

# АНАЛИЗ ДАННЫХ КАК ИНСТРУМЕНТ СОВЕРШЕНСТВОВАНИЯ ПРОЦЕССОВ

**Хаков Евгений Маратович**

Аспирант, Российская Академия Народного Хозяйства  
и Государственной Службы  
при Президенте Российской Федерации, г. Москва  
evgenykhakov@yandex.ru

## DATA ANALYSIS AS A TOOL FOR PROCESS IMPROVEMENT

**E. Khakov**

*Summary.* This article is devoted to the issues of expanding the use of data analysis to improve corporate processes. Within the framework of this study, the main principles and approaches to data analysis are considered, and ways to improve the methods already used by introducing artificial intelligence technology and using quantum computing are determined. The purpose of this work is to define the main tasks of analysis, methods for its implementation, and to identify ways to enhance analytical functions using end-to-end technologies. The author concludes that processes can be continuously improved by integrating new technologies and subsequently creating a synergistic effect between traditional management methods and intelligent analysis tools, providing the ability to quickly adapt to changes in the market environment due to the gradual transition from static to dynamic (executable) business processes.

*Keywords:* process improvement, data analytics, end-to-end technologies, big data, artificial intelligence, quantum communications.

Современные организации, оперирующие на мировом рынке информационных технологий, представляют собой сложную систему, огромный комплекс взаимосвязанных областей, в котором сосредоточено множество разнообразных видов деятельности, в значительной степени влияющих на экономику страны.

Формируя целевые параметры своего развития, гиганты отрасли добавляют, детализируют и стараются воплотить эти показатели в конкретных проектах цифровизации, модернизации, развития систем менеджмента качества, управления рисками, внутреннего контроля и многих других.

Обзор сложившихся и активно применяемых практик управления показывает, что лучшим способом связывать перечисленные задачи в единой основе является внедрение и расширение применения процессного подхода.

Методология процессного управления предполагает выстраивание процесса по жизненному циклу продукта или по управленческому циклу PDCA (Планирование —

*Аннотация.* Данная статья посвящена вопросам расширения применения анализа данных для совершенствования корпоративных процессов. В рамках настоящего исследования рассмотрены основные принципы и подходы к проведению анализа данных, а также определены способы улучшения уже применяемых методов за счет внедрения в них технологии искусственного интеллекта и использования квантовых вычислений. Целью данной работы является определение основных задач анализа, способов его проведения, а также обозначение путей усиления аналитических функций с помощью сквозных технологий. Автором делается вывод о непрерывном улучшении процессов за счет интеграции новых технологий и последующего создания синергетического эффекта между традиционными методами управления и инструментами интеллектуального анализа, обеспечивающими возможность быстро адаптироваться к изменениям в рыночной среде за счет постепенного перехода от статичных к динамичным (исполняемым) бизнес-процессам.

*Ключевые слова:* совершенствование процессов, анализ данных, сквозные технологии, большие данные, искусственный интеллект, квантовые коммуникации.

Действие — Проверка — Воздействие). Важно представлять цепочку процесса целиком и оценивать влияние всех входящих в него элементов, так как изменение даже одного из них может существенно повлиять на итоговый результат процесса — получаемый на выходе продукт. Важным инструментом комплексной оценки эффективности реализации процесса в таком случае становится анализ данных, в том числе интеллектуальный [4].

Возможность выполнения комплексного анализа подразумевает, что по каждому признаку необходимо иметь некоторый массив значений. Данные при этом становятся доступными либо как результат чьих-то натурных наблюдений, либо как сводка статистических отчетов. В силу этого обстоятельства, полученные исходные данные всегда требуют специальной предварительной обработки (подготовки), прежде чем их можно будет подвергнуть анализу.

Задача анализа данных состоит в том, чтобы:

1. Установить наличие имеющихся данных каких-либо общих скрытых закономерностей (паттернов);
2. Убедиться, что обнаруженная закономерность не является недостоверным результатом исследо-

вания, возникающим из-за дефектов метода исследования, а действительно существует в данных;

3. Понять, исходя из имеющихся знаний об объекте анализа, возможную причину возникновения паттерна;
4. Построить правдоподобную модель, описывающую выявленные закономерности.

Что касается цели анализа данных, то ее можно сформулировать следующим образом: уточнение или обогащение существующего знания об объекте анализа [3]. К основным способам такого обогащения относятся:

1. Порождение (вывод) новых понятий из данных, представляемых признаками объекта анализа (например, вычисление среднего значения ряда значений признака или кластеризация в пространстве признаков). Такой способ часто называют «агрегация» или «суммаризация»;
2. Обнаружение новых зависимостей (утверждений о связи признаков), подтвержденных данными (например, классификация для качественных признаков или регрессия для численных признаков). Такой способ обозначается термином «коррелирование».

При этом существующая в отрасли система аналитики устроена таким образом, что в большинстве случаев анализируются не сами процессы, а статистические показатели [2]. Поэтому самый простой вид анализа данных выражается в сопоставлении текущего значения показателя с каким-либо отчетным периодом либо же с его целевым значением. Риск такого подхода заключается в отсутствии возможности своевременного и взвешенного принятия корректирующих мер и воздействий.

Поэтому в случаях, когда речь идет об исполняемости и управляемости процессами, требуется прибегнуть к глубинному анализу процессов — технологии Process Mining. Она позволяет не только обеспечить более быстрый и точный анализ, но и оценить саму суть протекающих процессов, а не их модельные или обобщенные статистические представления. Данный подход позволяет перейти от констатации фактов о выполняемом процессе к выявлению причинно-следственных связей в самом механизме его реализации, определяя, почему процесс протекает именно таким образом [4].

Process Mining предполагает выполнение определенных шагов, позволяющих пройти путь от сбора данных до оптимизации процессов. В общем виде алгоритм проведения анализа по технологии Process Mining состоит из следующих этапов:

Обнаружение — фиксация операций в логах, а также сбор и консолидация данных;

Проверка соответствия — расчет характеристик и метрик, а также воссоздание реально протекающего процесса;

Улучшение — анализ и формирование выводов, а также проектирование и моделирование нового процесса;

Отслеживание — мониторинг корректности протекания обновленного процесса, а также контроль соответствия вмененных изменений целевому состоянию процесса.

Для проведения глубинного анализа процессов необходимы:

1. База моделей бизнес-процессов, которые будут способствовать определению взаимозависимостей исследуемых показателей в рамках единой процессной модели работы компании;
2. Наличие специализированных компьютерных программ, так как получение производных показателей из исходных данных невозможно посредством использования общепринятых офисных инструментов по причине их неадаптированности под работу с настолько большими информационными массивами;
3. Доступ к исходным источникам информации, а не к сводным статистическим отчетам для более корректного выстраивания причинно-следственных связей.

Анализ данных является базисом для обеспечения поэтапного перехода из исходного состояния организационных процессов в целевое, а его применение в качестве аналитического инструмента охватывает самые разные области и сегменты человеческой деятельности [4]. Подобные научно-технические направления, ориентированные на укрепление национальной экономики, прирост производственных мощностей и инновационное развитие, называются сквозными технологиями.

К ним, согласно Национальной технологической инициативе Российской Федерации, принято относить и технологии хранения и анализа больших данных (Big Data), в том числе с применением элементов искусственного интеллекта, а также квантовые технологии.

Искусственный интеллект (ИИ) сегодня — это не просто набор принципиально новых для мирового рынка технологий, а динамично развивающаяся область знаний, которая обещает трансформировать наше общество и экономику уже в ближайшие годы. По своей сути ИИ представляет собой набор программ и алгоритмов, главным отличительным свойством которых является способность обучаться, обобщать большой объем неструктурированной информации, а также выполнять

конкретные корректирующие действия без прямого вмешательства человека, что не только повышает эффективность реализации процесса за счет рационального математически выверенного использования имеющихся ресурсов (трудовых, финансовых, материальных и т.д.), но и позволяет переходить к предиктивной аналитике.

Интеллектуальный анализ данных в настоящее время находит активное применение в следующих сферах:

1. Прогнозирование — оценка потенциальных сценариев и расчет вероятности наступления определенных событий на основе отраслевых трендов и закономерностей в реализуемых процессах;
2. Классификация — разделение объектов на группы по характерным признакам для оптимизации информационных потоков компании;
3. Ассоциативный анализ — выявление взаимосвязей между различными переменными и оценка их влияния друг на друга;
4. Кластеризация — группировка объектов на основе схожести характеристик для выявления скрытых тенденций в данных.

Каждое из вышеперечисленных направлений акцентируется на принятие более обоснованных решений по процессам и закладывает основу для формирования искусственных разумных систем, которые на практике выражаются целым комплексом объединенных в единое целое интеллектуальных механизмов [2].

Хорошим примером подобного рода конфигураций могут служить интеллектуальные транспортные системы (ИТС), де-факто являющиеся совокупностью как аппаратных, так и программных решений, направленных на управление и оптимизацию транспортных потоков [3]. Они используют современные информационные технологии и результаты их вычислений для реализации ряда функций, таких как:

1. Управление дорожным движением: автоматизация светофоров и мониторинг загруженности дорог;
2. Информирование пользователей: предоставление актуальной информации о состоянии дорожной ситуации;
3. Оптимизация маршрутов: адаптация путей общественного транспорта и легковых автомобилей в зависимости от текущих условий и прогнозируемой ситуации на дорогах в заданном горизонте планирования.

Структурированная информация, получаемая из ИТС, не только улучшает качество транспортных услуг, но и способствует устойчивому развитию городской инфраструктуры, позволяя:

1. Сократить время в пути: благодаря адаптивному управлению светофорами и оптимизации маршрутов время ожидания уменьшается;

2. Увеличить безопасность: автоматизированные системы мониторинга помогают быстро реагировать на аварии и другие инциденты, накапливая при этом некоторую базу знаний и опыт, который в перспективе можно будет применить не только для ликвидации последствий аварии, но и для ее недопущения (например, более удачный вариант проектирования дорожных развязок);
3. Снизить экологическую нагрузку: оптимизация маршрутов и снижение пробок ведут к постепенному сокращению углеродного следа (уменьшение вредных выбросов).

Пути же расширения использования ИТС для совершенствования процессов за счет анализа данных в основном включают в себе те направления развития, которые на данный момент только начинают активно прорабатываться — например, интеграция ИТС с системами общественного транспорта, служб экстренной помощи и метеорологическими службами для обеспечения стандартизированного подхода к управлению движением [3]. Подобные объединенные аналитические платформы улучшат сам процесс сбора и обработки данных благодаря использованию датчиков и камер, ориентированных на комплексное отслеживание параметров, влияющих на движение (трафик, состояние дорог, наличие аварий, погодные условия), а также способствуют автоматизации управления транспортными потоками на основе анализа больших данных (создание в реальном времени исполняемых моделей движения и выявление на их основе наиболее проблемных участков за счет обновляемой нестатической информации).

Под большими данными понимаются огромные массивы неоднородной информации, которые могут как иметь, так и не иметь оформленную структуру, что нехарактерно для «классических» баз данных. В обобщенной трактовке «большие данные» являются социальным и экономическим феноменом, который связан с возникновением технологий анализа очень больших объемов информации в отдельных сферах и проявляющихся при этом сопутствующих (вытекающих) проблем [1]. Суть проблематики заключается не столько в создании громадных объемов данных, сколько в их структурном оформлении и выстраивании связей, что не соответствует общепринятому формату работы с информационными массивами. Одним из важнейших этапов работы с большими данными является очистка информации — процесс удаления ошибок и нерелевантной информации для повышения качества анализа и представления его результатов в максимально удобной для пользователя форме (информационная фильтрация).

Появление квантовых технологий может революционизировать подходы к обработке больших данных. Сам термин «квантовые технологии» относится к инноваци-

онным средствам, в которых применяются уникальные свойства квантовой механики — в первую очередь, квантовая запутанность. Квантовые компьютеры способны выполнять вычисления значительно быстрее, чем их классические аналоги, что открывает новые горизонты для улучшения моделей машинного обучения. Квантовые технологии могут повысить точность предсказаний за счет больших объемов проанализированной информации (в том числе — в реальном времени), а также заложить новые принципы хранения и интеграции данных из различных источников [1].

Если задать обычному компьютеру последовательность из тридцати 0 и 1, то на выходе мы получим миллиард возможных значений этой последовательности. Компьютер, использующий обычные биты, затратит на подобную обработку огромное количество времени, также используя значительный ресурс своей памяти, ведь каждая комбинация будет пройдена им по отдельности. С другой стороны, квантовый компьютер обладает способностью «видеть» все миллиарды по-

следовательностей одновременно, что при выполнении аналогичной операции многократно разгрузит память устройства, обеспечив единовременный доступ к любой из интересующих нас комбинаций [2].

Таким образом, квантовые технологии могут существенно изменить ландшафт работы с данными, предоставляя новые инструменты для анализа и принятия решений в условиях растущих объемов информации.

Все вышеуказанное является объективными вызовами в современном мире, где нужно уметь не только оперативно реагировать на внешние и внутренние изменения, но и обеспечивать непрерывное развитие по всем направлениям своей деятельности, эффективно управляя этими изменениями. Для этих целей важно понимать сущность процессов и образуемых ими систем. Ключом к этому пониманию (а как следствие и совершенствованию) был, есть и будет анализ данных — независимо от его видов, типов и форм реализации.

#### ЛИТЕРАТУРА

1. Р.В. Зыков «Роман с Data Science. Как монетизировать большие данные». — издательство «Питер», 2021. — 219–230 с.
2. Алексей Колоколов «Заставьте данные говорить». — издательство «Альпина ПРО», 2023. — 137–145 с.
3. Саманта Клейнберг «Почему. Руководство по поиску причин и принятию решений». — издательство «Манн, Иванов и Фербер», 2017. — 48–72 с.
4. Кирилл Еременко «Работа с данными в любой сфере: Как выйти на новый уровень, используя аналитику». — пер. с англ. — издательство «Альпина Паблишер», 2019. — 30–44 с.

© Хаков Евгений Маратович (evgenyhakov@yandex.ru)  
Журнал «Современная наука: актуальные проблемы теории и практики»