

Иордан Кирилл Иванович,  
Магистрант направления «Бизнес-информатика»,  
Кубанский Государственный Университет

## РОЛЬ DATA ENGINEERING ПОДХОДОВ В ОБЕСПЕЧЕНИИ КАЧЕСТВА ДАННЫХ ДЛЯ ЗАДАЧ РИСК-МЕНЕДЖМЕНТА

**Аннотация.** В статье рассматривается роль Data Engineering-подходов в обеспечении качества данных для задач риск-менеджмента в условиях цифровой трансформации организаций. Анализируется влияние качества данных на достоверность оценки рисков, устойчивость аналитических моделей и эффективность управленческих решений. Особое внимание уделяется современным практикам инженерии данных, включая ETL/ELT-процессы, построение data pipelines, механизмы data validation, monitoring, observability, data lineage и data governance. Исследуется взаимосвязь между надежностью инфраструктуры данных и способностью организаций минимизировать операционные, финансовые и информационные риски. Формируется вывод о том, что Data Engineering становится не только техническим инструментом обработки информации, но и стратегическим элементом формирования устойчивой системы риск-менеджмента.

**Ключевые слова:** Data Engineering, качество данных, риск-менеджмент, data governance, data pipelines, ETL, ELT, data lineage, observability, управление данными, аналитика данных, цифровая трансформация, инфраструктура данных, управление рисками.

Современные системы риск-менеджмента всё больше и больше зависят от качества данных, на основе которых формируются аналитические модели, прогнозируются сценарии развития событий и принимаются управленческие решения. В условиях цифровизации бизнеса объемы информации растут быстрее, чем возможности организации по ее структурированию и контролю. Это приводит к ситуациям, при которых даже технологически развитые компании сталкиваются с проблемами непостоянности, дублирования, неполноты и устаревания данных. Подобные недостатки особенно критичны для процессов управления рисками, поскольку искажение исходной информации непосредственно влияет на достоверность оценки вероятности угроз, финансовых потерь и операционной устойчивости.

На фоне усложнения архитектуры корпоративных информационных систем растёт и значение Data Engineering-подходов, обеспечивающих системную работу с данными на всех этапах их жизненного цикла. Инженерия данных перестаёт рассматриваться исключительно как техническая дисциплина, связанная с построением хранилищ или обработкой потоков информации. В современных условиях она становится фундаментом обеспечения качества данных, без которого невозможно формирование надежной среды риск-менеджмента. Именно поэтому исследование роли Data Engineering-подходов в повышении качества данных приобретает особую актуальность как для научного сообщества, так и для практики управления организационными рисками.

### Качество данных как основа эффективного риск-менеджмента

Система риск-менеджмента предполагает постоянный анализ информации, поступающей из внутренних и внешних источников. Качество этой информации определяет точность моделей оценки рисков, корректность прогнозов и своевременность управленческих реакций. Ошибки в данных способны привести не только к неверной интерпретации текущего состояния бизнеса, но и к стратегическим просчетам, последствия которых проявляются в финансовых потерях, репутационном ущербе и нарушении нормативных требований.



Особенность риск-менеджмента заключается в высокой чувствительности аналитических процессов к любым отклонениям в данных. Даже незначительные расхождения между фактическими и зарегистрированными показателями могут существенно повлиять на результаты расчетов кредитного, рыночного или операционного риска. В банковском секторе, страховании и промышленности качество данных становится не просто техническим параметром, а фактором устойчивости всей системы принятия решений.

Ключевыми характеристиками качественных данных традиционно считаются полнота, актуальность, достоверность, согласованность и доступность. Однако в условиях распределенной цифровой инфраструктуры обеспечение этих характеристик требует комплексного подхода. Простого хранения данных уже недостаточно. Организациям необходимо выстраивать механизмы контроля, верификации и непрерывного мониторинга информационных потоков. Именно здесь Data Engineering начинает играть определяющую роль, объединяя архитектурные, технологические и организационные инструменты обеспечения качества данных.

### **Data Engineering как инфраструктурная основа управления данными**

Развитие Data Engineering связано с переходом организаций к обработке больших объемов данных, поступающих из многочисленных разнородных источников. Современная корпоративная среда включает транзакционные системы, CRM-платформы, облачные сервисы, IoT-устройства и внешние аналитические каналы. В подобных условиях возникает проблема фрагментации данных, затрудняющая формирование единой картины рисков.

Data Engineering-подходы ориентированы на создание устойчивой инфраструктуры, обеспечивающей сбор, трансформацию, хранение и доставку данных для аналитических задач. В контексте риск-менеджмента такая инфраструктура должна гарантировать не только высокую производительность обработки информации, но и сохранение ее качества на каждом этапе движения данных.

Центральное место в инженерии данных занимают процессы ETL и ELT, обеспечивающие извлечение, преобразование и загрузку данных в аналитические системы. На практике именно этап трансформации становится критически важным для обеспечения качества информации. В процессе обработки устраняются дубликаты, корректируются аномальные значения, унифицируются форматы и восстанавливаются нарушенные связи между сущностями. Если эти процедуры реализованы некорректно, риск-аналитика начинает опираться на искаженные показатели, что приводит к ошибочным выводам.

Существенное значение имеют и data pipelines, представляющие собой автоматизированные цепочки обработки данных. Их использование позволяет снизить влияние человеческого фактора и обеспечить воспроизводимость процессов обработки информации. Для риск-менеджмента это особенно важно, поскольку ручные операции повышают вероятность ошибок и затрудняют аудит аналитических процедур. Автоматизация обработки данных способствует не только повышению скорости принятия решений, но и росту доверия к результатам анализа.

### **Практики обеспечения качества данных в инженерии данных**

Современные подходы к инженерии данных предполагают интеграцию механизмов контроля качества непосредственно в архитектуру обработки информации. Одним из наиболее значимых направлений становится data validation, обеспечивающая проверку корректности данных на этапе их поступления и трансформации. В системах риск-менеджмента подобные проверки позволяют своевременно выявлять аномалии, пропущенные значения и логические противоречия, способные исказить результаты аналитики.

Развитие потоковой обработки данных усилило потребность в инструментах monitoring и observability. Если ранее контроль качества часто осуществлялся постфактум, то



современные организации стремятся к непрерывному мониторингу информационных потоков в режиме реального времени. Такой подход позволяет оперативно фиксировать изменения в структуре данных, выявлять отклонения от нормативных параметров и предотвращать распространение ошибок по всей аналитической системе.

Особое место в обеспечении качества данных занимает концепция data lineage, ориентированная на отслеживание происхождения и трансформации данных. Для риск-менеджмента прозрачность движения информации имеет принципиальное значение, поскольку позволяет определить источник ошибок и оценить степень влияния некорректных данных на результаты анализа. Наличие механизмов lineage также облегчает выполнение требований регуляторов, предъявляющих высокие требования к прозрачности аналитических процессов.

Не менее важным элементом становится data governance, формирующий организационную модель управления данными. Даже наиболее совершенные технологические решения не способны обеспечить устойчивое качество информации без четкого распределения ответственности, единых стандартов и регламентов работы с данными. В контексте риск-менеджмента governance-подходы позволяют согласовать деятельность ИТ-подразделений, аналитиков и бизнес-структур, формируя единое понимание критериев качества данных.

### **Влияние качества данных на точность оценки рисков**

Связь между качеством данных и эффективностью риск-менеджмента носит прямой и системный характер. Аналитические модели, используемые для прогнозирования рисков, функционируют на основе статистических закономерностей, выявляемых в массивах информации. Если исходные данные содержат ошибки или искажения, модели начинают формировать некорректные зависимости, что приводит к снижению точности прогнозов.

Особенно отчетливо данная проблема проявляется при использовании технологий машинного обучения в риск-аналитике. Алгоритмы способны выявлять сложные паттерны и зависимости, однако качество результатов напрямую определяется качеством обучающих данных. При наличии систематических ошибок модель может воспроизводить и усиливать искажения, создавая ложное представление о характере рисков. В результате организация сталкивается с ситуацией, при которой формально сложная аналитическая система демонстрирует низкую практическую эффективность.

Качество данных оказывает влияние не только на математические модели, но и на устойчивость управленческих процессов в целом. Недостоверная информация затрудняет оценку текущего состояния бизнеса, снижает скорость реакции на угрозы и увеличивает вероятность принятия ошибочных решений. В условиях высокой неопределенности, характерной для современной экономики, подобные последствия способны существенно ухудшить конкурентные позиции организации.

В то же время внедрение Data Engineering-подходов позволяет сформировать более надежную аналитическую среду. Автоматизация контроля качества, стандартизация процессов обработки данных и развитие систем наблюдаемости повышают прозрачность информационной инфраструктуры и способствуют росту доверия к аналитическим результатам. Это особенно важно для компаний, работающих в высокорегулируемых отраслях, где качество данных становится частью системы корпоративной ответственности.

### **Ограничения и перспективы развития Data Engineering-подходов**

Несмотря на значительный потенциал инженерии данных, ее внедрение сопровождается рядом ограничений. Одной из ключевых проблем остается высокая сложность современных архитектур данных. По мере роста числа источников информации увеличивается вероятность возникновения несогласованности между системами, что усложняет процессы контроля качества. Кроме того, многие организации сталкиваются с дефицитом специалистов, способных одновременно понимать принципы Data Engineering и специфику риск-менеджмента.



Серьезным вызовом является и необходимость интеграции устаревших информационных систем с современными платформами обработки данных. В крупных организациях исторически накопленные ИТ-решения часто не соответствуют требованиям гибкой и масштабируемой архитектуры. Это ограничивает возможности автоматизации и затрудняет реализацию сквозного контроля качества данных.

Дополнительную сложность создает рост требований к скорости обработки информации. Современные системы риск-менеджмента во многом ориентированы на работу в режиме реального времени, что требует мгновенной проверки качества поступающих данных. В подобных условиях традиционные методы контроля оказываются недостаточно эффективными, а организации вынуждены развивать интеллектуальные механизмы автоматической диагностики и обнаружения аномалий.

Перспективы развития Data Engineering в сфере риск-менеджмента связаны с дальнейшим распространением облачных платформ, технологий DataOps и инструментов интеллектуального мониторинга данных. Усиливается тенденция к интеграции механизмов обеспечения качества непосредственно в процессы разработки аналитических систем. Подобный подход позволяет рассматривать качество данных не как отдельную функцию контроля, а как неотъемлемый элемент всей цифровой инфраструктуры организации.

#### **Заключение**

Качество данных становится одним из ключевых факторов эффективности современного риск-менеджмента. В условиях усложнения цифровой среды и роста объемов информации организации сталкиваются с необходимостью построения устойчивых механизмов контроля, обработки и анализа данных. Именно Data Engineering-подходы обеспечивают технологическую и организационную основу для решения этой задачи.

Инженерия данных позволяет формировать надежную инфраструктуру, способную поддерживать целостность, достоверность и актуальность информации на всех этапах ее жизненного цикла. Использование ETL/ELT-процессов, data pipelines, monitoring, lineage и governance-подходов способствует повышению прозрачности аналитических процедур и снижению вероятности ошибок в оценке рисков. При этом значение Data Engineering выходит за рамки технической поддержки аналитики и превращается в стратегический фактор устойчивости бизнеса.

Одновременно сохраняются проблемы, связанные со сложностью архитектур данных, интеграцией разнородных систем и необходимостью работы в режиме реального времени. Однако дальнейшее развитие технологий обработки данных и автоматизации контроля качества создает предпосылки для повышения надежности систем риск-менеджмента. В перспективе именно интеграция инженерии данных и риск-аналитики будет определять способность организаций эффективно функционировать в условиях растущей неопределенности и цифровой трансформации экономики.

#### *Список литературы:*

1. Головина Т. А., Авдеева И. Л., Суханов Д. А. Управление рисками организаций в условиях цифровой экономики // Вестник Академии знаний. 2022. №1 (48). URL: <https://cyberleninka.ru/article/n/upravlenie-riskami-organizatsiy-v-usloviyah-tsifrovoy-ekonomiki>
2. Шалтунович А. В. Нереляционные системы хранения в условиях проблемы больших данных и распределенных вычислений // Вестник НВГУ. 2013. №1. URL: <https://cyberleninka.ru/article/n/nerelyatsionnye-sistemy-hraneniya-v-usloviyah-problemy-bolshih-dannyh-i-raspredeleennyh-vychisleniy>
3. Голицына О.Л., Максимов Н.В., Попов И.И. Информационные системы : учебное пособие. - М.: Форум, 2021. - 496 с. URL: Издательство Форум - Информационные системы



4. Исмиханов З. Н., Гаджиева Н. А., Султанов Н. Г. Информационные модели и цифровые технологии в управлении: новые подходы и вызовы в бизнес-аналитике // РППЭ. 2024. №12 (170). URL: <https://cyberleninka.ru/article/n/informatsionnye-modeli-i-tsifrovye-tehnologii-v-upravlenii-novye-podhody-i-vyzovy-v-biznes-analitike>

5. Kleppmann M. Designing Data-Intensive Applications. - Sebastopol: O'Reilly Media, 2017. - 616 с. URL: O'Reilly - Designing Data-Intensive Applications.

