

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего образования
«Тольяттинский государственный университет»

Кафедра «Прикладная математика и информатика»
(наименование)

09.04.03 Прикладная информатика

(код и наименование направления подготовки / специальности)

Управление корпоративными информационными процессами

(направленность (профиль) / специализация)

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА (МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ)

на тему «Методы и модели управления эффективностью системы
аналитической обработки корпоративной информации»

Обучающийся

В.Г. Лымарь

(Инициалы Фамилия)

(личная подпись)

Научный руководитель

д.т.н., профессор, С.В. Мкртычев

(ученая степень (при наличии), ученое звание (при наличии), Инициалы Фамилия)

Тольятти 2024

Оглавление

Введение.....	3
Глава 1 Анализ состояния исследований в области управления эффективностью системы аналитической обработки корпоративной информации	7
1.1 Методы сбора информации.....	7
1.2 Особенности аналитической обработки корпоративной информации	11
1.3 Витрины данных	21
Глава 2 Анализ методов и моделей управления эффективностью системы аналитической обработки корпоративной информации	25
2.1 Метод управления эффективностью системы аналитической обработки корпоративной информации	25
2.2 Модель управления эффективностью системы аналитической обработки корпоративной информации	37
Глава 3 Разработка приложения с использованием предложенных методов и моделей управления эффективностью системы аналитической обработки корпоративной информации.	44
3.1 Разработка приложения с витринами данных.....	44
3.2 Оценка эффективности предложенных решения на принятие решений.....	53
Заключение	64
Список используемой литературы и используемых источников.....	66

Введение

Актуальность работы. Цифровая экономика диктует предприятиям и организациям новые подходы к обработке информации. Объемы данных растут, развиваются инструменты анализа, внедряются нейронные сети, от менеджеров требуется принимать решения в краткие сроки, используя аналитически обработанными данными. Улучшение методов и инструментов анализа большого количества данных для принятия оперативных и верных решений является актуальной задачей на стыке следующих предметов - управления, экономики и информационных систем.

«Управление эффективностью предприятия (ЕРМ, или Enterprise Performance Management) – это современный подход к бизнес-аналитике, который помогает организациям принимать стратегические решения, соответствующие их корпоративным целям и задачам» [2].

Преимущества внедрения ЕРМ систем на предприятия повышение точности прогнозов благодаря использованию современных аналитических инструментов и методов моделирования. Уменьшение времени на сбор и обработку данных, что позволяет сосредоточиться на их анализе и принятии решений.

В ходе магистерской работы исследуется математическое, информационно и техническое обеспечение, позволяющего проанализировать методы, используемые в компании для создания системы эффективного управления, определить точки роста и проблемные зоны, предложить практическое решение поставленных задач и внедрить новые инструменты в работу компании.

Методы и модели управления эффективностью системы аналитической обработки корпоративной информации расширят возможности мониторинга и анализа данных в процессе управления в компании, создадут улучшенную обратную связь всех подразделений компании и позволят в режиме реального времени идентифицировать и устранить проблему до ее развития.

Одним из перспективных методов является исследование и совершенствование витрин данных как важного элемента корпоративных аналитических систем.

Объект исследования – системы аналитической обработки корпоративной информации.

Предмет исследования – методы и модели управления эффективностью системы аналитической обработки корпоративной информации.

Цель исследования – исследование методов и разработка модели управления эффективностью системы аналитической обработки корпоративной информации.

Задачи исследования:

- «анализ современного состояния проблем управления эффективностью системы аналитической обработки корпоративной информации;
- проанализировать методы и разработать модели, обеспечивающие улучшение процессов управления эффективностью системы аналитической обработки корпоративной информации;
- «выполнить апробацию предлагаемых решений и оценить их эффективность» [34].

Анализ последних исследований и публикаций.

В научной литературе очень хорошо освещены вопросы Business Intelligence (BI) и управление эффективностью предприятия (EPM, или Enterprise Performance Management). В научных работах Козлова М. Г., Воронцов К. В., Борисова И. А. рассматриваются использование программных продуктов для реализации управления эффективностью бизнеса. Рассмотрим краткое содержание данных работ.

Вопросами интеллектуального анализа данных занимаются следующие авторы: Борисова И. А., Дюбанов В. В., Загоруйко Н. Г. и Кутненко О. А., Kaplan R.S., Norton D.P., Нивен П.Р., Подиновский В.В., Ногин В.Д.

Во многих работах как российских, так и зарубежных авторов приведено описание использования витрин данных.

Методы исследования. В процессе исследования темы магистерской работы применяются следующие подходы и методы:

- математический и статистический анализ данных;
- интеллектуализации обработки данных и поиска зависимости в больших данных;
- алгоритмы, применяемые при визуализации данных.

Гипотеза исследования.

Заключается в том, что для повышения эффективности системы аналитической обработки корпоративной информации необходимо использовать методы и модели управления, которые позволят проводить обзор корпоративной информации в режиме реального времени с помощью витрин данных.

Новизна исследования.

Разработана модель управления эффективностью системы аналитической обработки корпоративной информации для совершенствования витрин данных как важного элемента корпоративных аналитических систем.

Практическая значимость исследования заключается в применении витрин данных для работы с корпоративными данными.

Теоретической основой диссертационного исследования являются научные труды российских и зарубежных ученых, занимающихся проблемами моделирования и аналитики корпоративных данных и управления эффективностью организации.

На защиту выносятся:

- модели и методы управления эффективностью системы аналитической обработки корпоративной информации;
- результаты апробации и оценки эффективности предлагаемых проектных решений.

«Диссертация состоит из введения, трех глав, заключения, списка используемой литературы и используемых источников.

Во введении обоснована актуальность темы исследования, представлены объект, предмет, цели, задачи и положения, выносимые на защиту диссертации.

В первой главе дан анализ состояния исследований в области управления эффективностью системы аналитической обработки корпоративной информации.

Во второй главе дан анализ методов и моделей управления эффективностью системы аналитической обработки корпоративной информации.

Третья глава посвящена разработке приложения с использованием предложенных методов и моделей управления эффективностью системы аналитической обработки корпоративной информации, выполнена апробация предлагаемых проектных решений и оценка их эффективности» [34].

В заключении приводятся результаты исследования.

Работа изложена на 71 странице и включает 38 рисунков, 7 таблиц и 50 источников.

Глава 1 Анализ состояния исследований в области управления эффективностью системы аналитической обработки корпоративной информации

1.1 Методы сбора информации

Основным видом деятельности организации, на базе которой проводились исследования по теме магистерской диссертации является 62.02 - Деятельность консультативная и работы в области компьютерных технологий.

Дополнительные виды деятельности по ОКВЭД:

62.0 Разработка компьютерного программного обеспечения, консультационные услуги в данной области и другие сопутствующие услуги

62.09 Деятельность, связанная с использованием вычислительной техники и информационных технологий, прочая

63.1 Деятельность по обработке данных, предоставление услуг по размещению информации, деятельность порталов в информационно-коммуникационной сети Интернет

63.11.1 Деятельность по созданию и использованию баз данных и информационных ресурсов

95.1 Ремонт компьютеров и коммуникационного оборудования

Структура организации [8] – это упорядоченная совокупность взаимосвязанных элементов, обеспечивающая их функционирование и развитие как единого целого. Организационная структура компании, в которой проходила практика имеет линейно-функциональную форму (рисунок 1).

В организации выделяются следующие подсистемы: директор, заместители директора, начальники отделов, отделы.

Отделы подчиняются заместителям директора, по соответствующим направлениям деятельности, которые в свою очередь подчиняются стоящему во главе директору.

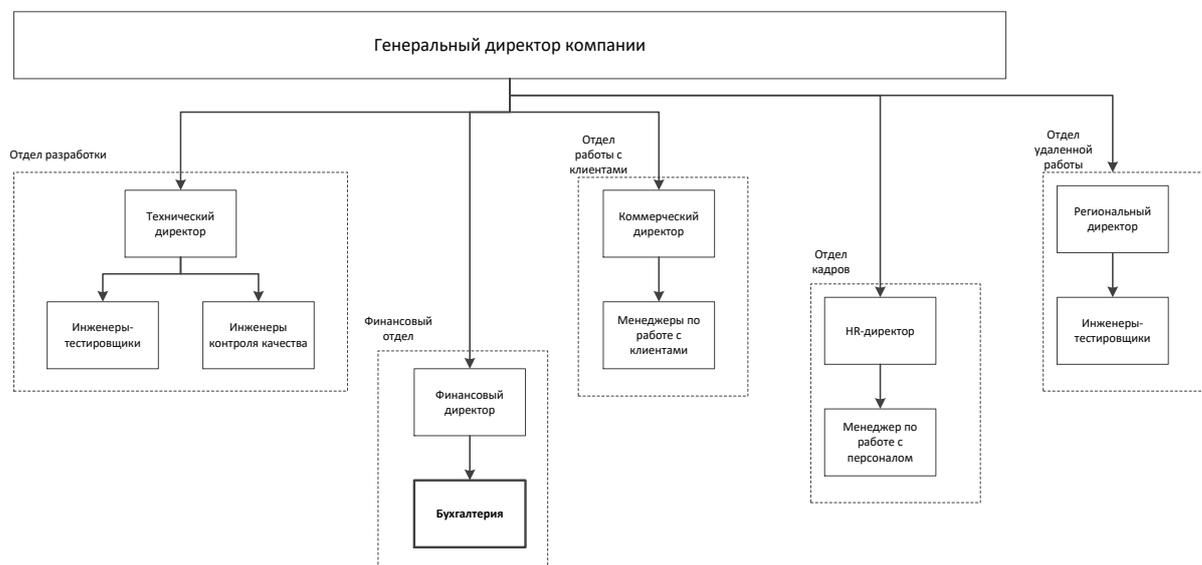


Рисунок 1 – Обобщенная организационная структура компании, в которой проходила учебная практика

Руководителем организации является Генеральный директор, выполняющий общее руководство и определяющий политику ее развития. В подчинении Генерального директора находятся следующие структурные подразделения:

- финансовый отдел;
- отдел кадров;
- отдел продаж;
- отдел разработки;
- отдел функционального и нагрузочного тестирования.

Организация напрямую работает с разработчиками программного обеспечения, вендорами, производителями аппаратных средств защиты и устройств и т.д., далее поставляя данный товар (программное обеспечение и

аппаратные средства) по подразделениям и заказчикам, а также осуществляя его тестирование.

Руководство финансовым отделом осуществляет финансовый директор, в подчинении которого находится бухгалтерия - подразделение, выполняющее ведение бухгалтерского, налогового и управленческого учета.

В состав отдела кадров входит HR-директор и менеджеры по работе с персоналом.

Отдел продаж выполняет функции по реализации, работе с клиентами, поиском новых клиентов. В состав отдела входят руководитель отдела и менеджеры по продажам, общее руководство осуществляет Коммерческий директор.

Отдел функционального и нагрузочного тестирования - туда входят технический директор и команда тестировщиков и инженеров по качеству.

Основной акцент в изучении методов сбора информации сделан на исследовании процесса сбора и обработки корпоративной информации. (рисунок 2).



Рисунок 2 – Сбор и анализ корпоративной информации

Процесс анализа данных включает четыре основных этапа.

Первый этап – сбор первичной информации.

Определяется цель анализа и источники данных: отделы компании, клиенты, сотрудники. Может потребоваться проведение опросов и сбор обратной связи за определённый период.

Второй этап – обработка данных.

На этом этапе выполняются обезличивание, обобщение и удаление некорректных или пустых данных. Также исключаются конфиденциальные данные без разрешения на их обработку.

«Третий этап – предоставление данных.

Используются витрины данных, позволяющие с помощью фильтров формировать различные представления корпоративной информации» [5].

«Четвертый этап – оценка эффективности.

Проводится анализ принятых решений на основе обработанных данных.

Реальный процесс работы с корпоративными данными и аналитический анализ с различными вариантами отображения будет интерактивным» [6] (рисунок 3).

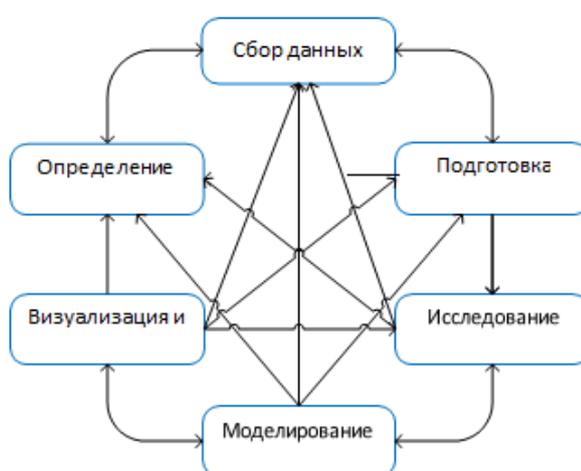


Рисунок 3 – Итеративный процесс управления эффективностью системы аналитической обработки корпоративной информации

В рамках магистерской работы, после детального изучения этапов функционирования системы аналитической обработки корпоративной информации и процессов сбора данных в организации, целесообразно сосредоточиться на анализе особенностей обработки корпоративной информации. Этот этап предполагает рассмотрение ключевых аспектов аналитической обработки, таких как использование методов и технологий бизнес-аналитики, инструментарий для визуализации данных, а также принципы обеспечения достоверности и конфиденциальности данных.

Кроме того, важным шагом является изучение существующих научных публикаций и исследований, посвящённых данной теме. Изучая литературу возможно оценить современные подходы к аналитической обработке данных, выявить тренды и перспективные направления в области корпоративной аналитики.

Сравнить различные методологии и инструменты, применяемые в системах управления данными, определить ключевые проблемы, с которыми сталкиваются организации при обработке данных, такие как неполнота данных, необходимость их предварительной очистки, а также вопросы, связанные с интеграцией данных из разных источников.

Рассмотрение научной литературы и исследований позволит сформировать теоретическую базу для дальнейшего анализа и предложений по оптимизации аналитической обработки корпоративной информации в организации. Таким образом, данный раздел станет важным связующим звеном между теоретическими аспектами темы и практическими рекомендациями, сформулированными в заключительной части работы.

1.2 Особенности аналитической обработки корпоративной информации

«Управление Эффективностью Предприятия (EPM, или Enterprise Performance Management) – следующее поколение бизнес-аналитики,

позволяющее компаниям принимать решения в системе бизнес-анализа и стратегически увязывать эти решения со своими корпоративными целями. ЕРМ включает в себя консолидированные данные из различных источников, запрос информации и анализ данных, а также использование результатов анализа на практике. Система расширяет процессы, создавая улучшенные цепи обратной связи. Непрерывные обзоры в режиме реального времени помогают идентифицировать и устранить проблему до ее развития» [10].

Система ЕРМ, основанная на платформе ВІ, объединяет возможности бизнес-аналитики с процессами планирования и контроля на предприятии. Система ЕРМ охватывает процессы планирования, стратегического анализа и моделирования сценариев «что если». Система ЕРМ позволяет управлять операциями компании в глобальном масштабе, консолидируя данные и стратегически используя их для достижения целей и реализации философии развития бизнеса.

Компании, использующие системы управления эффективностью предприятия (ЕРМ), получают комплексное представление о своей деятельности. Такие системы объединяют данные из множества источников, включая производственные операции, продажи, финансы и клиентскую информацию, обеспечивая мониторинг на уровне всей организации.

ЕРМ активно использует ключевые показатели эффективности (КРІ), которые помогают количественно оценивать достижение стратегических целей. КРІ, например, могут включать процент выручки от постоянных клиентов или стоимость процесса продаж. Современные ЕРМ-решения позволяют анализировать совокупности КРІ для оценки текущего состояния бизнеса и прогнозирования будущих тенденций, обеспечивая комплексный взгляд на компанию.

Одним из ключевых инструментов в рамках ЕРМ является Система сбалансированных показателей (Balanced Scorecard), которая связывает стратегическое развитие компании с тактическими действиями. Эта система

упрощает процесс стратегического планирования, помогая менеджерам преобразовывать цели в конкретные действия с измеримыми показателями.

ERP-системы способствуют повышению корпоративной эффективности за счёт интеграции данных, стратегического планирования и анализа. Они позволяют компаниям быстрее адаптироваться к изменениям на рынке, преобразуя большие объёмы данных в полезную информацию для принятия своевременных решений. Такой подход помогает организациям оставаться конкурентоспособными и реализовывать выигрышные бизнес-стратегии.

В научной литературе очень хорошо освещены вопросы Business Intelligence (BI) и управление эффективностью предприятия (EPM, или Enterprise Performance Management). В научных работах Козлова М. Г., Воронцов К. В., Колосков А. О., Богданов А. Б., Борисова И. А. рассматриваются использование программных продуктов для реализации управления эффективностью бизнеса. Рассмотрим краткое содержание данных работ.

«За последние годы многие поставщики средств Business Intelligence (BI) начали позиционировать свои продукты как решения для управления эффективностью бизнеса (Business Performance Management, BPM). Это связано с эволюцией BI-решений и консолидацией рынков прикладного ПО, а также с внедрением новых подходов к автоматизации управления, предложенных исследовательскими организациями» [12].

Современное управление бизнесом требует оперативного принятия решений на основе анализа значительных объемов данных. В этом помогают информационные системы, такие как ERP, CRM, SCM и BPM, каждая из которых автоматизирует определенные бизнес-процессы: управление ресурсами, маркетингом, продажами и цепочками поставок. Внедрение таких систем активно растет, так как они позволяют оптимизировать рутинные операции и обеспечивать над ними контроль.

«В рамках комплексных информационных систем автоматизации транзакционные системы (ERP, CRM, SCM) занимают базовый уровень. Они собирают и обрабатывают данные, которые затем анализируются системами бизнес-аналитики (BI), OLAP и другими инструментами аналитической обработки. Эти аналитические системы находятся на следующем уровне, предоставляя компании углубленный анализ для стратегического управления» [17].

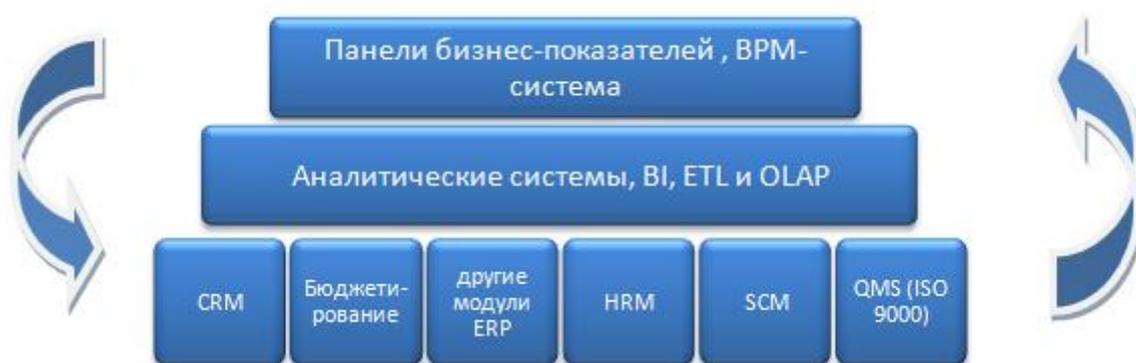


Рисунок 4 – Пирамида комплексной автоматизации и управления эффективностью бизнеса (BPM)

«Современные аналитические системы обладают инструментами для выборки, обработки и анализа данных на уровне предприятия. Они интегрируются с транзакционными системами, обеспечивая доступ к данным в необходимых разрезах. OLAP-системы, благодаря многомерному хранению данных и предрасчету агрегированных значений, позволяют выполнять анализ в режиме реального времени. Их цель — выявление закономерностей и формирование прогнозов для управления развитием компании» [19].

Ключевые требования к аналитическим системам:

- обеспечение сводной информации для управленческих решений;
- работа с внутренними и внешними данными, их консолидация и обработка;

- исключение посредников между системой и пользователями (топ-менеджментом) для ускорения анализа и повышения качества решений.

«На верхнем уровне таких систем находятся панели бизнес-показателей, которые используются для мониторинга, бизнес-моделирования, целевого управления и оценки результативности. Они отображают агрегированные данные, отслеживая прогресс в достижении стратегических целей» [1].

«Концепция ВРМ основывается на цикле управления: цель – планирование – выбор показателей – выполнение – мониторинг – анализ – корректировка целей. Оценка результативности осуществляется через сопоставление достигнутых результатов с целями» [2].

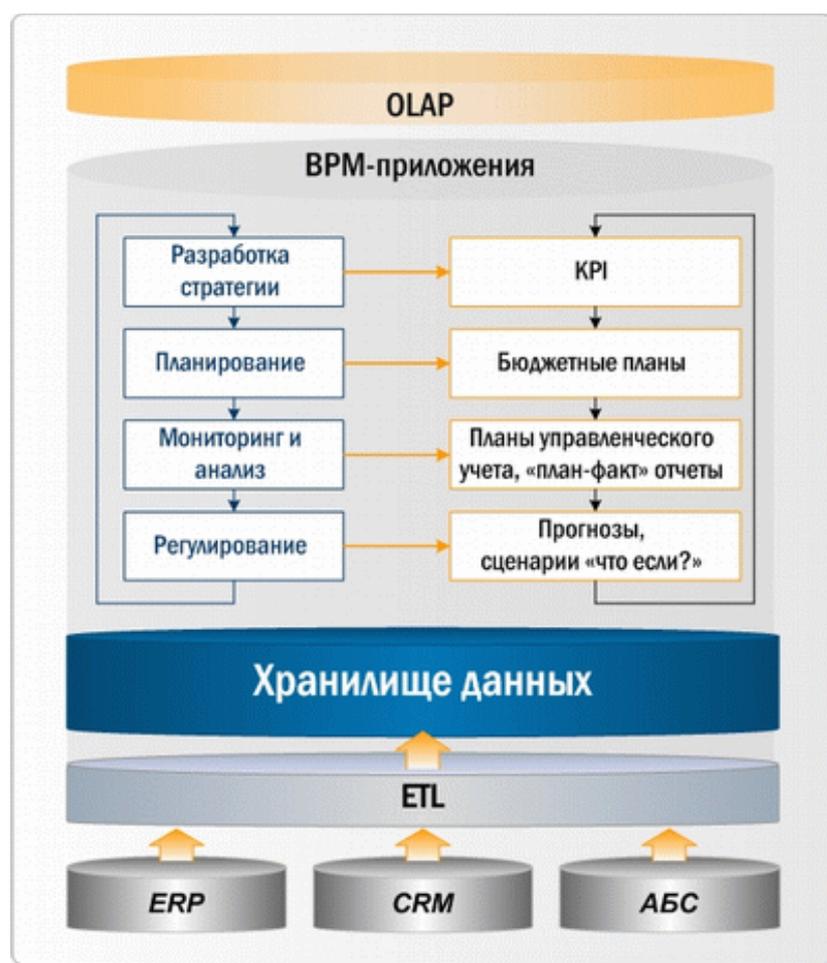


Рисунок 5 – Цикл управления ВРМ

На первом этапе управления предприятием формируется стратегия, устанавливаются цели и осуществляется планирование. Для этого определяется набор ключевых показателей, которые будут отражать степень достижения поставленных целей и обеспечивать контроль управляемости бизнеса. Эти показатели получают плановые значения, соответствующие установленным целям. В рамках BPM-системы функция бизнес-моделирования позволяет создать «формализованный образ будущего» и настроить механизмы мониторинга, чтобы отслеживать прогресс в реализации стратегии. На завершающем этапе анализируются достигнутые результаты, и принимаются решения о корректировке целей или изменении стратегических приоритетов, замыкая управленческий цикл.

Использование BPM-концепции способствует созданию интегрированной системы управления предприятием, обеспечивающей двустороннюю связь с транзакционными системами. Такая связь позволяет не только консолидировать данные и формировать показатели, но и направлять результаты анализа обратно в оперативный контур. Это дает возможность корректировать планы, адаптировать бизнес-процессы и, например, разрабатывать схемы мотивации персонала на основе достигнутых результатов.

«Комплексная интеграция BPM повышает прозрачность бизнеса, усиливает влияние на текущие процессы, помогает руководству глубже понимать ключевые факторы успеха и более эффективно делегировать полномочия. В результате система управления ориентируется на стратегические цели и становится более гибкой и адаптивной»[3].

«Контроль деятельности в BPM-системе осуществляется на трёх уровнях: стратегическом, тактическом и оперативном. Для каждого уровня формируются соответствующие матрицы KPI. Тактические и оперативные показатели включают функциональные KPI, отражающие работу подразделений или сотрудников, но при этом согласуются со стратегическими

показателями, которые характеризуют достижение общекорпоративных целей. Формирование стратегических»[4]

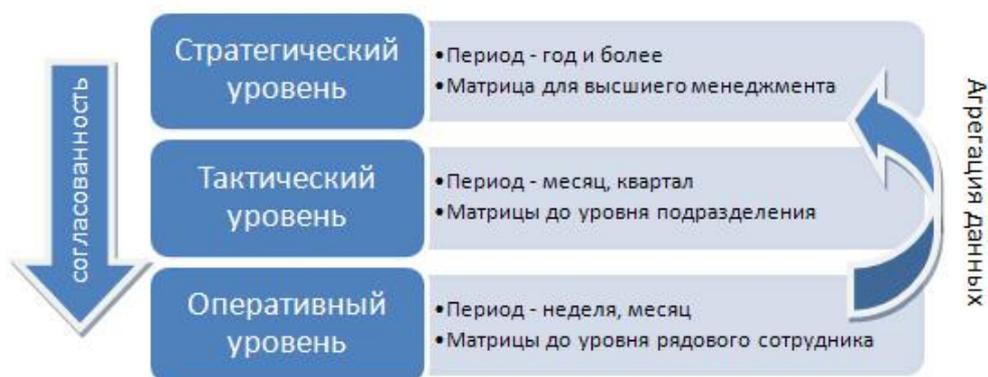


Рисунок 6 – Бизнес-показатели разного уровня

«Особое внимание следует уделить методам интеграции транзакционных систем и BPM-систем. Поскольку ERP-система может работать на платформе, отличной от BPM-системы, для их взаимодействия необходим специальный модуль. Этот модуль должен извлекать необходимые показатели из ERP-системы, преобразовывать их и передавать в BPM-систему. Он должен быть максимально универсальным, чтобы пользователи могли настраивать его самостоятельно, без привлечения технических специалистов, но при этом оставаться простым и удобным в использовании» [5] (рисунок 7).



Рисунок 7 – Система по трансформации данных

Для эффективного функционирования аналитических систем необходимо чётко определить и организовать работу с источниками данных. Важные аспекты включают:

«Перечисление источников данных. Формируется список баз данных, доступных для выборки информации. Эти источники должны быть совместимы с механизмами трансформации, встроенными в аналитическую систему, и предоставлять данные в структурированном виде» [4].

Алгоритм выборки данных. Для каждого показателя аналитической системы разрабатывается алгоритм выборки данных. Определить параметры запросов к базе данных, таких как фильтры, временные рамки или дополнительные условия.

Преобразование данных. Собранные данные должны быть обработаны и преобразованы к типу и формату, соответствующему требованиям базы-приемника. Этот процесс может включать агрегирование, фильтрацию, нормализацию или преобразование единиц измерения.

«Функция обратного контроля. Система обязана предоставлять пользователю возможность расшифровки итоговых данных. Это достигается через сформированные отчёты, где доступна ссылка на исходные документы или записи в базе-источнике. Пользователь может проверить корректность данных, а в случае изменения первичных данных система должна автоматически пересчитать итоговые показатели, обновляя только изменённые элементы без дублирования» [4].

«Коллективная работа и управление доступом. В условиях корпоративного управления, предполагающего коллективное принятие решений, система должна поддерживать работу большого числа пользователей» [1]. Это включает:

- распределение ролей и разграничение доступа на основе функциональных обязанностей. Настройку прав пользователей к данным в зависимости от их уровня доступа;

- обеспечение одновременного доступа к бизнес-информации без снижения производительности системы.

«Эти механизмы делают систему аналитики гибкой, удобной для использования и надёжной для управления корпоративными данными, что особенно важно в условиях динамичной бизнес-среды. Одной из самых популярных платформ для корпоративной учетной системы в России является 1С:Предприятие» [7].

«По определению Gartner, управление эффективностью корпорации (CPM) - это комбинация методик, показателей (финансовых и нефинансовых, долгосрочных и краткосрочных и др.), процессов и систем, используемых для контроля и управления эффективностью деловой деятельности организации» [44].

«В реальной жизни CPM-система объединяет те же функциональные блоки, что и BPM-решение. Это означает, что термины CPM и BPM не имеют существенной смысловой разницы. Хотя ряд авторитетных зарубежных экспертов предпочитает сокращение CPM, рассматривая концепцию корпоративного управления применительно к финансовой отрасли» [45].

«Как считают исследовательские компании, мировой рынок систем управления эффективностью бизнеса переживает период расцвета. Организации самых разных профилей деятельности получают реальную выгоду, внедряя BPM-инструменты. По оценкам IDC, в 2020 г. на аналитические приложения BPM в мире было потрачено 1,42 млрд. долл.» [9].

Вопросами интеллектуального анализа данных занимаются следующие авторы: Борисова И. А., Дюбанов В. В., Загоруйко Н. Г. и Кутненко О. А., Дайзенрот Марк Питер, Чен Сунь Он, Альдо Фейзал А., Kaplan R.S., Norton D.P., Нивен П.Р., Подиновский В.В., Ногин В.Д.

«Интеллектуальный анализ данных связан с отысканием шаблонов, трендов, ассоциаций, аномалий, важных атрибутов, структур и зависимостей в данных. Он является мультидисциплинарной областью, которая использует и совершенствует идеи из различных областей знаний, таких как обработка

сигналов и изображений, машинное обучение, распознавание образов, экспертные системы, оптимизация, статистика и другие» [46].

«ИАД анализирует большие объемы сложных необработанных данных, помогая принимать на их основе решения» [28, 29].

Интеллектуальный анализ данных (ИАД) является итеративным процессом, который можно разбить на следующие этапы [23-27]:

- хранение и управление данными;
- предварительная обработка и подготовка данных к анализу;
- процессы непосредственного анализа данных и извлечения знаний;
- оценка и интерпретация полученных результатов.

«Последний из перечисленных этапов в основном касается нетехнической работы, такой как документация и оценка результатов аналитиком» [47].

«Одним из направлений ИАД, перспективных с точки зрения исследований, является анализ временных рядов, Временной ряд – это последовательность собранных в разные моменты времени значений каких - либо параметров исследуемого процесса или объекта» [13].

«Методы интеллектуального анализа предназначены для изучения временных рядов с целью выявления скрытых закономерностей. Хотя эти методы основаны на разных математических подходах, их объединяет одна важная особенность: для их эффективного использования требуется предварительно обработанное, высокоуровневое представление данных, а не исходные необработанные данные» [44].

«Такое высокоуровневое представление необходимо как для выделения характеристик временного ряда, так и для его эффективного хранения, передачи и обработки.

Вейвлет–преобразование по своей природе призвано предоставить такое высокоуровневое представление временных рядов и поэтому может быть эффективно применено в следующих направлениях» [5].

«Поиск подобия во временных рядах предполагает отыскание полного или частичного совпадения с заданным рядом (шаблоном). При этом вводится параметр допустимого расстояния между рядами» [31].

«Решение данной задачи состоит из двух этапов: индексирования рядов и выполнения запросов. Индексирование – это процесс создания указателей для ускорения доступа к данным, предполагающий выделение признаков временного ряда и его сжатие. ДВП в данной задаче может быть применено для выполнения индексирования данных и создания метрики расстояния между рядами» [32].

«Классификация временных рядов состоит в назначении ряду одной из заранее известных меток класса. ДВП может быть интегрировано в классификацию временных рядов двумя путями: применение методов классификации к результатам вейвлет–преобразования и применение многомасштабного представления данных» [33].

«Кластеризация временных рядов состоит в их разбиение на группы по подобию. Кластеризация позволяет идентифицировать шаблоны и тренды для каждой из групп» [34].

«Выявление аномалий в поведении временных рядов включает: определение внезапных изменений в рядах, выявление аномалий путем сравнения расстояний между несколькими рядами и выявление нерегулярных шаблонов» [35].

«Прогнозирование значений временного ряда в будущем на основании исторических данных» [30].

1.3 Витрины данных

«Витрина данных (англ. Data Mart; специализированное хранилище данных, киоск данных) – подмножество (срез) хранилища данных, представляющее собой массив тематической, узконаправленной информации,

ориентированной, например, на пользователей одной рабочей группы или подразделения» [21].

Основные цели выделения витрин – приближение данных к конечным пользователям, ограничение только необходимыми данными.

«Витрина данных – это простая форма хранилища данных, которое ориентировано на определенную тему или направление деятельности, например на продажи, финансы или маркетинг» [43].

«С учетом этой узкой специализации получается, что витрины данных собирают данные из меньшего количества источников, чем хранилища данных. Источниками для витрины данных могут служить внутренние операционные системы, центральное хранилище данных и внешние данные» (рисунок 8) [18].

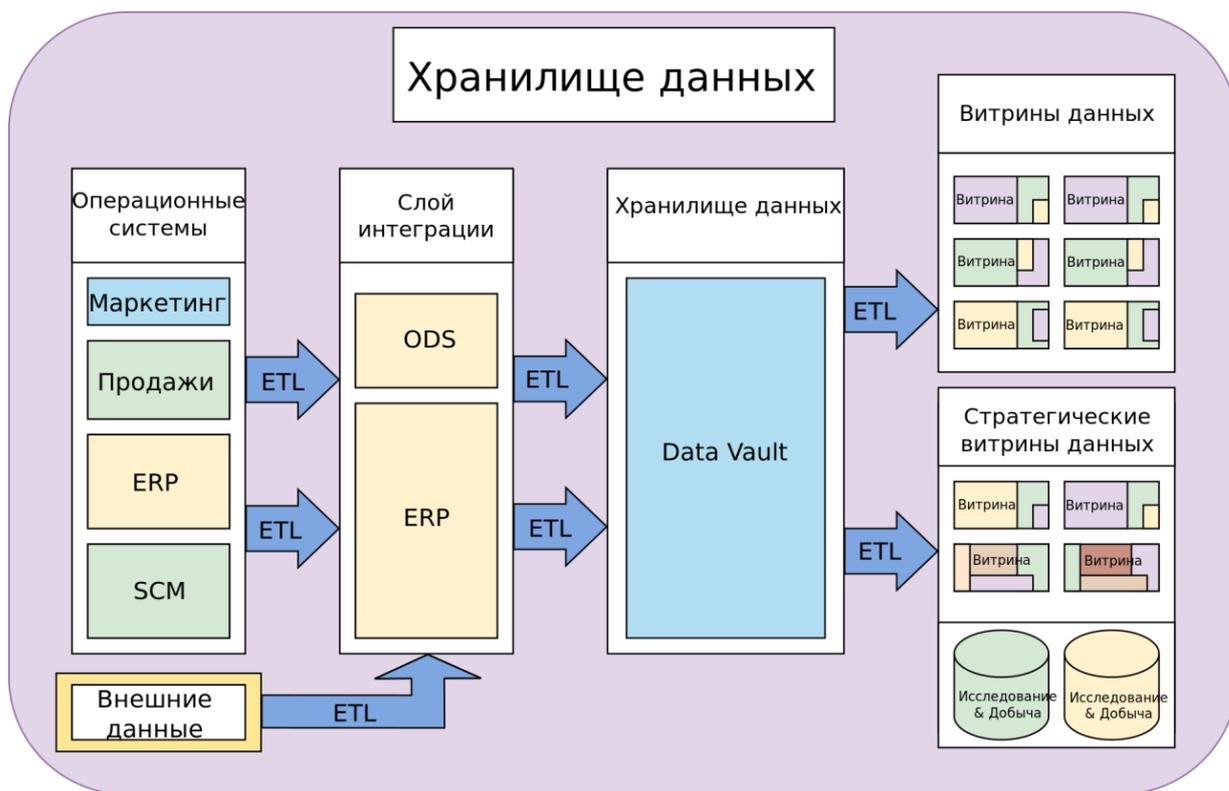


Рисунок 8 – Витрина данных

Витрина данных обеспечивает более простой доступ к данным, необходимым для определенного отдела или производственного направления внутри организации.

«Преимущества витрины данных:

- витрина данных, созданная для определенного отдела или направления деятельности, дает ряд преимуществ» [36].
- единый источник достоверных данных;
- более быстрый доступ к данным;
- быстрое получение статистических данных ускоряет принятие решений [37];
- более простое и быстрое применение;
- создание гибкого и масштабируемого решения для управления данными» [38];
- «анализ переходных процессов» [39].

«Витрины (витрина от англ. data mart) – это набор структурированных данных. Обычно это данные по определенной теме или задаче в компании. Ничего лишнего, только нужные и актуальные очищенные данные, полученные из других ИС предприятия. Таких витрин даже на одном предприятии может быть множество» [20].

«Витрина данных включает в себя подраздел общекорпоративных данных, который представляет ценность для определенной группы пользователей в организации. В отличие от хранилища данных, создание которого дорого и сложно, оно предлагает экономичную альтернативу» [48].

В процессе работы на магистерской работой предлагается использовать витрины данных для настройки и эффективной обработки корпоративной информации. Использование витрин данных позволит повысить эффективность работы аналитиков компании.

Таким образом, по результатам рассмотренных публикаций, учебников и диссертаций, можно сделать вывод, что по теме методы и модели

управления эффективностью системы аналитической обработки корпоративной информации необходимо сосредоточиться на следующих направлениях :

- управление эффективностью предприятия с помощью бизнес систем (CRM, ERP, BPM);
- технологии, где используется отображение информации с помощью витрин данных;
- методах сбора информации с разных источников корпоративной информации;
- моделях данных, которые используются в системах OLTP и OLAP;
- рассмотреть и определить показатели эффективности предприятия, необходимые в рамках магистерской работы.

«Управление эффективностью предприятия (EPM, или Enterprise Performance Management) – это бизнес-аналитика нового поколения, с помощью которой в компаниях принимают стратегические решения в соответствии с корпоративными целями» [40].

«Одним из перспективных методов является исследование и совершенствование витрин данных как важного элемента корпоративных аналитических систем. В ходе работы была сформулирована гипотеза исследования и сделана постановка проблемы, разработаны модели бизнес процесса» [33].

Выводы по главе 1

Большая работа в ходе научно-исследовательской работы проведена по работе с литературой и информационными ресурсами, по поиску статей, работ и других аналитических отчетов по теме диссертации.

Глава 2 Анализ методов и моделей управления эффективностью системы аналитической обработки корпоративной информации

2.1 Метод управления эффективностью системы аналитической обработки корпоративной информации

В ходе выполнения магистерской работы опишем методы управления эффективностью системы аналитической обработки корпоративной информации в методологии структурного подхода [33].

«Входной информацией бизнес-процессов процессов управления эффективностью системы аналитической обработки корпоративной информации являются множество первичных документов, которые организация собирает о своей деятельности (рисунок 9):

- данные Яндекс.метрики об активности на корпоративном сайте компании;
- налоговые отчеты о деятельности организации;
- внутренняя информация организации;
- данные с CRM-системы;
- аналитические отчеты аудиторских компаний о деятельности организации;
- запрос на данные с внешних систем;
- запрос о банковских операциях организации;
- форма обратной связи;
- статистические данные» [33].

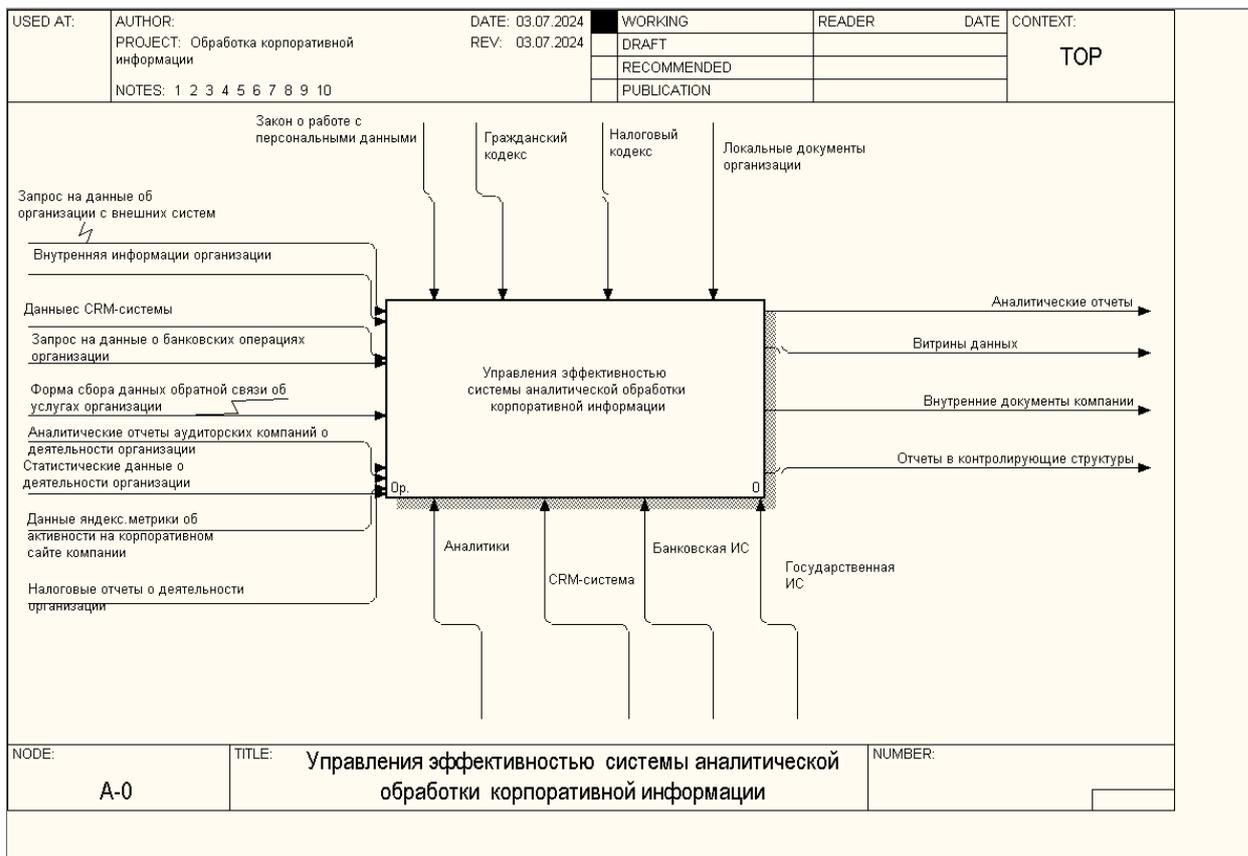


Рисунок 9 – Контекстная диаграмма «Управления эффективностью системы аналитической обработки корпоративной информации. Как есть»

«Выходной информацией бизнес-процессов процесса управления эффективностью системы аналитической обработки корпоративной информации являются, различные отчеты, а именно:

- аналитические отчеты;
- витрины данных;
- внутренние документы компании;
- отчеты в контролирующие структуры.

Управление – правовые документы и кодексы, определяющие работу с данными:

Механизмы – люди и программно-аппаратные средства» [33].

«Рассмотрим основные операции «Управления эффективностью системы аналитической обработки корпоративной информации» более подробно на рисунке 10» [33].

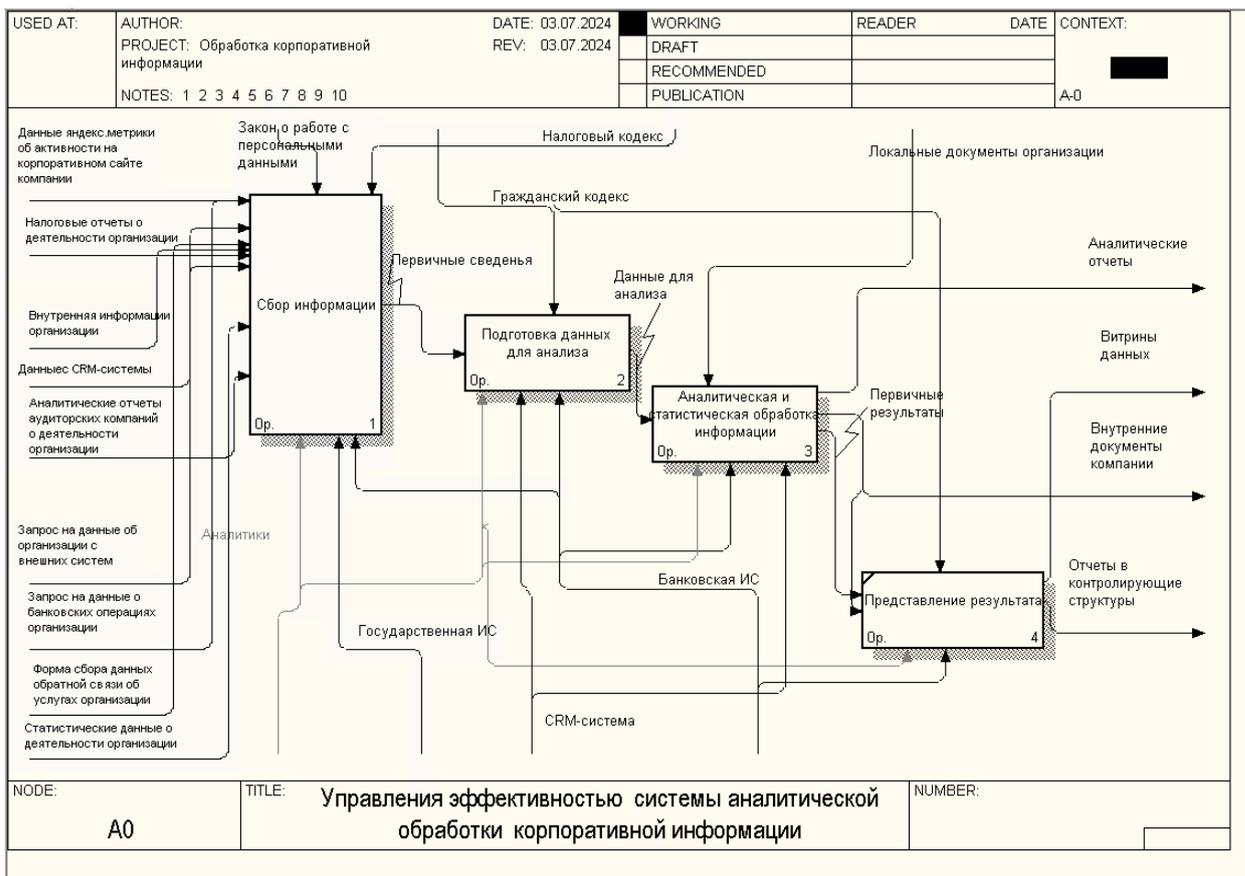


Рисунок 10 – Управления эффективностью системы аналитической обработки корпоративной информации. Как есть

Управления эффективностью системы аналитической обработки корпоративной информации состоит из следующих блоков:

- сбор информации;
- подготовка данных для анализа;
- аналитическая и статистическая обработка информации;
- представление результата.

Для более детального анализа состояния как есть, рассмотрим каждый блок более подробно и приведем результаты рассмотрения в таблице 1.

Таблица 1 – Информация управляющих блоков процесса управления эффективностью системы аналитической обработки корпоративной информации. Как есть

Название функционального блока	Входная информация функционального блока	Выходная информация функционального блока	Управление функционального блока	Механизм функционального блока
Сбор информации	<p>Аналитические отчеты аудиторских компаний о работе компании</p> <p>Форма для сбора обратной связи по предоставляемым услугам.</p> <p>Запрос данных о банковских операциях компании.</p> <p>Запрос информации об организации из внешних источников.</p> <p>Внутренняя информация компании.</p> <p>Налоговая отчетность о деятельности компании.</p> <p>Данные из CRM-системы.</p> <p>Статистическая информация о показателях деятельности компании.</p>	Первичные сведения	<p>Налоговый кодекс,</p> <p>Федеральный закон N 152-ФЗ «О персональных данных» регулирует отношения связанные с обработкой персональных данных</p>	<p>Аналитики, занимающиеся консультациям и области компьютерных технологий,</p> <p>Государственная ИС,</p> <p>Банковская ИС</p>

Продолжение таблицы 1

Название функционального блока	Входная информация функционального блока	Выходная информация функционального блока	Управление функционального блока	Механизм функционального блока
Подготовка корпоративных данных для анализа	Первичные сведения	Корпоративные данные для анализа	Гражданский кодекс — систематизированный законодательный акт, содержащий нормы гражданского права	Аналитики, занимающиеся консультациями и области компьютерных технологий, CRM-система, Банковская ИС
Аналитическая и статистическая обработка корпоративной информации	Корпоративные данные для анализа	Первичные результаты обработки корпоративных данных, Внутренние документы компании, Аналитические отчеты	Локальные документы организации	Аналитики, занимающиеся консультациями и области компьютерных технологий, Банковская ИС, CRM-система
Представление результата обработки корпоративных данных	Внутренние документы компании, Первичные результаты обработки корпоративных данных	Отчеты в контролирующей структуре, Витрины корпоративных данных	Федеральный закон N 152-ФЗ «О персональных данных» регулирует отношения связанные с обработкой персональных данных, Гражданский кодекс — систематизированный законодательный акт, содержащий нормы гражданского права, Локальные документы организации	Аналитики, занимающиеся консультациями и области компьютерных технологий

«Одним из перспективных методов является исследование и совершенствование витрин данных как важного элемента корпоративных аналитических систем, на котором сосредоточено исследование и изменения будут происходить в блоке «Представление результатов», на рисунке 11 процесс представлен «Как есть» и происходит исключительно отчетами, которые формируют аналитики» [33].

Функциональный блок «Представление результата. Как есть» состоит из следующий функциональных блоков следующего уровня декомпозиции (рисунок 11):

- интеграция первичных данных;
- формирования витрин данных;
- формирование стратегических витрин данных.

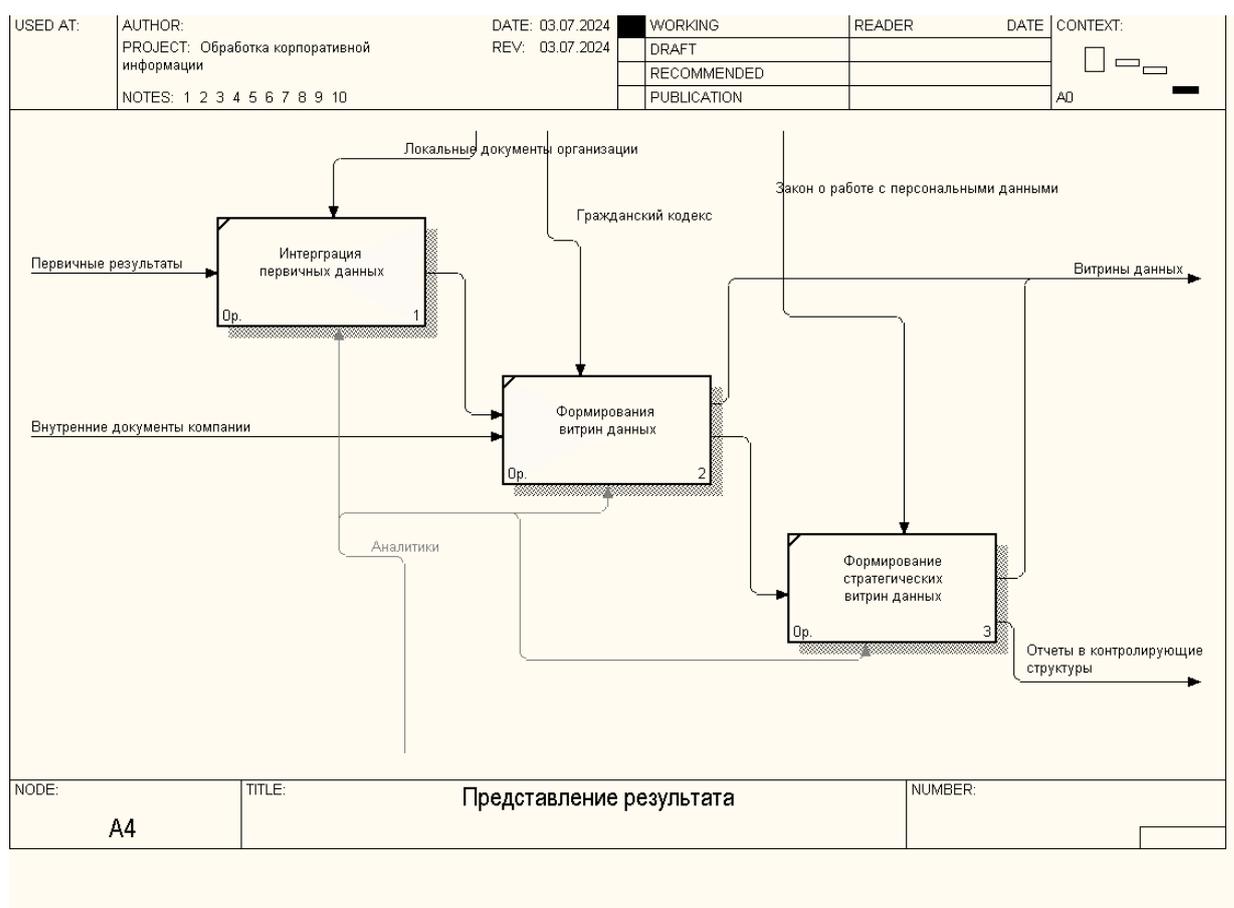


Рисунок 11 – Представление результатов. Как есть

Для более детального анализа состояния как есть бизнес процесса представление результата, рассмотрим каждый блок более подробно и приведем результаты рассмотрения в таблице 2.

Таблица 2 – Информация управляющих блоков процесса представление результата. Как есть

Название функционального блока	Входная информация функционального блока	Выходная информация функционального блока	Управление функционального блока	Механизм функционального блока
Интеграция первичных корпоративных данных	Первичные результаты обработки корпоративных данных	Список корпоративных данных	Локальные документы организации	Аналитики, занимающиеся консультациями и области компьютерных технологий
Формирования витрин корпоративных данных	список корпоративных данных	список вариантов отображения корпоративных данных, Витрины корпоративных данных	Гражданский кодекс — систематизированный законодательный акт, содержащий нормы гражданского права	Аналитики, занимающиеся консультациями и области компьютерных технологий
Формирование стратегических витрин корпоративных данных	список вариантов отображения корпоративных данных	Отчеты в контролирующей структуре, Витрины корпоративных данных	Федеральный закон N 152-ФЗ «О персональных данных» регулирует отношения связанные с обработкой персональных данных	Аналитики, занимающиеся консультациями и области компьютерных технологий

В работе предлагается внедрить управление эффективностью предприятия BPM- систему, которое позволит компании принимать решения в системе бизнес-анализа и стратегически увязывать эти решения со своими корпоративными целями. И по результатам обзора литературы принято решение о необходимости добавления BPM-системы.

Контекстная диаграмма «Управления эффективностью системы аналитической обработки корпоративной информации. Как должно быть» показана на рисунке 12.

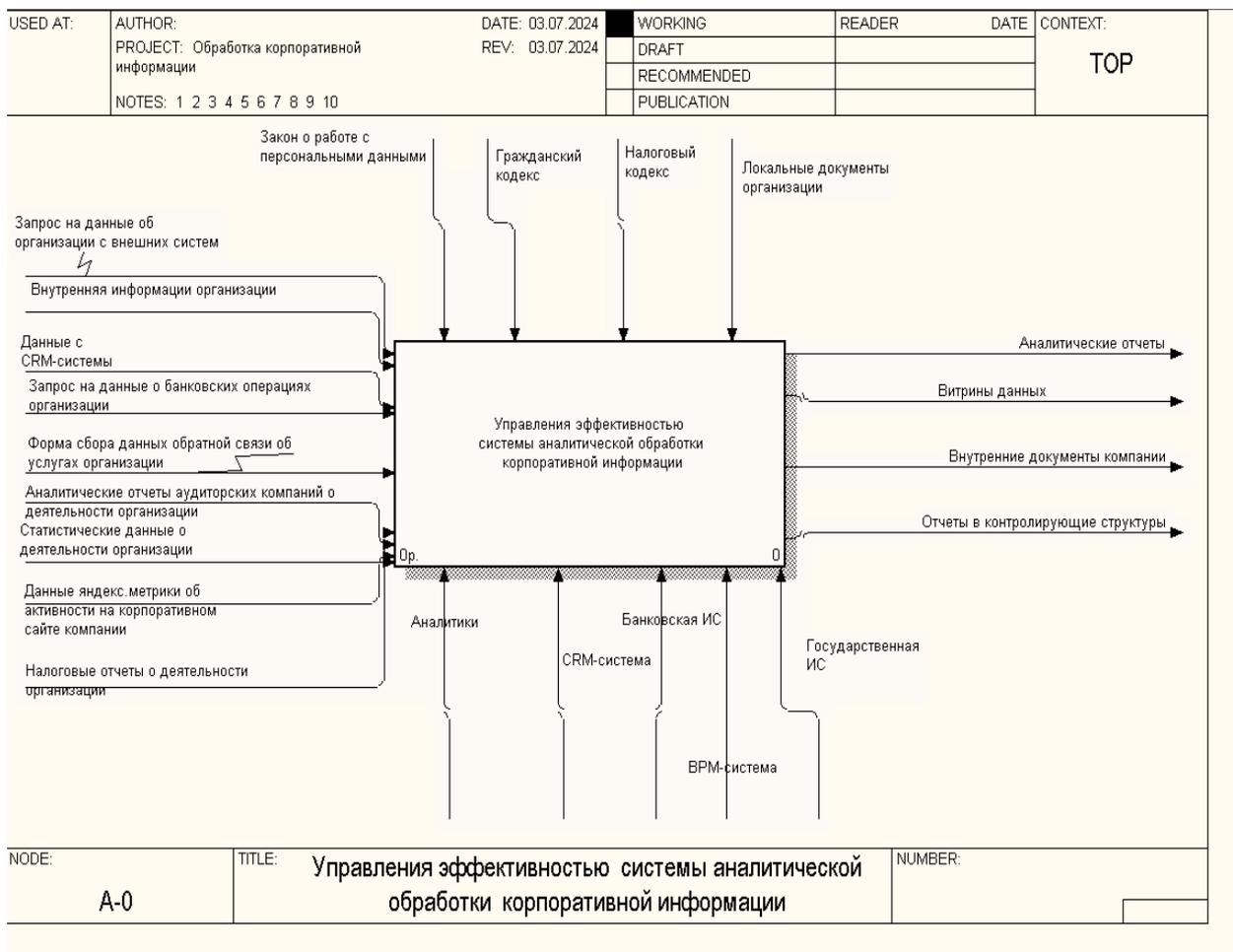


Рисунок 12 – Контекстная диаграмма «Управления эффективностью системы аналитической обработки корпоративной информации. Как должно быть»

Обратим внимание, в какой функциональный блок добавлена новая система (рисунок 13).

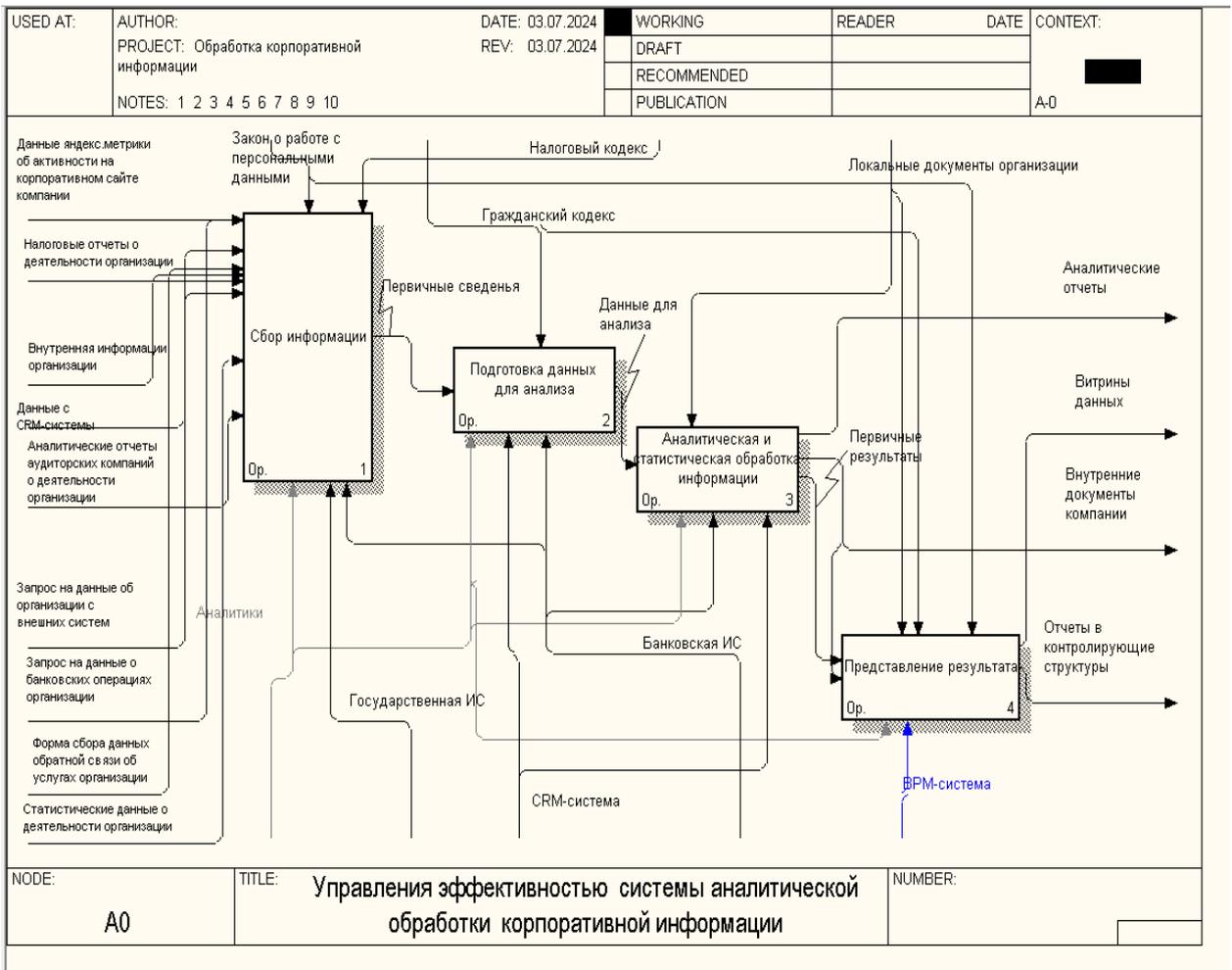


Рисунок 13 – Управления эффективностью системы аналитической обработки корпоративной информации. Как должно быть

Более детальный анализ изменений, произошедших в бизнес- процессе показан в таблице 3.

Таблица 3 – Информация управляющих блоков процесса управления эффективностью системы аналитической обработки корпоративной информации. Как должно быть

Название функционального блока	Входная информация функционального блока	Выходная информация функционального блока	Управление функционального блока	Механизм функционального блока
Сбор корпоративной информации	<p>Аналитические отчеты аудиторских компаний о работе компании</p> <p>Форма для сбора обратной связи по предоставляемым услугам.</p> <p>Запрос данных о банковских операциях компании.</p> <p>Запрос информации об организации из внешних источников.</p> <p>Внутренняя информация компании.</p> <p>Налоговая отчетность о деятельности компании.</p> <p>Данные из CRM-системы.</p> <p>Статистическая информация о показателях деятельности компании.</p>	Первичные сведения	Налоговый кодекс, Закон о работе с персональными данными	Аналитики, занимающиеся консультациям и области компьютерных технологий, Государственная ИС, Банковская ИС

Продолжение таблицы 3

Название функционального блока	Входная информация функционального блока	Выходная информация функционального блока	Управление функционального блока	Механизм функционального блока
Подготовка корпоративных данных для анализа	Первичные сведения	Данные для анализа	Гражданский кодекс — систематизированный законодательный акт, содержащий нормы гражданского права	Аналитики, занимающиеся консультациями и области компьютерных технологий, CRM-система, Банковская ИС
Аналитическая и статистическая обработка корпоративной информации	Данные для анализа	Первичные результаты, Внутренние документы компании, Аналитические отчеты	Локальные документы организации	Аналитики, занимающиеся консультациями и области компьютерных технологий, Банковская ИС, CRM-система
Представление результата обработки корпоративной информации	Внутренние документы компании, Первичные результаты	Отчеты в контролирующей структуре, Витрины данных	Федеральный закон N 152-ФЗ «О персональных данных» регулирует отношения связанные с обработкой персональных данных, Гражданский кодекс — систематизированный законодательный акт, содержащий нормы гражданского права, Локальные документы организации	Аналитики, занимающиеся консультациями и области компьютерных технологий и BPM-система

Витрины (витрина от англ. data mart) – это набор структурированных данных. Обычно это данные по определенной теме или задаче в компании.

Ничего лишнего, только нужные и актуальные очищенные данные, полученные из других ИС предприятия. Таких витрин даже на одном предприятии может быть множество » [33] (рисунок 14).

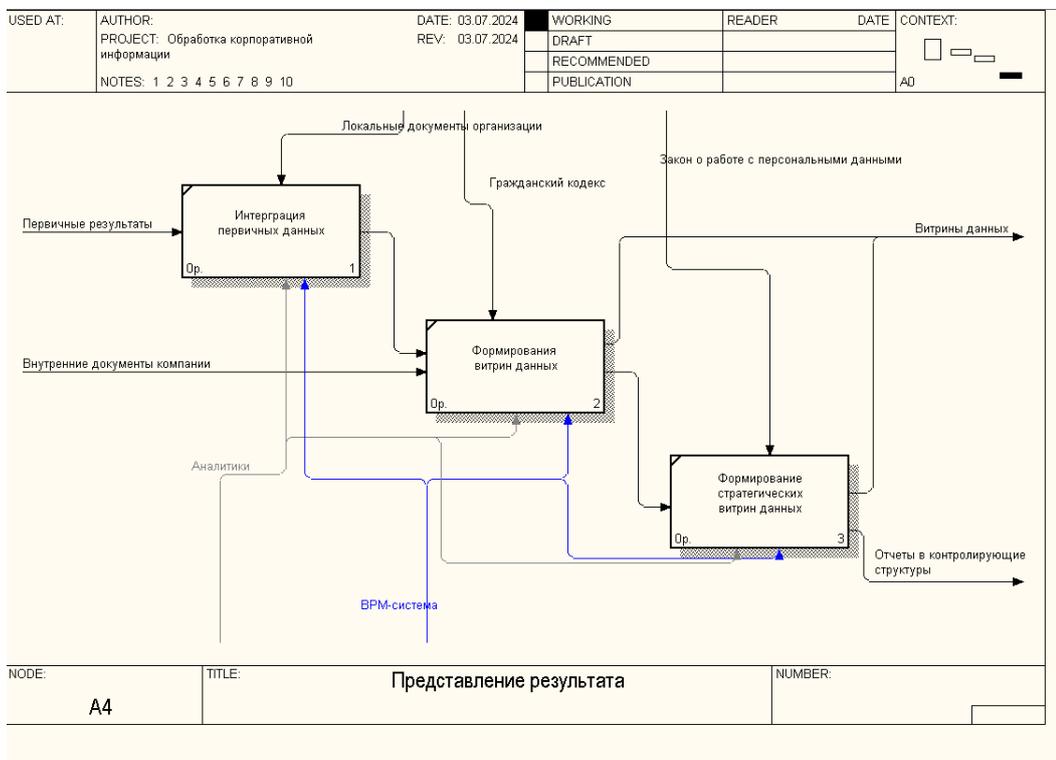


Рисунок 14 – Представление результатов. Как должно быть

Более детальный анализ изменений показан, произошедших в бизнес процессе показан в таблице 4.

Таблица 4 – Информация управляющих блоков процесса представление результата. Как есть

Название функционального блока	Входная информация функционального блока	Выходная информация функционального блока	Управление функционального блока	Механизм функционального блока
Интеграция первичных корпоративных данных	Первичные результаты обработки корпоративных данных	список данных	Локальные документы организации	Аналитики, занимающиеся консультациям и области компьютерных технологий, BPM-система

Продолжение таблицы 4

Название функционального блока	Входная информация функционального блока	Выходная информация функционального блока	Управление функционального блока	Механизм функционального блока
Формирования витрин корпоративных данных	список корпоративных данных	список вариантов отображения данных, Витрины данных	Гражданский кодекс — систематизированный законодательный акт, содержащий нормы гражданского права	Аналитики, занимающиеся консультациями и области компьютерных технологий, ВРМ-система
Формирование стратегических витрин корпоративных данных	список вариантов отображения корпоративных данных	Отчеты в контролирующей структуре, Витрины данных	Федеральный закон N 152-ФЗ «О персональных данных» регулирует отношения связанные с обработкой персональных данных	Аналитики, занимающиеся консультациями и области компьютерных технологий, ВРМ-система

«Результатом выполнения задания по практике является проектирования бизнес-процессов, которые отражают текущее состояние и состояние после внедрения новых методов управления эффективностью системы аналитической обработки корпоративной информации. Внедрение новых методов и алгоритмов приведет к единому формату данных и позволит взаимодействовать с различными источниками входных данных, а также позволит получать различные витрины данных (простые и сложные)» [34].

2.2 Модель управления эффективностью системы аналитической обработки корпоративной информации

Модели управления эффективностью системы аналитической обработки корпоративной информации опишем в методологии ВРМН. Вначале подробно опишем процесс получения первичной информации и источники откуда

можно получить корпоративную информацию (рисунок 15).

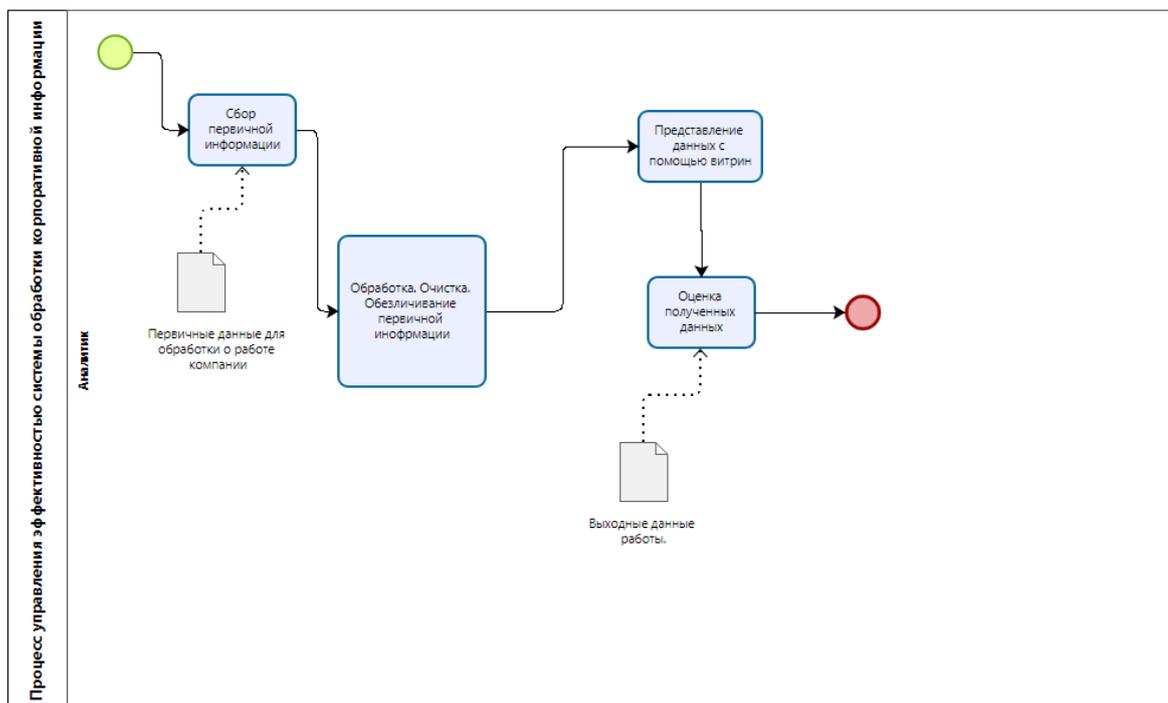


Рисунок 15 – Процесс управления эффективностью системы обработки корпоративной информации. Как есть

«Основными операциями верхнего уровня являются (рисунок 16):

- сбор первичной информации;
- обработка, очистка, обезличивание первичной информации;
- представление данных с помощью витрин;
- оценка полученных данных» [34].

Входной информацией является – первичные данные для обработки о работе компании.

Выходной информацией – выходные данные работы.

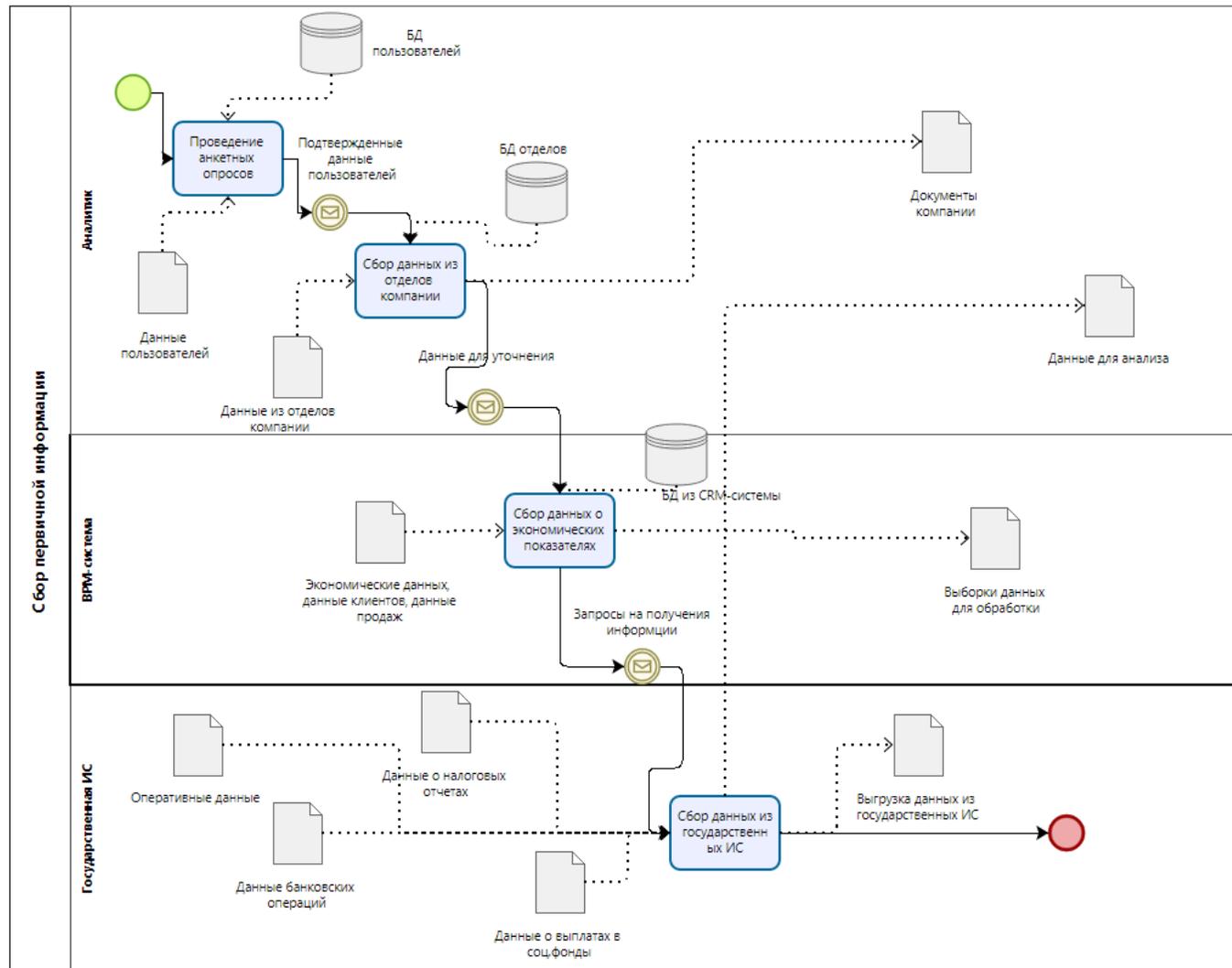


Рисунок 16 – Работа по сбору первичных данных. Как есть

Рассмотрим подробно модель по сбору первичных данных. Данная модель является ключевой в магистерской работе. На этом этапе важно определить следующие источники:

- «база данных;
- данные, из каких источников, будут загружены в систему.

Система подключается к следующим базам данных:

- база данных отделов;
- база данных из CRM-системы;
- база данных пользователей.

Источниками с данными являются следующие документы и сводные отчеты:

- данные пользователей;
- данные из отделов компании;
- экономические данные клиентов, данные продаж;
- оперативные данные;
- данные банковских операций;
- данные о налоговых отчетах;
- данные о выплатах из социальных фондов» [33].

После этапа сбора данных формируются интегрированные первичные данные, которые становятся основой для дальнейшего анализа. Эти данные включают:

- документы компании – внутренние отчёты, акты, договоры и другие документы, отражающие ключевые бизнес-процессы;
- данные для анализа – структурированная информация, собранная из различных источников, используемая для аналитической обработки;
- выборки данных для обработки – специально отобранные наборы данных, предназначенные для расчётов или моделирования;

- выгрузка данных из государственных информационных систем – сведения, полученные из внешних источников, таких как налоговые службы, реестры и другие регулирующие органы.

После определения источников данных и их обработки необходимо перейти к идентификации ключевых показателей эффективности (KPI), которые являются важными для оценки качества обработки данных и принимаемых решений. В рамках данной магистерской работы рассматриваются следующие показатели:

Уровень доступности аналитических отчётов - этот показатель отражает, насколько оперативно и в каком объёме пользователи могут получить доступ к аналитическим данным. Доступность отчётов является основой для работы аналитических систем, так как она напрямую влияет на своевременность принятия решений.

Своевременность обработки запросов - показатель отражает скорость обработки запросов на получение данных и формирования аналитических отчётов. Этот KPI характеризует эффективность работы системы и её способность предоставлять необходимые данные для принятия управленческих решений.

Удовлетворённость представлением данных - показатель определяет, насколько полно и понятно представлена информация для конечных пользователей, включая топ-менеджмент и аналитиков. Удобство визуализации данных и возможность детализации напрямую влияют на принятие решений.

Взаимосвязь ключевых показателей - уровень доступности аналитических отчётов и своевременность обработки запросов тесно связаны между собой, поскольку доступность отчётов зависит от того, насколько быстро система может обработать запрос. В общем случае эта метрика рассчитывается как отношение числа своевременно обработанных запросов к общему количеству запросов за определённый период:

Для уровня доступности – время задержки при получении данных, процент запросов, выполненных без ошибок.

Для своевременности обработки – среднее время ответа системы на запрос, максимальное время ожидания.

Для удовлетворённости представлением данных – отзывы пользователей, результаты опросов и метрики использования интерфейсов системы.

Подробное рассмотрение этих показателей позволит выявить узкие места в работе системы и предложить решения для их устранения. Таким образом, анализ этих КРІ станет ключевой частью исследования и позволит оценить эффективность работы аналитической системы в целом. (рисунок 17).

Низкие значения	Меньше 80%	Анализировать причины низкой своевременности решения: <ul style="list-style-type: none"> • низкая оперативность реагирования • возвраты на доработку • нехватка ресурсов • отсутствие контроля исполнителей • нехватка знаний, прав, инструментов • внешние зависимости
Высокие значения	Больше 98%	<ul style="list-style-type: none"> • Выдающаяся поддержка • Завышенные сроки в нормативах • Злоупотребления механизмом останова таймера • Нерезультативные решения

Рисунок 17 – Метрики уровня доступности аналитических отчетов

Удовлетворенность представлением данных, на этом этапе предполагается использование витрин данных. Общее представление, где используются витрины показано на рисунке 18.

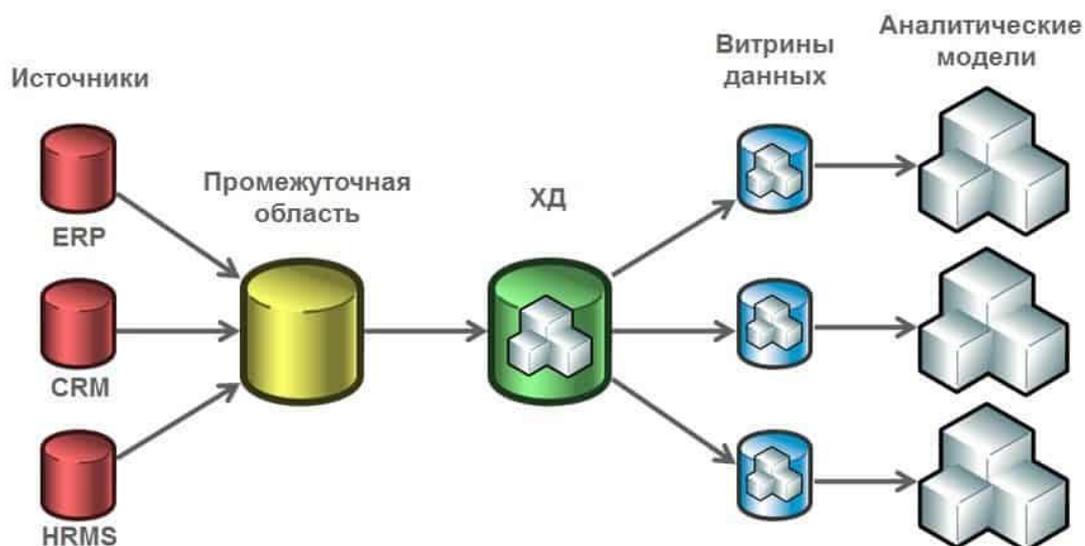


Рисунок 18 – Представлением данных с помощью витрин

Процесс отображения данных с помощью витрин данных предполагает создание фильтров по данным и использование различного отображения данных. Это позволяет пользователям получать доступ к информации, которая им необходима для принятия обоснованных решений.

Выводы по главе 2

Во второй главе подробно рассмотрены методы и модели управления эффективностью системы аналитической обработки корпоративной информации. Определены источники информации для каждого бизнес-процесса модели, выявлены зависимости каждого информационного потока на метрику, которая в дальнейшем используется для определения эффективности процесса.

В результате анализа данных компании могут принимать обоснованные решения о том, как улучшить свою деятельность, снизить издержки и повысить конкурентоспособность. Таким образом, управление эффективностью системы аналитической обработки корпоративной информации является важным инструментом для успешного развития бизнеса.

Глава 3 Разработка приложения с использованием предложенных методов и моделей управления эффективностью системы аналитической обработки корпоративной информации

3.1 Разработка приложения с витринами данных

«Последовательность использования витрин данных для управления эффективностью системы аналитической обработки корпоративной информации» [34] показана на рисунке 19.



Рисунок 19 – Последовательность использования витрин данных для управления эффективностью системы аналитической обработки корпоративной информации

Рассмотрим последовательность использования витрин данных для управления эффективностью системы аналитической обработки корпоративной информации, на первом этапе это – «Сбор данных». На этом этапе витрины данных, представляющие очищенные и агрегированные данные из разных источников, служат основой для экспертных оценок.

На втором этапе это – «Построение экспертных правил». На этом этапе на основе данных из витрин могут быть созданы правила, используемые для принятия решений. Эти правила основаны на исторических данных и опыте корпоративного аналитика [14-16].

На третьем этапе это – «Анализ и оценка». На этом этапе витрины данных облегчают процесс создания и тестирования моделей оценки.

В работе применён язык моделирования UML. Для создания диаграмм использовался редактор Microsoft Visio, который упрощает процесс разработки и минимизирует ошибки.

«Диаграммы помогают сосредоточиться на результатах работы системы, определить входные и выходные данные, а также сформулировать более точные задачи для разработки. Основные элементы таких диаграмм представлены ниже» [11].

Действующее лицо (Actor) – представляет пользователя или внешнюю систему, взаимодействующую с системой.

Вариант использования (Use Case) – описывает функциональность, предоставляемую системой [22].

Связь (Association) – показывает взаимодействие между действующими лицами и вариантами использования. (рисунок 20).

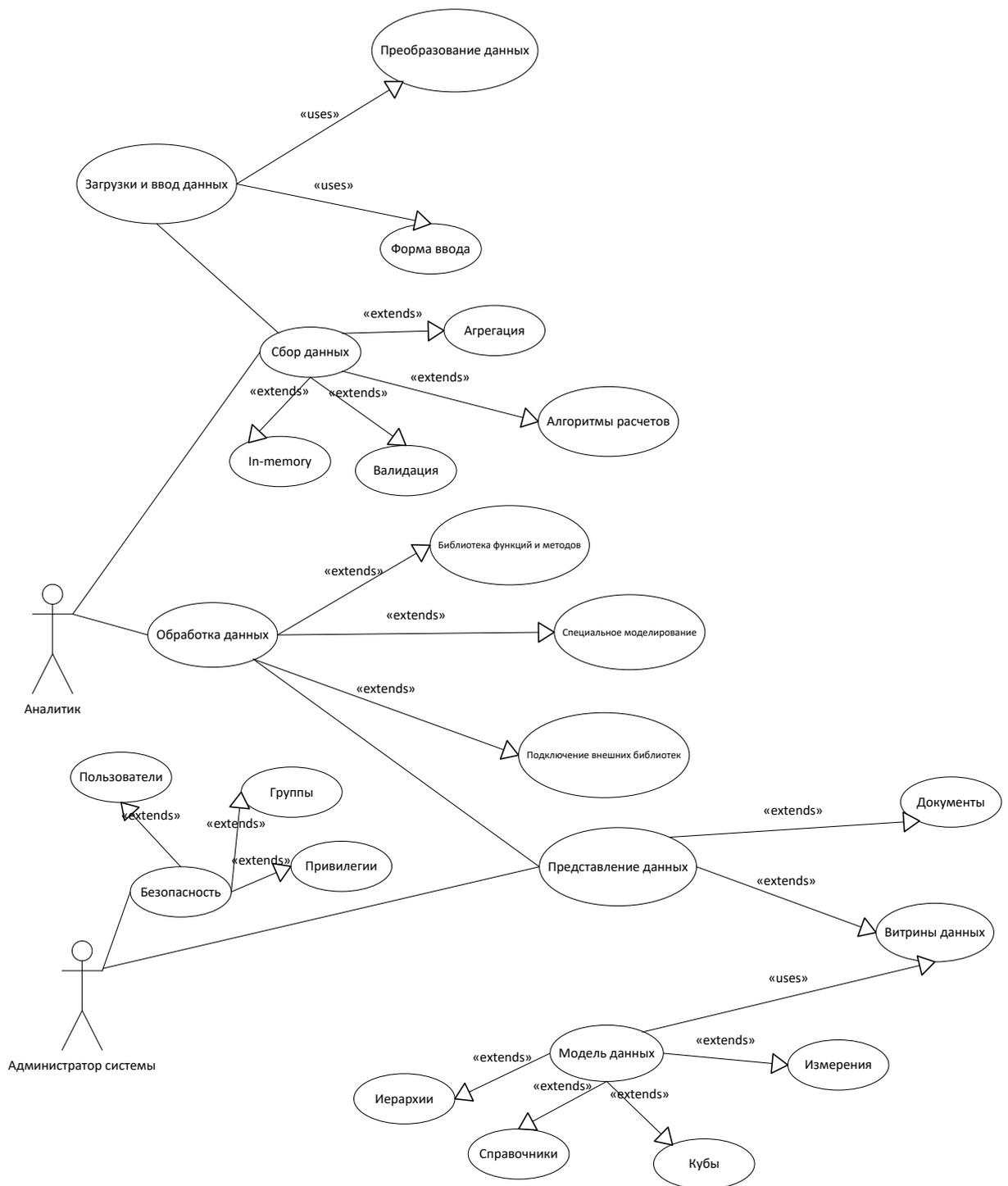


Рисунок 20 – Диаграмма вариантов использования системы

Чтобы раскрыть содержания действий, представленных на диаграмме, требуется разработать текстовый сценарий для следующего этапа проектирования. Сценарий действий системы и действующих лиц представлен в таблице 5.

Таблица 5 – Описание диаграммы вариантов использования

Элемент диаграммы	Описание элемента диаграммы вариантов использования
Основные акторы	Аналитик, Администратор системы
<p>Варианты использования</p> <p>Каждый из перечисленных вариантов использования и функций играет важную роль в реализации системы, обеспечивая её целостность, производительность и удобство для пользователей. Это позволяет эффективно управлять данными на всех этапах их обработки — от ввода до анализа и представления.</p>	<p>Система предусматривает ключевые варианты использования</p> <p>«Сбор данных - включает функции, необходимые для эффективного управления и консолидации информации:</p> <ul style="list-style-type: none"> – In-memory — технология оперативной обработки данных, обеспечивающая высокую производительность и доступ к данным в режиме реального времени»[50]. – Валидация — проверка данных на корректность, что исключает ошибки и несоответствия. – Алгоритмы расчетов — выполнение сложных вычислений на основе заданных формул и методов. – OLAP — инструменты аналитической обработки данных, позволяющие выполнять многомерный анализ. – Агрегация — объединение данных с целью получения сводной информации для аналитики. <p>«Загрузка и ввод данных - включает функции, обеспечивающие подготовку и ввод информации в систему</p> <p>Обработка данных реализуется с использованием следующих функций»[49]:</p> <ul style="list-style-type: none"> – Библиотека функций и методов — набор стандартных и пользовательских функций для обработки данных. – «Специальное моделирование — инструменты для создания сложных моделей обработки и анализа. – Подключение внешних библиотек — интеграция с дополнительными библиотеками для расширения функциональности системы» [34]. <p>Безопасность</p> <p>Функции безопасности обеспечивают защиту данных и контроль доступа:</p> <ul style="list-style-type: none"> – Пользователи — управление учетными записями и правами доступа. – Группы — объединение пользователей с одинаковыми правами для упрощения управления доступом. – Привилегии — настройка уровня доступа к различным функциям и данным системы. <p>Представление данных</p> <p>Этот вариант использования включает функции для визуализации и представления информации:</p>

Продолжение таблицы 5

Элемент диаграммы	Описание элемента диаграммы вариантов использования
	<ul style="list-style-type: none"> – Документы — формирование отчетов, таблиц и других форматов представления данных. – Витрины данных — специализированные интерфейсы для отображения сводной информации. <p>Витрины данных и модель данных</p> <p>Вариант использования «Витрины данных» опирается на:</p> <ul style="list-style-type: none"> – Иерархии — структурирование данных для отображения их взаимосвязей. – Справочники — унифицированные базы данных для хранения стандартной информации. – Кубы — многомерные структуры данных для аналитики. <p>Измерения — параметры, по которым осуществляется анализ данных.</p>
<p>Краткое описание.</p> <p>Взаимодействие между аналитиком и администратором системы обеспечивает плавный процесс работы с данными: от их загрузки и преобразования до аналитической обработки. Эта связка позволяет системе быть не только гибкой и адаптивной, но и максимально продуктивной для принятия управленческих решений.</p>	<p>Процесс работы с данными в системе начинается с этапа загрузки информации из внешних источников, что выполняет аналитик. Аналитик собирает данные из различных систем, баз данных или других внешних источников, чтобы обеспечить полноценный информационный поток для анализа.</p> <p>После загрузки данных в систему может возникнуть необходимость в их преобразовании в формат, соответствующий требованиям системы. «За выполнение этого этапа отвечает администратор системы. Он использует встроенные инструменты для модификации представления данных и настройки модели данных»[34].</p> <p>После ввода и преобразования информации данные проходят этап обработки и консолидации. На основе этой информации аналитик начинает выполнять свою основную задачу — проведение аналитической работы. Для этого он использует различные инструменты системы, включая витрины данных. С помощью этих инструментов аналитик может выполнять следующие задачи:</p> <p>Оценка текущих бизнес-показателей на основе данных в реальном времени.</p> <p>Выявление трендов и закономерностей в деятельности компании.</p> <p>Формирование прогнозов на основе имеющейся информации.</p>

Диаграмма классов (англ. class diagram) — структурная диаграмма языка моделирования UML (рисунок 21).

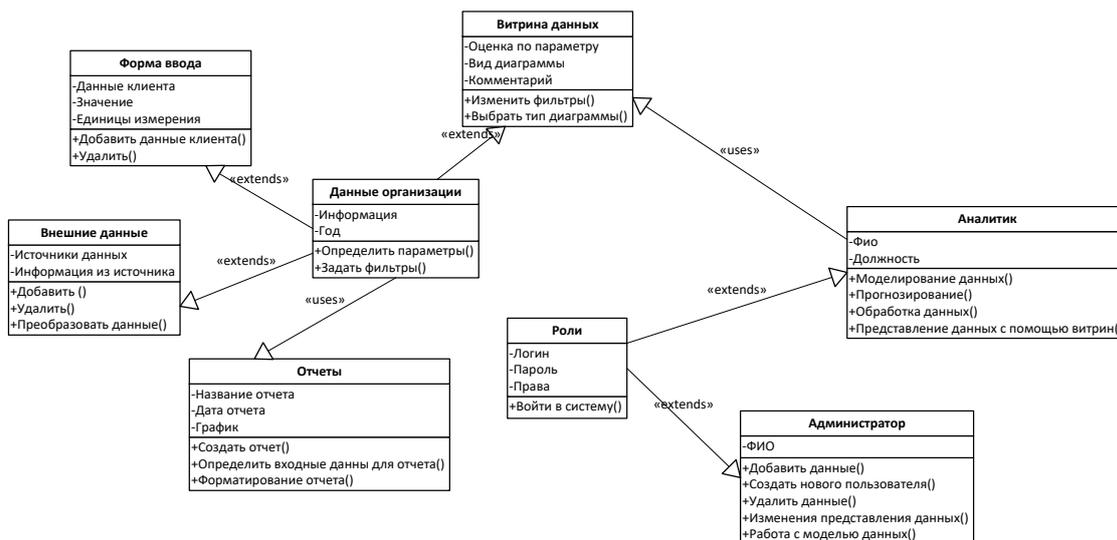


Рисунок 21 – Диаграмма классов

Администратор – имеет доступ ко всем элементам программы.

Аналитик – имеет возможность поставить оценку оперативной информации, т.е. работать с первичными источниками информации по клиенту.

Далее рассмотрим на примере оценки кредитного риска компании при работе с определенным клиентом, как проходит процесс сбора данных, построения экспертных правил и анализ и оценка с помощью витрины данных.

Сбор данных. Создание витрины, которая содержит информацию о клиенте (рисунок 22).

```

import pandas as pd
from sqlalchemy import create_engine

# Подключение к базе данных и загрузка витрины данных о клиентах
engine = create_engine('sqlite:///data_mart.db')
customer_df = pd.read_sql('customer_data_mart', engine)

# Проверка структуры данных витрины
print(customer_df.head())
  
```

Рисунок 22 – Создание витрины данных

Построение экспертных правил. В примере по определению кредитного риска компании при работе с определенным клиентом, важно учитывать следующие значимые факторы (рисунок 23):

- возраст клиента;
- доходы клиента;
- кредитная история;
- тип занятости клиента.

```
# Определение весов и правил для экспертной оценки
weights = {
    'age': 0.2,
    'income': 0.3,
    'credit_history': 0.4,
    'employment_type': 0.1
}

def score_age(age):
    if age < 25:
        return 0.1
    elif 25 <= age <= 40:
        return 0.4
    else:
        return 0.5

def score_income(income):
    if income < 50000:
        return 0.2
    elif 50000 <= income <= 100000:
        return 0.5
    else:
        return 0.7

def score_credit_history(credit_score):
    return credit_score / 100 # Нормализация кредитного балла (от 0 до 1)

def score_employment(employment_type):
    return 0.5 if employment_type == "постоянная" else 0.1
```

Рисунок 23 – Определение значимых параметров

Витрина данных содержит очищенные данные по этим критериям для каждого клиента (рисунок 24).

```

# Функция для расчета итоговой экспертной оценки для каждого клиента
def calculate_risk_score(row):
    return (
        weights['age'] * score_age(row['age']) +
        weights['income'] * score_income(row['income']) +
        weights['credit_history'] * score_credit_history(row['credit_score']) +
        weights['employment_type'] * score_employment(row['employment_type'])
    )

# Добавление колонки с расчетом риска в витрину данных
customer_df['risk_score'] = customer_df.apply(calculate_risk_score, axis=1)

# Просмотр данных с учетом оценки риска
print(customer_df[['customer_id', 'risk_score']].head())

```

Рисунок 24 – Определение риска по каждому клиенту

Анализ и оценка. Основное преимущество использования витрин данных, возможность визуализации полученных данных с помощью различных видов графиков.

В Python можно строить графики и визуализировать данные с помощью библиотек (Matplotlib и Seaborn). Данные библиотеки поддерживают простые графики, такие как линейные и столбчатые, а также работают с тепловыми картами, графиком плотности распределения и т.д.

Для наглядного представления уровня риска можно построить гистограмму, показывающую распределение рисков среди клиентов (рисунок 25).

```

import matplotlib.pyplot as plt
import seaborn as sns

# Построение гистограммы для визуализации риска
plt.figure(figsize=(10, 6))
sns.histplot(customer_df['risk_score'], bins=20, kde=True)
plt.title("Распределение кредитного риска среди клиентов")
plt.xlabel("Оценка риска")
plt.ylabel("Количество клиентов")
plt.grid()
plt.show()

```

Рисунок 25 – Визуализация данных

Интерпретация результатов по определению кредитного риска компании при работе с определенным клиентом (рисунок 26).

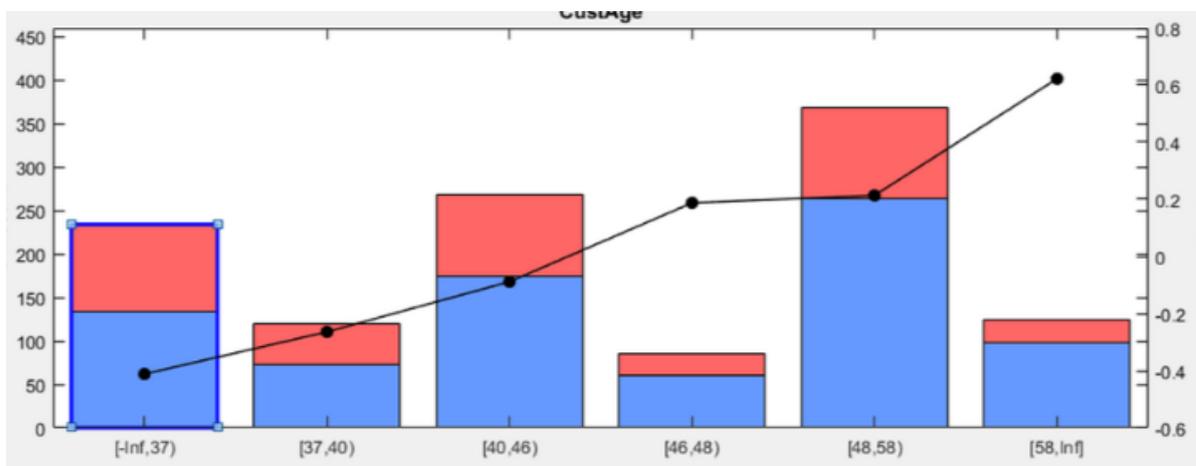


Рисунок 26 – Результаты анализа

Клиенты компании с высокими оценками риска могут получить отказ в предоставлении отсрочки оплаты заказов или получить выгодные предложения по работе на специальных условиях.

Клиенты с низким риском могут получить более выгодные условия по работе с компанией.

Используя витрины данных в методах и моделях управления эффективностью системы аналитической обработки корпоративной информации, компании могут автоматизировать процесс оценки и принимать решения на основе структурированных данных, что повышает эффективность и снижает вероятность ошибки.

Для оценки эффективности предложенных решений и моделей, проведено исследование временных и количественных характеристик до внедрения и после, а также проведен анализ характеристик получившихся витрин данных.

Вначале были проанализированы источники информации, из которых обрабатывается корпоративные данные. В ходе работы над магистерской диссертацией было предложены механизмы взаимодействия с распределенными базами данных, источники которых находятся вне организации, но в связи с ограничениями на количество запросов и не возможностью получать данные в реальном времени, в ходе практического эксперимента от таких данных пришлось отказаться.

3.2 Оценка эффективности предложенных решения на принятие решений

В ходе практической апробации результатов работы и для получения данных в режиме реального времени, было принято решение, что данные полученные ранее из сторонних источников о деятельности компании (банковские выписки, налоговые отчеты и т.д.) будут размещены в таблицу базу данных на локальном сервере. Данное решение позволило получать оперативный доступ к данным, накапливать данные для ретроспективного анализа и прогнозирования и максимально эффективно использовать витрины данных для аналитической обработки корпоративной информации.

В таблице 6 показано количество источников информации, используемых до передоложенных решений.

Таблица 6 – Количество источников информации о деятельности компании

Структура источников информации о деятельности компании		
	ДО внедрения количество баз данных	ПОСЛЕ внедрения количество баз данных/таблиц
Общее количество источников корпоративной информации	1	5
Внутренние данные компании	0	1
Данные из Федеральной налоговой службы (ФНС России)	0	1
Данные с портала Госзакупки	0	1
Данные из банка	0	1

В ходе практической апробации и интеграции данных получилась следующая структура баз данных организации, теперь все данные со сторонних источников о деятельности компании собираются в распределенную базу данных, но находящуюся на сервере компании.

Анализирую полученные данные, можно сделать следующий вывод: количество источников информации не изменилось, была изменена структура хранения данных, необходимых для аналитической обработки, благодаря такой системе изменилась скорость получения информации за счет хранения данных на внутреннем сервере, а также появилась возможность оперативного доступа к ретроспективным данным.

Данное изменение структуры хранения данных отражает полностью предложенную модель по сбору корпоративной информации (рисунок 27).

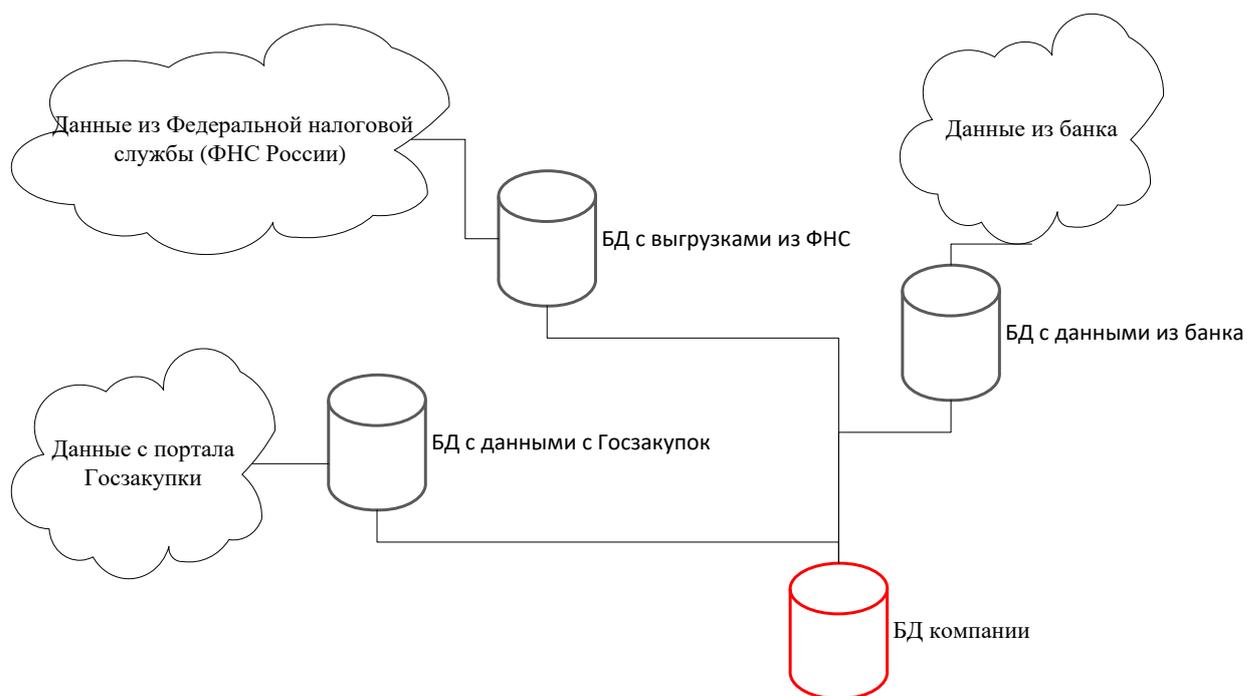


Рисунок 27 – Графическое представление структуры источников информации о деятельности компании

Проанализируем второй этап предложенной модели, а именно создание экспертных правил для анализа.

До внедрения нового решения аналитику компании необходимо было написать запрос к данным, хранящимся в базе данных компании, далее получить данные из сторонних источников и срок получение некоторых данных был от 3х рабочих дней, такой подход был приемлем для среднесрочного и долгосрочного прогнозирования деятельности компании.

При текущем состоянии компании возникла необходимость оперативного анализа данных, и предыдущие методы экспертного анализа предложено было заменить с помощью информационной системы, которая поддерживает графическое создание правил для анализа (позволяя выбирать источник данных, таблицы для анализа, поля для анализа и выбирать правила агрегирования и условия вывода данных) (рисунок 28).

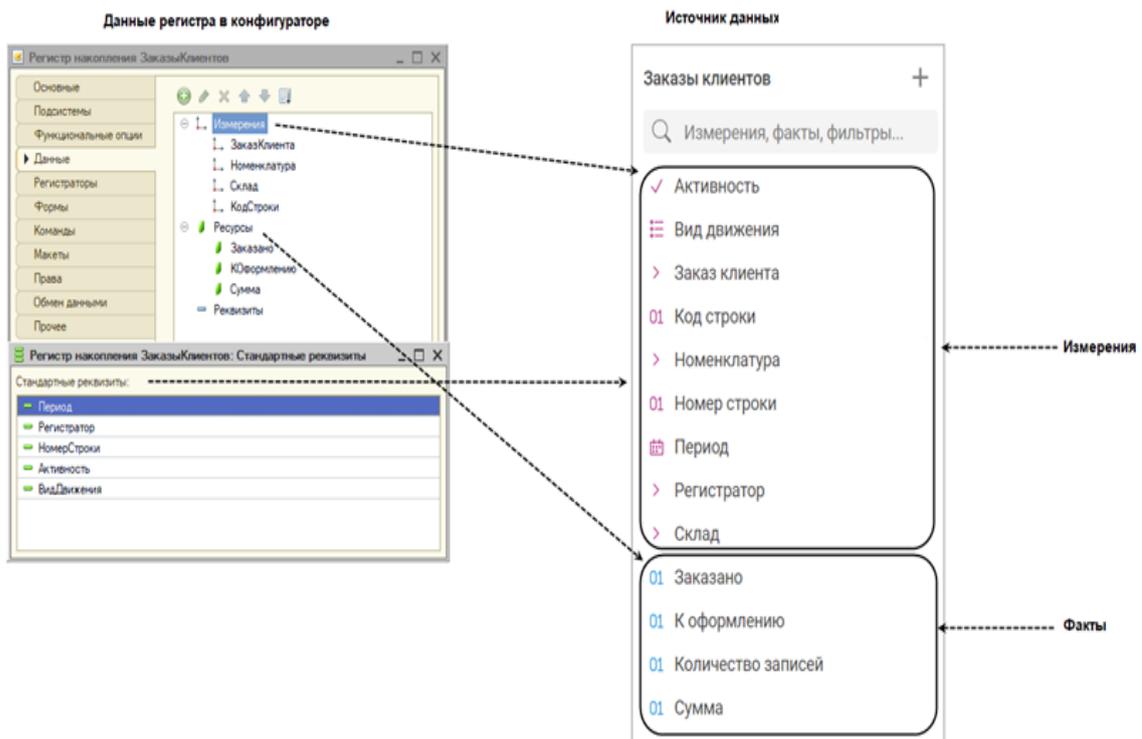


Рисунок 28 – Создание полей данных для витрины

Для удобства источники данных отличаются не только цветом своей роли, но и иконкой слева от имени поля, указывающей на тип поля.

Выбирая с помощью конструктора данные для диаграммы, аналитик каждый раз создает новые экспертные правила для анализа, тем самым рассматривает корпоративную информацию под разным углом (рисунок 29).

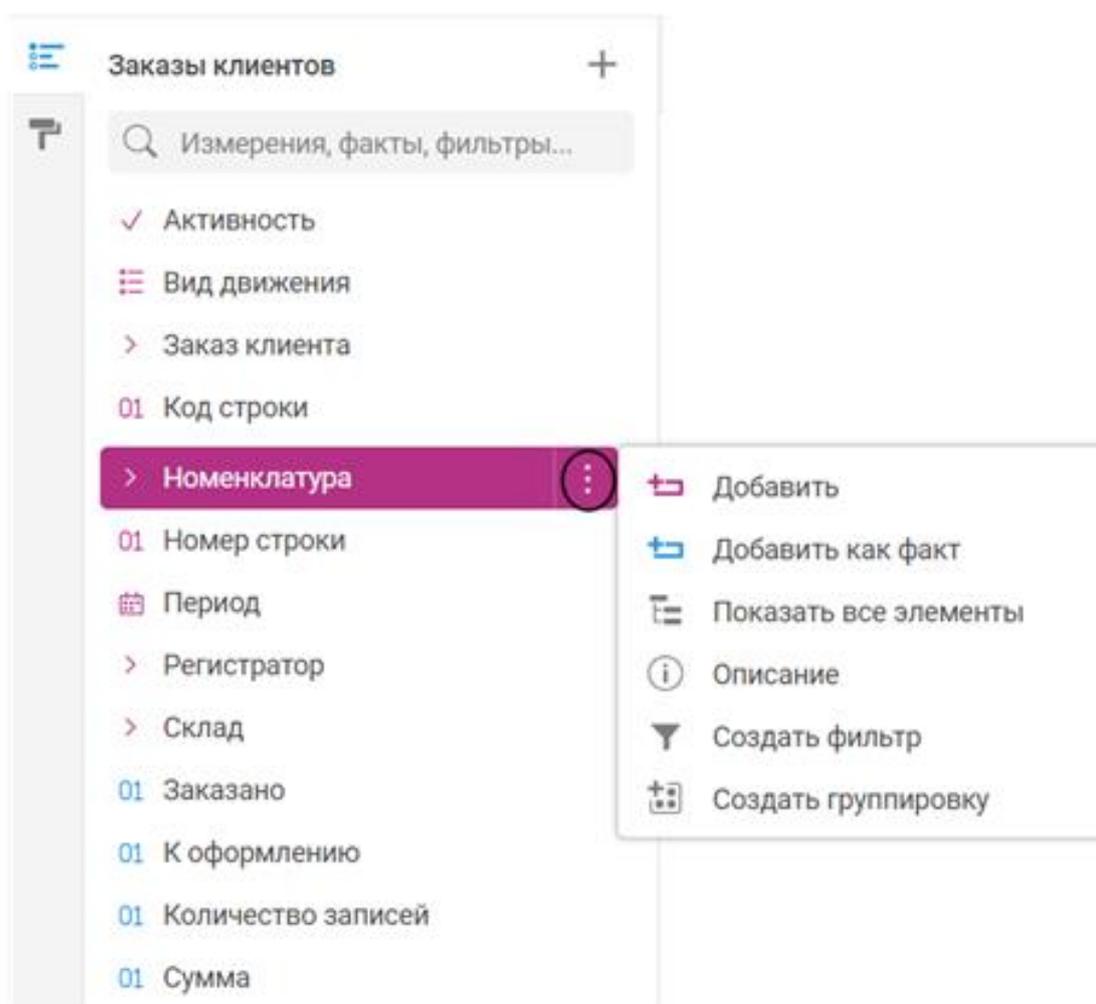


Рисунок 29 – Выбор полей таблицы

Важным этапом работы с правилами – это создание фильтров. Рассмотрим, как проходит процесс, создания фильтров.

Как было описано выше, главное достоинство – это создание ретроспективного анализа данных, т.е. работа с различными временными периодами или возможность работы с фильтрами дат (рисунок 30).

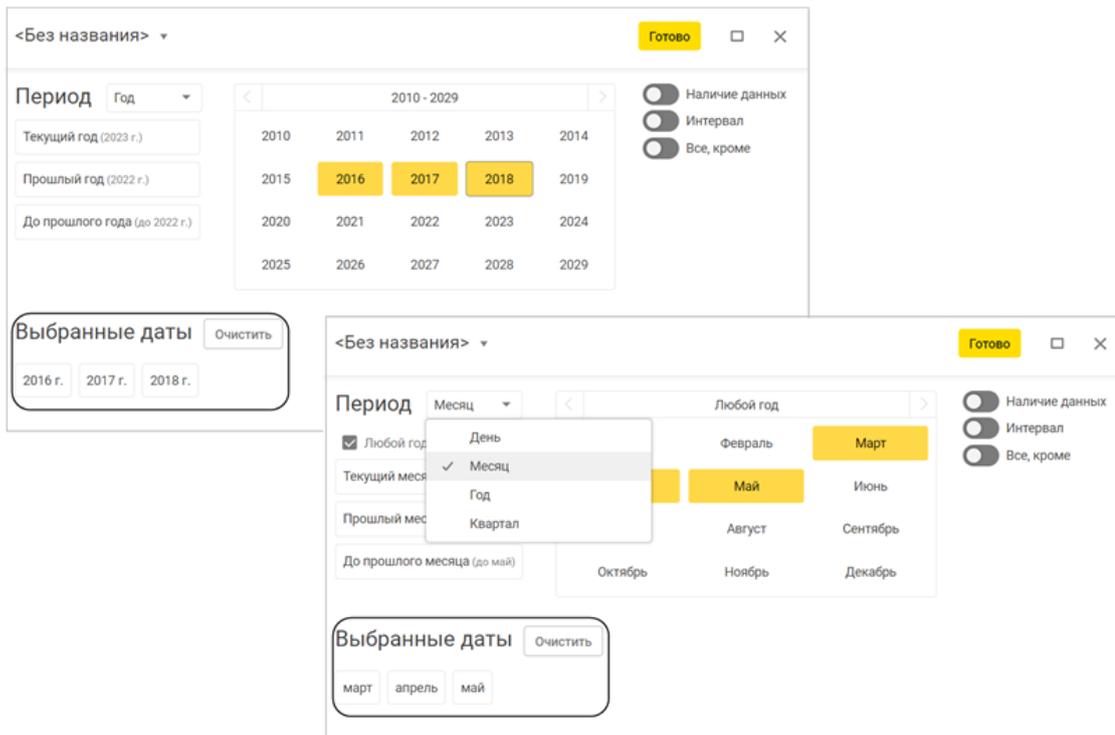


Рисунок 30 – Создание фильтра на даты

При создании правил, возможно, создавать операции необходимые для отображения данных (рисунок 31).

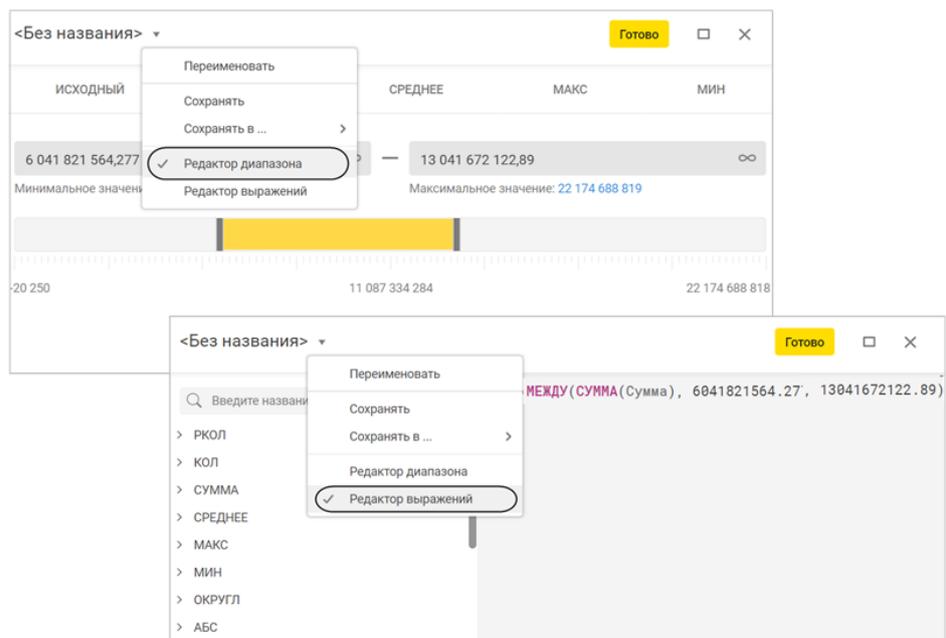


Рисунок 31 – Создание операций на данные

Анализирую работу этого этапа по созданию экспертных правил, можно привести следующие результаты (таблица 7).

Таблица 7 – Принятие решений по клиентам

Создание экспертных правил		
	ДО внедрения моделей	ПОСЛЕ внедрения моделей
Время на создание запроса, ч	3	0,5
Процент запросов, используемых для анализа, %	67%	75%
Процент запросов, которые необходимо переделать, %	85%	30%

Полученные данные представлены в виде графика, для наглядности результата и визуализация решений по разрабатываемым аналитиком запросам представлена на рисунках 32 и 33.



Рисунок 32 – Графическое представление времени на создания запроса

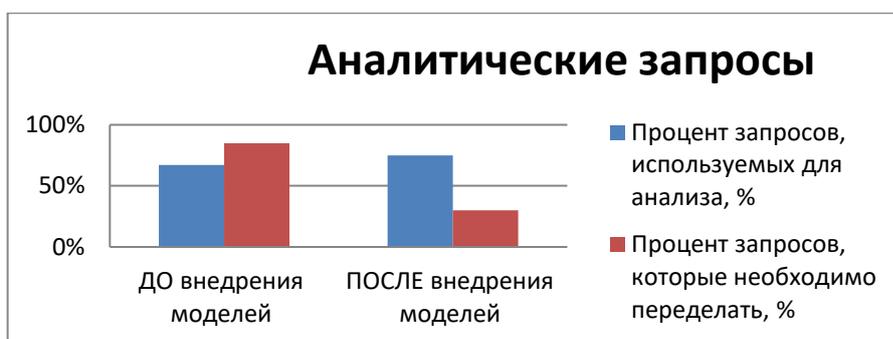


Рисунок 33 – Графическое представление эффективности второго этапа

Как видно из приведенного анализа после внедрения предложенных моделей, за счет наличия в системе различных фильтров, время создания запроса уменьшается, процент запросов, которые в дальнейшем используются при оперативном анализе выше.

На третьем этапе предложенной модели является возможность представления данных с помощью витрин данных. Количество доступного отображения данных после внедрения равняется тридцати видам диаграмм.

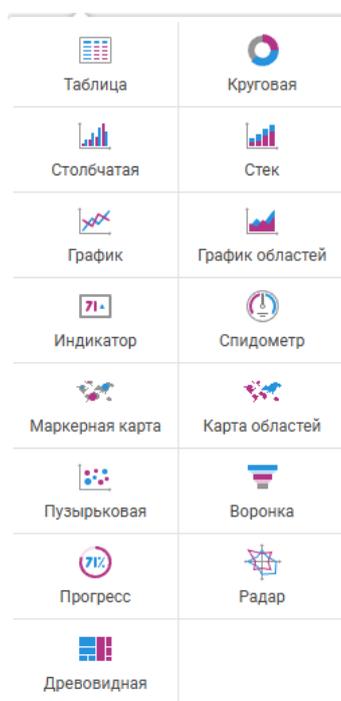


Рисунок 34 – Типы отображения данных на третьем этапе

Приведем несколько примеров отображения (рисунок 34-35)

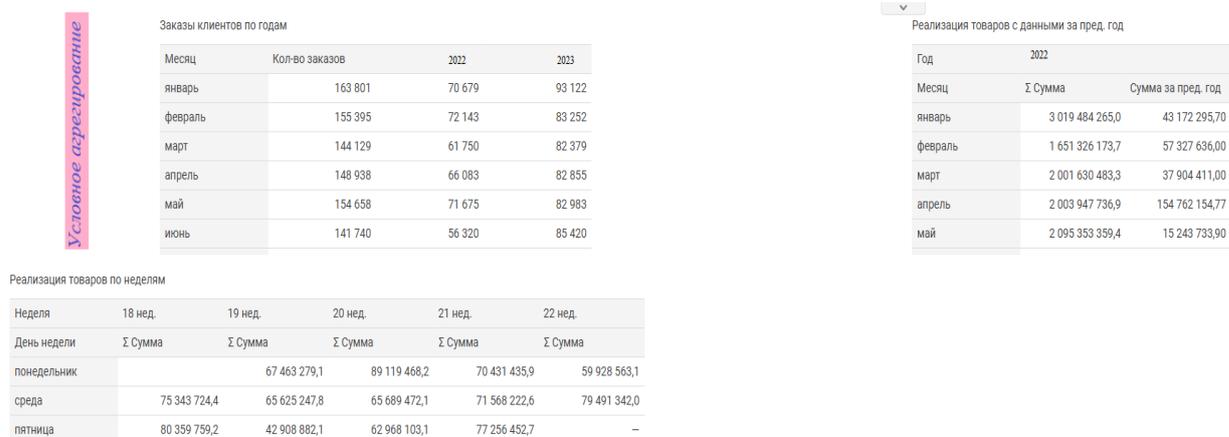


Рисунок 35 – Типы отображения данных: в виде таблицы

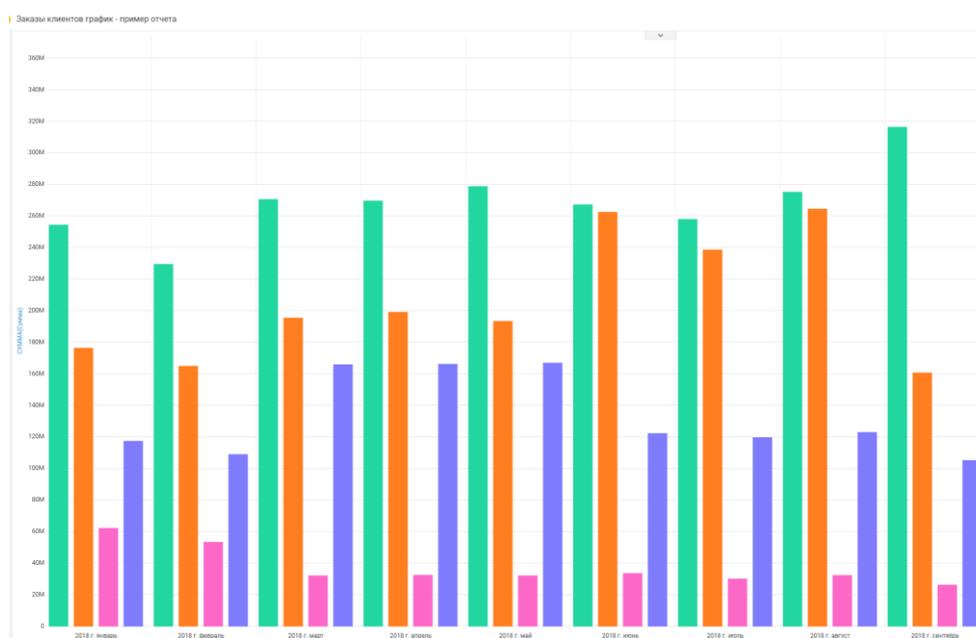


Рисунок 36 – Типы отображения данных в виде столбчатой диаграммы

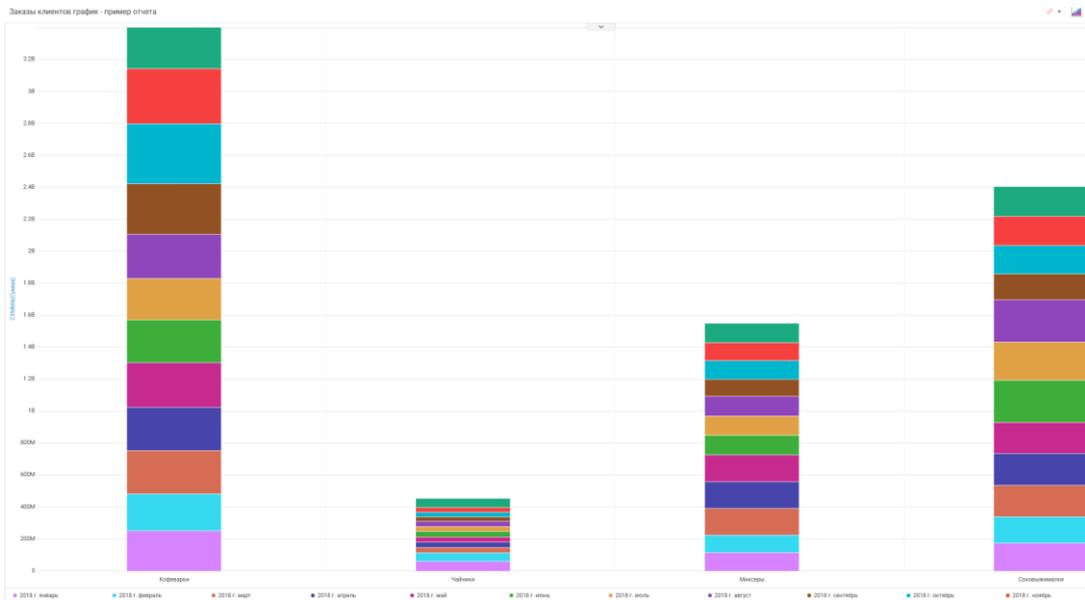


Рисунок 37 – Типы отображения данных: в виде стека

При наличии в данных географической информации имеется возможность отображения информации на картах (рисунок 38).

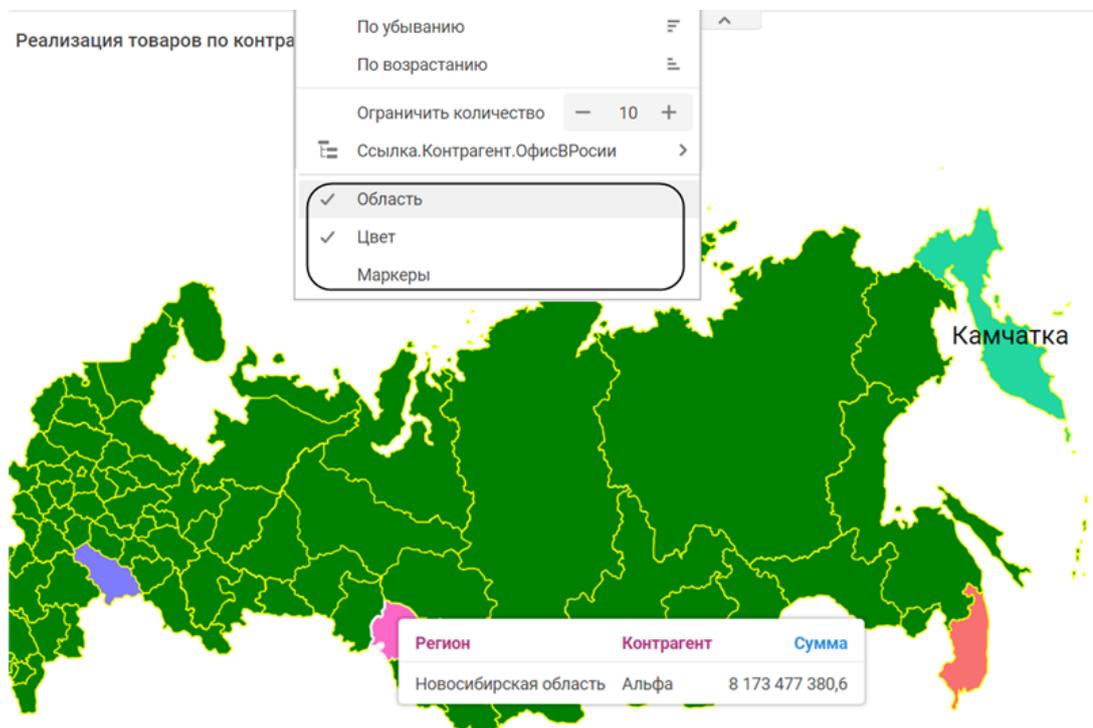


Рисунок 38 – Типы отображения данных: в виде карты

Выводы по главе 3

«В результате проделанной работы было выявлено, что предложенные методы и модели управления эффективностью системы аналитической обработки корпоративной информации способствуют повышению эффективности работы аналитиков организации.

Они также позволяют компании получить качественно новый подход к принятию решений.

Проведена апробация и оценка эффективности проектных решений по качественным и количественным характеристикам.

Результаты показали, что использование витрин данных для оценки корпоративных данных значительно повышает эффективность работы аналитиков компании» [34].

Заключение

«В ходе выполнения магистерской диссертации по теме «Методы и модели управления эффективностью системы аналитической обработки корпоративной информации» получены выводы по всем задачам, которые были поставлены в рамках работы.

В первой главе «Теоретический раздел. Основы управления бизнес-информацией» был проведён анализ современного состояния исследований в области управления эффективностью системы аналитической обработки корпоративной информации. Этот анализ показал, что данная тема является актуальной и требует более глубокого изучения» [34].

В этой главе были рассмотрены основополагающие вопросы аналитической обработки корпоративной информации, а также показано практическое применение и характеристика каждой метрики процесса. Были определены основные принципы и методы аналитической обработки данных, которые позволяют повысить эффективность работы компании.

Кроме того, был проведён анализ современных программных средств в области аналитической обработки корпоративной информации. Особое внимание было уделено российским программным продуктам, их особенностям и преимуществам. Это позволило определить наиболее подходящие инструменты для решения задач аналитической обработки корпоративных данных.

Также была выполнена апробация и оценка эффективности проектных решений по качественным и количественным характеристикам. Результаты показали, что использование витрин данных для оценки корпоративных данных значительно повышает эффективность работы аналитиков компании.

Во второй главе магистерской работы «Аналитический раздел. Метод и модель управления эффективностью системы аналитической обработки корпоративной информации» рассмотрены методы и модели управления

эффективностью системы аналитической обработки корпоративной информации.

В рамках исследования были определены источники данных для каждого бизнес-процесса модели, а также выявлены зависимости между каждым информационным потоком и метрикой, которая в дальнейшем используется для оценки эффективности процесса.

В третьей главе «Практический раздел. Разработка приложения с использованием предложенных методов и моделей управления эффективностью системы аналитической обработки корпоративной информации» Выполнена апробация и оценка эффективности проектных решений по качественным и количественным характеристикам, которые показали увеличение эффективности работы аналитиков компании при оценке корпоративных данных, используя витрины данных

«Для оценки эффективности методов и моделей управления эффективностью системы аналитической обработки корпоративной информации, проведено исследования временных характеристик до внедрения моделей в организацию и после, а также проведен анализ работы внедренных витрин данных и отображения с помощью различного рода диаграмм» [34].

Гипотеза исследования подтверждена.

Список используемой литературы и используемых источников

1. Витрина данных. Проектирование [Электронный ресурс] – Режим доступа: <https://habr.com/ru/articles/650237/> (дата обращения 13.06.2024.)
2. Витрина данных [Электронный ресурс] – Режим доступа: <https://www.oracle.com/cis/autonomous-database/what-is-data-mart/> (дата обращения 13.06.2024.)
3. Управление Эффективностью Предприятия [Электронный ресурс] – Режим доступа: <https://optimacros.com/blog-ru/enterprise-performance-management/> (дата обращения 13.06.2024.)
4. Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Классификация и снижение размерности. – М.: Финансы и статистика, 1989. – 607 с.
5. Барсегян, А.А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP / А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод. СПб.: БХВ–Петербург, 2007. – 384 с.
6. Богданов А. Б., Борисова И. А. и др. Интеллектуальный анализ спектральных данных // Автометрия, 2009, № 1. –С. 92–101.
7. Борисова И. А., Дюбанов В. В., Загоруйко Н. Г. и Кутненко О. А. Сходство и компактность // Всеросс. конф. «Математические методы распознавания образов» (ММРО–14), г. Суздаль, 21–23 сентября 2023 г. – С. 1–4.
8. Воронцов К. В., Колосков А. О. Профили компактности и выделение опорных объектов в метрических алгоритмах классификации // Искусственный Интеллект. – 2006. –С. 30–33.
9. Дайзенрот Марк Питер, Чен Сунь Он, Альдо Фейзал А. Математика в машинном обучении. – СПб.: Питер, 2024. – 512 с.: ил. – (Серия «Для профессионалов»)

10. Донской В. И. Алгоритмы поиска аналогий в булевых таблицах эмпирических данных // В. И. Донской, А. В. Ильченко. – Искусственный интеллект. – 2002. – №2. – С. 99–107.
11. Дорофеев Ю. А. Структурно–классификационные методы анализа и прогнозирования в системах управления // Таврический вестник информатики и математики. – 2008. – №1. – С.166–170.
12. Дубров А.М., Мхитрян В.С., Трошин Л.И. Многомерные статистические методы: Учебник. – М.: Финансы и статистика, 2020.
13. Ежов А.А., Шумский С.А. Нейрокомпьютинг и его применения в экономике и бизнесе. Оптимизация с помощью сети Кохонена.
14. Журавлев Ю.И., Рязанов В.В., Сенько О.В. Распознавание. Математические методы. Программная система. Практические применения. – М.: Фазис, 2005. –159 с.
15. Кривенко М. П. Обучаемая классификация данных с учётом анализа главных компонент // Информатика и её применение. 2018. Т 12, №3. – 59–61.
16. Люгер Д.Ф. Искусственный интеллект: стратегии и методы решения сложных проблем. – М.: Издательский дом «Вильямс», 2003 – 864 с.
17. Малла, С. Вэйвлеты в обработке сигналов [Текст]/С. Малла. – М.: Мир, 2005. – 671 с.
18. Михнев И.П. Большие данные (Big Data) и новые технологии будущего для обработки глобальной информации: сборник трудов конференции. / И.П. Михнев, А.А. Новикова, М.К. Петросян // Научные исследования и современное образование : материалы II Междунар. науч.-практ. конф. (Чебоксары, 26 март 2018 г.) / редкол.: О.Н. Широков [и др.] – Чебоксары: ЦНС «Интерактив плюс», 2018. – С. 235-239. – ISBN 978-5-6040859-4-3. doi:10.21661/r-470260
19. Михнев И.П. Информационная безопасность в современном экономическом образовании // Международный журнал прикладных и фундаментальных исследований. – 2019. – №4. – С. 111–113.

20. Михнев И.П. Технологии Big Data и их применение в сфере современного высшего образования / И.П. Михнев, А.Д. Челнокова, А.Д. Реут // Развитие современного образования: от теории к практике: Материалы IV Междунар. науч.-практ. конф. (Чебоксары, 19 март 2018 г.) / Редкол.: О.Н. Широков [и др.]. – Чебоксары: ЦНС «Интерактив плюс», 2018.
21. Моррисон А. Большие Данные: как извлечь из них информацию. Технологический прогноз / Ежеквартальный журнал. – 2010. – №3. – С. 22–29.
22. Нивен П.Р. Сбалансированная система показателей: шаг за шагом. Максимальное повышение эффективности и закрепление полученных результатов. – М.: Баланс–Клуб, 2004. – 328 с.
23. Ногин В.Д. Сужение множества Парето: аксиоматический подход. – М.: ФИЗМАТЛИТ, 2016, – 272 с. ISBN 978–5–9221–1638–1.
24. Петровский А.Б. Пространства множеств и мультимножеств. – М.: УРСС, 2003. – 248 с. [http://www.raai.org/about/persons/petrovsky/pages/Petrovsky 2003.pdf](http://www.raai.org/about/persons/petrovsky/pages/Petrovsky%2003.pdf)
25. Подиновский В.В., Ногин В.Д. Парето–оптимальные решения многокритериальных задач. – М.: ФИЗМАТЛИТ, 2007, – 256 с. ISBN 978–5–9221–0821–6.
26. Пойа Дж. Математика и правдоподобные рассуждения // Дж. Пойа. – Москва : Наука, 1975. – 464 с.
27. Толмачев И.Л., Хачумов М.В. Бинарная классификация на основе варьирования размерности пространства признаков и выбора эффективной метрики. – Искусственный интеллект и принятие решений, № 2, 2010, с.3–10.
28. Фрэнкс Б. Революция в аналитике. Как в эпоху Big Data улучшить ваш бизнес с помощью операционной аналитики / Б. Фрэнкс. - М.: Альпина Диджитал, 2014. - 370 с.
29. Харламов А.А. Нейросетевой подход к интегрированному представлению и обработке информации в интеллектуальных системах. – Автореферат докторской диссертации. – М.: 2009. – 31 с.

30. Хачумов М.В. Задача кластеризации текстовых документов. Информационные технологии и вычислительные системы, № 2, 2010, с.42–49.
31. Хачумов М.В. Модели представления данных в задачах распознавания образов и кластеризации. – Тезисы докладов Всероссийской конференции с международным участием «Информационно–телекоммуникационные технологии и математическое моделирование высокотехнологических систем» (18–22 апреля 2011 Москва). – М.: РУДН, 2011, с. 146–148.
32. Черняк Л. Большие Данные – новая теория и практика // Открытые системы. СУБД. – 2021. – №10. – С. 36–41.
33. Лымарь В.Г. Модель управления эффективностью системы аналитической обработки корпоративной информации// Международный научный журнал «Вестник науки» №7(76)– 2024. - С. 522–529.
34. Лымарь В.Г. Использование витрин данных для аналитической обработки корпоративной информации// Международный научный журнал «Вестник науки» №10(79) т.3 – 2024. - С. 843–849.
35. Addison, P.S. Illustrated wavelet transform handbook. Introductory Theory and Applications in Science, Engineering, Medicine and Finance / Paul S. Addison. – Bristol: Institute of Physics Publishing, 2002. – 400 p.
36. Arunalatha, G. A Survey of Wavelet Compression Methods / G. Arunalatha, S. Sarathraj, K.Manodurga // International Journal of Scientific & Engineering Research. – 2013. – Vol. 4. Is. 3. – P. 1–4.
37. Azure Cosmos DB. Мультимодельная, глобально распределенная служба базы данных для любого масштаба [Электронный ресурс] – Режим доступа: <https://azure.microsoft.com/ru-ru/services/cosmos-db/> (дата обращения 13.06.2024.).
38. Big Data Analytics Landscape 2019. [Электронный ресурс] – Режим доступа: <https://www.learnbigdatatools.com/big-data-analytics-landscape-2019> (дата обращения 13.06.2024)

39. Ergen, B. Signal and Image Denoising Using Wavelet Transform / Burhan Ergen // Signal and Image Denoising Using Wavelet Transform, Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology. InTech, 2012. – P. 495–514.
40. Han, J. Data Mining: Concepts and Techniques / J. Han, M. Kamber. – Morgan Kaufmann Publishers, 2000. – 550 p. https://www.prodsysde/Service/Downloads/bin/DMC2009_Evgebnisliste.pdf
41. Kaplan R.S., Norton D.P. Using the Balanced Scorecard as a Strategic Management System // Harvard Business Review, 1996, January/February, p. 76.
42. Kira, K. & Rendell, L. A., The. Feature Selection Problem: Traditional Methods and a New Algorithm // Proceedings of the Tenth National Conf. Artificial Intelligence (AAAI-92) – 1992. – Pp. 129–134
43. Osipov G. Strategies for Stabilization Behaviour of Intelligent Dynamic Systems. – Proc. of 20th European Meeting on Cybernetics and Systems, Vienna, 2010, pp. 195–197.
44. Remya, S. Wavelet Based Compression Techniques: A Survey / S. Remya, V.A. Dilshad Rasheed // Advances in Communication, Network, and Computing.– 2012. – Vol.108. P. 394 397
45. Shahabi, C. 2D TSA–tree: a wavelet–based approach to improve the efficiency of multi–level spatial data mining / C. Shahabi, S. Chung; M.Safar, G. Hajj// Proceedings of 13th Int. Conf. ofScientificandStatisticalDatabaseManagement. 2001. – P. 59–68.
46. Shahabi,C. TSA–tree: A Wavelet–Based Approach to Improve the Efficiency of Multi–Level Surprise and Trend Queries / C. Shahabi, X. Tian, W. Zhao // Proceedings of 12th Int. Conf. ofScientificandStatisticalDatabaseManagement, 2000. P. 55–68.
47. Sheikholeslami G. WaveCluster: A multi–resolution clustering approach for very large spatial databases / G. Sheikholeslami, S. Chatterjee, A. Zhang // Proc. 24th Int. Conf. VeryLargeDataBases. – 1998. – P. 428 439.

48. Tripathy, A. Dimensionality Reduction of Data Warehouse Using Wavelet Transformation: An Enhanced Approach for Business Process / A.Tripathy, D.Kaberi, S. Tripti// Computer Networks and Information Technologies. – 2011. – Vol. 142. P. 523–525.

49. Witten,I.H. Data Mining: Practical Machine Learning Tools and Techniques / I.H. Witten, E.Frank, M.A. Hall.–Morgan Kaufmann, 2011. – 664.

50. Zagoruiko, N.G., Borisova, I.A., Dyubanov, V.V., Kutnenko, O.A.: Methods of Recognition Based on the Function of Rival Similarity // Pattern Recognition and Image Analysis. – 2008. – V.18. – Pp. 1–6.