения, знакомые программистам.

1.3.2 Шаблоны

Визуальная составляющая сайта состоит из множества шаблонов, которые специализируются для разных типов страниц. Всего для сайта были созданы 11 шаблонов, из которых составляется вся картина веб ресурса, например шапка сайта, раздел с новостями, раздел с популярными статьями и так далее. Шаблон для каждой страницы был создан собственноручно с нуля без использования изначальных заготовок.

1.3.3 *Search Automatic Plugin*

Плагин, выполняющий задачи парсера и публикующего звена, позволяющий искать новости из выбранных пользователем ресурсов, а также автоматически публиковать их у себя на сайте. В нём существует большое количество фильтров поиска, от таких как есть ли в публикации какое-либо изображение до количества символов в публикации.

					ПГУ 1.090301.06.001 ПЗ	Лист
						11
Изм.	Лист	№ докум.	Подпись	Дата		

Парсер позволяет забрать какой-либо файл в одном исходном формате и в дальнейшем преобразовать его данные в более удобную для нас форму, которую далее мы сможем использовать для собственных целей. Например, изначально может быть файл *HTML* под рукой. С помощью парсинга информация может преобразовываться в обычный текст и стать понятной обычному человеку, а также конвертироваться в другой формат и сделать файл понятным уже для приложения или скрипта.

Но в нашем случае для определения парсинга подходит более узкое определение. Этот процесс можно назвать как метод обработки данных на различных сайтах. Парсер включает в себя анализ текста, изъятие оттуда необходимых материалов и преобразование их в соответствующую форму (такую, которая может быть полезна и использована для конкретных, поставленных за). Благодаря парсингу мы можем найти небольшие кусочки полезной для нас информации на страницах и автоматически извлекать их оттуда, чтобы использовать их позже.

1.4 Инструменты разработки

1.4.1 *HTML*

HTML код, используемый для структурирования и отображения веб-страницы и ее содержимого. Например, контент может быть структурирован в несколько абзацев, отмеченных списков или с использованием изображений и таблиц данных.

HTML - это язык разметки и используется, чтобы пояснить вашему браузеру, как отображать веб-страницы, которые посещает пользователь. Это может быть сложным или простым процессом, в зависимости от того, как хочет веб-дизайнер. *HTML* состоит из серии элементов, которые используют для вставки или оберты различных частей контента, чтобы содержимое отображалось или действовало определенным образом. Ограждающие теги

могут сделать слово или изображение ссылки на что-то еще, могут сделать слова курсивом, сделать шрифт больше или меньше и так далее.

1.4.2 *JavaScript*

JavaScript это язык скриптов, который основывает весь *frontend* веб разработки. Это позволяет перехватывать события и выполнять различные действия. Например, пользователь нажал кнопку событие срабатывает. И, связавшись с ним, мы можем выполнить нужную нам функцию - открыть какое-нибудь окно или изменить цвет элемента.

JavaScript используется, чтобы сделать страницы интерактивными, а конкретно позволяет пользователю взаимодействовать с различными элементами на сайте. В случае, когда страницы и элементы на них могут реагировать на какие-то действия, сайт приобретает свой шарм и индивидуальность.

В *JavaScript* можно работать с переменными, объектами, функциями, условиями, математическими и логическими операторами.

1.4.3 *PHP*

PHP (*Hypertext Preprocessor*) является распространенным языком программирования общего назначения с открытым исходным кодом. *PHP* специально разработан для веб-разработки, его код может быть встроен непосредственно в *HTML*.

PHP отличается от *JavaScript* тем, что *PHP* скрипты работают на сервере и генерируют *HTML*, который отправляется клиенту. Если бы нас был скрипт, размещенный на сервере, то клиент получил бы только результат его выполнения и не смог бы определить, какой именно код его произвел. Мы даже можем настроить свой сервер так, чтобы обычные *HTML* файлы обрабатывались процессором *PHP*, чтобы клиенты даже не могли понять, получают ли они обычный *HTML* файл или результат выполненного скрипта.

1.4.4 CSS

CSS (*Cascading Style Sheets*, каскадные таблицы стилей) – язык, описывающий внешний вид *HTML* файла. Эта технология является одной из основных в современной разработке. Почти невозможно найти сайт без *CSS* вёрстки, поэтому *CSS* и *HTML* действуют в единой идеологии.

Каскадные таблицы стилей работают с *HTML*, но эти языки совершенно разные. *HTML* упорядочивает информацию и создаёт структуру документа, а *CSS* работает напрямую с браузером, чтобы оформить внешний вид страницы.

Поэтому берем за фундамент *HTML* файл и подключаем к нему *CSS* стиль. С помощью *CSS* уже создаем оформление для всего, что необходимо, то есть задаём уникальные свойства для элементов *HTML*.

В каскадной таблице стилей мы прописываем стили, либо же правила, желаемого оформления. В состав каждого правила входит селектор и блок объявления.

Селектор определяет, какие части документа следуют правилам. В фигурные скобки загоняется блок объявлений, в состав его входит одно или более объявление.

1.4.5 Хостинг *ISPmanager*

ISPmanager - одна из самых популярных в России и СНГ панелей управления виртуальными серверами или выделенным *Linux* сервером. Платформа предоставляет удобные инструменты для создания пользователей, сайтов, доменов, адресов и многое другое. Можно создать свой собственный плагин, используя *API*.

Основные возможности *ISPmanager*:

- создание и управление неограниченным количеством пользователей (шаблонов пользователей);
- установить пользовательские ограничения;

- создание и управление web доменами. Возможность выбора режима работы php, установка сертификата *SSL*, быстрая установка популярного *cms* на домен, настройка переадресации и *SSL* переадресации;
- управление *DNS*. Создание доменов и редактирование записей;
- создание почтовых доменов и почтовых ящиков, управление почтовой системой. Установка отправлений и почтовых псевдонимов. Почтовые фильтры. Поддержка *DKIM*. Системы защиты от спама;
- управление базами данных. Создание баз данных и их пользователей с установкой прав доступа;
- создание, включение, отключение *FTP* и управление пользователями. Возможность установки домашней директории пользователя;
- редактирование файлов с помощью панели управления. Установка прав доступа к файлам и каталогам;
- управление брандмауэром;
- система резервного копирования;
- статистика потребляемых ресурсов;
- разграничение прав доступа пользователей.

1.4.6 Sublime text

Sublime Text – это текстовый редактор, предназначенный для верстальщиков и разработчиков. Он позволяет взаимодействовать с кодом различных языков программирования, от *Erlang* до *C++*. Он приобрел популярность благодаря кроссплатформенной поддержке и настройкам, которые легко расширяются, благодаря этим достоинствам разработчик может легко менять параметры программы. Существует достаточно много вариаций - также можно изменить внешний вид и дополнить функциональность различными дополнениями и плагинами.

Инструменты *sublime text*.

Текстовый редактор *sublime* имеет большую функциональность, которая упрощает написание кода и компиляцию. Функции заключаются в следующем:

goto Anything - удобная функция для быстрого доступа к желаемым файлам;

сопоставление скобок – программа быстро определяет неправильное совпадение. Программное обеспечение само выделяет такие наборы скобок и указывает на них;

выделение в нескольких местах - полезная функция для быстрого изменения кода в нескольких местах. Поставив курсор в пару участков программы, можно легко изменять написанный код;

мощный *Python API*, из-за которого *Sublime* обходит конкурентов. Это позволяет добиться высокой производительности с помощью дополнительных плагинов;

отдельное редактирование - благодаря этой функции можно разделить окно на несколько таких же окон, например, в одном редактировать *HTML*, другой для *JavaScript* и оставшийся третий под *CSS* или что-то еще;

сочетание клавиш – позволяет собственноручно назначать горячие клавиши для различных функций.

1.4.7 Cron

Чтобы новостные статьи искались и публиковались регулярно и с определенным интервалом, используется утилита *Cron*.

Cron является планировщиком задач. Если более подробно, то это утилита, которая позволяет запускать скрипты на сервере в определенное время с заданной частотой.

Например, у нас есть определенный скрипт, который собирает какую-нибудь информацию с ресурса каждый день в 8 часов утра. Такие скрипты называются регулярными заданиями и их способ применения описывается в отдельных файлах, называемых *crontab*.

Crontab – это некая таблица с графиком запуска скриптов и определённых сценариев, оформленная в своеобразной форме, которую может читать компьютер. У каждого пользователя системы на его персональном компьютере генерируется собственный отдельный *crontab*-файл с индивидуальным расписанием.

Обычно *cron* нужен для того, чтобы регулярно выполнять создание резервных копий системы, в случае случайной утраты доступа к ней, но также используется для следующих задач:

- можно использовать планировщик для настройки системного времени. На большинстве компьютерах настройка времени происходит через *Network Time Protocol*. И поскольку этот протокол настраивает только время в операционной системе, нужно дополнительно настроить время и для комплектующих сервера, установленное время для данных комплектующих может варьироваться. *Cron* позволяет регулярно корректировать время, установленное для оборудования, в зависимости от времени, настроенном для работы операционной системы;

Другим популярным способом использования является создание заранее настроенных оповещений, которые появляются, допустим, каждый вечер и оповещают о времени, проведённом за компьютером. Эти оповещения могут включать в себя любую полезную информацию для пользователя, которую он хочет видеть регулярно.

С помощью *cron* автоматизируются самые разнообразные задачи, уменьшая вмешательство в работу сервера системного администратора.

1.5 Реализация

В данной работе реализована стандартная последовательность для разработки веб-сайта, которая разделяет проектирование на несколько этапов:

- разработка макета;

- разработка дизайна;
- разработка сайта;
- разработка и добавление различных плагинов и дополнений;
- настройка системы хранения данных.

1.5.1 Создание сайта

После определения тематики сайта необходимо собрать тематические текстовые и графические исходные данные, в качестве них будет выступать наполнение самих статей и новостей.

Далее нужно определиться со структурой проекта, она подразумевает разделы сайта, блоки, исходя из которых будет формироваться навигационное меню и в дальнейшем создаваться дизайн проекта.

Для наглядности была создана схема расположения блоков на странице (рисунок 1.1).



Рисунок 1.1 – Схема структуры сайта

Роль главного контейнера у меня выполняет элемент *body*, ширина данного элемента может быть резиновой (*fluid*) или фиксированной (*fixed*), разница этих параметров продемонстрирована на рисунке 1.2. В данном

проекте блок *body* реализован с параметром *fixed*. Этот этап очень важен, так как на нём будет строиться весь веб сайт, нужно это для того, чтобы на мониторах с любым расширением демонстрация контента была корректна, будь это 4:3 или же 21:9.

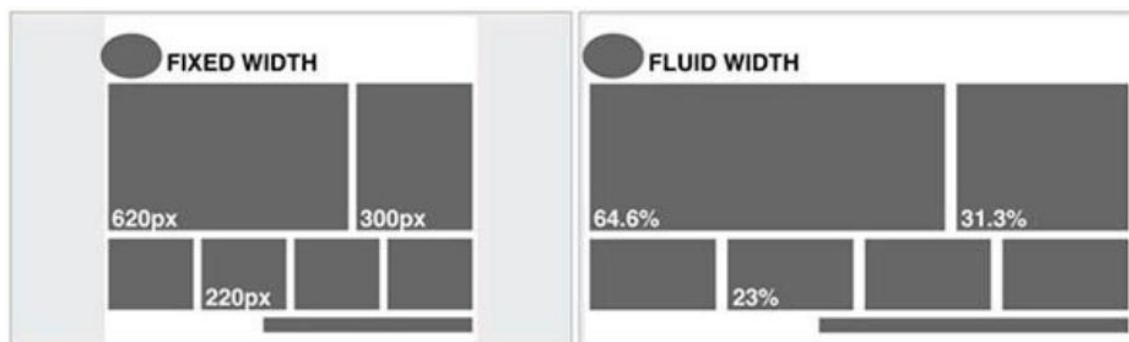


Рисунок 1.2 – Разница между *fixed* и *fluid*

“Разделы” выступают в роли навигационной составляющей сайта, эта панель содержит ссылки на основные подразделы.

Контент - это основа сайта, поэтому занимает наибольшее пространство на странице, помимо текста, все новости подкреплены графическим дополнением.

Нижний колонтитул состоит из ссылки на полезное программное обеспечение, которое можно скачать и установить для личного пользования, а также ссылка на контакты, для решения деловых вопросов, а также возможного указания на баги и ошибки.

Также сайт следует правилу Z-схемы для более удобного и понятного ознакомления с наполняемым контентом.

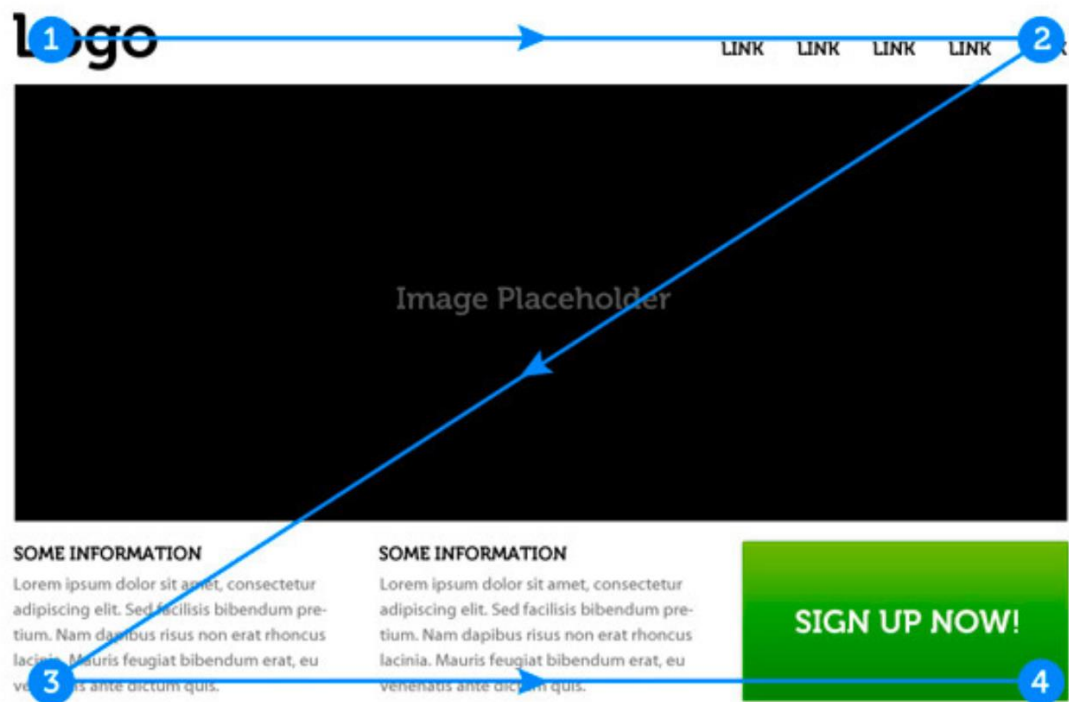


Рисунок 1.3 – Z-схема просмотра страницы

1.5.2 Создание и использование плагина автонаполнения

Встала задача создать плагин ищущий (парсящий), а также автоматически выкладывающий новостные статьи на веб-сайт.

Анализатор новостей или же парсер - программа, которая собирает все необходимые данные, которые есть в программе. Давайте представим ситуацию, когда необходимо отслеживать изменения в новостях и появление новых на дюжине различных веб сайтов по определенным заданным параметрам и переводить полученную информацию на транслирование своей аудитории.

Основной задачей для написания парсера, а в моём случае, парсера новостных статей - заполучить страницу из стороннего источника, извлечь из нее необходимое наполнение и поработать с ним, пример подобной страницы показана рисунке 1.4.



Рисунок 1.4 – Страница новостей в “Яндекс”

Основное, что нужно сделать, это проанализировать и обработать *HTML* разметку и понять, где располагаются необходимые новости (рисунок 1.5):

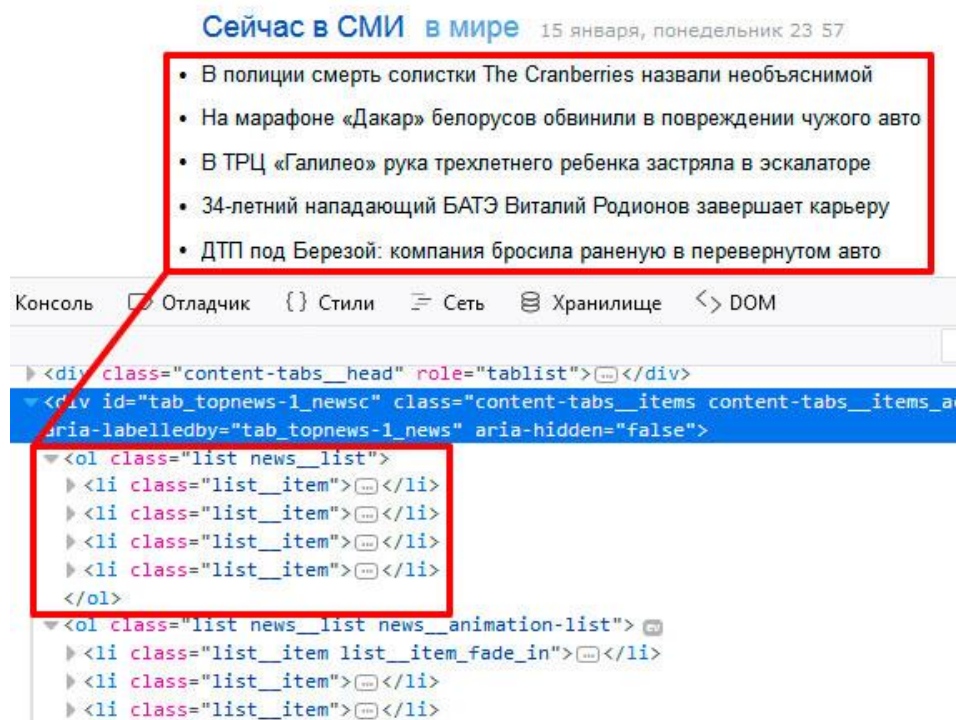


Рисунок 1.5 – Расположение новостей в вёрстке

Все новостные статьи находятся в двух списках. Первый из них это видимый список, а оставшийся – невидимый (скрытый), нужно это для эффекта смены новостей. Сама же новостная ссылка на публикацию выглядит в виде ссылки (рисунок 1.6):

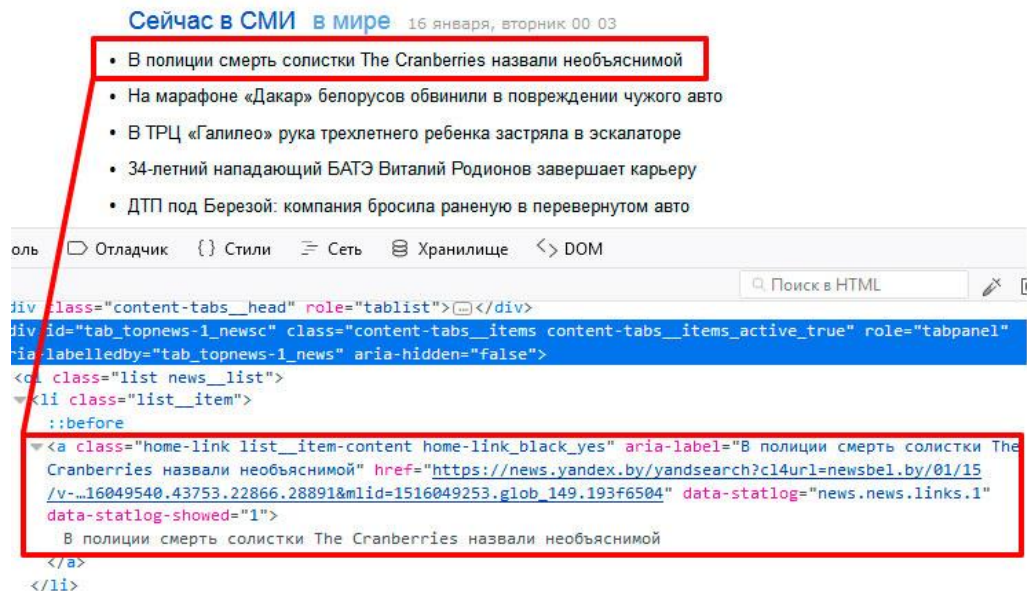


Рисунок 1.6 – Ссылка на новостную статью

После получения страницы, нам нужно вытащить все эти ссылки из каждой новости, из которой в дальнейшем возьмем содержимое в виде текста, картинки, а также вытянем значение атрибута *href* [1].

При парсинге и механизме выкачивания контента я использую функцию *file_get_contents()*. Также для разбора контента потребуется библиотека – *phpQuery*, очень востребованная и современная, так как парсинг страниц с каждым годом приобретает все большие обороты.

За сервер будет отвечать *opensever*, это локальный WAMP/WMMP сервер, на котором требуется создать сайт *parser.loc*, где будут отображаться новости, далее с помощью *composer* устанавливаю библиотеку *phpQuery*. Запуск сервера представлен на рисунке 1.7:

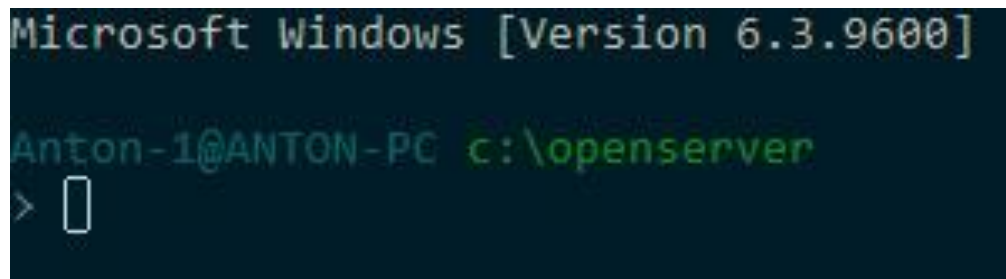
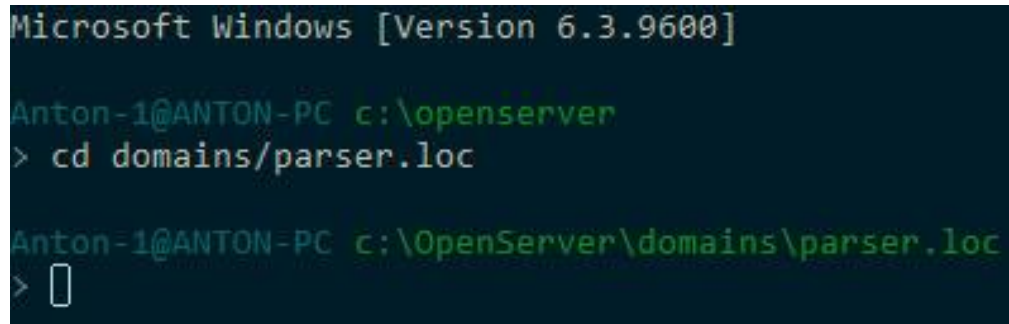


Рисунок 1.7 – Запуск сервера

Переход в папку с сайтом, процесс представлен на рисунке 1.8:

```
cd domains/parser.loc
```



```
Microsoft Windows [Version 6.3.9600]

Anton-1@ANTON-PC c:\openserver
> cd domains/parser.loc

Anton-1@ANTON-PC c:\OpenServer\domains\parser.loc
> 
```

Рисунок 1.8 – Открытие папки с сайтом

Установка *phpQuery* представлена на рисунке 1.9:

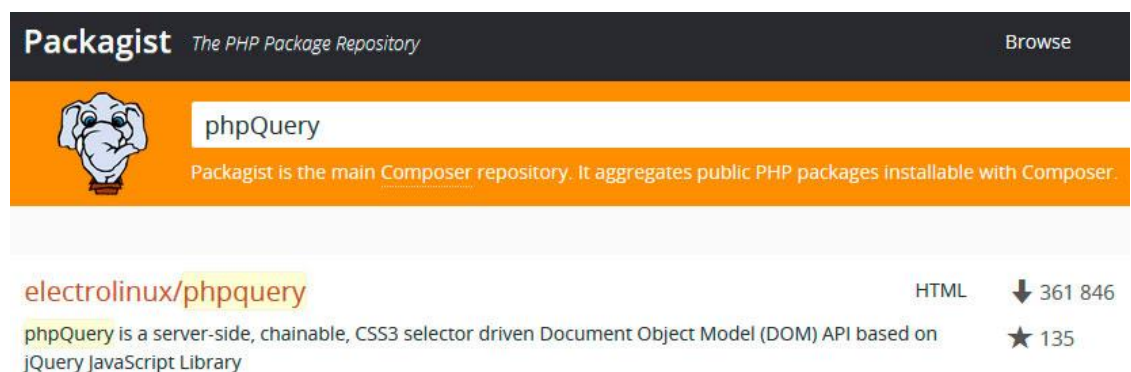


Рисунок 1.9 – Библиотека *phpQuery*

В консоли была выполнена установка пакета, а далее был создан файл *index.php* и после закончена работа с консолью, процесс показан на рисунке 1.10:

```

Microsoft Windows [Version 6.3.9600]

Anton-1@ANTON-PC c:\openserver
> cd domains/parser.loc

Anton-1@ANTON-PC c:\OpenServer\domains\parser.loc
> composer require electrolinux/phpquery
Using version ^0.9.6 for electrolinux/phpquery
./composer.json has been created
Loading composer repositories with package information
Updating dependencies (including require-dev)
Package operations: 1 install, 0 updates, 0 removals
  - Installing electrolinux/phpquery (0.9.6): Loading from cache
Writing lock file
Generating autoload files

Anton-1@ANTON-PC c:\OpenServer\domains\parser.loc
> copy con index.php
<?php
^Z
Скопировано файлов:      1.

Anton-1@ANTON-PC c:\OpenServer\domains\parser.loc
> |

```

Рисунок 1.10 – Создание *index.php*

Этап с установкой библиотеки в исполняемый файл, а также этап написания кода:

```
require_once("vendor/autoload.php");
```

Далее вытянем страницу и сразу отобразим её для проверки, отображение страницы представлено на рисунке 1.11:

```
$html = file_get_contents("https://yandex.ru/");
echo($html);
```

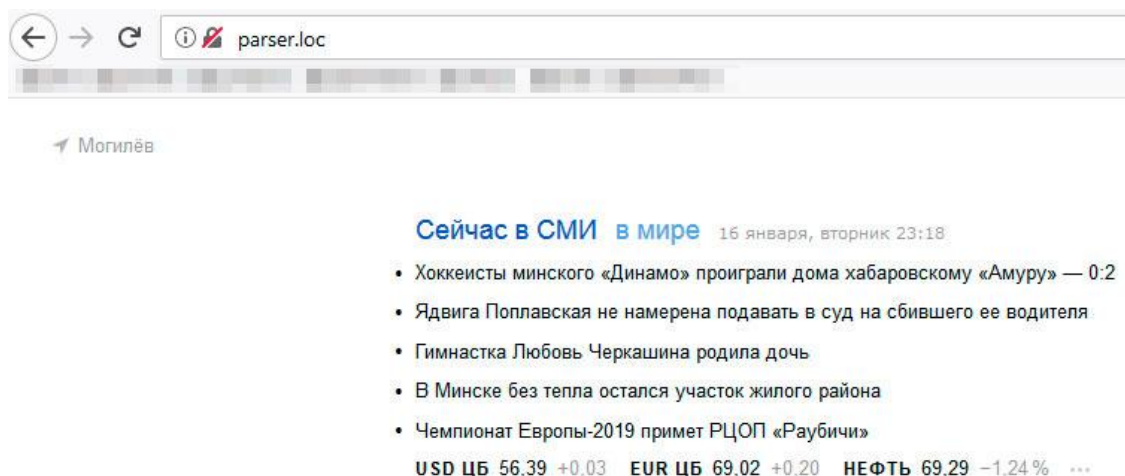


Рисунок 1.11 – Отображение страницы

Теперь нужно спарсить имена новостных статей и ссылки на них, чтобы было удобнее результат парсинга занесу в массив.

```

phpQuery::newDocument($html);
$links = pq("#tabnews_newsc")->find("a");
$tmp = array();
foreach($links as $link){
    $link = pq($link);

    $tmp[] = array(
        "text" => $link->text(),
        "url"  => $link->attr("href")
    );
}
phpQuery::unloadDocuments();

```

Когда получили страницу, её нужно передать методу *newDocument()*, в результате получился объект файла, в дальнейшем его нужно обрабатывать. Важным этапом является удаление этого файла после всего алгоритма. Удаление происходит с помощью методики *unloadDocuments()*. Чистка памяти осуществляется для того, чтобы при парсинге сотен и более страниц, процесс шёл гладко и правильно, так как в случае хранения всех объектов в памяти, весь алгоритм с каждым разом будет работать все медленнее [2]. Теперь по аналогии вытягиваю все ссылки и отображаю их для наглядности (рисунок 1.12):

```

<ul>
    <?php foreach($tmp as $value): ?>
    <li>
        <a      href="<?php      echo($value["url"]);      ?>"
target="_blank">
            <?php echo($value["text"]); ?>
        </a>
    </li>
    <?php endforeach; ?>
</ul>

```

- [Хоккеисты минского «Динамо» проиграли дома хабаровскому «Амuru» — 0:2](#)
- [Ядвига Поплавская не намерена подавать в суд на сбившего ее водителя](#)
- [Гимнастка Любовь Черкашина родила дочь](#)
- [Двух витебских милиционеров уволили за изготовление и продажу самогона](#)
- [В Минске без тепла остался участок жилого района](#)
- [Арина Соболенко проиграла в первом круге Australian Open](#)
- [Софья Лапина отказалась от участия в отборе на «Евровидение»](#)
- [Чемпионат Европы-2019 примет РЦОП «Равбичи»](#)
- [Минские спасатели помогли достать застрывшего под киоском кота](#)
- [В Минске пьяный водитель избил мужчину, который снимал его на телефон](#)

Рисунок 1.12 – Готовые ссылки на новости

В парсинге существует такой нюанс, что при большом количестве собирания ссылок с одного ресурса, он может вас заблокировать по юзер агенту, в проекте это предусмотрено, поэтому каждый процесс парсинга юзер агент меняется с помощью функции *replace* перед процессом собирания информации *agent* меняется:

```
$chrome_version = rand ( 31, 81 ) . '.0.' . rand ( 1, 4044 )
. '.' . rand ( 0, 138 );
// Chrome
if ($os_type == 1) {
    $agent = "Mozilla/5.0 (Windows NT 10.0; Win64;
x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/81.0.4044.138
Safari/537.36";
} elseif ($os_type == 2) {
    $agent = "Mozilla/5.0 (Macintosh; Intel Mac OS X
10_15_4) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/81.0.4044.138 Safari/537.36";
    $macversion = "10" . rand ( 6, 15 ) . "_" . rand
( 1, 4 );
    $agent = str_replace ( '10_15_4', $mac_version,
$agent );
} elseif ($os_type == 3) {
    $agent = "Mozilla/5.0 (X11; Linux x86_64)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/81.0.4044.138
Safari/537.36";
}
// main version 81
```

```
$agent = str_replace ( "81.0.4044.138",
$chrome_version, $agent );
```

Происходит это за счет смены браузеров и прокси. Алгоритмы первоисточника считают, что его страницу каждый раз парсят разные пользователи.

Новость на сайте также не появляется без сопутствующего изображения, изображение из первоисточника скачивается и хранится на хостинге:

```
// saving image
    if (! isset ( $idental )) {
        if (wp_mkdir_p ( $upload_dir ['path'] ))
            $file = $upload_dir ['path'] . '/' . $filename;
        else
            $file = $upload_dir ['basedir'] . '/' . $filename;
        $f = file_put_contents( $file,$image_data );
    }
```

После того как изображение появилось на хостинге, его нужно выложить вместе с текстом в виде новостной публикации, делается это при помощи нескольких строчек и функции *post_content()* :

```
    if (in_array ( 'OPT_REPLACE', $camp_opt )) {
foreach ( $keywords as $keyword )
    $keyword = trim ( $keyword );
    if (trim ( $keyword != '' )) {
$post_content = str_replace ( $keyword, '<a href="' . $camp-
>camp_replace_link . '">' . $keyword . '</a>', $post_content );
    }
    }
}
```

Ещё не стоит забывать об удалении гиперссылок в текстовой публикации, их на сайте агрегаторах быть не должно, поэтому этот момент тоже был предусмотрен [3]:

```
function strip_urls($content) {
return preg_replace ( '{http[s]?://[^\s]*}', '', $content );
}
```

Функция *strip_urls* находит в контенте гиперссылки и снимает их с текста.

Загрузка новости происходит по обычному алгоритму:

```
$upload_dir = wp_upload_dir ();
$file_path = $upload_dir ['basedir'] . '/wp_automatic_temp'
. $ext;
$file_url = $upload_dir ['baseurl'] . '/wp_automatic_temp' .
$file_ext;
file_put_contents ( $file_path, $exec );
return $file_url;
```

Файл выгружается из директории хостинга и следом грузится на сайт с изображением и правильно выставленной категорией.

1.6 Описание работы сайта

Главная страница выполнена в виде стандартной стартовой страницы любого новостного веб-ресурса (рисунок 1.13).

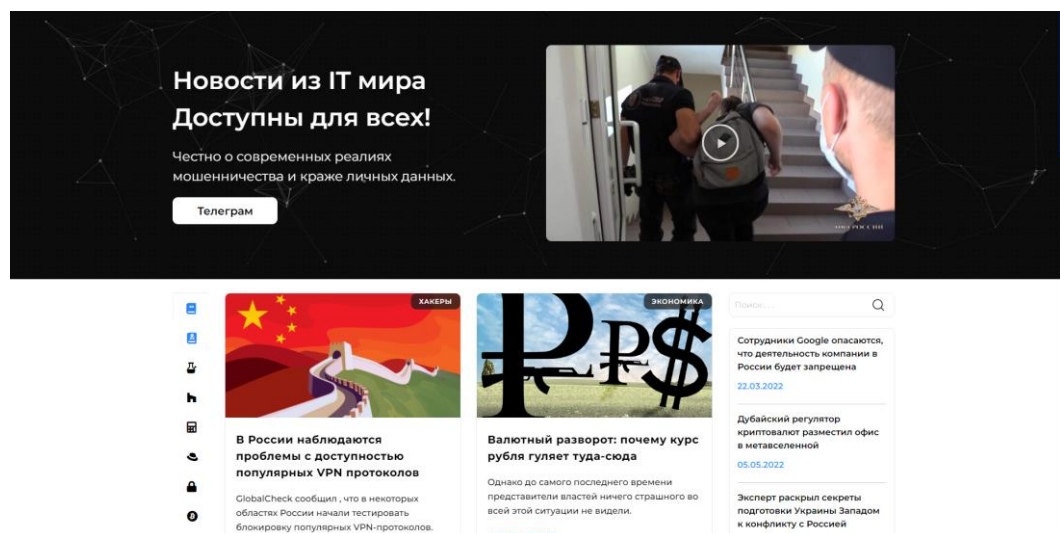


Рисунок 1.13 – Готовые ссылки на новости

Расположение зон на главной странице интуитивно понятно. Вверху расположена шапка сайта. Слева представлена область с разделами веб-ресурса. Справа собраны новости, набравшие наибольший обхват среди

пользователей и поиск статей по горячим словам. Пролистав страницу до конца, мы наткнёмся на ссылку, ведущую к полезным программам для персонального компьютера и контакты, с помощью которых можно связаться для решения различных вопросов.

Далее было создано разбиение на разделы новостей и статей (рисунок 1.14). Все разделы выполнены в виде тематических иконок на ту сферу, с которой связана область. Одним кликом мыши пользователь может выбрать для себя интересующую его подборку статей или сферы новостей.

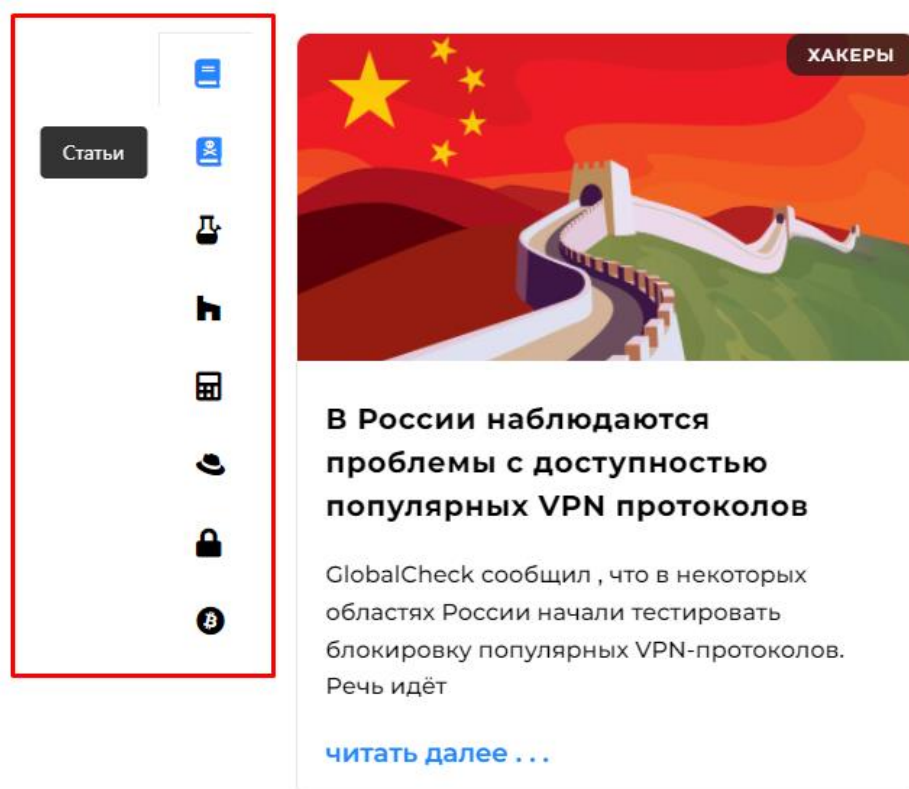


Рисунок 1.14 – Разбиение на разделы

Существуют такие разделы как статьи, где собраны обучающие публикации на тему того, как нужно вести себя в интернете. Также присутствуют ответвления на тематику:

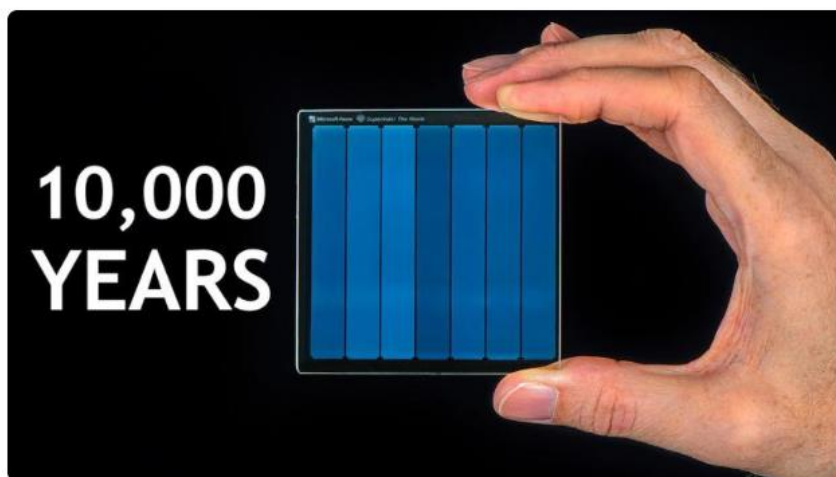
- криминал;
- политика;
- экономика;

- хакерство;
- интернет-безопасность;
- блокчейн.

Справа расположена боковая панель взаимодействия. Подборка с популярными новостями (рисунок 1.15), тут опубликованы новости или статьи, набравшие наибольший охват среди других публикаций за определенный период времени.

Microsoft работает над созданием музыкального “хранилища судного дня”

13 июня, 2022 · 16:30



Microsoft работает над созданием всемирного музыкального хранилища, которое будет хранить музыку тысячи лет. Музыкальное хранилище будет расположено там же, где и всемирное хранилище семян – на Шпицбергене.

Проект Microsoft получил название Project Silica и представляет собой стеклянный жесткий диск, способный вмещать до 100 Гб цифровых данных. Кварцевое стекло имеет симметричную молекулярную структуру, что делает его более устойчивым к высоким температурам и давлению.

«На этих дисках будут сохранены самые важные музыкальные файлы. Они отправятся в хранилище и останутся навсегда в истории», – говорит компания Elire Group занимающаяся проектированием хранилища.

Вместе с Microsoft, компания тестирует систему хранения на основе кварцевого стекла, которая может сохранить музыку дольше, чем любой доступный нам носитель информации.

Рисунок 1.15 – Страница новостной публикации

Перейдя на любую из новостей, откроется стандартная страница с новостной статьей (рисунок 1.15), изображением, временем и датой добавления на сайт. В конце статьи расположена гиперссылка на источник.

Также расположен поиск по сайту, пользователь может воспользоваться отбором по горячим словам, чтобы найти интересующую его новость или статью.

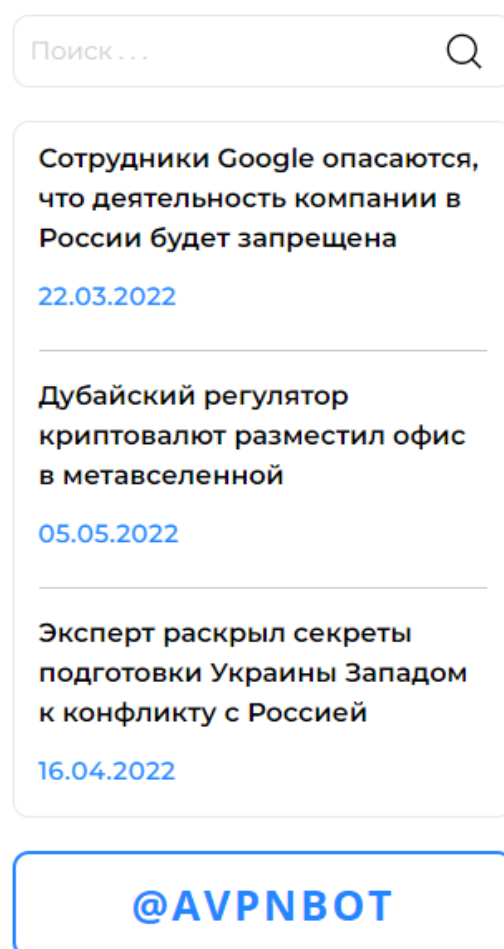


Рисунок 1.16 – Самые интересные новости

Введя в поиск, к примеру “Илон Маск”, откроется страница со всеми последними новостями, связанными с горячим словом, введенным в поиск, показано на рисунке 1.17.

Илон Маск передал Украине солнечные панели Tesla Powerwall

Глава ведомства уточнил, что Маск прислал солнечные панели для обеспечения работы объектов критической инфраструктуры Киева и области. Они будут подавать резервное питание

[читать далее »](#)

Илон Маск купил Twitter за 44 млрд долларов

Американский предприниматель Илон Маск выкупил социальную сеть Twitter, выплатив в общей сложности 44 млрд долларов, говорится в заявлении компании. «Twitter,

[читать далее »](#)

Илон Маск открыл доступ к спутниковому интернету Starlink на территории Украины

До недавнего времени ближайшей к Украине западной страной, на территории которой оператор спутниковой связи Starlink осуществлял официальную деятельность, была Польша.

[читать далее »](#)

Илон Маск рекламирует новую криптовалютную аферу

Мошенник использовал дипфейки Илона Маска и других известных личностей для продвижения мошенничества с торговой платформой BitVex, которая крадет депонированную криптовалюту.

[читать далее »](#)

В следующем году **Илон Маск** начнет производство роботов-гуманоидов

Tesla может начать производство робота-гуманоида, известного как Optimus, уже в следующем году, заявил в четверг генеральный директор Илон Маск. Производитель

[читать далее »](#)

Microsoft и Amazon готовы спасти сотрудников Tesla от **Илона Маска**

После заявлений Илона Маска о запрете удаленного режима работы, а также планируемом сокращении штата компании на 10%, рекрутеры нескольких высокотехнологических

[читать далее »](#)

Илон Маск хочет ввести сквозное шифрование в twitter сообщениях

Через несколько дней после покупки сервиса микроблогов Илон Маск заявил о необходимости введения функции шифрования. Генеральный директор SpaceX и Tesla

[читать далее »](#)

Илон Маск попросил использовать Starlink в Украине «с осторожностью» из-за риска атак

Через несколько дней после отправки интернет-терминалов SpaceX Starlink в Украину Илон Маск предупредил власти использовать устройства с «осторожностью». Являясь не

[читать далее »](#)

Рисунок 1.17 – Поиск по горячим словам

Если пользователь пролистает страницу до конца, то наткнётся на раздел с полезным программным обеспечением. Здесь представлены программы, упрощающие работу в разных сферах деятельности:

- антивирусы;
- фото и видео редакторы;
- операционные системы;
- мессенджеры;
- архиваторы;
- другое полезное ПО (рисунок 1.18)

Графика

[Photoshop](#) 

Браузеры

[Brave](#) [Chrome](#) [Firefox](#) [Firefox / Import](#) [TOR](#) [Opera GX](#) 

Мессенджеры

[Telegram](#) [Signal](#) [Wickr](#) [QTox](#) 

Рисунок 1.18 – Страница с полезным ПО

Изм.	Лист	№ докум.	Подпись	Дата

2 Охрана труда

2.1 Требования к организации рабочего места с ПК

Нужно соблюдать расстояние в 2 метра между столами не менее 1 метр между мониторами при расстановки рабочих мест в помещении.

Согласно СанПиН 2.2.2/2.4.1340-03, площадь одного рабочего места пользователя с монитором на базе электроннолучевой трубки (ЭЛТ) должна составлять более 6 квадратных метров, в помещениях культурно-развлекательных учреждений и с монитором на базе плоских дискретных экранов (жидкокристаллические, плазменные) - 4,5 квадратных метра. При использовании персонального компьютера с монитором перового типа с продолжительностью работы не более 4-х часов в день допускается минимальная площадь 5 квадратных метров на одно рабочее место пользователя старше 18 лет.

Рабочие места с персональным компьютером для работы, требующей умственного напряжения или высокой концентрации внимания, рекомендуется изолировать друг от друга перегородками высотой 1,5-2,0 метра.

Расстояние между монитором и глазами пользователя должно быть 60-70 сантиметра, но не менее 50 сантиметров с учётом размеров знаков и символов. Расположения клавиатуры должно быть 10-30 сантиметров от края, обращённого к пользователю, или на регулируемой по высоте рабочей поверхности, отделённой от основной столешницы.

Пользовательские стул (кресло) должно обеспечивать поддержание правильной позы при работе, чтобы в районе шейно-плечевой области и спины не развивалась усталость и избежании развития искривления позвоночника. Рабочий стул (кресло) следует выбирать с учётом роста пользователя, продолжительности и характера работы с персональным компьютером.

					ПГУ 1.090301.06.001 ПЗ	Лист
						34
Изм.	Лист	№ докум.	Подпись	Дата		

Рабочий стул (кресло) должен быть подъемно-поворотным, с возможностью изменения угла наклона спинки и изменением высоты подлокотников. При этом каждая регулировка должна быть независимой, легко осуществляемой и иметь надёжную фиксацию. Поверхности сиденья, спинки и других элементов стула (кресла) должны быть полумягкими, с нескользящим, слабо электризующимся и воздухопроницаемым покрытием.

2.2 Требования к освещению помещений и рабочих мест

Компьютерные залы должны иметь искусственное и естественное освещение. Световой поток из оконного проёма должен падать на рабочее место пользователя с левой стороны и с коэффициентом естественного освещения не ниже 1,2% в зонах со снежным покровом и не ниже 1,5% на остальной территории. Искусственное освещение в помещениях с компьютерами должно быть равномерным. Освещённость на столах должна быть 300-500 лк. Допускается установка светильников местного освещения для подсветки документов. Местное освещение не должно создавать бликов на поверхности экрана и увеличивать освещённость экрана более 300 лк. Яркость светящихся поверхностей, находящиеся в поле зрения, должна быть не более 200 кд/м². Яркость бликов на экране монитора не должна превышать 40 кд/м². Для их уменьшения стоит повернуть монитор от попадания естественного или местного освещения.

Для искусственного освещения помещений с персональными компьютерами следует применять светильники с зеркализированными решётками. Допускается применять светильники прямого свет с люминисцентными лампами или светильники местного освещения с лампами накаливания. Светильники должны располагаться в виде сплошных или прерывистых линий сбоку от рабочих мест параллельно линии зрения пользователя при разном расположении компьютеров. Защитный угол светильников должен быть не менее 40 градусов [4].

					ПГУ 1.090301.06.001 ПЗ	Лист 35
Изм.	Лист	№ докум.	Подпись	Дата		

2.3 Режим труда и отдыха при работе с компьютером

Режим труда и отдыха предусматривает соблюдение определённой длительности непрерывной работы на ПК и перерывов, регламентированных с учётом продолжительности рабочей смены, видов и категории трудовой деятельности.

Виды трудовой деятельности на ПК разделяются на 3 группы: работа по считыванию информации с экрана, работа по вводу информации, творческая работа на ПК с высоким умственным напряжением или высокой концентрации внимания.

Если в течение рабочей смены пользователь выполняет разные виды работ, то его деятельность относят к той группе работ, на выполнение которой тратится не менее 50% времени рабочей смены.

Количество и длительность регламентированных перерывов, их распределение в течение рабочей смены устанавливается в зависимости от категории работ на ПК и продолжительности рабочей смены.

При 8-часовой рабочей смене и работе на ПК регламентированные перерывы следует устанавливать:

- для первой категории работ через 2 часа от начала смены и через 2 часа после обеденного перерыва продолжительностью 15 минут каждый;
- для второй и третьей категории работ — через 1,5- 2,0 часа от начала рабочей смены и через 1,5-2,0 часа после обеденного перерыва продолжительностью 20 минут каждый или продолжительностью 15 минут через каждый час работы.

При 12 часовой рабочей смене регламентированные перерывы должны устанавливаться в первые 8 часов работы аналогично перерывам при 8-часовой рабочей смене, а в течение последних 4 часов работы, независимо от категории и вида работ, каждый час продолжительностью 15 минут. Продолжительность непрерывной работы на ПК без регламентированного перерыва не должна превышать 2 часа. При работе на ПК в ночную смену

					ПГУ 1.090301.06.001 ПЗ	Лист
						36
Изм.	Лист	№ докум.	Подпись	Дата		

продолжительность регламентированных перерывов увеличивается на 60 минут независимо от категории и вида трудовой деятельности.

Эффективными являются нерегламентированные перерывы длительностью 1-3 минуты. Регламентированные перерывы целесообразно использовать для выполнения комплекса упражнений и гимнастики для глаз, пальцев рук, а также массажа. Комплексы упражнений целесообразно менять через 2- 3 недели.

Пользователям ПК, выполняющим работу с высоким уровнем напряженности, показана психологическая разгрузка во время регламентированных перерывов и в конце рабочего дня в специально оборудованных помещениях.

Досуг рекомендуется использовать для пассивного и активного отдыха (занятия на тренажерах, плавание, езда на велосипеде, бег, игра в теннис, футбол, лыжи, аэробика, прогулки по парку, лесу, экскурсии, прослушивание музыки и т.п.). Дважды в год (весной и поздней осенью) рекомендуется проводить курс витаминотерапии в течение месяца. Следует отказаться от курения. Категорически должно быть запрещено курение на рабочих местах и в помещениях с ПК.

					ПГУ 1.090301.06.001 ПЗ	Лист
						37
Изм.	Лист	№ докум.	Подпись	Дата		

Заключение

В рамках данной выпускной квалификационной работы был создан новостной веб портал, публикующий новости из других источников автоматически.

На основании проделанной работы веб сайт обладает такими возможностями:

- автоматически проверять обновления новостей на сайтах первоисточниках;
- автоматически публиковать новостную статью в правильную тематическую категорию;
- правильное настроенное семантическое ядро и большая ссылочная масса веб ресурса;
- сайт на любом расширении и устройстве демонстрируется без каких-либо недочетов.

Во время выполнения данной работы были получены навыки разработки сайтов на языке *PHP* и *JS* с использованием библиотеки *PhpQuery* и системой регулярно выполняемых задач *Cron*, а также реализации парсера любых новостных сайтов.

За время работы с данным проектом были изучены новые способы и механизмы веб программирования и проектирования веб ресурсов.

В дальнейшем это можно усовершенствовать до создания мобильного приложения. Также добавить систему авторизации пользователя, а как следствия, ведения личных блогов различных авторов, составления и написания статей. Пользователи смогут знакомиться, комментировать и следить за личными блогами авторов.

Список использованных источников

1. Codegeekz.com [Электронный ресурс]. URL: <https://codegeekz.com/> (дата обращения: 20.04.2022)
2. Php.net [Электронный ресурс]. URL: <https://www.php.net/> (дата обращения: 31.04.2022)
3. W3schools.com [Электронный ресурс]. URL: <https://www.w3schools.com/Js/> (дата обращения: 2.05.2022)
4. 1posvetu.ru [Электронный ресурс]. URL: <https://1posvetu.ru/montazh-i-nastrojka/raschet-osveshheniya-rabochego-mesta.html> (дата обращения: 14.05.2022)

Приложение А

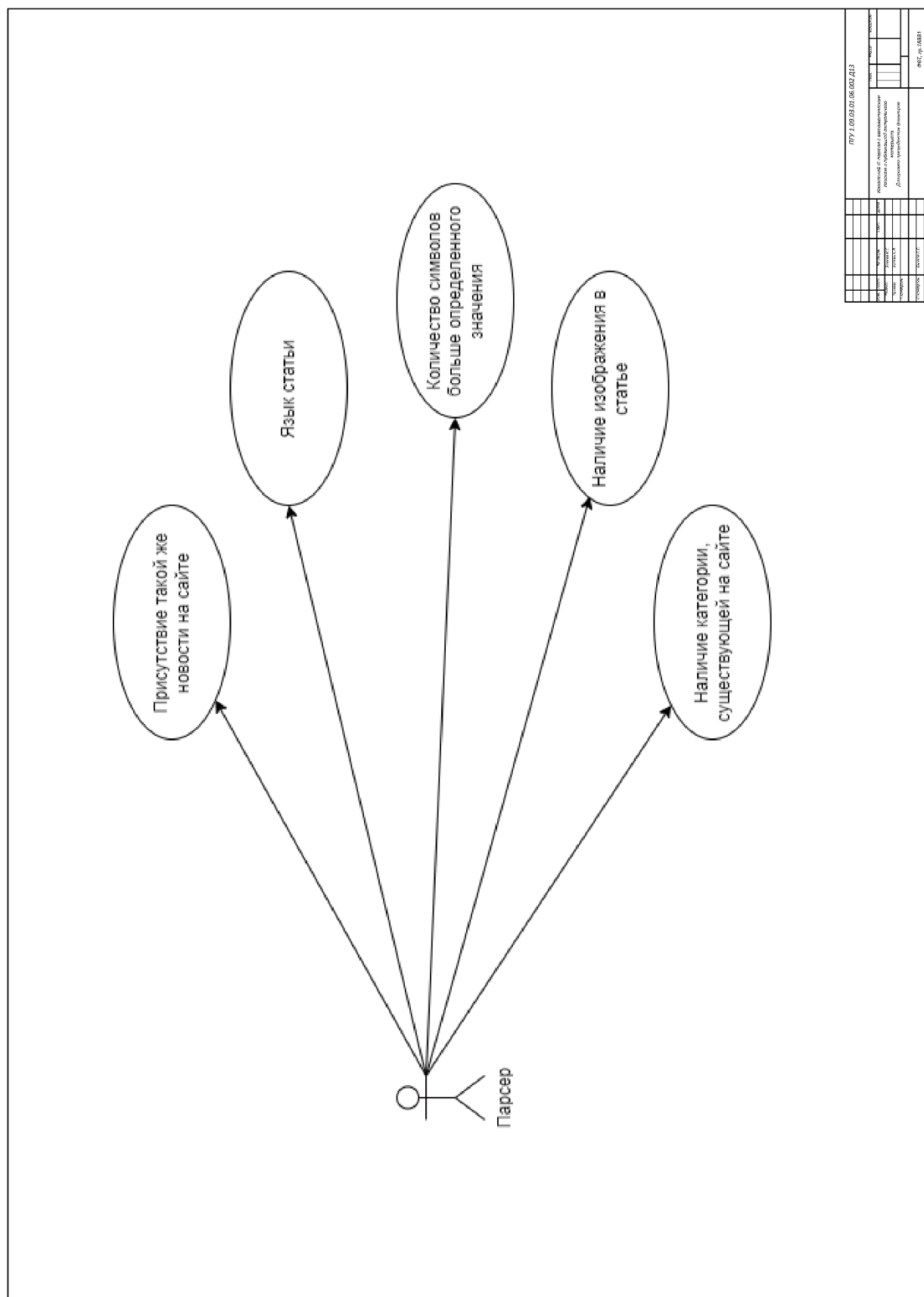
(Обязательное)

Новостной IT портал с автоматическим поиском и публикацией актуального
материала

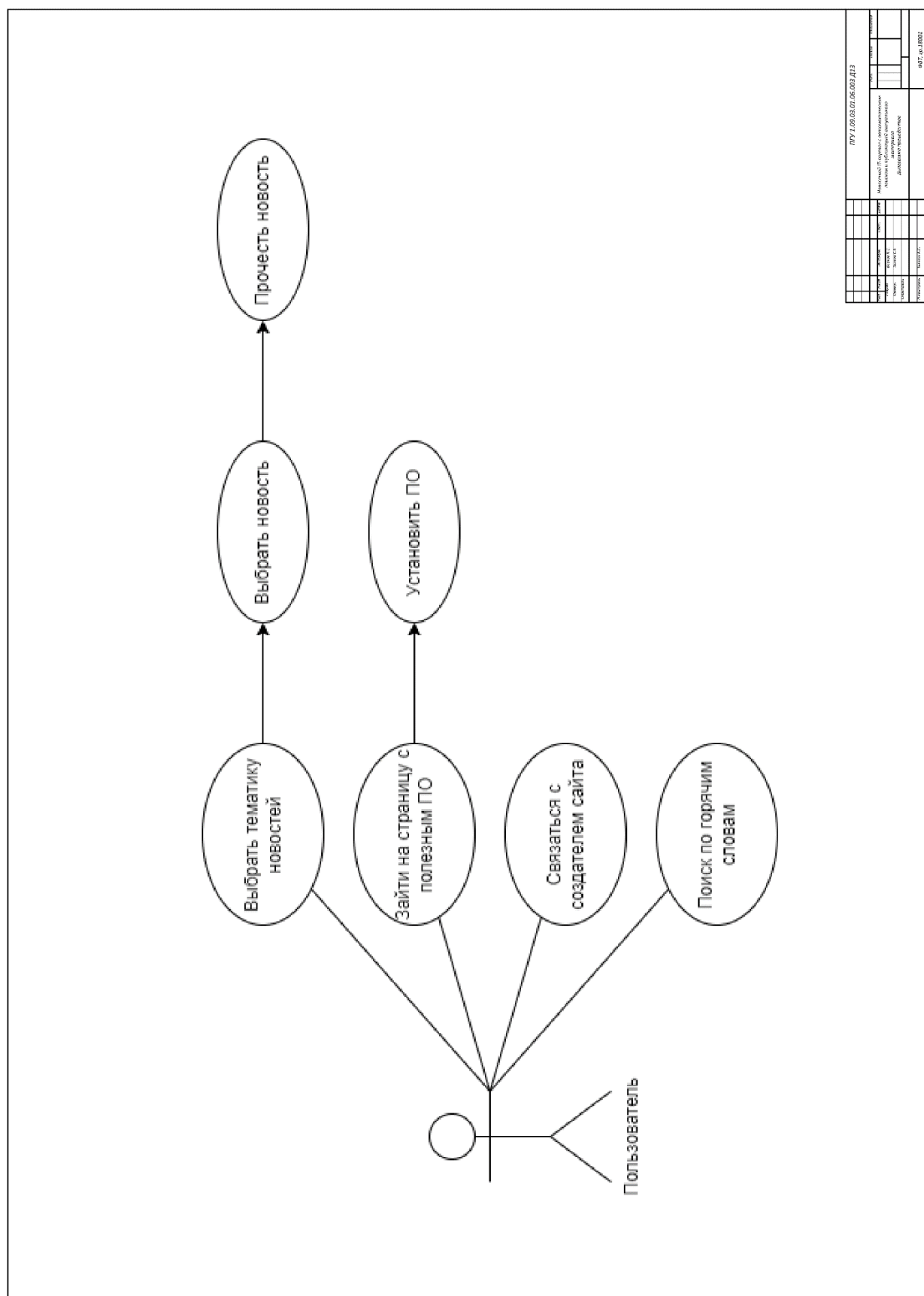
UML - диаграммы

					ПГУ 1.090301.06.001 ПЗ	Лист
						40
Изм.	Лист	№ докум.	Подпись	Дата		

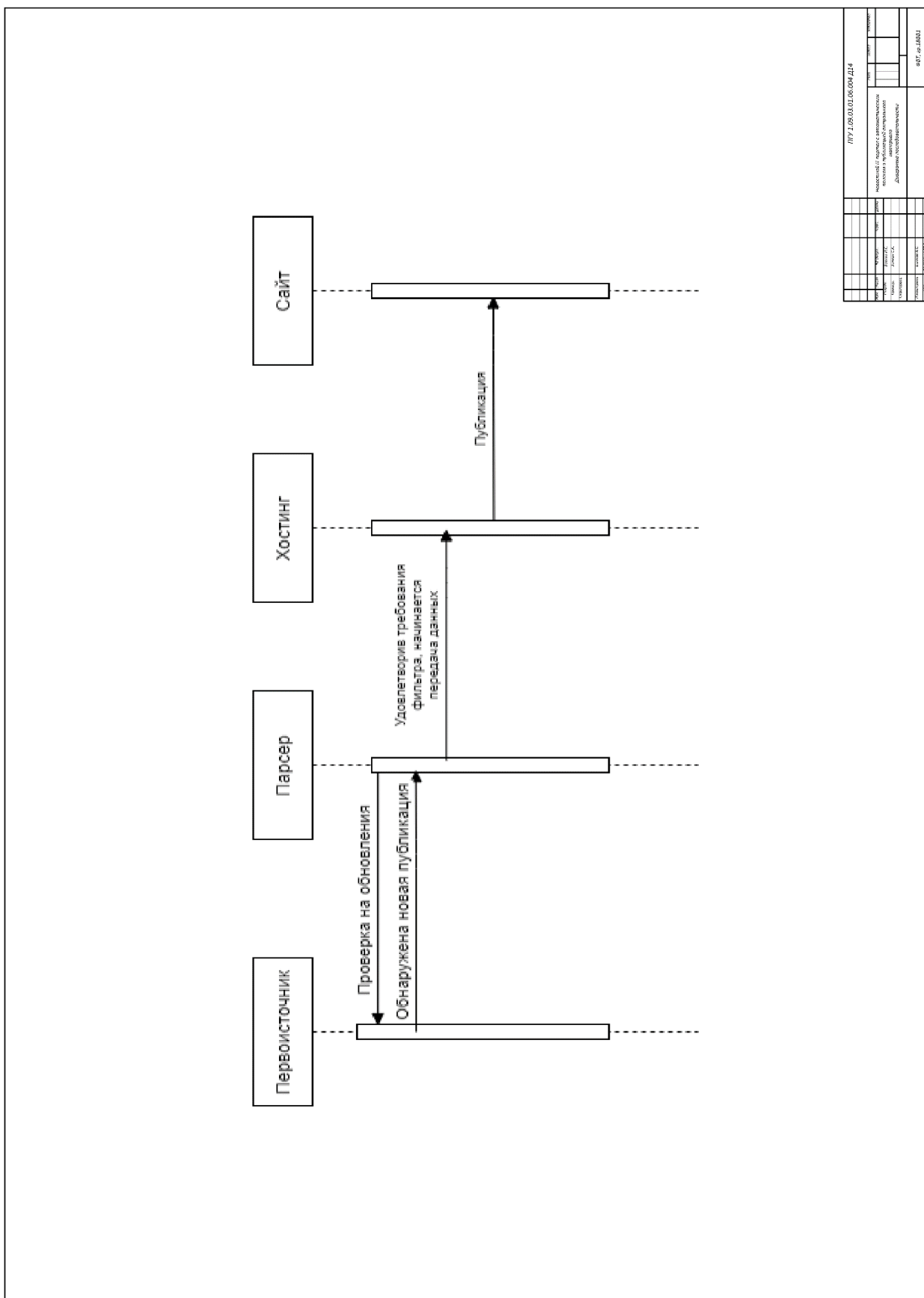
Приложение А.1 – Диаграмма прецедентов фильтров



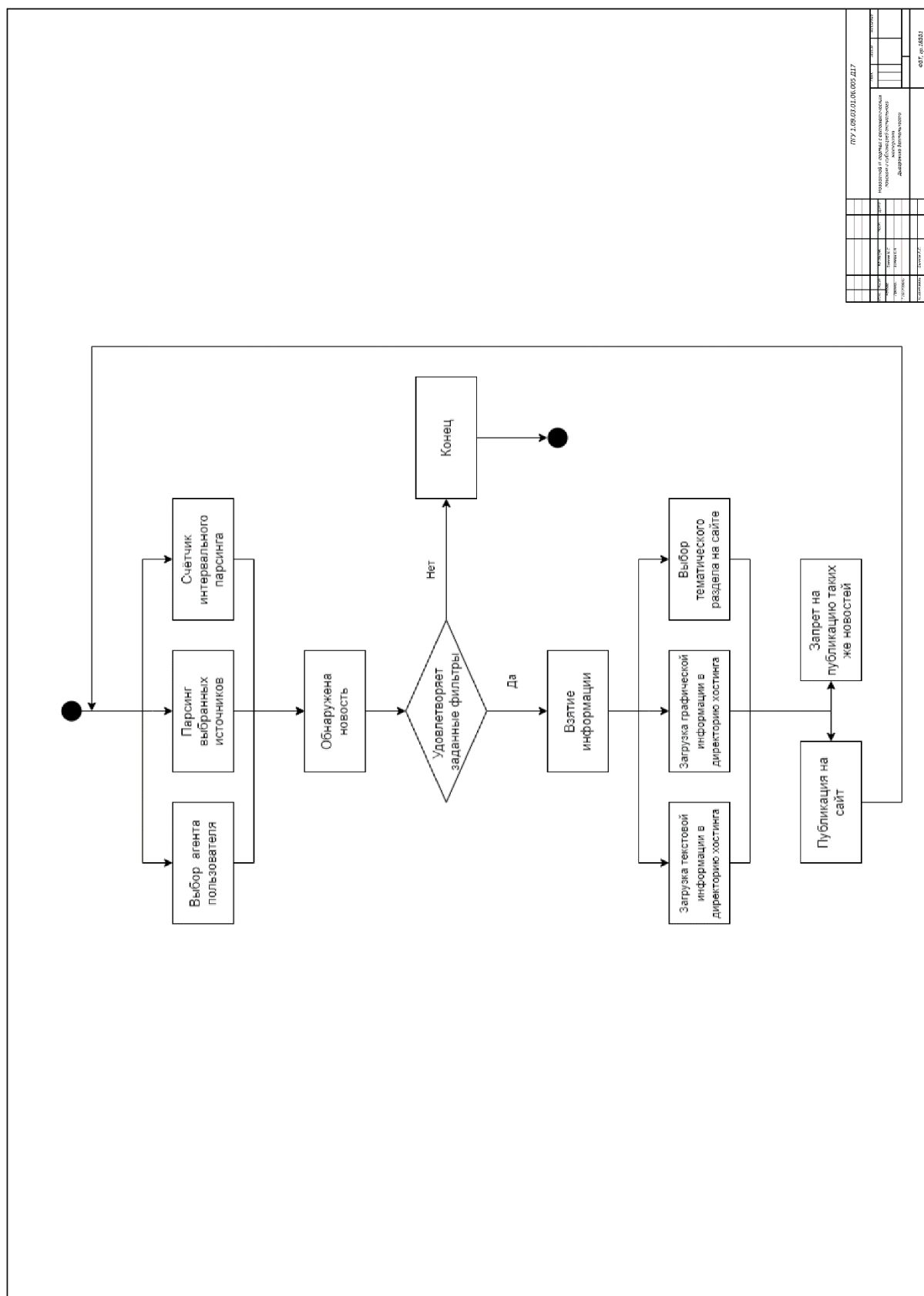
Приложение А.2 – Диаграмма прецедентов



Приложение А.3 – Диаграмма последовательности



Приложение А.4 – Диаграмма деятельности



Приложение Б

(обязательное)

Новостной IT портал с автоматическим поиском и публикацией актуального материала

Листинг программы

Файл “download”

```
function download_file($url, $ext = '') {

    // curl get
    $x = 'error';

    curl_setopt ( $this->ch, CURLOPT_HTTPGET, 1 );
    curl_setopt ( $this->ch, CURLOPT_URL, trim ( $url ) );
    $exec = curl_exec ( $this->ch );
    $x = curl_error ( $this->ch );

    if (trim ( $x ) == '' && trim ( $exec ) != '') {

        $upload_dir = wp_upload_dir ();
        $filePath = $upload_dir ['basedir'] . 'wp_automatic_temp' .
$ext;

        $fileUrl = $upload_dir ['baseurl'] . 'wp_automatic_temp' . $ext;
        file_put_contents ( $filePath, $exec );
        return $fileUrl;
    }

    return false;
}
```

Файл “Non-block_parsing”

```
function randomUserAgent() {
    $os_type = rand ( 1, 3 );

    $chrome_version = rand ( 31, 81 ) . '.0.' . rand ( 1, 4044 ) . '.' .
rand ( 0, 138 );

    // Chrome
    if ($os_type == 1) {

        $agent = "Mozilla/5.0 (Windows NT 10.0; Win64; x64)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/81.0.4044.138 Safari/537.36";
    } elseif ($os_type == 2) {

        $agent = "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_4)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/81.0.4044.138 Safari/537.36";
        $macversion = "10" . rand ( 6, 15 ) . "_" . rand ( 1, 4 );
        $agent = str_replace ( '10_15_4', $mac_version, $agent );
    } elseif ($os_type == 3) {
        $agent = "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36
(KHTML, like Gecko) Chrome/81.0.4044.138 Safari/537.36";
    }
}
```

					ПГУ 1.090301.06.001 ПЗ	Лист
						45
Изм.	Лист	№ докум.	Подпись	Дата		

```

// main version 81
$agent = str_replace ( "81.0.4044.138", $chrome_version, $agent );

return $agent;
}

```

Файл “*Hosting_Download*”

```

// saving image
if (! isset ( $idental )) {
    if (wp_mkdir_p ( $upload_dir ['path'] ))
        $file = $upload_dir ['path'] . '/' . $filename;
    else
        $file = $upload_dir ['basedir'] . '/' . $filename;

    $f = file_put_contents ( $file, $image_data );
}

```

Файл “*Delete_hyperlink*”

```

// function for stripping inline urls
function strip_urls($content) {
    return preg_replace ( '{http[s]?://[^\s]*}', '', $content );
}

```

Файл “*add_image_in_article*”

```

// add featured image
if (in_array ( 'OPT_REPLACE', $camp_opt )) {
    foreach ( $keywords as $keyword ) {

        $keyword = trim ( $keyword );

        if (trim ( $keyword != '' )) {
            $post_content = str_replace ( $keyword, '<a href="' .
$camp->camp_replace_link . '">' . $keyword . '</a>', $post_content );
        }
    }
}

```

Файл “*Translate*”

```

//PolyLang
if(in_array('OPT_POLY',    $camp_opt)    &&    function_exists    (
'pll_set_post_language' )){

    $language_code = trim ( $camp_general ['cg_poly_lang'] );

    echo '<br>Setting language of Polylang plugin to ' .
$language_code;

    pll_set_post_language($id,$language_code);

    //connect older posts?
    $cleanSource = preg_replace ( '{.rand=.}', '', $source_link
);

    $sourceMD5 = md5 ( $cleanSource );
    $keyName = $sourceMD5 . '_poly';

    $Qresults = $this->db->get_results ( "SELECT FROM {$this->db-
>prefix}postmeta WHERE meta_key = '$keyName' limit 1", ARRAY_A );

    if (count ( $Qresults ) != 0) {

```

					ПГУ 1.090301.06.001 ПЗ	Лист
						46
Изм.	Лист	№ докум.	Подпись	Дата		

```

        // we have a previous instance
        $metaRow = $Qresults [0];

        $metaValue = $metaRow ['meta_value'];

        echo '<br>Found a previous instance of the post in a
different langauge:' . $metaValue . ' connecting...';

        PLL()->model->post->save_translations(          $id          ,
array($metaValue => $metaRow ['post_id']) );

    }else{

        //mark the post for next translation
        add_post_meta      (      $id,      $sourceMD5      .      '_poly',
$language_code );

    }
}

```

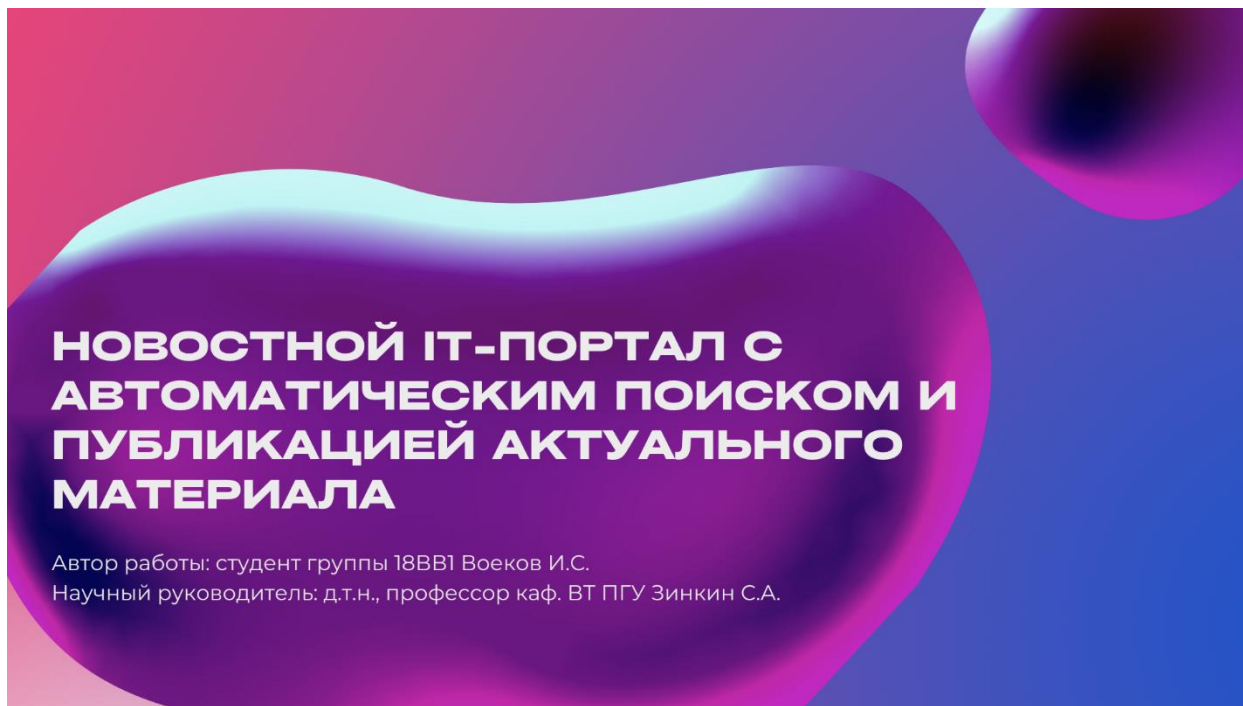
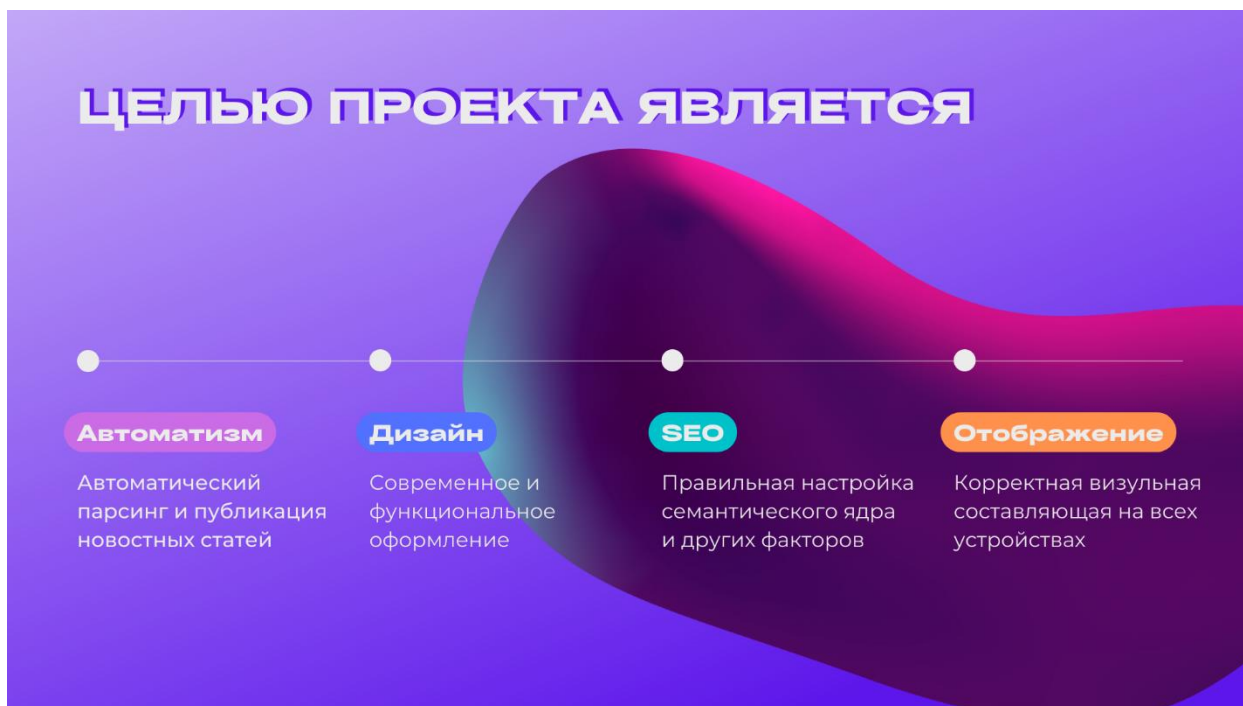
					ПГУ 1.090301.06.001 ПЗ	Лист
						47
Изм.	Лист	№ докум.	Подпись	Дата		

Приложение В

(Обязательное)

Новостной IT-портал с автоматическим поиском и публикацией актуального материала

Презентация



ОТСУТСТВИЕ НОВОСТНЫХ ИТ-ПОРТАЛОВ С
НАЛИЧИЕМ СОВРЕМЕННОГО
ФУНКЦИОНАЛЬНОГО ДИЗАЙНА

НЕУДОБНАЯ АДАПТАЦИЯ САЙТА НА
МОБИЛЬНЫХ УСТРОЙСТВАХ

ИТ-СООБЩЕСТВО РАЗВИВАЕТСЯ С
КАЖДЫМ ДНЁМ И ВСЁ БОЛЬШЕЕ
КОЛИЧЕСТВО ЛЮДЕЙ ХОТЯТ БЫТЬ В
КУРСЕ НОВЫХ СОБЫТИЙ В ЭТОЙ СФЕРЕ

АКТУАЛЬНОСТЬ ПРОЕКТА

АНАЛОГАМИ
ЯВЛЯЮТСЯ



IXBT.COM



3Dnews



Cnews

Изм.	Лист	№ докум.	Подпись	Дата

ПГУ 1.090301.06.001 ПЗ

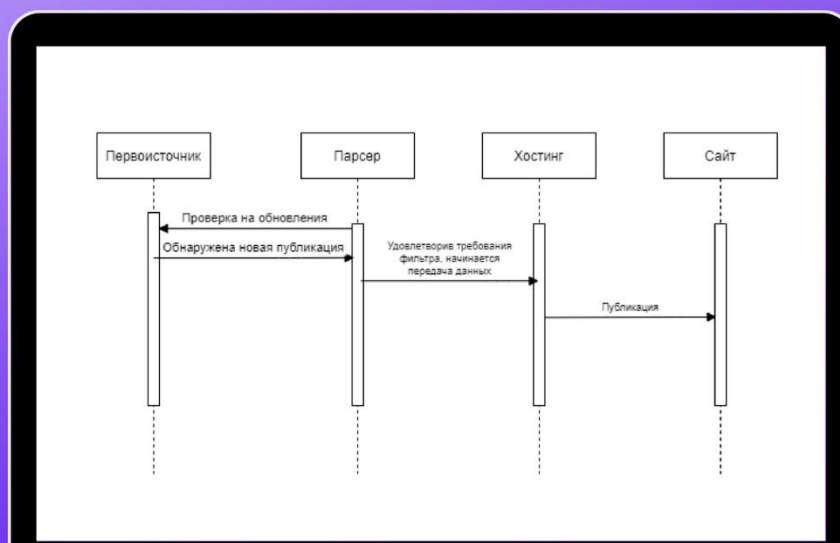
Лист

49

Стек технологий



ДИАГРАММА ПОСЛЕДОВАТЕЛЬНОСТИ



Изм.	Лист	№ докум.	Подпись	Дата

ПГУ 1.090301.06.001 ПЗ

Лист

50

ДИАГРАММА ПРЕЦЕДЕНТОВ

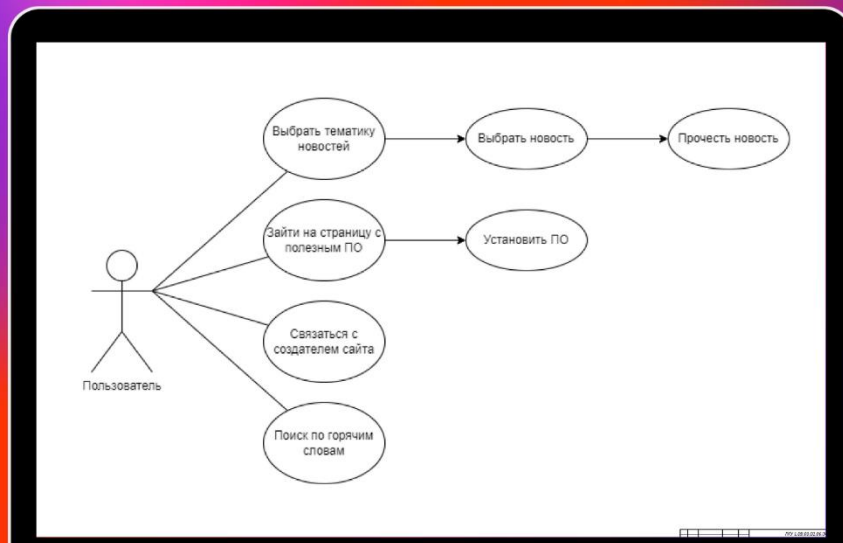
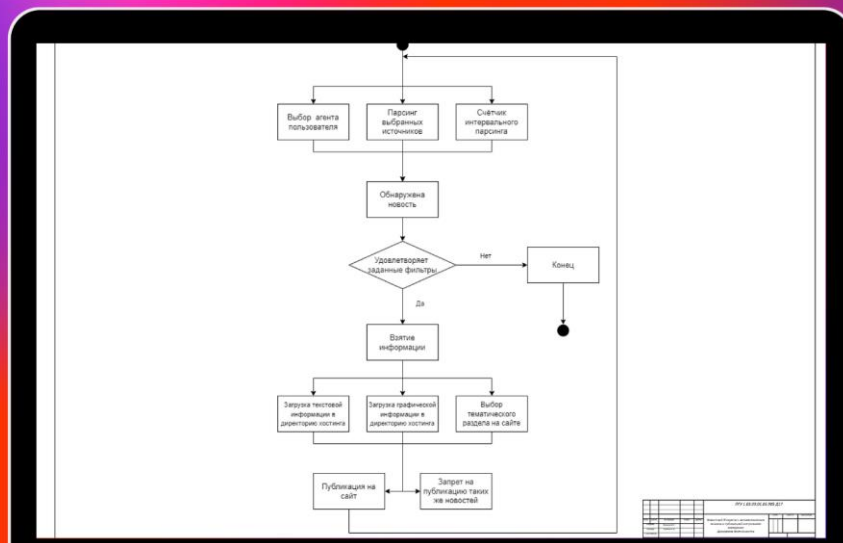
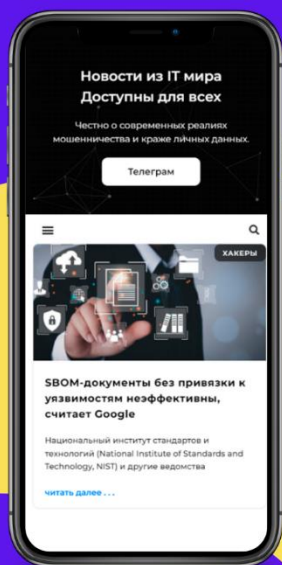
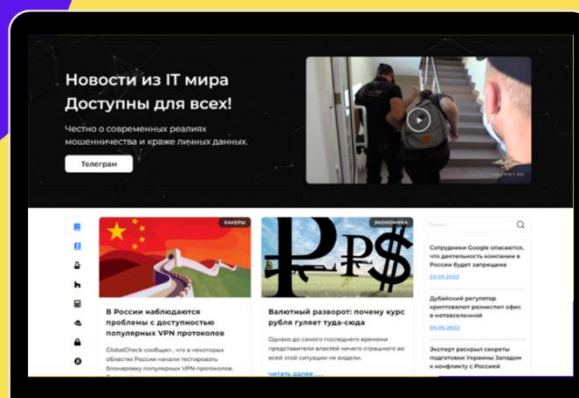


ДИАГРАММА ДЕЯТЕЛЬНОСТИ



Демонстрация



ЗАКЛЮЧЕНИЕ

В ХОДЕ РЕАЛИЗАЦИИ ПРОЕКТА БЫЛИ ГЛУБОКО ИЗУЧЕНЫ PHP, JAVASCRIPT, РАБОТА С ВНЕШНЕЙ И ВНУТРЕННЕЙ НАСТРОЙКОЙ ВЕБ-РЕСУРСОВ

В ХОДЕ РАЗРАБОТКИ ПРОЕКТА БЫЛИ РЕШЕНЫ АКТУАЛЬНЫЕ ПРОБЛЕМЫ СОВРЕМЕННЫХ НОВОСТНЫХ IT-ПОРТАЛОВ

БЫЛ ПРОВЕДЕН АНАЛИЗ ДЕЙСТВУЮЩИХ АНАЛОГОВ НОВОСТНЫХ ПОРТАЛОВ, А ТАКЖЕ ИХ ПЛЮСЫ И МИНУСЫ

БЫЛИ ПРОДУМАНЫ ДАЛЬНЕЙШИЕ ВЕКТОРЫ РАЗВИТИЯ ДАННОГО ПРОЕКТА

Выйти



Изм.	Лист	№ докум.	Подпись	Дата

ПГУ 1.090301.06.001 ПЗ

Лист

52