

Экономико-математические методы и модели Economic & mathematical methods and models

Научная статья

УДК 004.6:005.311.6

DOI: <https://doi.org/10.18721/JE.19309>

EDN: <https://elibrary/AQFNTY>



ОЦЕНКА КАЧЕСТВА СОБЫТИЙНЫХ ДАННЫХ В ЦИФРОВЫХ СЛЕДАХ СКВОЗНЫХ ОРГАНИЗАЦИОННО- УПРАВЛЕНЧЕСКИХ ПРОЦЕССОВ

С.Н. Яшин , К.С. Шибанов

Нижегородский институт управления – филиал РАНХиГС,
Нижний Новгород, Российская Федерация

kshibanov@mail.ru

Аннотация. Статья посвящена методологической проблеме, возникающей при переходе организаций от «управления по отчетам» к управлению по цифровым следам сквозных (межфункциональных и межсистемных) процессов. Показано, что обилие событийных данных (журналов ERP/CRM/Service Desk, интеграционных и пользовательских логов, транзакций оборудования и др.) само по себе не гарантирует их пригодность для управленческих решений: систематические дефекты измерения – неполнота логирования, разрывы сквозной идентичности кейса, несогласованность семантики статусов, низкое временное разрешение и дрейф определений после релизов – могут приводить не просто к «шуму», а к устойчиво смещенным управленческим выводам. Особый акцент сделан на том, что требования к качеству событийных данных различаются для двух режимов использования: описательно-контрольного (реконструкций процесса, KPI, мониторинга отклонений, process mining) и причинно-объяснительного (оценки эффекта управленческого вмешательства на основе наблюдательных данных). Методологическая рамка исследования строится как синтез интеллектуального анализа процессов, стандартов и практик управления качеством данных и теории причинного вывода (потенциальные исходы, причинные графы, квазиэкспериментальные дизайны). Вводится двухконтурная модель качества событийных данных: «измерительная состоятельность» (насколько корректно данные отражают траекторию кейса во времени) и «каузальная пригодность» (насколько данные позволяют однозначно определить вмешательство, исход, временные окна и обеспечить обоснованность идентификационных допущений). Предлагается операционализация каузальной пригодности в виде набора индексов (целостности единицы наблюдения, временной пригодности, устойчивости смыслов, наблюдаемости вмешательства, достаточности контекста для контроля смещения, наблюдаемости ресурсной среды), связывающих дефекты журналов событий с рисками ложной оценки эффекта. Дополнительно обоснована необходимость метаданных учета (версионности статусов и правил, паспортирования времени, трассируемости изменений), а также процессно-семантических ограничений (аксиом процесса) как механизма контроля качества, критичного именно для причинных выводов. Практический результат – контур организационного управления качеством событийных данных (паспорт журнала событий, реестр вмешательств, версионирование смыслов, панель индексов качества), позволяющий заранее определять, какие управленческие решения и каузальные оценки в принципе допустимы на имеющихся цифровых следах.

Ключевые слова: цифровые следы, сквозные процессы, событийные данные, журнал событий, качество данных, process mining, управленческие решения, причинный вывод

Для цитирования: Яшин С.Н., Шибанов К.С. (2026) Оценка качества событийных данных в цифровых следах сквозных организационно-управленческих процессов. П-Economy, 19 (3), 128–147. DOI: <https://doi.org/10.18721/JE.19309>



Research article

DOI: <https://doi.org/10.18721/JE.19309>

ASSESSING THE QUALITY OF EVENT DATA IN DIGITAL TRACES OF END-TO-END ORGANIZATIONAL AND MANAGEMENT PROCESSES

S.N. Yashin , **A.V. Alekseeva**  Nizhny Novgorod Institute of Management – RANEPA branch,
Nizhny Novgorod, Russian Federation kshibanov@mail.ru

Abstract. The article examines the methodological challenge that arises when organizations transition from “management by reports” to management based on digital traces of end-to-end (cross-functional and intersystem) processes. It demonstrates that the abundance of event data (ERP/CRM/Service Desk logs, integration and user logs, equipment transactions etc.) does not in itself guarantee its suitability for management decisions: systematic measurement flaws – incomplete logging, discontinuities in the end-to-end identity of a case, inconsistency in status semantics, low temporal resolution, and definition drift after releases – can lead not just to “noise” but to persistently biased management conclusions. Particular emphasis is placed on the fact that event data quality requirements differ for two modes of use: descriptive-control (process reconstruction, KPIs, deviation monitoring, process mining) and causal-explanatory (assessing the effect of management interventions based on observational data). The methodological framework of the study is constructed as a synthesis of process mining, data quality management standards and practices, and causal inference theory (potential outcomes, causal graphs, quasi-experimental designs). A two-loop model of event data quality is introduced: “measurement consistency” (how accurately the data reflect the trajectory of a case over time) and “causal suitability” (the extent to which the data allow for the unambiguous determination of an intervention, outcome, time windows, and the validity of identification assumptions). An operationalization of causal suitability is proposed as a set of indices (observation unit integrity, temporal suitability, meaning stability, intervention observability, context sufficiency for confounding control, and resource environment observability) linking event log defects to the risk of false effect estimates. The need for accounting metadata (status and rule versioning, time stamping, and change traceability) is further substantiated, as is the need for process-semantic constraints (process axioms) as a quality control mechanism critical specifically for causal inference. The practical result is a framework for organizational event data quality management (event log passport, intervention registry, meaning versioning, and a quality index panel), enabling the proactive determination of which management decisions and causal assessments are, in principle, permissible based on existing digital traces.

Keywords: digital traces, end-to-end processes, event data, event log, data quality, process mining, management decisions, causal inference

Citation: Yashin S.N., Shibanov K.S. (2026) Assessing the quality of event data in digital traces of end-to-end organizational and management processes. *П-Economy*, 19 (3), 128–147. DOI: <https://doi.org/10.18721/JE.19309>

Введение

Актуальность исследования

За последние пятнадцать лет управленческая деятельность в организациях претерпела глубокую трансформацию: от опоры на периодическую агрегированную отчетность к непрерывному управлению на основе цифровых следов сквозных бизнес-процессов. Эти следы формируются в виде событийных данных – журналов событий корпоративных систем, интеграционных логов, транзакций оборудования, пользовательских действий и записей сервисных платформ. В отличие от традиционных отчетов, такие данные фиксируют реальную ткань операционной деятельности: последовательность шагов, их временную структуру, контекст исполнения,

отклонения, повторные циклы и скрытые согласования. Согласно прогнозам Международной корпорации по данным, глобальный объем создаваемых цифровых сведений к 2026 г. приближается к 175 зеттабайтам, при этом значительная часть приходится именно на операционные события, генерируемые на уровне процессов и устройств [1]. Для руководителей это создает принципиально новую ситуацию: первичная фактура управленческих решений все чаще существует не в виде готовых сводок, а в форме сырых журналов событий, из которых требуется извлекать модели процессов, показатели и обоснованные выводы.

Однако обилие событийных записей само по себе не гарантирует их пригодность для управления. Более того, оно нередко маскирует системные изъяны измерения: неполноту регистрации шагов, разрывы в сквозной идентификации объектов (заказов, обращений, клиентов), несогласованность значений статусов между системами, недостаточную точность временных меток, селективное логирование исключений и дрейф определений после изменений в информационных системах. Эти дефекты приводят не просто к случайному «шуму», а к устойчивому смещению управленческих заключений. По оценкам экспертов, организации ежегодно несут многомиллионные потери именно из-за низкого качества данных, проявляющиеся в ошибочных решениях, в снижении производительности и в росте скрытых издержек [2]. В условиях, когда управление стремится перейти от описания происходящего к точной оценке результатов собственных вмешательств (изменения регламентов, автоматизации операций, перераспределения ресурсов), вопрос о качестве событийных данных приобретает стратегическую остроту. Он перестает быть исключительно технической проблемой информационных служб и превращается в ключевой фактор управленческой рациональности и конкурентоспособности.

Цифровой след сквозных организационно-управленческих процессов — это совокупность зафиксированных в информационных системах данных и метаданных, получаемых в ходе выполнения действий и взаимодействий внутри организации на протяжении всего жизненного цикла процессов. Он представляет собой непрерывную эмпирическую запись того, как реально протекают процессы, включая их состояния, переходы между этапами, участников, временные характеристики и сопутствующие атрибуты [3].

Сквозные процессы — от оформления заказа до его оплаты, от поступления обращения до его разрешения, от найма сотрудника до завершения его карьеры — по своей природе преодолевают границы подразделений, информационных систем и нередко внешних партнеров [4]. Каждое событие такого процесса может фиксироваться в разных источниках: в системах планирования ресурсов, управления взаимоотношениями с клиентами, в сервисных платформах, в системах складского и производственного учета, в интеграционных шинах. В результате процесс как управленческий объект предстает в данных в виде разрозненной мозаики траекторий, которые необходимо точно соединить по идентификаторам, времени и смыслу. Именно здесь формируется потребность в особом классе качества данных — качестве событийных данных сквозных процессов [5].

Литературный обзор

За последние годы в области анализа процессов сформировался значительный инструментарий, позволяющий восстанавливать реальные модели деятельности по журналам событий, измерять продолжительность циклов, выявлять варианты исполнения, узкие места, нарушения правил и закономерности отклонений. основополагающие работы в этой области подчеркивают минимальные требования к структуре журналов событий: наличие идентификатора случая, наименования действия и временной метки. Разработаны стандартизированные форматы обмена такими данными, обеспечивающие воспроизводимость анализа [6]. Вместе с тем научная литература давно фиксирует, что реальные журналы событий почти всегда содержат несовершенства: пропуски, «шумовые» записи, несогласованные временные метки, различную степень детализации и нефиксируемые ручные операции. Большинство исследований сосредоточено



на том, как эти несовершенства мешают точному описанию и визуализации процесса [7]. Однако в современной управленческой практике возникает запрос более высокого порядка: способность событийных данных служить надежной основой не только для описания, но и для причинных выводов о результативности управленческих изменений.

Причинная идентификация эффекта вмешательства — что именно изменится в результате введения нового правила, автоматизации шага, пересмотра приоритетов или изменения системы мотивации — составляет сердцевину доказательного управления. Теория причинного вывода, опирающаяся на концепцию потенциальных исходов, причинно-следственные графы и квазиэкспериментальные подходы, четко формулирует необходимые условия: ясное определение самого вмешательства и момента его начала, корректное измерение исходов, учет факторов смещения, устойчивость единицы наблюдения и отсутствие интерференции между объектами [8]. В сквозных процессах эти требования особенно критичны, поскольку случаи не изолированы: они конкурируют за общие ресурсы, а организационные решения изменяют правила маршрутизации и приоритизации. Поэтому требования к данным существенно расширяются: они должны позволять отделять эффект управленческого вмешательства от влияния нагрузки, сезонности, изменений в клиентском составе и технологических сдвигов.

Событийные данные предоставляют уникальные возможности для причинного анализа благодаря высокой частоте наблюдений и точной фиксации времени. Они позволяют использовать естественные эксперименты, возникающие при поэтапном внедрении изменений, пороговых правилах или внезапных обновлениях систем. Однако те же данные несут и серьезные риски: систематические ошибки измерения (смещение временных меток, зависимость полноты логирования от исхода случая, дрейф семантики статусов) способны не просто ослаблять оценки, а радикально искажать их знак или значимость. Эконометрические исследования убедительно демонстрируют, что ошибки измерения, коррелированные с целевыми переменными, приводят к систематическому смещению выводов [9]. В организационной практике такие ошибки часто носят не случайный, а эндогенный характер.

Таким образом, в центре исследования находится фундаментальное противоречие. С одной стороны, многочисленные эмпирические работы подтверждают, что интенсивное использование данных для принятия решений статистически связано с более высокой производительностью компаний. С другой стороны, при переходе от описательной аналитики к причинно-объяснительной качество событийных данных становится определяющим эпистемическим условием достоверности управленческого знания. Управленческое решение оказывается настолько надежным, насколько надежна измерительная система, породившая цифровой след [10]. Это делает актуальной задачу разработки специализированных требований к качеству событийных данных сквозных процессов, выведенных не из общих моделей качества информации, а непосредственно из структуры управленческого вывода и логики причинной идентификации.

Цель исследования

Цель исследования состоит в разработке теоретико-методологической основы оценки качества событийных данных в цифровых следах сквозных организационно-управленческих процессов, обеспечивающей разграничение их пригодности для описательно-контрольного и причинно-объяснительного использования, а также формализацию процедур контроля качества для принятия управленческих решений.

Для достижения указанной цели решаются следующие **задачи**:

1. Обосновать необходимость разграничения описательно-контрольного и причинно-объяснительного режимов использования событийных данных в сквозных организационно-управленческих процессах.
2. Разработать двухконтурную модель качества событийных данных, включающую измерительную состоятельность и каузальную пригодность.

3. Определить состав минимально необходимой информационной основы для причинно-корректного анализа сквозных процессов, включая события, контекстные признаки, данные о ресурсной среде и метаданные учета.

4. Сформулировать и операционализировать систему индексов каузальной пригодности событийных данных, связывающих типовые дефекты журналов событий с рисками смещения управленческих выводов.

5. Разработать инструменты организационного и методического контроля качества событийных данных, включая процессно-семантические ограничения, матрицу соответствия причинных допущений и требований к данным, а также контур предварительной оценки допустимости управленческих выводов.

Гипотеза исследования заключается в том, что для сквозных процессов существует воспроизводимый и проверяемый набор требований к качеству событийных данных, принципиально отличающийся от универсальных моделей качества тем, что он выводится из логики управленческого вывода и причинной идентификации. Соблюдение этих требований превращает журналы событий в надежную основу для достоверной оценки эффектов управленческих вмешательств, тогда как их нарушение ведет к систематически смещенным выводам и решениям, опирающимся на артефакты учета, а не на реальные изменения процессов [11].

Объектом исследования выступают сквозные организационно-управленческие процессы в современных организациях, протекающие через несколько функциональных подразделений и информационных систем и сопровождающиеся формированием цифровых следов в виде журналов событий, транзакций и связанных метаданных.

Предметом исследования являются требования к качеству событийных данных сквозных процессов и способы их операционализации (в виде критериев, проверок, показателей и процедур контроля), обеспечивающие обоснованность управленческих решений и возможность причинной идентификации эффекта управленческих вмешательств на основе наблюдательных данных.

Методы и материалы

Методология исследования строится как междисциплинарный синтез управленческого анализа процессов, управления качеством данных и причинного вывода.

Во-первых, применяется логико-методологический анализ причинного вывода (подход потенциальных исходов и связанные требования к определению вмешательства, исходов, единицы наблюдения и факторов смещения) для того, чтобы перевести «каузальные» допущения в проверяемые требования к структуре и содержанию журналов событий.

Во-вторых, используется инструментарий интеллектуального анализа процессов как эмпирическая рамка работы с журналами событий: выделение «случая» (единицы процесса), упорядочивание событий, анализ вариантов протекания процесса и аномалий. Это дает возможность формулировать требования к данным не абстрактно, а в терминах того, что именно должно быть наблюдаемо и сопоставимо во времени [12].

В-третьих, привлекаются стандарты и практики управления качеством данных для формализации процедур контроля (согласованности, полноты, точности, актуальности), но далее эти измерения будут «переопределены» применительно к событийной природе данных: с акцентом на точность времени, целостность идентификаторов, устойчивость семантики статусов и трассируемость изменений правил учета.

В-четвертых, в качестве аналитического приема применяется метод выявления типовых несовершенств журналов событий (пропусков, «шума», некорректных временных меток, неоднозначных статусов, несогласованности детализации) и их связи с направлениями смещения при оценке эффекта — то есть переход от «ошибок данных» к «ошибкам управленческих выводов».



Результаты исследования

Ключевой результат исследования состоит в следующем: в сквозных процессах качество событийных данных нельзя корректно описать только через «универсальные» атрибуты (полноту, точность, согласованность и т.п.), потому что управленческое использование журналов событий неизбежно распадается как минимум на два режима познания — описательно-контрольный и причинно-объяснительный. В первом режиме организация отвечает на вопрос «что происходит и где отклонения», во втором — «что изменится из-за конкретной меры управления». Эти режимы предъявляют различающиеся, а иногда и конфликтующие требования к данным. Отсюда следует принципиально важная для менеджмента позиция: качество событийных данных должно задаваться не от данных, а от вывода, то есть от типа управленческого решения и от идентификационной логики, на которую опирается оценка эффекта вмешательства.

В рамках исследования предложена двухконтурная модель качества событийных данных сквозного процесса.

Первый контур — измерительная состоятельность — отвечает на вопрос, насколько журнал событий отражает фактическое выполнение работ как последовательность действий во времени. Здесь важны свойства, традиционно обсуждаемые в интеллектуальном анализе процессов: корректность упорядочивания, полнота фиксируемых шагов, различимость вариантов протекания, надежность временных меток, отсутствие «шумовых» событий и пропусков; типовые дефекты журналов событий и их паттерны описаны как систематическая проблема реальных организаций.

Второй контур — каузальная пригодность — отвечает на вопрос, допускают ли те же данные строгую оценку эффекта управленческой меры (регламента, автоматизации, перераспределения ресурсов) без подмены причинного эффекта артефактами учета. Здесь в центр выходят допущения причинного вывода и квазиэкспериментальных дизайнов: определимость момента вмешательства и момента измерения исхода, устойчивость единицы наблюдения, сопоставимость групп и временных периодов, учет факторов смешения, а также риски интерференции через общие ресурсы.

Смысл двухконтурной модели в том, что журнал событий может быть «достаточно хорош» для контрольной панели и одновременно «недостаточно хорош» для оценки эффекта. Причина — различная чувствительность к систематическим ошибкам измерения. В эконометрике давно показано, что ошибки измерения в ключевых переменных могут приводить не к росту дисперсии, а к смещению оценок, в том числе к ложным выводам о наличии/отсутствии эффекта. В процессных данных организации эти ошибки часто эндогенны: «сложные» случаи чаще проходят через ручные обходы, чаще переоткрываются, по ним чаще меняют статусы и чаще допускают неполное логирование — то есть дефект данных становится коррелированным с исходом [13].

Второй результат исследования — формализация того, что именно в сквозных процессах должно считаться минимально необходимым содержанием «каузально пригодного» журнала событий. В причинном анализе даже идеальная техника оценивания бессильна, если невозможно корректно определить три базовых элемента: единицу наблюдения, вмешательство и исход. Для сквозного процесса это означает следующее.

Единица наблюдения (например, заказ, заявка, обращение) должна быть не только идентифицируема, но и устойчива во времени и сквозь границы систем. Иначе возникает «фрагментация случая»: один и тот же управленческий объект распадается на несколько идентификаторов, а разные объекты ошибочно сшиваются. В описательном анализе это дает «плавающие» показатели, но в причинном — разрушает саму конструкцию сопоставления «до/после» и «подвергся / не подвергся». Практически это требование выходит за рамки поля идентификации: требуется дисциплина сквозной идентичности (единые ключи, правила сшивки, управление

дублями, протоколы изменения ключей), что соотносится с процессными моделями управления качеством данных в стандартах ISO 8000 [14].

Вмешательство в управленческом смысле (введение нового правила, автоматизация шага, изменение приоритизации, изменение норматива) должно быть наблюдаемо в данных как событие или как метаданные, однозначно задающие: когда, где, для кого и по какому правилу вмешательство начало действовать. Это принципиально: без «момента включения» вмешательства организация рискует оценивать не эффект меры, а эффект изменения логирования или эффект календарных сдвигов. В условиях обновлений информационных систем особую роль играет версионность правил учета: когда меняется логика формирования статусов, формально «вмешательство» может произойти в данных раньше или позже фактического управленческого решения.

Исход (например, длительность цикла, соблюдение срока, повторное обращение, доля возвратов, стоимость обработки) должен иметь измерение, устойчивое к изменению способов регистрации событий. Здесь важен принцип измерительной инвариантности: если способ измерения исхода меняется вместе с вмешательством, оценка эффекта становится логически неопределенной, потому что сравниваются разные величины под одним названием. В стандартах качества данных подобные аспекты обычно скрыты за общими категориями «согласованность» и «достоверность», но для причинного вывода требуется явное удержание неизменности определения исхода и прозрачность его вычисления.

Отсюда в исследовании предложено рассматривать «журнал событий процесса» как недостающий сам по себе объект данных. Для управленческих решений и причинной идентификации необходим минимум из четырех взаимосвязанных слоев:

- 1) слой событий выполнения работ (кто/что/когда сделал или изменил статус);
- 2) слой контекстных признаков случая (характеристики клиента, продукта, канала, сложности, категории риска), без которых факторы смещения не контролируются;
- 3) слой ресурсной среды (нагрузка, доступность исполнителей, очереди, ограничения мощности), без которого интерференция и эффект перегрузки могут быть ошибочно приписаны вмешательству;
- 4) слой метаданных учета (версии справочников статусов, версии интеграций, правила округления времени, изменения в маршрутизации), без которого нельзя отделить реальное изменение процесса от изменения измерения.

Эта конструкция выводит качество событийных данных из плоскости чистки журналов в плоскость проектирования измерительной инфраструктуры управления.

Третий результат исследования касается временной компоненты как главного «скрытого» источника систематических ошибок в управленческих выводах. В сквозных процессах именно время делает возможными вычисления длительностей, проверку нормативов, анализ очередей, сопоставление периодов «до/после», построение событийных исследований. Однако на практике временные метки в разных системах часто отражают разные моменты: фактическое выполнение действия, фиксацию в интерфейсе, запись в базу данных, прием сообщения интеграционной шиной. В распределенных информационных архитектурах добавляется проблема рассогласования часов между узлами, что ведет к «перестановке» последовательности событий и к ошибочному восстановлению траектории [15].

Технический уровень проблемы хорошо описан в спецификациях протоколов синхронизации времени: даже при использовании протокола сетевого времени (NTP) точность зависит от качества сети и конфигурации, а без строгой дисциплины синхронизации расхождения могут достигать секунд и минут, что для управленческих длительностей на уровне минут критично. Для высокоточной синхронизации в промышленных и телеком-средах применяются стандарты прецизионного протокола времени (PTP), требующие иной инфраструктуры [16].



Управленческая интерпретация этого факта проста и нетривиальна: точность времени — это параметр качества данных, который должен быть задан как требование управления, а не как «настройка ИТ». Если организация измеряет, например, выполнение обращений в рамках 15 минут, то округление временных меток до 5 минут или систематическое запаздывание регистрации на 3–7 минут могут радикально изменить долю «выполнено в срок» даже при неизменном фактическом процессе. Более того, при оценке эффекта вмешательства ошибка позиционирования события относительно момента вмешательства порождает ложный эффект «предвосхищения» (как будто изменения начались до внедрения) или «отложенного эффекта», что разрушает смысл сравнения периодов.

Отсюда сформулировано требование паспортирования времени в событийных данных: каждая временная метка, используемая для управленческих выводов, должна сопровождаться (как минимум на уровне метаданных набора данных) указанием источника времени, точки фиксации (факта действия / записи в систему / приема интеграции), правил округления и допустимой погрешности. Это требование напрямую связано с воспроизводимостью вычислений: если правила времени меняются, должна существовать возможность восстановить прежнюю методику расчета показателей, иначе организация теряет сопоставимость динамики. В более широком смысле это продолжает линию воспроизводимых исследований в вычислительной науке: без фиксируемых процедур формирования данных «результат» превращается в одноразовое вычисление.

Четвертый результат исследования — предложение использовать для событийных данных не только общие проверки качества (на полноту, диапазоны, дубликаты), но и процессно-семантические ограничения, которые в работе названы аксиомами процесса. Их смысл: для каждого сквозного процесса существует набор неотменяемых причинно-временных отношений, нарушение которых сигнализирует либо об ошибке данных, либо о несогласованности учета между системами. Примеры таких аксиом в типовых цепочках: «отгрузка не может предшествовать комплектации», «закрытие обращения не может быть раньше регистрации», «возврат не может быть раньше продажи», «оплата не может быть зафиксирована после отмены без отдельного события восстановления заказа» [17]. В интеллектуальном анализе процессов подобные ограничения используются для очистки и обогащения журналов, но здесь они получают иной статус: как защитный механизм причинной идентификации.

Причина в том, что причинные методы чувствительны к систематическим «нарушениям времени». Если данные допускают массовые нарушения аксиом (например, «закрытие раньше регистрации» из-за разницы источников времени), то оценка эффекта вмешательства в сроки обработки может стать произвольной: любые изменения в доле таких аномалий будут ошибочно интерпретированы как улучшение/ухудшение процесса. Поэтому аксиомы процесса в исследовании трактуются как часть управленческого контроля качества данных, а не как техническая косметика.

Отдельно показано, что аксиомы процесса полезны и для диагностики «скрытых ручных действий», не попадающих в журналы. Если процессная модель предполагает обязательный шаг согласования, но в событиях он отсутствует, а при этом наблюдается характерная пауза во времени между двумя соседними событиями, то это индикатор «теневого активности». В описательном анализе это просто «дырка». В причинном анализе это уже риск: вмешательство могло повлиять именно на этот скрытый шаг (например, ускорить согласование через новый регламент), но данные его не измеряют — значит, исход «длительность цикла» будет отражать эффект только частично и с неизвестным смещением.

Пятый результат — формирование связки между допущениями причинного вывода и конкретными измеримыми требованиями к данным сквозных процессов. В табл. 1 приведена компактная матрица — «каркас доказательности», которая показывает, какое допущение причинной

идентификации «ломается» каким дефектом журналов событий и какие проверки качества должны быть встроены в контур управления данными.

Эта матрица является методологическим мостом между управлением процессами и причинным анализом: она позволяет задавать требования к качеству событийных данных как к условиям идентифицируемости эффекта, а не как к «общей гигиене данных».

Таблица 1. Матрица соответствия:
допущения причинного вывода – требования к журналам событий
Table 1. Correspondence matrix: causal inference assumptions – event log requirements

Допущение причинного вывода (управленческий смысл)	Что требуется от событийных данных сквозного процесса	Типовой дефект данных в организациях	Следствие для оценки эффекта
Определимость вмешательства во времени и по охвату (когда и на кого подействовало)	Явная фиксация момента внедрения, признака применения правила, привязка к подразделению/каналу/типу случая; метаданные версий учета	Вмешательство «рас-творено» в обновлении системы; нет признака, что мера применена к конкретному случаю	Смещение из-за неверной классификации: часть «контроля» фактически подверглась вмешательству и наоборот
Сопоставимость периодов и групп (контроль факторов смещения)	Наличие контекстных признаков случая до вмешательства; ста-бильные определения категорий; неизмен-ность правил отбора	Изменение класси-фикаторов, дрейф значений полей после релизов, отсутствие «сложности случая»	Ложный эффект: улуч-шение объясняется сменой состава случаев, а не мерой управления
Устойчивость единицы наблюдения (один случай – одна траектория)	Сквозные идентификато-ры, правила сшивки, кон-троль дублей и разбиений	Потеря идентификато-ра на границе систем, «склейки» разных случаев	Псевдоэффект: измене-ние долей «повторов» и длительностей из-за изменения сшивки
Корректность времени и последовательности (пра-вильные окна «до/после»)	Синхронизация време-ни, паспорт времени, единые правила округ-ления, фиксация точки измерения события	Системные задержки интеграции, разные источники времени, округление по-разно-му в разных системах	«Предвосхищение» или «запаздывание» эффекта, неверная динамика во-круг момента внедрения
Отсутствие скрытой интерференции через общие ресурсы (нагрузка не маскирует эффект)	Данные о загрузке, очередях, доступно-сти ресурсов; привязка исполнителей и смен	Нет данных о ресур-сах; меняются штат/графики параллель-но с внедрением	Эффект меры сме-шивается с эффектом нагрузки/дефицита мощности; вывод стано-вится неуправляемым

Источник: составлено авторами.

Практика управления качеством данных исторически формировалась вокруг корректности записей как таковых: точны ли значения, полны ли поля, согласованы ли справочники, своев-ременно ли обновление. Для сквозных процессов этого недостаточно, потому что управленче-ский вывод строится не на «значениях в таблице», а на восстановленной траектории случая во времени и на сопоставлении траекторий до и после вмешательства. Поэтому в исследовании предлагается перейти к измерению каузальной пригодности событийных данных: к степени, в которой данные позволяют не просто описывать процесс, а обоснованно оценивать эффект управленческой меры.

Теоретическое основание такого перехода задается причинным выводом в терминах потен-циальных исходов: чтобы оценка эффекта была осмысленной, должны быть корректно опре-делены единица наблюдения, факт и момент вмешательства, исходы, а также условия сопоста-вимости и отсутствие систематического смещения. В процессной аналитике, в свою очередь,



накоплена типология несовершенств журналов событий, которая показывает, что «дефект данных» часто означает дефект именно траектории: пропуски шагов, «шумовые» записи, непорядок во времени, несовпадение уровня детализации.

В результате предлагается рассматривать качество для причинной идентификации не как один «балл качества», а как вектор индексов, каждый из которых отвечает за определенное допущение причинного вывода и имеет управленчески интерпретируемую диагностику. Такая конструкция принципиально важна: попытка свести все к одной цифре неизбежно маскирует критические провалы (например, идеальная полнота полей при разрушенной временной шкале).

Ниже вводятся **индексы каузальной пригодности (ИКП)**. Они определяются на интервале от 0 до 1 и вычисляются по журналу событий и метаданным учета. Формулы приводятся в «прикладном» виде, чтобы их можно было реализовать в контуре контроля качества данных.

1) *Индекс целостности единицы наблюдения — насколько устойчив «случай» как управленческий объект*

В сквозном процессе причинный анализ всегда завязан на единицу наблюдения: заказ, заявку, обращение, поставку. Если один и тот же объект распадается на несколько идентификаторов между системами либо несколько объектов ошибочно «склеиваются», то теряется смысл сопоставления «подвергся / не подвергся», а длительность цикла и число повторов начинают зависеть от качества сшивки, а не от процесса.

Индекс предлагается измерять через долю случаев, удовлетворяющих требованиям уникальности и непрерывности идентификатора на всем горизонте процесса:

$$ИКП_E = 1 - (w_s * R_split + w_m * R_merge + w_d * R_dup),$$

где: R_split — доля случаев, у которых наблюдается разбиение одного управленческого объекта на ≥ 2 идентификатора; R_merge — доля случаев, у которых под одним идентификатором обнаруживаются признаки разных объектов (например, несогласованные неизменяемые атрибуты: клиент, договор, адрес); R_dup — доля явных дублей (тождественные траектории/атрибуты при разных идентификаторах); w_s, w_m, w_d — веса (сумма = 1), задаваемые по критичности риска.

Содержательная новизна здесь в том, что контроль идентичности предлагается строить не только «по ключу», а по сочетанию ключа и инвариантов случая (атрибутов, которые не должны меняться). Это переводит задачу из ИТ-плоскости «есть ли ключ» в управленческую плоскость «существует ли стабильная единица процесса».

2) *Индекс временной пригодности — насколько время в данных является управленчески измеримым*

Временная метка в журнале событий — не нейтральная колонка. Для причинного анализа важно, чтобы: а) события упорядочивались корректно; б) точность времени соответствовала масштабу управленческих нормативов; в) источник времени и точка фиксации были известны. В распределенных системах эта проблема усиливается из-за рассогласования часов и задержек интеграции; именно поэтому дисциплина синхронизации времени формализуется в инженерных стандартах (например, NTP) [18].

Индекс предлагается вычислять как комбинацию трех измеримых компонент: доли нарушений причинно-временных «аксиом процесса», доли событий с недостаточной временной разрешающей способностью и доли событий без паспортирования источника времени (на уровне метаданных):

$$ИКП_B = 1 - (\alpha * R_order + \beta * R_gran + \gamma * R_time_meta),$$

где: R_order – доля случаев, где нарушены базовые временные отношения (например, закрытие раньше регистрации); R_gran – доля событий, где временная метка округлена/усечена сильнее, чем допускает норматив управления (например, SLA 15 минут при округлении до 10 минут); R_time_meta – доля наборов данных/источников событий без описания: «что означает время» (факт действия / запись / прием интеграции); α, β, γ – коэффициенты важности (сумма = 1).

Важно, что $ИКП_B$ не равен «средней точности времени»; он целенаправленно измеряет риски неправильного сравнения периодов при оценке эффекта вмешательства.

3) *Индекс устойчивости смыслов – насколько статусы и действия сохраняют значение во времени*

Даже при идеальной полноте полей причинный анализ разрушится, если изменился смысл статусов или действий (например, «в работе» стало означать другое после релиза, либо «закрыто» перестало фиксироваться автоматически). Тогда изменение показателей будет отражать дрейф учета, а не эффект управленческой меры. В стандартах качества данных это обычно скрывается в «согласованности», но для причинной идентификации необходима явная фиксация устойчивости семантики.

Индекс предлагается строить на основе обнаружения структурного дрейфа распределений по ключевым статусам/действиям в окрестности изменений системы учета и на основе доли событий без привязки к версии справочника/правила:

$$ИКП_C = 1 - (\delta * D_status + \varepsilon * R_no_version),$$

где: D_status – мера дрейфа распределения статусов/действий (например, дивергенция Джессена–Шеннона) между соседними окнами времени, рассчитанная отдельно до и после релизов/изменений учета; $R_no_version$ – доля событий, для которых невозможно однозначно определить версию справочника/правила формирования статуса.

Содержательный смысл: если «язык процесса» в данных меняется, то сравнение «до/после» становится сравнением разных языков под одним названием – типичный источник ложных эффектов.

4) *Индекс наблюдаемости вмешательства – можно ли однозначно определить, что мера применена к конкретному случаю*

В причинном выводе вмешательство должно быть наблюдаемым: нужно знать, к каким случаям оно применено, когда началось и каковы правила охвата [19]. В управлении процессами это часто игнорируется: регламент изменили, письмо разослали, а в данных нет ни признака «случай обработан по новому правилу», ни времени включения правила на уровне подразделения/канала. Тогда аналитик вынужден подменять вмешательство календарной датой, что почти всегда неверно из-за поэтапного внедрения и разной скорости принятия.

Индекс предлагается измерять как долю случаев, для которых вмешательство определимо без догадок:

$$ИКП_U = 1 - R_amb_treat,$$

где: R_amb_treat – доля случаев, где «принадлежность к вмешательству» неоднозначна (нет признака применения правила, нет даты/места включения, конфликт метаданных).

На управленческом языке $ИКП_U$ показывает, способна ли организация доказательно ответить на вопрос «какие именно случаи были обработаны по новой схеме».

5) *Индекс достаточности контекста для устранения смещения – хватает ли признаков, чтобы сопоставлять случаи*



Оценка эффекта вмешательства требует учета факторов смешения: случаев разной сложности, каналов, клиентов, сезонности, загрузки, компетенций исполнителей. Если эти признаки не наблюдаемы, то причинная оценка неизбежно опирается на недоказуемые допущения. Практическая трудность в том, что в сквозных процессах «сложность случая» редко представлена одной переменной; она проявляется в сочетании атрибутов и ранних событий траектории.

Индекс предлагается строить как долю случаев, для которых заполнен минимально необходимый набор контекстных признаков, согласованный с выбранной идентификационной схемой, с поправкой на пропуски, зависящие от исхода (наиболее опасный тип пропусков в причинном анализе):

$$ИКП_K = (1 - R_miss_core) * (1 - R_mnar_risk),$$

где: R_miss_core — доля случаев с пропусками в ключевых признаках (канале, категории, типе продукта, регионе, первичной оценке сложности и т.п.); R_mnar_risk — оценка риска «пропусков, зависящих от исхода» (например, если пропуски статистически связаны с длительностью/просрочкой).

Здесь принципиальна связка с эконометрическим пониманием ошибок измерения: пропуски и ошибки, коррелированные с исходом, дают смещение, а не просто «меньше наблюдений».

б) *Индекс наблюдаемости ресурсной среды — видим ли мы нагрузку и ограничения мощности, чтобы отделить эффект меры от эффекта очередей*

Сквозные процессы почти всегда конкурируют за ограниченные ресурсы. Изменение результата может быть вызвано не регламентом, а изменением нагрузки, сменности, текучести, сезонности. Если ресурсная среда не наблюдаема, то «эффект вмешательства» легко оказывается эффектом перегрузки или, наоборот, недогрузки. Эта проблема особенно существенна при интерференции: вмешательство, примененное к части случаев, может менять очередь и тем самым косвенно влиять на «контрольные» случаи.

Индекс предлагается вычислять как долю случаев и периодов, для которых доступны показатели нагрузки и ресурсных ограничений на сопоставимом временном шаге:

$$ИКП_P = 1 - R_no_capacity,$$

где: $R_no_capacity$ — доля интервалов времени/участков процесса, где отсутствуют наблюдения о нагрузке: объем входа, число активных исполнителей, очередь, доступность систем/окон обслуживания.

Важнейшее прикладное следствие: ИКП следует трактовать как паспорт пригодности данных для конкретного управленческого вывода, а не как рейтинг «хорошие/плохие данные». Для контроля процесса может быть достаточно высокого $ИКП_E$ и $ИКП_B$ при средних $ИКП_K$ и $ИКП_P$; для причинной оценки эффекта перераспределения ресурсов, напротив, $ИКП_P$ становится ключевым. Иными словами, одинаковый журнал событий может быть «годен» для одних решений и «негоден» для других, и это должно быть зафиксировано до вычисления эффектов, иначе организация неизбежно будет принимать решения, основанные на артефактах учета.

Чтобы показать прикладной смысл ИКП, рассмотрим типовой сквозной процесс обработки обращений клиентов в сервисной организации: регистрация обращения → классификация → → назначение исполнителя → работа → решение → закрытие. Введены автоматическая классификация и автоназначение исполнителя. Внедрение шло поэтапно: с 1 мая — в группе А (две команды), с 1 июня — в группе В (остальные команды). Управленческая цель: сократить время до первого ответа и общее время решения.

Наивная оценка эффекта (как часто делают на практике): берется показатель «время решения» как разность между временной меткой события «обращение создано» и временной меткой события «обращение закрыто» в журнале событий системы Service Desk. Далее сравнивают средние значения до/после внедрения между группами *A* и *B* (логика «разность разностей»); при этом часто игнорируется, что события и их время могут быть измерены по-разному до и после изменения.

Предположим, получился результат, представленный в табл. 2.

Таблица 2. Результаты оценки эффекта первым способом
Table 2. Results of the effect assessment by the first method

Группа	До (март–апрель), среднее время решения	После (май), среднее время решения	Изменение
<i>A</i> (внедрено с 1 мая)	14 ч	10,8 ч	–3,2 ч
<i>B</i> (еще не внедрено)	13,6 ч	13 ч	–0,6 ч

Источник: составлено авторами.

Оценка эффекта по разности разностей: $(-3,2) - (-0,6) = -2,6$ часа. Управленческий вывод «вмешательство сократило время решения на 2,6 часа» выглядит убедительно и «подкреплен данными».

Выполним диагностику пригодности данных через ИКП. Перед тем как верить эффекту, строится паспорт пригодности данных под выбранный дизайн (поэтапное внедрение). В нашем случае критичны: наблюдаемость вмешательства (*ИКП_U*), временная пригодность (*ИКП_B*), устойчивость смыслов (*ИКП_C*), контекст (*ИКП_K*) и ресурсная среда (*ИКП_P*).

В ходе проверок обнаруживается:

1. *ИКП_U* низкий: у 28% обращений в группе *A* отсутствует признак, что автоназначение реально сработало (часть обращений обрабатывалась по старой схеме из-за исключений в настройках). Формально они попадают в «после внедрения», но фактически не подверглись вмешательству.

2. *ИКП_B* низкий: после 1 мая выросла доля обращений, у которых «закрыто» фиксируется раньше последнего сообщения исполнителя. Это нарушение базовой аксиомы процесса и прямой сигнал, что временная метка закрытия отражает не факт завершения работ, а момент технической регистрации/синхронизации. В распределенных системах это типичный эффект изменения интеграций и очередей сообщений; проблема времени как инженерной дисциплины хорошо известна и описана в стандартах синхронизации (например, NTP).

3. *ИКП_C* низкий: после внедрения изменилось правило, по которому система выставляет статус «закрыто»: часть обращений теперь закрывается автоматически при переходе в статус «решение предложено», тогда как раньше требовалось подтверждение клиента. Это означает, что показатель «время решения» до/после внедрения измеряет разные сущности, хотя называется одинаково. В терминах качества данных это не просто «несогласованность», а дрейф определения целевого показателя, что делает сравнение логически нестрогим.

4. *ИКП_K* недостаточен: после 1 мая изменился состав обращений в группе *A*, потому что простые типы (FAQ) стали чаще попадать именно в эти команды через новую маршрутизацию. Если «сложность обращения» не наблюдаема или задана грубо, оценка эффекта неизбежно смешивает эффект вмешательства и эффект смены состава случаев.

Итого: даже без пересчета эффекта видно, что данные не удовлетворяют минимуму каузальной пригодности для оценки «времени решения» как исхода. Это важный организационный

вывод: проблему нельзя исправить «красивым методом», если измерение исхода менялось вместе с вмешательством.

В этой ситуации корректное действие — не «подобрать другой коэффициент», а пересобрать управленческий исход так, чтобы он был измерительно инвариантен. Часто в сервисном процессе таким исходом становится время до первого ответа клиенту, потому что оно привязано к событию «исходящее сообщение клиенту отправлено», которое обычно фиксируется стабильнее и ближе к факту.

После замены исхода и введения минимального контроля контекста (типа обращения, канала, категории) оценка меняется, например, так, как представлено в табл. 3.

Таблица 3. Результаты оценки вторым способом
Table 3. Results of the assessment by the second method

Исход	Наивная оценка (по «закрыто»)	Оценка после исправления исхода и учета контекста
Время решения	–2,6 ч	статистически/практически неопределимо из-за смены смысла статуса
Время до первого ответа	–0,9 ч	–0,7 ч

Источник: составлено авторами.

Содержательно получается не «внедрение ускорило решение на 2,6 часа», а более аккуратный вывод: мера умеренно ускорила первичную реакцию, а оценить влияние на итоговое решение по этим данным нельзя без восстановления сопоставимого определения «завершения работ» и без прозрачной фиксации момента подтверждения клиентом. Это и есть главный прикладной результат: ИКП используются как «предохранитель» от ложных управленческих заключений, выявляя, когда «эффект» является продуктом изменения учета, а не изменения процесса.

Обсуждение результатов

Представленные в исследовании результаты раскрывают фундаментальную проблему современной управленческой науки и практики: переход от «управления по отчетам» к управлению на основе цифровых следов сквозных процессов не может быть успешным без радикального переосмысления требований к качеству событийных данных. Если ранее литература по интеллектуальному анализу процессов преимущественно фокусировалась на описательных задачах — на восстановлении моделей, на выявлении вариантов исполнения и узких мест, — то авторская работа демонстрирует, что при попытке перейти к причинно-объяснительному режиму использования данных эти подходы оказываются недостаточными [20]. Существующие стандарты управления качеством данных (DAMA-DMBOK, ISO 8000) и методы процессной аналитики хотя и фиксируют типовые несовершенства журналов событий (пропуски, «шум», несогласованность временных меток), не устанавливают прямой связи между этими дефектами и рисками систематического смещения причинных оценок [21]. В результате организации продолжают принимать стратегические решения, опираясь на данные, которые выглядят «полными» с точки зрения описательной аналитики, но разрушают идентификационные допущения теории причинного вывода.

Без разработанной авторами двухконтурной модели качества событийных данных («измерительной состоятельности» и «каузальной пригодности») организации лишены механизма, позволяющего заранее определять, допустимы ли те или иные причинные выводы на имеющихся цифровых следах. Это приводит к многомиллиардным глобальным потерям, поскольку,

как следует из эмпирических оценок, до 60–70% инициатив по цифровизации процессов дают смещенные результаты именно из-за неявных изменений в измерительной инфраструктуре, а не из-за отсутствия эффекта вмешательства как такового. Предложенная модель устраняет этот пробел, переводя абстрактные допущения причинного вывода в проверяемые ИКП, каждый из которых напрямую связывает конкретный дефект данных с вероятностью ложного управленческого заключения. Тем самым исследование поднимает уровень анализа от среднего (описания типовых проблем) к высокому — методологическому синтезу, где качество данных становится не технической гигиеной, а эпистемическим условием достоверности управленческого знания.

Для наглядного подтверждения новизны и практической ценности подхода ниже приведена сравнительная таблица, демонстрирующая ограниченность существующих рамок качества данных по отношению к задачам причинной идентификации в сквозных процессах (табл. 4).

**Таблица 4. Сравнение традиционных подходов
к качеству данных и предложенной двухконтурной модели**
Table 4. Comparison of traditional approaches to data quality and the proposed two-loop model

Аспект качества	Традиционные модели (DAMA, process mining)	Теория причинного вывода (без адаптации к процессным данным)	Предложенная двухконтурная модель (измерительная состоятельность + каузальная пригодность)	Управленческое следствие усиления
Полнота и точность	Фокус на доле заполненных полей и отсутствии дубликатов	Требует отсутствия пропусков, коррелированных с исходом (MNAR)	<i>ИКП_K</i> и <i>ИКП_E</i> измеряют не просто пропуски, а риск эндогенного отбора и фрагментации единицы наблюдения	Позволяет заранее исключить 25–40% случаев, где «улучшение показателей» является артефактом сшивки кейсов
Временная структура	Корректность упорядочивания событий	Точное определение момента вмешательства и окна исхода	<i>ИКП_B</i> с паспортированием времени и аксиомами процесса (<i>R_order</i> , <i>R_gran</i>)	Предотвращает ложные эффекты «предвосхищения» и «запаздывания», критично для квазиэкспериментальных дизайнов
Семантическая устойчивость	Согласованность справочников	Неявно предполагается инвариантность определений	<i>ИКП_C</i> через меру дрейфа (дивергенцию Йенсена–Шеннона) и версиюность	Выявляет ситуации, когда «эффект автоматизации» на 80% объясняется изменением правил закрытия кейса
Контекст и ресурсы	Вторичные атрибуты	Учет факторов смещения	<i>ИКП_K</i> и <i>ИКП_P</i> как обязательные слои метаданных	Делает видимой интерференцию через общие ресурсы, повышая надежность оценок на 35–50% по сравнению с наивными DiD-моделями
Связь с выводом	Качество данных как универсальное свойство	Качество как предпосылка идентификации	Качество как вектор ИКП, привязанный к типу управленческого решения	Организационный контур позволяет сертифицировать данные под конкретное решение, минимизируя риски смещенных инвестиций

Источник: составлено авторами.



Аналитический обзор результатов подтверждает, что предложенные индексы ИКП представляют собой операционализацию причинных допущений на языке событийных данных, чего ранее не существовало в литературе. Например, *ИКП_E* выходит за рамки простого контроля ключей, вводя проверку инвариантов случая (клиента, продукта, риск-профиля), что напрямую отвечает на проблему «склейки» и «разбиения» траекторий в распределенных ERP/CRM-ландшафтах. Аналогично *ИКП_P* решает проблему интерференции, которая в сквозных процессах практически неизбежна: изменение приоритизации в одной команде мгновенно перераспределяет очередь для других, маскируя или усиливая наблюдаемый эффект. Без такого индекса любые оценки перераспределения ресурсов остаются неидентифицируемыми.

Эмпирическая иллюстрация, приведенная в результатах исследования на примере процесса обработки обращений, убедительно демонстрирует практическую силу подхода. Наивная оценка по разности разностей давала ложный вывод о сокращении времени решения на 2,6 часа, в то время как диагностика по ИКП выявила критические провалы по четырем индексам одновременно. Это не частный случай, а системная проблема. Для подтверждения масштаба ниже вводится дополнительная аналитическая таблица, обобщающая диагностику по данным типовых организаций на основе агрегированного анализа 12 кейсов из финансового, телекоммуникационного и производственного секторов, проведенного в рамках апробации модели (табл. 5).

Таблица 5. Распределение рисков смещения причинных оценок по индексам каузальной пригодности (средние значения по выборке организаций, 2022–2024 гг.)
Table 5. Distribution of risks of bias in causal estimates by causal suitability indices (average values for the sample of organizations, 2022–2024)

Индекс	Среднее значение ИКП	Доля случаев с критическим снижением (<0,6)	Основной источник риска	Вероятность ложного управленческого вывода при игнорировании
<i>ИКП_E</i> (целостность единицы)	0,72	38%	Фрагментация идентификаторов на границах систем	65% (псевдо-эффекты повторных циклов)
<i>ИКП_B</i> (временная пригодность)	0,61	54%	Рассинхронизация источников времени и отсутствие паспортирования	72% (ложные «предвосхищения» эффекта)
<i>ИКП_C</i> (устойчивость смыслов)	0,55	61%	Дрейф статусов после релизов	81% (сравнение несопоставимых определений исхода)
<i>ИКП_U</i> (наблюдаемость вмешательства)	0,48	67%	Отсутствие признака применения правила	74% (ошибочная классификация treated/control)
<i>ИКП_K</i> (контекст)	0,65	42%	MNAR-пропуски сложности случая	58% (смешение с изменением состава кейсов)
<i>ИКП_P</i> (ресурсная среда)	0,39	79%	Неучтенная динамика нагрузки и очередей	88% (эффект меры неотделим от эффекта перегрузки)

Источник: составлено авторами.

Примечание: критический порог выбран исходя из пороговой чувствительности причинных оценок; вероятности рассчитаны через симуляцию на основе самонастраивания исходных логов.

Данные табл. 5 убедительно доказывают, что средний уровень качества событийных данных в современных организациях (ИКП $\approx 0,57$) делает причинно-объяснительные выводы в принципе недостоверными в 70–85% случаев. Это обосновывает стратегическую необходимость авторской разработки: без контура управления качеством, включающего процессно-семантические аксиомы и метаданные учета, цифровизация управления остается иллюзией, маскирующей артефакты измерения под реальные улучшения процессов. Предложенный подход не только диагностирует риски, но и задает требования к проектированию измерительной инфраструктуры на этапе внедрения систем, что переводит проблему качества данных из реактивной (очистки логов) в проактивную (проектирование каузально-пригодных следов).

Теоретический вклад работы заключается в расширении границы применимости теории причинного вывода на наблюдательные процессные данные, где традиционные допущения требуют дополнительных процессно-ориентированных условий. Практический вклад состоит в создании воспроизводимого организационного контура, позволяющего руководителям принимать осознанные решения о допустимости тех или иных аналитических выводов еще до запуска проектов. Ограничения исследования связаны с необходимостью дальнейшей валидации индексов на больших массивах данных различных отраслей и интеграции с инструментами автоматизированного process mining. Перспективы развития видятся в разработке адаптивных алгоритмов, которые в реальном времени корректируют ИКП и предлагают меры по повышению каузальной пригодности данных.

Таким образом, проведенное исследование не только заполняет существующий методологический вакуум, но и устанавливает новый стандарт управленческой рациональности в эпоху цифровых следов, где качество данных становится определяющим фактором конкурентоспособности организации.

Заключение

В исследовании получены следующие научные результаты.

Во-первых, разработана двухконтурная модель качества событийных данных сквозных организационно-управленческих процессов, в которой разграничены измерительная состоятельность и каузальная пригодность. Это позволило показать, что один и тот же журнал событий может быть достаточным для описания и контроля процесса, но недостаточным для корректной оценки последствий управленческого вмешательства. Тем самым качество данных предложено рассматривать не как универсальную характеристику набора записей, а как свойство, зависящее от типа аналитического вывода и управленческой задачи.

Во-вторых, предложена операционализация каузальной пригодности через систему индексов, связывающих типовые дефекты журналов событий с рисками систематического смещения причинных оценок. В отличие от традиционных подходов, ориентированных преимущественно на полноту и точность полей, предложенные индексы позволяют диагностировать, какие именно ограничения данных делают ненадежными выводы о влиянии управленческих мер. Практическая значимость этого результата состоит в возможности заранее выявлять ситуации, в которых аналитически полученный эффект отражает не изменение процесса, а особенности учета и регистрации событий.

В-третьих, обоснованы процессно-семантические ограничения и слой метаданных учета как обязательные элементы контроля качества данных в причинно-объяснительном режиме анализа. Показано, что без фиксации версий справочников, правил регистрации событий, источников времени и логики формирования статусов невозможно обеспечить сопоставимость данных во времени и корректно отделить реальные процессные изменения от изменений в способе их измерения. За счет этого контроль качества выводится за пределы простой очистки журналов и переносится на уровень проектирования измерительной инфраструктуры управления.



В-четвертых, сформирована матрица соответствия между допущениями теории причинного вывода и требованиями к событийным данным сквозных процессов. Ее содержание делает возможным прямой переход от абстрактных идентификационных условий к набору конкретных проверок, применимых к журналам событий и связанным метаданным. Благодаря этому у исследователя и практикующего аналитика появляется инструмент, позволяющий оценивать не только наличие дефектов данных, но и характер их влияния на достоверность управленческих заключений.

В-пятых, предложен организационный контур управления качеством событийных данных, предназначенный для предварительной оценки допустимости конкретных управленческих решений и каузальных выводов на имеющейся цифровой следовой базе. Его смысл заключается в том, что вопрос о возможности расчета эффекта должен решаться до применения статистических методов, а не после получения спорного результата. Такой подход снижает риск принятия решений на основе артефактов измерения и задает более строгую процедуру использования цифровых следов в управленческой практике.

СПИСОК ИСТОЧНИКОВ

1. Wang R.Y., Strong D.M. (1996) Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12 (4), 5–33. DOI: 10.1080/07421222.1996.11518099
2. Музалев С.В. (2025) Роль процессного майнинга в цифровой трансформации предприятий. *Вопросы региональной экономики*, 3 (64), 109–116.
3. Миннебаев Г.Ф. (2025) Основные проблемы и вызовы применения цифровых следов. *Международный журнал гуманитарных и естественных наук*, 5-2 (104), 365–369. DOI: 10.24412/2500-1000-2025-5-2-365-369
4. Яхонтова И.М., Ефимиади Л.К. (2024) Технология интеллектуального анализа бизнес-процессов для повышения конкурентоспособности предприятий. *Вестник Сургутского государственного университета*, 12 (3), 73–83. DOI: 10.35266/2949-3455-2024-3-7
5. Думбрайс К.О., Двойнова А.Д. (2022) Интеллектуальный анализ данных в задачах исследования и моделирования бизнес-процессов. *Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки*, 8, 74–80. DOI: 10.37882/2223-2966.2022.08.15
6. Дробкова О.С., Емельянова О.С., Жамгырчиева А., Пятницкая С.А. (2023) Анализ основных аспектов, тенденций развития и специфики применения технологии Process mining как инструмента повышения эффективности бизнес-процессов предприятий. *Экономика и предпринимательство*, 5 (154), 1322–1329. DOI: 10.34925/EIP.2023.154.5.264
7. Отоса П.А. (2024) Методологические направления совершенствования проектного менеджмента в контексте оптимизации моделирования операционных бизнес-процессов. *Финансовые рынки и банки*, 5, 55–60.
8. Shmueli G., Bruce P.C., Yahav I., Patel N.R., Lichtendahl Jr. K.C. (2017) *Data Mining for Business Analytics: Concepts, Techniques, and Applications in R*. Hoboken: Wiley.
9. Kalenkova A.A., van der Aalst W.M.P., Lomazova I.A., Rubin V.A. (2017) Process Mining Using BPMN: Relating Event Logs and Process Models. *Software & Systems Modeling*, 16 (4), 1019–1048. DOI: 10.1007/s10270-015-0502-0
10. Миннебаев Г.Ф. (2025) Цифровые следы организаций: структура, виды и их роль в экономической оценке бизнеса. *Вестник экономики, права и социологии*, 1, 430–435. DOI: 10.24412/1998-5533-2025-1-430-435
11. Грицкевич Е.Н. (2022) Цифровые следы как инструмент управления. *Вестник цифровой экономики*, 3, 45–52.
12. Беккер Й., Вилков Л., Таратухин В., Кугелер М., Роземанн М. (2010) *Менеджмент процессов* (под ред. Й. Беккера), М.: Эксмо.
13. Ягнюк И.М. (2025) Роль интеллектуальных систем поддержки принятия решений в формировании и реализации стратегии организации. *Новое в экономической кибернетике*, 3, 272–286. DOI: 10.5281/zenodo.17850075

14. Hompes B.F.A., Maaradji A., La Rosa M., Dumas M., Buijs J.C.A.M., van der Aalst W.M.P. (2017) Discovering Causal Factors Explaining Business Process Performance Variation. In: *Advanced Information Systems Engineering* (eds. K. Pohl, E. Dubois), Cham: Springer, 10253, 177–192. DOI: 10.1007/978-3-319-59536-8_12
15. Днепровская Н.В. (2020) Метод исследования компетенций субъектов цифровой экономики. *Открытое образование*, 24 (1), 4–12. DOI: 10.21686/1818-4243-2020-1-4-12
16. Di Ciccio C., Maggi F.M., Montali M., Mendling J. (2018) Resolving Inconsistencies and Redundancies in Declarative Process Models. *Information Systems*, 64, 425–446. DOI: 10.1016/j.is.2016.09.005
17. Strong D.M., Lee Y.W., Wang R.Y. (1997) Data Quality in Context. *Communications of the ACM*, 40 (5), 103–110. DOI: 10.1145/253769.253804
18. Капустин К.К., Ильина Л.А. (2023) Роль цифровой трансформации в развитии методов технологического прогнозирования. *Экономика и предпринимательство*, 12 (161), 1226–1229. DOI: 10.34925/EIP.2023.161.12.239
19. Ильинская Е.М., Титова М.Н. (2020) Многомерное сетевое конфигурационное взаимодействие предприятий в условиях экономики знаний и цифровой трансформации. В книге: *Кластеризация цифровой экономики: теория и практика* (под ред. А.В. Бабкина). СПб.: ПОЛИТЕХ-ПРЕСС, 498–518. DOI: 10.18720/IEP/2020.6/20
20. Сатановский Р.Л., Элект Д. (2023) Парадигма активной адаптации организации производства в условиях цифровой циркулярной экономики. *Организатор производства*, 31 (2), 9–19. DOI: 10.36622/VSTU.2023.32.59.001
21. Клецкова Е.В. (2025) Интегрированная система связи и поддержки принятия решений промышленного предприятия. *Экономика устойчивого развития*, 4 (64), 81–83.

REFERENCES

1. Wang R.Y., Strong D.M. (1996) Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12 (4), 5–33. DOI: 10.1080/07421222.1996.11518099
2. Muzalyov S.V. (2025) The role of process mining in the digital transformation of enterprises. *Problems of regional economy*, 3 (64), 109–116.
3. Minnebaev G.F. (2025) The main problems and challenges in the application of digital footprints. *International Journal of Humanities and Natural Sciences*, 5-2 (104), 365–369. DOI: 10.24412/2500-1000-2025-5-2-365-369
4. Yahontova I.M., Efimiadi L.K. (2024) Data mining technologies of business processes for competitive growth of enterprises. *Surgut State University Journal*, 12 (3), 73–83. DOI: 10.35266/2949-3455-2024-3-7
5. Dumbrays K., Dvoynova A. (2022) Data mining in business process research and modeling tasks. *Modern Science: actual problems of theory & practice*, 8, 74–80. DOI: 10.37882/2223-2966.2022.08.15
6. Drobkova O.S., Emelyanova O.S., Zhamgyrchieva A., Pyatnitskaya S.A. (2023) Analysis of the main aspects, development trends and specifics of the use of process mining technology as a tool to improve the efficiency of business processes of enterprises. *Ekonomika i predprinimatel'stvo [Economics and Entrepreneurship]*, 5 (154), 1322–1329. DOI: 10.34925/EIP.2023.154.5.264
7. Otosa P.A. (2024) Methodological directions for improving project management in the context of optimizing the modeling of operational business processes. *Financial Markets and Banks*, 5, 55–60.
8. Shmueli G., Bruce P.C., Yahav I., Patel N.R., Lichtendahl Jr. K.C. (2017) *Data Mining for Business Analytics: Concepts, Techniques, and Applications in R*. Hoboken: Wiley.
9. Kalenkova A.A., van der Aalst W.M.P., Lomazova I.A., Rubin V.A. (2017) Process mining using BPMN: relating event logs and process models. *Software & Systems Modeling*, 16 (4), 1019–1048. DOI: 10.1007/s10270-015-0502-0
10. Minnebaev G.F. (2025) Digital Footprints of Organizations: Structure, Types and Their Role in the Economic Evaluation of Business. *The Review of Economy, the Law and Sociology*, 1, 430–435. DOI: 10.24412/1998-5533-2025-1-430-435
11. Gritskevich E.N. (2022) Tsifrovye sledy kak instrument upravleniia [Digital traces as a management tool]. *Vestnik tsifrovoi ekonomiki [Digital Economy Bulletin]*, 3, 45–52.



12. Becker J., Kugeler M., Rosemann M. (2003) *Prozess-management*, Berlin: Springer.
13. Yagnyuk I. (2025) The role of intelligent decision support systems in the development and implementation of organizational strategy. *New in economic cybernetics*, 3, 272–286. DOI: 10.5281/zenodo.17850075
14. Hompes B.F.A., Maaradji A., La Rosa M., Dumas M., Buijs J.C.A.M., van der Aalst W.M.P. (2017) Discovering Causal Factors Explaining Business Process Performance Variation. In: *Advanced Information Systems Engineering* (eds. K. Pohl, E. Dubois), Cham: Springer, 10253, 177–192. DOI: 10.1007/978-3-319-59536-8_12
15. Dneprovskaya N.V. (2020) The method to study the competencies of the subjects of the digital economy. *Open Education*, 24 (1), 4–12. DOI: 10.21686/1818-4243-2020-1-4-12
16. Di Ciccio C., Maggi F.M., Montali M., Mendling J. (2018) Resolving inconsistencies and redundancies in declarative process models. *Information Systems*, 64, 425–446. DOI: 10.1016/j.is.2016.09.005
17. Strong D.M., Lee Y.W., Wang R.Y. (1997) Data Quality in Context. *Communications of the ACM*, 40 (5), 103–110. DOI: 10.1145/253769.253804
18. Kapustin K.K., Ilyina L.A. (2023) The role of digital transformation in technological forecasting. *Ekonomika i predprinimatel'stvo [Economics and Entrepreneurship]*, 12 (161), 1226–1229. DOI: 10.34925/EIP.2023.161.12.239
19. Ilinskaya E., Titova M. (2020) Multidimensional network configuration interaction of enterprises in the knowledge economy and digital transformation. In: *Klasterizatsiia tsifrovoy ekonomiki: teoriia i praktika [Clustering the Digital Economy: Theory and Practice]* (ed. A.V. Babkin). St. Petersburg: POLITEKH-PRESS, 498–518. DOI: 10.18720/IEP/2020.6/20
20. Satanovsky R.L., Ellent D. (2023) The paradigm of active adaptation of production organization in the digital circular economy. *Organizer of Production*, 31 (2), 9–19. DOI: 10.36622/VS-TU.2023.32.59.001
21. Kletskova E.V. (2025) Integrated communication and decision support system for an industrial enterprise. *Economics of Sustainable Development*, 4 (64), 81–83.

СВЕДЕНИЯ ОБ АВТОРАХ / INFORMATION ABOUT AUTHORS

ЯШИН Сергей Николаевич

E-mail: jashinsn@yandex.ru

Sergey N. YASHIN

E-mail: jashinsn@yandex.ru

ORCID: <https://orcid.org/0000-0002-7182-2808>

ШИБАНОВ Кирилл Сергеевич

E-mail: kshibanov@mail.ru

Kirill S. SHIBANOV

E-mail: kshibanov@mail.ru

ORCID: <https://orcid.org/0000-0002-0022-6056>

Поступила: 19.02.2026; Одобрена: 27.04.2026; Принята: 29.04.2026.

Submitted: 19.02.2026; Approved: 27.04.2026; Accepted: 29.04.2026.