

УДК 004.9

1.8. Методика управления корпоративными данными для применения искусственного интеллекта в торгово-промышленном предприятии

Никитчук С.С.

Аспирант РЭУ им. Г.В. Плеханова

В статье представлена авторская методика управления корпоративными данными для применения искусственного интеллекта (AI-driven) в торгово-промышленных предприятиях. Актуальность исследования обусловлена ростом объемов и разнородности данных, необходимостью интеграции с государственными информационными системами (ФГИС) и развитием платформенной экономики. В отличие от классических подходов к управлению данными, методика ориентирована на практическую реализацию и включает пять взаимосвязанных компонентов: протоколы и интеграции (data integration), обработка (моделирование) данных (data modelling), оркестрацию сервисов и прослеживаемость (data orchestration and lineage), управление и безопасность данных (data governance and security), а также слой бизнес-аналитики и моделей искусственного интеллекта (self-service analytics и artificial intelligence).

В обзоре литературы рассмотрены современные международные подходы (DAMA-DMBOK, Gartner Data Fabric/Data Mesh, DataOps/MLOps) и российская специфика цифровой экономики, включающая национальную модель бизнеса, проект цифрового рубля и развитие ФГИС. Согласно анализу, ключевым вызовом становится необходимость повышения вычислительных ресурсов и усложнение технологий обработки данных при снижении их структурированности.

Практическая апробация методики проведена в ряде крупных проектов цифровой трансформации для торгово-промышленных компаний. Результаты показали, что использование методики позволяет повысить качество данных, обеспечить прозрачность сквозных процессов и снизить проектные риски. В перспективе развитие методики связано с углублением процессов управления данными, автоматизацией интеграций с государственными системами и применением моделей искусственного интеллекта для построения интеллектуальных и адаптивных цифровых архитектур.

Введение

В предыдущих исследованиях автора была предложена концепция фреймворка цифровой трансформации, где данные и искусственный интеллект занимают ключевое место в развитии организации [1]. В основе фреймворка лежит системное соединение принципов организационного развития и цифровых доменов, среди которых именно управление данными становится центральным элементом, обеспечивающим устойчивость цифровизации. В дальнейшем, в статье о целевой архитектуре цифрового торгово-промышленного предприятия, было сформировано прикладное направление работы с процессами, их цифровизацией и увязкой с архитектурой ИТ-систем [2].

Ранее предприятия строили монолитные системы, где каждая функция была тесно связана с остальными, а изменения становились всё сложнее и дороже. Перевнедрения ERP новых версий позволял частично снизить риски, но ядро оставалось монолитным. Со временем компании осознали, что именно данные становятся ключевым ИТ-активом, и начали внедрять новые сервисы вокруг ERP (RPA, OCR, EDI, BI и др.). Это потребовало переосмысления архитектуры и перехода к микросервисному подходу, в котором автоматизация бизнес-функций строится как набор сервисов, интегрированных с корпоративным ERP. Такой сдвиг позволил повысить гибкость, ускорить обновления и обеспечить развитие сквозных процессов в условиях растущих требований цифровой экономики.



Рисунок 1 – Виды архитектур

Такой подход позволяет не только согласовать стратегию и технологии, но и создавать условия для накопления больших объемов корпоративных данных, которые в перспективе становятся базой для

применения методов управления на основе искусственного интеллекта (AI-driven). Таким образом, результаты предшествующих исследований заложили основу для перехода от стратегического проектирования к практическому управлению данными с использованием технологий искусственного интеллекта (ИИ).

Современные торгово-промышленные предприятия развиваются в условиях комплексных вызовов цифровой экономики. Ключевую роль в формировании новой бизнес-среды играют государственные цифровые инициативы, включающие Налоговый мониторинг, Системы операторов фискальных данных, Федеральные государственные информационные системы. Эти инструменты задают новые стандарты полноты, качества и актуальности корпоративных данных.

Дополнительное влияние оказывает переход к платформенной экономике, предполагающий обмен не только транзакционными, но и аналитическими данными между предприятиями. В этих условиях усиливаются требования к информационной безопасности: компании обязаны учитывать, как защиту персональных, так и коммерчески значимых данных в соответствии с национальными законами и корпоративными регламентами.

Согласно прогнозам Gartner, будущее управление данными будет определяться концепциями Фабрика данных (Data Fabric) и Сетка данных (Data Mesh), внедрением техники обработки данных (DataOps) и операций машинного обучения (MLOps), а также активным использованием генеративного ИИ для автоматизации процессов. Всё это формирует необходимость поиска новых методик управления корпоративными данными, в которых применения подходов управления на основе искусственного интеллекта (AI-driven) становится не просто возможностью, а необходимостью для устойчивого развития предприятия.

Научная новизна предлагаемого исследования заключается в разработке методики управления корпоративными данными, которая интегрирует возможности искусственного интеллекта в практику цифровой трансформации торгово-промышленных предприятий. В отличие от существующих подходов, ориентированных преимущественно на технологическую инфраструктуру или отдельные процессы, данная методика опирается на целостное представление об архитектуре данных, включающее их сбор, хранение, обработку и применение для поддержки управленческих решений. Новизна также состоит в использовании механизмов искусственного интеллекта (AI-driven), позволяющих автоматизировать задачи очистки, категоризации и анализа больших массивов информации, а также в создании условий для адаптивного управления в реальном времени. Предлагаемый подход обеспечивает не только повышение качества и достоверности данных, но и расширяет возможности их стратегического применения – от прогнозирования и оптимизации бизнес-процессов до формирования новых моделей взаимодействия с партнёрами и государственными платформами.

1. Цифровая экономика

В современной российской экономике цифровизация рассматривается не только как технологический процесс, но и как инструмент формирования новой модели взаимодействия государства и бизнеса. Центральным элементом этой модели становится работа с данными, которые предприятия обязаны предоставлять в государственные системы, а государство использует для управления экономикой и регулирования рынков.

На федеральном уровне сформулирована концепция национальной модели бизнеса, включающей систему ключевых показателей эффективности (КПЭ) для федерации и регионов. Эта модель нацелена на повышение прозрачности хозяйственной деятельности и унификацию принципов оценки эффективности на разных уровнях управления [3].

Для сбора информации и ее стандартизации государственные органы активно развивают федеральные государственные информационные системы (ФГИС). Наиболее известные примеры включают АИС «Налог-3» для администрирования налогов [4], «Честный знак» для маркировки и прослеживаемости товаров [5], а также систему операторов фискальных данных (ОФД), предназначенную для передачи кассовых чеков в ФНС [6]. Наряду с этим в ряде отраслей действуют специализированные ФГИС, такие как «ОРД» для прослеживаемости древесины [7], «Зерно» для контроля качества зерна [8], «Меркурий» для ветеринарного надзора [9] и «ЕГАИС» для оборота алкогольной продукции [10].

В последние годы значимость корректного и полного обмена данными с ФГИСами напрямую увязывается с доступом к получению субсидий и возможностью предприятий участвовать в программах господдержки (СПИК, СЗПК, Индустриальные парки, Особые экономические зоны, др.). Например, при рассмотрении заявок на специальные инвестиционные контракты проверяется налоговая дисциплина компаний, основанная на данных из АИС «Налог-3» и операторов фискальных данных (ОФД) [11].

Параллельно разрабатывается проект цифрового рубля [12], который рассматривается как один из ключевых элементов новой финансово-экономической инфраструктуры России. Его внедрение направлено на то, чтобы объединить данные о хозяйственной деятельности предприятий с информацией о фактических расчетах, формируя сквозную систему прослеживаемости денежных потоков. Таким образом, цифровой рубль становится инструментом, объединяющим финансовые и производственные данные в едином цифровом контуре экономики.

Интеграционные платформы становятся ключевым элементом цифровой трансформации предприятий, поскольку позволяют объединять не только внутренние ИТ-системы (ERP, CRM, MES), но и обеспечивать интеграцию с внешними государственными сервисами – ФГИС, системами маркировки, отраслевыми реестрами и платёжной инфраструктурой. Их применение упрощает построение архитектуры

взаимодействия между бизнесом и государством, сокращает сроки внедрения новых цифровых решений и повышает управляемость данными в режиме реального времени [13].

Такая интеграция делает государство и бизнес всё более взаимосвязанными: качество передаваемых данных напрямую влияет не только на точность налоговых расчетов и макроэкономическую статистику, но и на возможность компаний участвовать в программах государственной поддержки. Таким образом, управление корпоративными данными приобретает стратегическое значение, определяя не только эффективность внутреннего управления, но и устойчивость взаимодействия с государственными институтами цифровой экономики.

2. Платформенная экономика

Платформенная экономика становится одним из ключевых направлений цифровой трансформации, радикально меняя способы взаимодействия предприятий, потребителей и государства. В её основе лежит идея объединения участников вокруг цифровых платформ, которые обеспечивают стандартизированный обмен данными и услугами. В современной практике можно выделить несколько типов таких платформ. Прежде всего это маркетплейсы (Ozon, Wildberries, AliExpress), которые позволяют предприятиям расширять каналы продаж и формировать новые модели дистрибуции [14].

Вторую группу составляют тендерные площадки и B2B-порталы, обеспечивающие электронные закупки, взаимодействие с поставщиками и повышение прозрачности цепочек поставок [15].

Важную роль играют социальные сети (ВКонтакте, RuTube, мессенджер MAX), которые выступают не только каналами продвижения бренда, но и источниками данных о предпочтениях и поведении клиентов [16].

Наряду с этим активно развиваются отраслевые интеграционные платформы, ориентированные на создание сквозных цифровых процессов в рамках отрасли, от агропромышленных экосистем до промышленных кластеров [17]. Цели использования платформ разнообразны: повышение эффективности продаж, оптимизация закупок, поддержка бренда и коммуникаций, а также формирование новых бизнес-моделей, основанных на совместном использовании данных и сервисов.

Как отмечает Роджерс, цифровая трансформация требует не только внедрения новых технологий, но и фундаментального переосмысления логики бизнеса в условиях сетевых рынков [18]. В традиционной модели предприятия создавали ценность в рамках замкнутой цепочки – от разработки продукта до его продажи конечному потребителю. Однако в цифровую эпоху ценность всё чаще формируется во взаимодействии множества участников платформ: производителей, поставщиков, партнёров, потребителей и даже конкурентов. Такие платформы становятся не просто каналами сбыта, а центрами экосистем, где происходит обмен данными, формируются новые сервисы и появляются дополнительные источники прибыли. При этом конкурентное преимущество определяется не столько наличием уникального продукта, сколько способностью компании встроиться в сетевую структуру, управлять потоками данных и поддерживать долгосрочные отношения с участниками экосистемы.

Таким образом, платформенная экономика трансформирует конкурентную среду, превращая цифровые платформы в точки концентрации отраслевых экосистем и создавая предпосылки для интегрированного управления данными между предприятиями.

3. Международные исследования

В международных исследованиях проблематика управления данными связывается прежде всего с ростом их объёмов, многообразием форматов и снижением степени структурированности. Согласно DAMA-DMBOK [19], данные становятся центральным корпоративным активом, однако именно их разнообразность создаёт новые вызовы: компании вынуждены увеличивать вычислительные ресурсы и внедрять более сложные технологии для их сбора, интеграции и анализа.

Gartner отмечает, что классический подход на основе реляционных баз данных уже не обеспечивает требуемой гибкости, что привело к переходу сначала к архитектуре «Озера данных» (Data lakes), а затем к новым моделям «Фабрика данных» (Data fabric) и «Сетка данных» (Data mesh) [20, 21]. Эти концепции позволяют объединять источники разнородных данных и создавать распределённые структуры, где управление качеством и доступностью информации становится системным процессом. Важную роль начинают играть AI-driven решения: интеллектуальная интеграция данных, автоматическая генерация процессов (Data lineage) и применение алгоритмов в области управления данными (Data governance).

Современные исследования также подчёркивают значимость внедрения техник обработки данных (DataOps) и операций машинного обучения (MLOps) как методологий, обеспечивающих управляемую разработку и эксплуатацию аналитических моделей [22].

Всё это формирует основу для появления новых методов бизнес-аналитики, которые выходят за пределы традиционных отчётов и дашбордов, обеспечивая возможность прогнозного анализа, интеллектуальной поддержки управленческих решений и построения адаптивных бизнес-моделей в условиях высокой динамики внешней среды [23].

4. Результаты исследования

Практические аспекты методики управления корпоративными данными

Предлагаемая методика – это практический каркас, который переводит стратегию управления корпоративными данными в управляемые артефакты, роли и процедуры. Она охватывает полный жизненный цикл данных (от источников до потребителей), стандартизирует взаимодействие систем и людей, а также связывает ИТ-механику с бизнес-целями: КПЭ процессов, нормативные требования (в том числе ФГИСы), риск-контроль и экономический эффект. Методика опирается на принципы DAMA, DataOps/MLOps и ориентируется на прозрачность (lineage), качество и безопасность данных. На выходе предприятие получает воспроизводимый набор спецификаций, политик, моделей и сервисов, который обеспечивает масштабирование интегрированных процессов и ускоряет внедрение ИИ-сценариев.

Источником корпоративных данных каждого предприятия выступает, прежде всего, ERP-система, являющаяся ядром цифровизации процессов. В дополнение к ней используются внешние системы и сервисы (MDM, BPM, BI и другие), которые расширяют возможности автоматизации и управления. Существенный объём информации формируется также из внешних разнородных источников – государственных (например, ФГИС, налоговые сервисы) и коммерческих (EDI-провайдеры, маркетплейсы), что обеспечивает комплексное наполнение корпоративного контура данными для принятия управленческих решений.

Методика управления данными состоит из следующих компонент:

1) **Протоколы и интеграции (data integration)** – формирование стандартов взаимодействия ИТ-решений, что обеспечивает целостность информационного пространства и согласованность транзакционных процессов.

2) **Обработка (моделирование) данных (data modelling)** – описание ключевых сущностей предприятия (клиенты, заказы, продукты, партнёры и др.) и их взаимосвязей, что позволяет формализовать правила работы с данными и заложить основу для стандартизированной аналитики.

3) **Оркестрация сервисов и прослеживаемость данных (data orchestration and lineage)** – построение механизмов, обеспечивающих прослеживаемость движения и трансформации данных от источников до аналитических витрин.

4) **Управление данными и безопасность (data governance and security)** – моделирование сквозных цепочек, которые объединяют данные из разных систем, обеспечивая прозрачность и управляемость бизнес-процессов над системным уровнем.

5) **Аналитика и модели искусственного интеллекта (analytics & artificial intelligence)** – применение алгоритмов машинного обучения и крупных языковых моделей (ML, LLM) для автоматизации очистки, категоризации, прогнозирования и интеллектуальной поддержки управленческих решений.



Рисунок 2 – Концептуальная схема методики

Протоколы и интеграции (data integration)

Сегодня распространены десятки методов интеграции: от классических файловых обменов и работы через общую СУБД до более современных веб-сервисов и API-вызовов. В зависимости от зрелости инфраструктуры системы могут взаимодействовать напрямую или через специализированное промежуточное ПО – шину данных. Такая шина обеспечивает транспорт данных, контроль ошибок и логирование операций, что особенно важно в условиях взаимодействия с государственными органами и соблюдения регуляторных требований. По опыту автора в российском рынке ключевыми решениями выступают платформы RabbitMQ, Kafka, 1C, Датареон и другие.

Функционально интеграционные платформы, соответствующие современным международным практикам (iPaaS, Data Fabric), включают:

1) **Сбор и загрузку данных (Data Ingestion)**. Поддерживается широкий спектр сценариев – от пакетных загрузок до потоковой передачи в реальном времени. Зрелые платформы обеспечивают бесшовный доступ к данным как из внутренних систем (ERP, CRM, WMS), так и из внешних источников: государственных реестров, облачных сервисов, IoT-датчиков и транзакционных API.

2) **Моделирование и трансформация данных (Data Mapping & Transformation)**. Декларативные ETL/ELT-конвейеры автоматически сопоставляют разнородные схемы с целевой аналитической моделью. При этом выполняется валидация, нормализация и обогащение справочниками, а также применение бизнес-правил. На современном этапе критически важна встроенная поддержка трансформаций бизнес-сущностей (metadata-driven) и возможностей для самообслуживания бизнес-пользователей (self-service).

3) **Мониторинг и управление качеством (Data Monitoring & Controls)**. Встроенные механизмы мониторинга (data observability) отслеживают полноту, своевременность и достоверность, а система оповещений автоматически реагирует на инциденты и нарушения SLA. Такой уровень прозрачности обеспечивает доверие к данным и позволяет использовать их для аналитики и AI-сценариев.

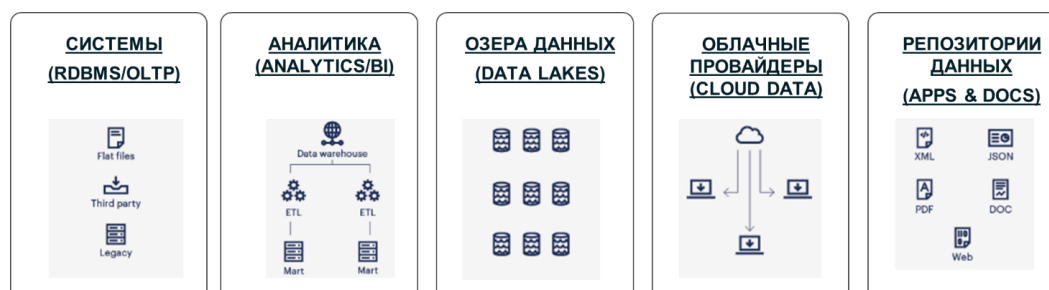


Рисунок 3 – Концептуальная схема интеграции разнородных данных

Обработка (моделирование) данных (data modelling)

Фундаментом управления корпоративными данными выступает обработка (моделирование) данных, которая обеспечивает единое представление информации для всех участников бизнес-процессов. Согласно мировой практике (DAMA, Gartner), моделирование данных должно охватывать весь спектр корпоративных потребностей:

1) **Поддержка многоуровневых моделей**. Включает концептуальный, логический и физический уровни описания, обеспечивающие связь между стратегическим планированием и конкретными системными реализациями.

2) **Нормализованные и многомерные схемы**. Предусматривается возможность работы как с нормализованными структурами для транзакционных систем, так и с многомерными моделями для аналитики (звезда, снежинка).

3) **Гибкость и масштабируемость**. Важной функцией является способность модифицировать модели с учётом изменений в бизнес-процессах, росте объёмов данных и появлении новых источников.

4) **Импорт и экспорт моделей**. Современные инструменты должны поддерживать обмен моделями в различных форматах (ERwin, PowerDesigner, UML, BPMN), что обеспечивает интеграцию с внешними системами и международными стандартами.

5) **Учёт аналитических требований**. Логическая модель должна включать показатели, измерения и бизнес-правила, необходимые для построения BI-отчётов, прогнозной аналитики и применения AI-моделей.

Особое значение имеет организация данных по слоям, что обеспечивает структурированность и управляемость:

1) **Справочники и метаданные (DICT)**. Базовый слой, который хранит корпоративные справочники и описания метаданных. Он обеспечивает единый словарь понятий, унифицированные классификаторы и коды, что позволяет согласовать данные между различными системами. Слой играет ключевую роль в управлении качеством и интерпретацией данных, формируя основу для процессов управления данными (data governance).

2) **Сырые данные (RAW)**. Хранилище первичной информации, поступающей из источников без существенной обработки. Данные сохраняются в максимально близком к источнику виде, включая дубликаты, ошибки и избыточные атрибуты. Цель слоя Сырых данных – обеспечить полноту и воспроизводимость информации, которая при необходимости может быть перепроверена или перезагружена.

3) **Оперативное хранилище (ODS)**. Оперативное хранилище, где данные очищаются, нормализуются и приводятся к согласованному виду. Здесь выполняется первичная стандартизация, агрегирование и консолидация информации из разных источников. Слой обеспечивает доступ к данным, пригодным для ежедневной операционной аналитики и интеграции бизнес-процессов.

4) **Аналитический слой (DDS)**. Финальный аналитический слой, включающий витрины данных, оптимизированные для построения отчётов, BI-панелей и прогнозных моделей. В DDS данные агрегированы, обогащены бизнес-правилами и готовы для применения инструментов бизнес-аналитики и AI-сервисов.

Такой подход позволяет выстраивать единую архитектуру управления данными, которая одновременно удовлетворяет требованиям транзакционных систем, аналитических платформ и AI-сервисов, сохраняя целостность и масштабируемость корпоративной информационной среды.



Рисунок 4 – Концептуальная схема обработки данных

Оркестрация сервисов и прослеживаемость данных (data orchestration and lineage)

Оркестрация сервисов в управлении корпоративными данными направлена на то, чтобы обеспечить согласованное выполнение всех процессов по загрузке, обработке и доставке информации. В мировой практике этот аспект связывается с концепцией наблюдаемости данных (data observability) – способностью платформ отслеживать движение данных в реальном времени, фиксировать отклонения и гарантировать соблюдение SLA.

Ключевым элементом здесь выступает data lineage – прослеживаемость данных от момента их поступления из источника до формирования аналитического отчёта или применения в модели искусственного интеллекта. Линия происхождения позволяет определить, какие системы участвовали в обработке, какие бизнес-правила применялись, а также зафиксировать все преобразования и агрегирования. Это критически важно как для соответствия требованиям регуляторов (например, ФГИС, НСУД), так и для обеспечения доверия со стороны бизнеса и внешних пользователей.

Оркестрация процессов реализуется через DAG-конвейеры (Directed Acyclic Graphs), где все шаги обработки данных описываются в виде последовательности задач. Такие системы автоматически запускают процессы, контролируют их выполнение, формируют события (alerts) при сбоях и обеспечивают воспроизводимость. На практике используются специализированные инструменты (Apache Airflow, Датарееон, др.), позволяющие централизованно управлять обработкой данных.

Технологии искусственного интеллекта (AI-driven) усиливают этот слой методики – алгоритмы машинного обучения применяются для автоматического выявления узких мест, прогнозирования возможных сбоев и адаптивного перераспределения ресурсов. Это превращает прослеживаемость данных из статического журнала преобразований в активный инструмент мониторинга качества, поддерживающий надежность и прозрачность всего контура управления корпоративными данными.

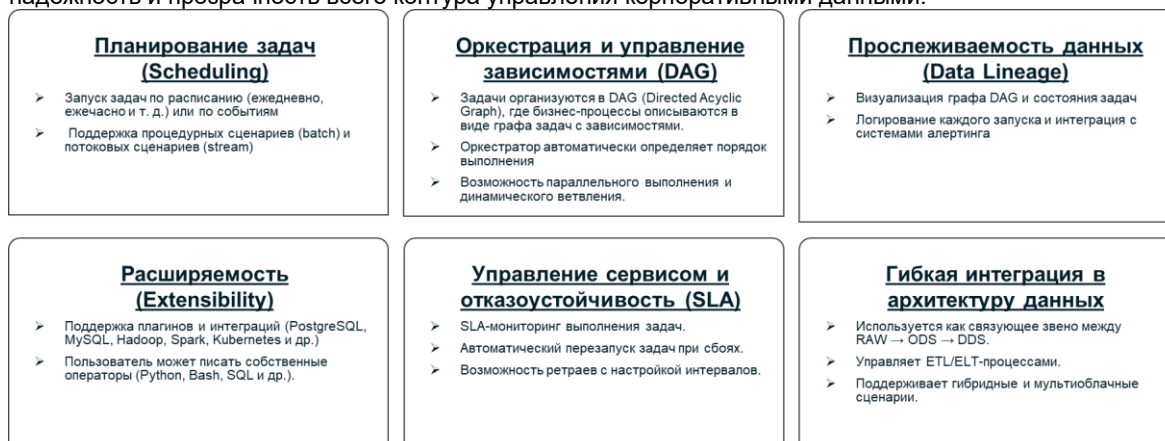


Рисунок 5 – Концептуальная схема оркестрации данных

Управление и безопасность данных (data governance and security)

Современная методика управления корпоративными данными должна объединять два взаимосвязанных направления: управление качеством и обеспечение безопасности. Такой подход соответствует мировым практикам, где управление данными рассматривается не только как средство контроля качества, но и как основа для защиты информации и соблюдения нормативных требований. Ключевые элементы включают:

1) **Централизованный каталог данных.** Поддержка корпоративного каталога для упрощения поиска, совместного использования и понимания данных различными подразделениями. Каталог становится инструментом коллаборации, объединяющим бизнес и ИТ.

2) **Мониторинг и оценка качества.** Применение инструментов профилирования данных и встроенных контролей качества. Система позволяет отслеживать метрики (полнота, достоверность, своевременность) и автоматически формировать отчёты об отклонениях.

3) **Управление схемами.** Введение схем для прозрачного обмена между системами и контроль эволюции схем, что обеспечивает согласованность при интеграции и снижает риски ошибок.

4) **Классификация данных и персональные данные.** Автоматическая классификация корпоративной информации, включая выявление персональных данных и другой чувствительной информации, требующей особого режима защиты.

5) **Маскирование и анонимизация.** Реализация механизмов скрытия и обобщения данных для использования в тестовой среде и аналитике без риска раскрытия конфиденциальной информации.

6) **Шифрование.** Обеспечение защиты данных при передаче и хранении с помощью современных криптографических методов, что соответствует требованиям ФЗ-152, GDPR и других регуляторов.

7) **Контроль доступа.** Внедрение ролевых (RBAC) и атрибутивных (ABAC) моделей доступа, которые позволяют дифференцированно управлять правами пользователей в зависимости от их функций, контекста и уровня доверия.

Такое сочетание процессов управления и безопасности данных позволяет выстроить доверенную информационную среду, где данные не только точные и актуальные, но и защищённые на всех стадиях жизненного цикла.

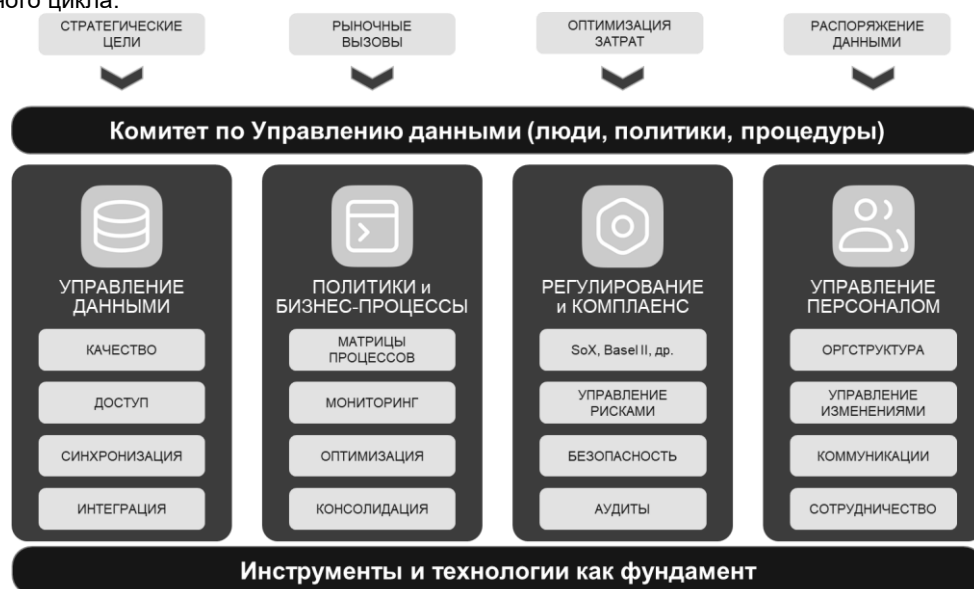


Рисунок 6 – Концептуальная схема управления данными

Самообслуживаемая аналитика и искусственный интеллект
(*self-service analytics & artificial intelligence*)

Финальным элементом методики является слой самообслуживаемой аналитики (self-service) и искусственного интеллекта (artificial intelligence), где данные превращаются в бизнес-ценность. Ключевую роль здесь играет слой витрин данных (QuickMart layer), оптимизированных под конкретные аналитические задачи. Он позволяет пользователям работать в самостоятельном режиме, создавая собственные отчёты и визуализации без постоянного вовлечения ИТ-подразделений. Такой подход ускоряет принятие решений, снижает нагрузку на ИТ и формирует культуру работы с данными внутри предприятия.

Современные платформы, такие как Dataiku, AnyLogic или LogiDom, предоставляют богатый набор инструментов для бизнес-аналитики и моделирования. Среди ключевых возможностей:

1) **Интерактивные дашборды.** Пользователи могут самостоятельно строить визуализации, использовать фильтры и интерактивные панели для анализа ключевых метрик.

2) **Регламентированные отчёты.** Поддержка создания формализованных отчётов в различных форматах с автоматизированной рассылкой заинтересованным сторонам.

3) **Геоинформационный анализ.** Включение GIS-функциональности для пространственного анализа и визуализации территориально распределённых данных (например, склады, торговые точки, логистические маршруты).

4) **Сценарный анализ и анализ чувствительности.** Проведение сценарного моделирования (What-if) для оценки влияния изменения отдельных переменных на бизнес-результаты. Анализ

чувствительности показывает, какие параметры оказывают наибольшее влияние на результаты, помогая сосредоточиться на критичных факторах.

5) **Модели искусственного интеллекта (ML/LLM).** Использование алгоритмов машинного обучения для детектирования аномалий, распознавания паттернов и прогнозирования ключевых показателей. Поддержка батчевых и потоковых сценариев (near real-time). Классификация и категоризация данных, интеллектуальное ценообразование, прогнозирование спроса, оптимизация цепочек поставок и автоматизация рутинных операций.

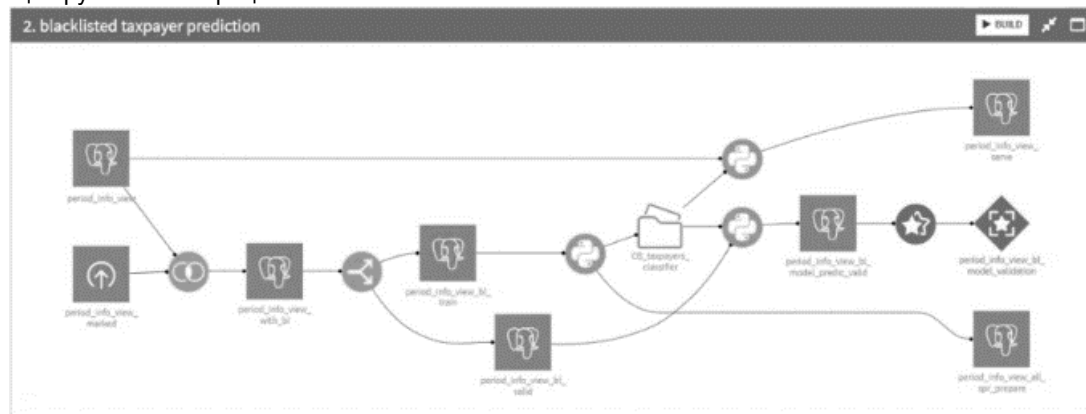


Рисунок 7 – Пример построения прогнозной модели в режиме self-service (составлено автором в решении Dataiku)

Таким образом, интерактивные дашборды и модели искусственного интеллекта обеспечивают переход от описательной аналитики к предиктивной и предписывающей (predictive & prescriptive analytics), где система не только показывает текущее состояние, но и формирует рекомендации для действий.

Выводы

Цифровизация становится неотъемлемым условием устойчивого развития современных торгово-промышленных предприятий, где особое значение приобретает системное управление корпоративными данными. В настоящей работе представлена авторская методика, ориентированная на интеграцию принципов управления данными и технологий искусственного интеллекта (AI-driven), что позволяет формировать прикладные и воспроизводимые механизмы цифровой трансформации. В отличие от классических подходов, сосредоточенных на ИТ-архитектуре или отдельных процессах, методика делает акцент на практическую реализацию протоколов интеграции данных, обработку и оркестрацию данных, сервисы управления и безопасности, самостоятельного использования данных пользователями и применения моделей искусственного интеллекта (AI/ML).

Предложенный подход был апробирован автором в ряде крупных проектов цифровой трансформации, включая:

- 1) ведущую торговую испанскую компанию в фэшн-индустрии;
- 2) одного из крупнейших отечественных аграрных холдингов;
- 3) крупного производителя бытовой техники.

По опыту автора построение методики управления корпоративными данными невозможно без применения комплексного технологического стека, объединяющего системы-источники, средства обработки и моделирования, оркестрации и управления, протоколов интеграции, а также инструменты аналитики и искусственного интеллекта. Примером подхода является сочетание отечественных решений (1C, ELMA, Галактика, отечественные BI и ИИ-платформы) и международных инструментов (PostgreSQL, ClickHouse, MongoDB, Airflow, Kafka, Spark). Такое сочетание позволяет обеспечивать соответствие нормативным требованиям, расширять возможности автоматизации и поддерживать высокую гибкость цифровой архитектуры предприятия.



Рисунок 8 – Пример Технологического стека (по опыту автора)

Практические результаты подтверждают, что внедрение методики управления корпоративными данными с компонентами искусственного интеллекта обеспечивает рост качества информации, прозрачность сквозных процессов и снижение проектных рисков.

Список использованных источников

1. Фреймворк цифровой трансформации. ИТиММ-2024. Сборник статей XIII Международной научно-практической конференции имени А.И. Китова. Москва, 2024. С. 250-255.
2. Целевая ИТ-архитектура цифрового торгово-промышленного предприятия. ИТиММ-2025. Сборник статей XIV Международной научно-практической конференции имени А.И. Китова. Москва, 2025. С. 169-174.
3. <https://ac.gov.ru/news/page/v-rossii-poavitsa-nacionalnaa-model-celevyh-uslovij-vedenia-biznesa-28064> (дата обращения: 06.09.2025).
4. https://www.nalog.gov.ru/rn77/about_fts/gos_inf/aisnalog3/ (дата обращения: 06.09.2025).
5. <https://честныйзнак.рф> (дата обращения: 06.09.2025).
6. <https://ofd.nalog.ru> (дата обращения: 06.09.2025).
7. <https://lesegais.ru/> (дата обращения: 06.09.2025).
8. <https://fgisgrain.ru> (дата обращения: 06.09.2025).
9. <https://fsvps.gov.ru/fsvps/mercury> (дата обращения: 06.09.2025).
10. <https://egais.ru> (дата обращения: 06.09.2025).
11. <https://minpromtorg.gov.ru> (дата обращения: 06.09.2025).
12. <https://cbr.ru/fintech/dr/> (дата обращения: 06.09.2025).
13. https://www.tadviser.ru/index.php/Статья:Российские_платформы_для_интеграции_данных_и_приложений._Рейтинг_TAdviser (дата обращения: 06.09.2025).
14. <https://www.economy.gov.ru> (дата обращения: 06.09.2025).
15. Высшая школа экономики. Платформенная экономика в России: потенциал развития. Москва, 2023. 72 с.
16. McKinsey & Company. The Rise of Digital Marketplaces. 2021.
17. Gartner. Market Guide for Digital Marketplaces. Gartner Research, 2022.
18. Rogers, D. L. The Digital Transformation Playbook: Rethink Your Business for the Digital Age. New York: Columbia University Press, 2016. – 278 с.
19. DAMA International. The DAMA Guide to the Data Management Body of Knowledge (DAMA-DMBOK2). Technics Publications, 2017. – 628 с.
20. Gartner. Top Trends in Data and Analytics 2024. Stamford: Gartner Research, 2024.
21. <https://www.gartner.com/en/documents/4005998> (дата обращения: 06.09.2025).
22. Skoulis, I., Mitropoulos, D., Spinellis, D. MLOps: Continuous Delivery and Automation Pipelines in Machine Learning // IEEE Software. 2020. Vol. 37, No. 5. P. 28-35.
23. Saggi, M. K., Jain, S. A survey towards an integration of big data analytics to big insights for value-creation // Information Processing & Management. 2018. Vol. 54, No. 5. P. 758–790.

Ключевые слова:

Цифровая экономика, Платформенная экономика, Искусственный интеллект, Управление данными, Протоколы и интеграции, Обработка и моделирование, Оркестрация, Прослеживаемость, Безопасность, Самообслуживаемая аналитика.

Никитчук Сергей Сергеевич,

аспирант, РЭУ им. Г.В. Плеханова (115054, Россия, Москва, Стремянный переулок 36, корпус 9),

ORCID: 0009-0006-3422-6456,

snitchuk002@yandex.ru

Sergei S. Nikitchuk. Methodology of AI-driven Data Management for Digital Enterprise

Keywords

Digital economy, Platform economy, Artificial intelligence, Data management, Protocols and integrations, Processing and modeling, Orchestration, Lineage, Security, Self-service analytics

DOI: 10.34706/DE-2025-04-08

JEL classification: C02 – C4 Эконометрические и статистические методы: специальные темы; C45 Нейронные сети и смежные темы

Abstract

The article presents an original methodology for corporate data management aimed for applying artificial intelligence (AI-driven) in industrial and trade enterprises. The relevance of the research is determined by the growing data volume and heterogeneity, the necessity of integration with federal state information systems (FGIS), and development of platform economics. Unlike classical approaches to data management, the proposed methodology is focused on practical implementation and includes five interrelated components: data

protocols and integrations, data processing and modeling, data services orchestration and data lineage, data governance and security, as well as the business self-service analytics and artificial intelligence layer.

The literature review considers modern international approaches (Data Fabric/Data Mesh, DataOps/MLOps) and the Russian specifics of the digital economy, including the national business model, the digital ruble project, and the development of FGISes. According to the analysis, the key challenge is the increasing demand for computational resources and the growing complexity of data processing technologies, while the degree of data structure decreases.

Practical testing of the methodology was carried out in several large-scale digital transformation projects, including trade, industrial, agricultural, and financial companies. The results demonstrated that the use of AI components improves data quality, ensures transparency of end-to-end processes, and reduces project risks. In the future, the methodology is expected to be further developed through the enhancement of data governance practices, the automation of integrations with state systems, and the application of generative AI models to build intelligent and adaptive digital architectures.