

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Кафедра теории функций и стохастического анализа

АВТОМАТИЗАЦИЯ ПРОЦЕССА ПОДГОТОВКИ SMM-ОТЧЕТНОСТИ
НА ПРИМЕРЕ СТАТИСТИКИ VK-СООБЩЕСТВА
АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 451 группы
направления 38.03.05 — Бизнес-информатика

механико-математического факультета
Привалова Ивана Сергеевича

Научный руководитель
доцент, к. ф.-м. н.

Д. В. Мельничук

Заведующий кафедрой
д. ф.-м. н., доцент

С. П. Сидоров

Саратов 2026

Выпускная квалификационная работа посвящена автоматизации процесса подготовки SMM-отчетности на основе статистики VK-сообщества. Тема относится к области бизнес-информатики, поскольку объединяет анализ прикладного бизнес-процесса, формализацию требований, моделирование данных, разработку информационной системы и оценку эффекта от внедрения цифрового инструмента. В работе рассматривается не абстрактная задача построения отчетов, а конкретный регулярный процесс: получение выгрузок статистики из VK, проверка их состава, нормализация таблиц, расчет показателей, подготовка графиков, формирование выводов и экспорт результата для дальнейшего использования руководителем или SMM-специалистом.

Актуальность темы связана с тем, что социальные сети остаются важным каналом коммуникации организаций с аудиторией. Для российских компаний, образовательных проектов, локальных медиа и небольших сообществ VK часто выполняет сразу несколько функций: новостной площадки, витрины, канала обратной связи и источника данных о реакции пользователей. При этом регулярная оценка эффективности SMM-деятельности требует не только просмотра встроенной статистики, но и подготовки сопоставимого отчета за период. Такой отчет должен показывать динамику охвата, вовлеченность, рост аудитории, эффективность форматов контента, активность по дням и часам, а также ключевые выводы, пригодные для корректировки контент-плана.

На практике подготовка подобных отчетов часто выполняется вручную. Специалист выгружает файлы из кабинета VK, открывает их в электронных таблицах, проверяет кодировку и структуру, переносит значения в рабочий шаблон, обновляет формулы и диаграммы, после чего собирает итоговый файл для передачи руководителю. При небольшом количестве данных этот порядок может казаться приемлемым, однако при регулярной работе он становится источником ошибок. Риски возникают из-за неодинаковых названий колонок, пустых строк, разных типов выгрузок, ручного копирования формул и отсутствия единой истории обработки. В результате один и тот же показатель может быть рассчитан по-разному в разных отчетных периодах, а восстановить ход подготовки отчета через несколько месяцев становится затруднительно.

Цель работы состоит в разработке и обосновании прототипа информационной системы, автоматизирующей подготовку SMM-отчетности на основе выгрузок статистики VK-сообщества. Для достижения цели решены задачи анализа предметной области, описания исходного процесса AS-IS, определения целевого процесса TO-BE, формирования требований, проектирования архитектуры, разработки алгоритмов обработки данных, реализации прототипа, проверки пользовательского сценария и оценки экономического эффекта. Объект исследования — процесс подготовки SMM-отчетности по статистике VK-сообщества. Предмет исследования — методы и программные средства автоматизации импорта, нормализации, расчета KPI, визуализации и экспорта аналитического отчета.

Методическая и проектная основа

Методическая основа работы включает анализ бизнес-процессов, требования к информационным системам, принципы построения аналитических отчетов и подходы к оценке качества программного продукта. Процесс подготовки отчетности рассматривается как последовательность операций, в ходе которой первичные данные превращаются в управленческую информацию. В этом смысле SMM-отчетность представляет собой не только набор диаграмм, но и регламентированный информационный поток: от исходных файлов до решения о корректировке публикаций, форматов и времени размещения материалов.

В работе используется логика разделения процесса на состояние AS-IS и состояние TO-BE. В состоянии AS-IS отчет подготавливается вручную: специалист выбирает период, выгружает несколько файлов, переносит данные в таблицы и вручную формирует итоговые графики. В состоянии TO-BE повторяющиеся операции переносятся в программный контур. Пользователь создает проект, загружает выгрузки VK, запускает обработку, получает диагностику файлов, расчетный набор данных, KPI, визуализации, рекомендации и файлы экспорта. Управленческая интерпретация при этом остается за специалистом: система не заменяет SMM-эксперта, а снижает трудоемкость и повышает воспроизводимость расчета.

Ключевым показателем в работе выступает вовлеченность публикации.

Она определяется как отношение суммы реакций к охвату публикации:

$$ER = \frac{Likes + Comments + Shares + Saves}{Reach} \cdot 100.$$

Такая формула используется потому, что абсолютное число реакций плохо сопоставимо между публикациями с разным охватом. Например, публикация со 100 реакциями при охвате 1000 пользователей имеет более высокий относительный отклик, чем публикация со 120 реакциями при охвате 10000 пользователей. Поэтому ER отражает не только количество действий, но и качество контакта аудитории с конкретной публикацией. Наряду с ER в работе учитываются охват, просмотры, подписки, отписки, динамика аудитории, распределение по форматам и временные слоты публикаций.

Для оценки качества данных выделены пять критериев: полнота, согласованность, точность, актуальность и прослеживаемость. Полнота показывает, загружены ли необходимые выгрузки за период. Согласованность отражает отсутствие противоречий между типами файлов, датами и значениями метрик. Точность связана с корректным чтением чисел, дат и строк. Актуальность означает соответствие выгрузок выбранному периоду. Прослеживаемость дает возможность связать итоговый показатель с конкретным исходным файлом и этапом обработки. Эти критерии важны для доверия к отчету: автоматизация должна не только ускорять расчет, но и показывать ограничения данных, на которых построены выводы.

Краткое содержание первой главы

Первая глава посвящена анализу предметной области и постановке задачи автоматизации. В начале главы раскрывается роль SMM-отчетности в управлении коммуникациями. Показано, что регулярный отчет отличается от разовой аналитической справки: он должен сохранять сопоставимость между периодами, использовать устойчивые правила расчета и предоставлять результаты в понятной форме. Для SMM-специалиста важна детализация, для руководителя — краткий итог, для аналитика — качество и проверяемость данных. Поэтому информационная система должна поддерживать разные уровни представления: управленческий, аналитический и технический.

Далее в главе описан состав типовой отчетности по VK-сообществу.

Встроенный кабинет статистики VK уже содержит визуальные представления данных, однако регулярная отчетность требует объединения нескольких выгрузок и подготовки единого документа. В типовой отчет включаются показатели сообщества, аудитории, публикаций и контента. По сообществу важны охват, просмотры, посещаемость и динамика подписчиков. По аудитории учитываются демографические и географические признаки. По публикациям рассчитываются реакции, комментарии, репосты, сохранения и ER. По контенту анализируются форматы, темы, время публикации и отклик пользователей.

Отдельный раздел первой главы посвящен описанию процесса AS-IS. В исходном процессе специалист входит в кабинет статистики VK, выбирает период, выгружает несколько файлов и далее работает с ними в электронных таблицах. Часть визуализаций доступна уже в кабинете VK, однако итоговый отчет обычно требует объединения данных из разных разделов, приведения таблиц к единому виду и подготовки файлов, пригодных для передачи руководителю. Основные проблемы AS-IS-процесса связаны не только с длительностью, но и с управляемостью: формулы могут изменяться от отчета к отчету, история преобразований не сохраняется, а качество данных проверяется вручную.

В первой главе также сформулирован целевой процесс TO-BE. В нем пользователь работает не с набором разрозненных файлов, а с проектом и запусками анализа. Система сохраняет исходные данные, определяет типы файлов, выполняет нормализацию, формирует расчетный payload, строит графики, создает рекомендации и предоставляет экспорт. Такое разделение стадий позволяет сделать процесс более прозрачным. Если в отчете отсутствует определенный график или показатель, причина должна быть объяснима: например, не загружена нужная выгрузка, отсутствует дата или не найден обязательный столбец.

Важным результатом первой главы является определение границ исследования. В работе рассматривается автоматизация подготовки отчета по выгрузкам статистики VK-сообщества, а не универсальная система управления социальными сетями. Прототип не публикует контент, не управляет рекламными кампаниями, не модерировать комментарии и не заменяет кабинет

VK. Его задача уже и конкретнее: преобразовать экспортированные статистические данные в структурированный аналитический отчет с KPI, графиками, рекомендациями и файлами экспорта. Такое ограничение делает проект достижимым и позволяет сосредоточиться на наиболее трудоемком участке регулярной отчетности.

Первая глава показывает, что выбранная задача имеет прикладной характер и подходит для направления бизнес-информатики. Она требует не только программирования, но и анализа процесса, выделения заинтересованных сторон, описания требований и оценки организационного эффекта. В результате глава подводит к выводу, что автоматизация должна охватывать полный путь от исходной выгрузки до готового управленческого отчета, включая диагностику данных, расчет ER, построение графиков и подготовку итоговых файлов.

Ключевой вывод первой главы состоит в том, что ручной процесс подготовки SMM-отчетности плохо масштабируется и зависит от конкретного исполнителя. При смене специалиста организация теряет часть неформализованных знаний: какие файлы нужны, какие столбцы учитывать, какие строки исключать, как интерпретировать пустые значения и какие формулы применять. Программный прототип переводит эти правила в воспроизводимый контур. Это повышает сопоставимость отчетов между периодами и снижает риск ошибок, связанных с ручной обработкой.

Таким образом, первая глава выполняет постановочную функцию. В ней определены предметная область, состав показателей, проблемы исходного процесса, требования к целевому процессу и границы разработки. Эти результаты используются во второй главе для проектирования информационной системы: определения функциональных и нефункциональных требований, выбора архитектуры, описания модели данных и разработки алгоритмов обработки выгрузок.

Краткое содержание второй главы

Вторая глава посвящена проектированию системы автоматизации SMM-отчетности. В начале главы сформулированы функциональные требования. Система должна создавать проект, загружать несколько файлов ста-

тики, распознавать типы выгрузок, выполнять диагностику, нормализовать строки, рассчитывать KPI, строить графики, формировать текстовые выводы, сохранять историю запусков и экспортировать результат. Эти требования связаны с целевым процессом ГО-ВЕ и отражают полный пользовательский сценарий от загрузки файлов до получения отчета.

Нефункциональные требования описывают качество выполнения этих функций. Для прототипа особенно важны надежность, удобство использования, сопровождаемость, переносимость и прозрачность обработки. Надежность означает, что система должна корректно реагировать на неполные или неоднородные загрузки. Удобство использования связано с тем, что пользователь должен понимать состояние запуска: какие файлы загружены, какие предупреждения выявлены и доступен ли отчет. Сопровождаемость обеспечивается разделением интерфейса, аналитической логики и локального хранения. Переносимость достигается за счет local-first архитектуры, при которой приложение работает в браузере без обязательной серверной базы данных.

Архитектурно прототип реализован как local-first веб-приложение в папке `platform-local`. Актуальная версия не использует backend-хранилище: файлы читаются в браузере, проекты и запуски анализа сохраняются в IndexedDB, а экспорт формируется на стороне пользователя. Для интерфейса выбран Next.js App Router, React и TypeScript. Для чтения файлов используются Web File API и SheetJS. Такой стек позволяет собрать рабочий прототип, который можно запускать локально и разворачивать без настройки отдельной серверной инфраструктуры.

В модели данных выделены сущности проекта, запуска анализа, загруженной таблицы, нормализованных строк, фильтров, расчетного payload и журнала событий. Проект соответствует одному VK-сообществу или аналитическому направлению. Запуск анализа фиксирует конкретный период или набор файлов. Загруженные таблицы сохраняют исходную информацию о файле и его типе. Нормализованные строки служат промежуточным слоем между сырой таблицей и расчетными метриками. Payload содержит итоговые KPI, графики, рекомендации и данные для экспорта. Журнал событий фиксирует этапы обработки и предупреждения.

Вторая глава подробно описывает алгоритм импорта и нормализации. Обработка начинается с чтения файла в памяти приложения. Для Excel используется библиотека SheetJS, для CSV выполняется определение структуры строк. После чтения формируется унифицированное представление, в котором каждая строка сохраняет исходные названия колонок. Далее выполняется классификация файла: система определяет, относится ли он к статистике сообщества, аудитории, публикаций или контента. Классификация строится по набору признаков: названиям колонок, наличию дат, типичным метрикам и структуре строк.

Нормализация необходима, потому что выгрузки VK могут отличаться по составу колонок и формату данных. В названиях встречаются пробелы, скобки, русские слова, разные варианты обозначения дат и числовых показателей. Алгоритм приводит эти значения к каноническим полям: дата, охват, просмотры, реакции, комментарии, репосты, сохранения, подписки, отписки, формат контента и другие признаки. Если обязательное поле отсутствует, система не должна скрывать проблему: предупреждение фиксируется в диагностике и журнале событий.

Отдельно во второй главе рассматривается расчетный слой. На его вход поступают нормализованные строки, а на выходе формируются KPI, агрегированные наборы данных для графиков, текстовые выводы и рекомендации. Расчетный слой отделен от интерфейса, поэтому его можно тестировать независимо от страниц Next.js. Это важно для сопровождаемости: изменения в визуальном представлении не должны ломать формулы и правила обработки. При дальнейшем развитии проекта аналитические функции можно перенести в backend или использовать в другом интерфейсе.

Во второй главе также раскрыт вопрос качества программного продукта. Для оценки применима модель ISO/IEC 25010:2023. В контексте дипломного прототипа наиболее значимы функциональная пригодность, надежность, удобство использования, сопровождаемость и переносимость. Производительность учитывается, но ключевой риск связан не с обработкой миллионов строк, а с неоднородностью реальных выгрузок. Поэтому система должна быть устойчива к неполным данным, разным типам файлов и частичному отсутствию показателей.

Итог второй главы — проектное описание системы, достаточное для реализации прототипа. В ней определены требования, архитектура, модель данных, алгоритмы импорта, нормализации, расчета, визуализации и экспорта. Эти решения создают основу для третьей главы, в которой описывается фактическая реализация приложения, пользовательский сценарий, интерфейс, проверка работы и оценка эффекта автоматизации.

Краткое содержание третьей главы

Третья глава посвящена реализации и проверке прототипа. В ней описана актуальная local-first версия приложения, расположенная в папке `platform-local`. Пользовательский сценарий начинается с открытия каталога проектов. Пользователь может создать проект, выбрать существующий проект, запустить демо-анализ или загрузить собственные выгрузки. Внутри проекта хранится история запусков, каждый из которых имеет статус, период, набор файлов и результат обработки. Такая структура заменяет разрозненные папки с Excel-файлами на понятную историю аналитических запусков.

Загрузка файлов реализована на стороне браузера. Пользователь выбирает несколько выгрузок, приложение получает объекты `File`, читает содержимое, определяет формат и передает строки в модуль импорта. После обработки нормализованные таблицы сохраняются в записи запуска анализа в `IndexedDB`. Диагностика проверяет расширение файла, наличие строк, обнаружение даты, тип выгрузки и конфликты колонок. В интерфейсе выводится подробный результат диагностики вместо формального сообщения об успехе. Это помогает пользователю понять, какие данные приняты системой, а какие ограничения могут повлиять на отчет.

Обработка запуска включает несколько стадий: классификацию файлов, нормализацию строк, расчет метрик, построение аналитического payload и сохранение результата. Состояние каждой стадии фиксируется в журнале событий. Даже для локального прототипа журнал важен, потому что обработка файлов является отдельным процессом с собственными состояниями. Пользователь должен видеть, что файлы загружены, классификация выполнена, нормализация завершена, расчет готов, а экспорт доступен. Журнал

также полезен при отладке и демонстрации прототипа, поскольку подтверждает прослеживаемость процесса.

В аналитическом отчете представлены KPI, графики, рекомендации, таблицы и блок экспорта. В верхней части отчета выводятся основные показатели: охват, просмотры, ER, число публикаций, динамика подписчиков и другие агрегаты. Далее расположены графики динамики сообщества, эффективности форматов, ER публикаций, временных слотов и географии аудитории. Каждая визуализация связана с конкретным управленческим вопросом: выявить спад посещаемости, определить сильные форматы, найти удачное время публикации или оценить структуру аудитории.

Рекомендательный слой прототипа формируется детерминированно на основе расчетных метрик. Система не пытается заменить SMM-стратегию, но выделяет закономерности, которые помогают специалисту быстрее перейти от таблиц к действиям. Например, прототип может отметить лучший пост по ER, сильный формат контента, временные интервалы с повышенной вовлеченностью, возможные проблемы с охватом или отрицательную динамику подписчиков. Рекомендации остаются объяснимыми: они строятся на тех же расчетных данных, что и графики, поэтому пользователь может сопоставить вывод с исходными показателями.

Экспорт завершает пользовательский сценарий. Прототип поддерживает CSV для выгрузки таблицы публикаций, XLSX для табличного отчета, PDF через печатную версию страницы и JSON-backup для резервного копирования локальных данных. Экспорт строится на том же расчетном payload, что и экран отчета. Это снижает риск расхождения между интерфейсом и итоговым файлом: если показатель рассчитан один раз в аналитическом ядре, он одинаково используется на странице и в экспортируемом документе. Для регулярной отчетности это принципиально, поскольку отчет должен быть не только красивым, но и воспроизводимым.

Проверка прототипа проводилась на тестовом наборе выгрузок VK. Проверялись модульные и интеграционные сценарии: распознавание файлов, нормализация колонок, обнаружение даты, расчет KPI, запуск демо-режима, сохранение данных в IndexedDB, применение фильтров и получение отчета. Отдельно оценивалось соответствие интерфейса рабочему процессу. Если по-

сле загрузки файлов следующий шаг неочевиден, автоматизация становится хрупкой. Поэтому кнопка обработки, статусы и диагностические сообщения размещены рядом с контекстом запуска.

В третьей главе приведена оценка экономического эффекта. В исходном ручном процессе подготовка одного типового отчета занимает около 360 минут. После автоматизации основная часть времени переносится с ручной обработки таблиц на интерпретацию результатов и проверку итогового отчета. Оценочная трудоемкость автоматизированного процесса составляет около 70 минут. Экономия времени сопровождается качественным эффектом: снижается риск ошибок формул, появляется история запусков, повышается прозрачность обработки, а отчеты становятся сопоставимыми между периодами.

Основные результаты работы

В результате выполнения выпускной квалификационной работы разработан прототип информационной системы VK Analytics, предназначенный для автоматизации подготовки SMM-отчетности по выгрузкам статистики VK-сообщества. Прототип поддерживает создание проекта, загрузку файлов, диагностику, нормализацию данных, расчет KPI, построение графиков, формирование рекомендаций и экспорт отчета. В системе реализовано разделение исходных, нормализованных и расчетных данных, что повышает прозрачность и воспроизводимость обработки.

Первым результатом является формализация предметной области. В работе определен состав данных, необходимых для регулярного отчета: статистика сообщества, аудитории, публикаций и контента. Описаны ключевые показатели, включая охват, просмотры, ER, реакции, динамику подписчиков, эффективность форматов и временные слоты публикаций. Показано, что регулярный отчет должен быть сопоставимым между периодами и пригодным для управленческого обсуждения.

Вторым результатом является описание процесса AS-IS и целевого процесса TO-BE. Исходный процесс основан на ручной работе с выгрузками и электронными таблицами. Целевой процесс переносит повторяющиеся операции в программный контур: диагностику, нормализацию, расчет, построение визуализаций и экспорт. Это позволяет снизить зависимость от конкретного

исполнителя и сделать подготовку отчета более устойчивой.

Третьим результатом является проектирование архитектуры и модели данных. Актуальная версия реализована как local-first веб-приложение, работающее в браузере. Такое решение соответствует задачам дипломного прототипа: оно демонстрирует полный цикл подготовки отчета без настройки внешней инфраструктуры. При этом архитектура сохраняет возможность дальнейшего развития: подключения VK API, внешнего backend/storage, многопользовательского режима и планировщика регулярных отчетов.

Четвертым результатом является реализация пользовательского сценария и проверка прототипа. Пользователь может запустить демо-анализ, загрузить собственные выгрузки, увидеть диагностику, получить расчетный отчет, применить фильтры и экспортировать результат. Скриншоты ключевых разделов подтверждают наличие реализованного сценария: каталог проектов, экран загрузки, блок KPI и рекомендаций, графики, журнал событий и экспорт.