Сборник научных статей
по итогам работы
Международного научного форума

НАУЧНЫЙ ДИАЛОГ: ТЕОРИЯ И ПРАКТИКА

Москва 2026

Коллектив авторов

Сборник научных статей по итогам работы
Международного научного форума

НАУЧНЫЙ ДИАЛОГ: ТЕОРИЯ И ПРАКТИКА

Том 2 из 2

Москва, 2026

УДК 330
ББК 65
С56

С56 Сборник научных статей по итогам работы Международного научного форума НАУЧНЫЙ ДИАЛОГ: ТЕОРИЯ И ПРАКТИКА (г. Москва, 22 октября 2026 г.). / Отв. ред. Д. Р. Хисматуллин. — Москва: Издательство Инфинити, 2026. — 129 с.
Том 2 из 2.

DOI 10.34660/INF.2026.16.87.034

Сборник материалов включает в себя доклады российских и зарубежных участников, предметом обсуждения которых стали научные тенденции развития, новые научные и прикладные решения в различных областях науки.

Предназначено для научных работников, преподавателей, студентов и аспирантов вузов, государственных и муниципальных служащих.

Все статьи, представленные в сборнике, проходят процедуру рецензирования. Рецензирование осуществляется членами редакционной коллегии и приглашенными экспертами по соответствующим научным направлениям.

УДК 330
ББК 65

© Издательство Инфинити, 2026
© Коллектив авторов, 2026

СОДЕРЖАНИЕ

Науки о Земле

Сопоставимость оценок выбросов CO_2 от авиационной деятельности в различных системах учета

Камеристая Мария Андреевна ····· 6

Влияние зон с особыми условиями использования территории на кадастровую стоимость земельных участков на территории г.Москвы

Фролова Софья Владимировна, Волнышева Алина Владимировна, Антонцева Анастасия Игоревна, Синенко Виктория Александровна ····· 13

Технические науки

Промпт-инжиниринг генеративных систем: парная приёмка русско-китайских результатов и бюджет проверки

Савельева Олеся Вячеславовна, Массаковский Владимир Михайлович, Штыренко Александр Сергеевич ····· 17

Операционно-семантическая модель производственного комплекса с использованием нечетких графовых отношений

Котеленко Сергей Анатольевич, Кузнецов Денис Сергеевич ····· 87

Justification of optimal discretization of geometric measurements of vertical steel tanks based on spline interpolation

Bazina Kristina Alekseevna, Yanchushka Anna Pavlovna ····· 101

Герметизация и гидроизоляция шарнирных сочленений судовых робототехнических систем в условиях агрессивных сред

Дружинина Н.В. ····· 112

Юридические науки

Ответственность маркетплейсов перед потребителями: проблемы и перспективы

Коногорова Альбина Олеговна ····· 117

Исторические науки

Российско-китайское экономическое сотрудничество в Арктике: оценки оборонных и разведывательных ведомств США с точки зрения безопасности (2024–2026 гг.)

Чэнь Шэнь ····· 123

DOI 10.34660/INF.2026.55.11.055

СОПОСТАВИМОСТЬ ОЦЕНОК ВЫБРОСОВ CO_2 ОТ АВИАЦИОННОЙ ДЕЯТЕЛЬНОСТИ В РАЗЛИЧНЫХ СИСТЕМАХ УЧЕТА

Камеристая Мария Андреевна
аспирант
Санкт-Петербургский государственный университет,
Санкт-Петербург, Россия

Аннотация. *Сопоставлены подходы к учету прямых выбросов CO2 от авиации в CORSIA, Руководящих принципах МГЭИК и методике Минприроды России № 371. Сравнение охватывает состав учитываемых полетов, данные о расходе топлива и расчетные коэффициенты. На примере показано, что при одинаковом расходе авиационного керосина оценки CORSIA и авиационного раздела российской методики совпадают. Расчет по базовым параметрам МГЭИК дает результат на 0,22 % ниже. При согласованных коэффициентах массовая и энергетическая формы расчета дают одинаковый результат. Для сопоставления отчетных данных предложено последовательно проверять границы учета, исходные данные и параметры расчета, а также различать прямые выбросы и показатели для определения компенсационных обязательств в рамках CORSIA. При анализе многолетней динамики авиационных выбросов такой порядок позволяет отделять изменения расхода топлива от влияния пересчета и изменения охвата полетов.*

Ключевые слова: *авиация, парниковые газы, CORSIA, МГЭИК, Методика № 371, авиационное топливо, коэффициенты выбросов.*

Abstract. *This study compares approaches to accounting for direct CO_2 emissions from aviation under CORSIA, IPCC Guidelines and the methodology approved by Order No. 371 of the Russian Ministry of Natural Resources and Environment. The comparison covers the scope of flights included, fuel consumption data and the factors used in the calculations. An example shows that CORSIA and the aviation section of the Russian methodology yield identical estimates for the same amount of aviation kerosene consumed. Calculations using the IPCC default parameters give an estimate that is 0.22% lower. Mass-based and energy-based calculations yield identical results when consistent factors are used. A procedure for comparing reported data is proposed, involving sequential checks of accounting boundaries, input data and*

calculation parameters, while distinguishing direct emissions from the metrics used to determine offsetting requirements under CORSIA. When analysing long-term trends in aviation emissions, this procedure helps distinguish changes in fuel consumption from the effects of recalculation and changes in flight coverage.

Keywords: *aviation, greenhouse gases, CORSIA, IPCC, Methodology No. 371, aviation fuel, emission factors.*

Введение

Оценки авиационных выбросов можно сравнивать, если известно, какие полеты учтены, как определен расход топлива и для каких целей выполнен расчет.

В работе рассмотрены три системы учета: Схема компенсации и сокращения выбросов углекислого газа для международной авиации (CORSIA), реализуемая Международной организацией гражданской авиации (ИКАО) [1, 2]; национальная инвентаризация по Руководящим принципам Межправительственной группы экспертов по изменению климата (МГЭИК) [3, 4]; расчет выбросов организаций по методике Минприроды России № 371 [5].

В расчет включены только прямые выбросы CO_2 от сжигания ископаемого авиационного керосина. Влияние оксидов азота, конденсационных следов и другие климатические эффекты авиации [6] количественно не оценивались.

Данные и методы

Для сравнения использованы документы CORSIA [1, 2], раздел 3.6 и таблица 3.6.4 тома 2 Руководящих принципов МГЭИК [3], таблица 1.2 того же тома [4], пункт 18.7 и таблица 1.1 приложения № 2 к российской методике определения объема выбросов [5].

Для однородного топлива расчет по массе имеет вид:

$$Q = M \cdot k \qquad (1)$$

где Q – выбросы, т CO_2; M – масса израсходованного топлива, т; k – коэффициент выбросов на единицу массы топлива, т CO_2/т топлива.

При расчете по энергетическому содержанию топлива используется формула:

$$Q = H \cdot EF = M \cdot NCV \cdot EF = M \cdot k_{eq} \qquad (2)$$

где $H = M \cdot NCV$ – энергетическое содержание топлива, ТДж; NCV – низшая теплота сгорания, ТДж/т; EF – коэффициент выбросов на единицу энергии, т CO_2/ТДж; $k_{eq} = NCV \cdot EF$ – коэффициент выбросов, пересчитанный на единицу массы топлива, т CO_2/т топлива. В обоих расчетах принято полное окисление углерода топлива.

Если $k = NCV \cdot EF$, обе формулы дают одинаковый результат.

При одной и той же массе топлива расхождение возникает из-за выбора параметров, а не способа записи формулы. При использовании нескольких видов топлива выбросы рассчитывают для каждого из них и суммируют.

Для сравнения принят условный расход 10 000 т топлива Jet-A/Jet-A1. Во всех основных вариантах перечень полетов и отчетный период одинаковы. Это позволяет отдельно оценить влияние коэффициентов. Их отклонение от значения CORSIA $k_0 = 3{,}16$ т CO_2/т топлива рассчитано по формуле:

$$\delta = \frac{k - k_0}{k_0} \cdot 100\ \% \qquad (3)$$

В энергетическом расчете вместо k в формулу (3) подставляют k_{eq}.

Результаты и обсуждение

Границы учета. В CORSIA международные полеты, для которых ведутся мониторинг, отчетность и верификация (MRV), следует отличать от полетов, на которые распространяются обязательства по компенсации [1, 2]. В инвентаризации МГЭИК внутренние и международные полеты разделяются по пунктам отправления и прибытия каждого этапа, а не по национальности перевозчика. Выбросы от международной авиации представляются отдельно от национального итога [3].

В пункте 18.7 российской методики предусмотрен расчет для внутренних и международных полетов с установленными исключениями. Этот пункт отсылает к процедурам тома IV Приложения 16 ИКАО [5]. Поэтому при сравнении с CORSIA следует учитывать общие расчетные процедуры и различия в охвате полетов (табл. 1).

Таблица 1. Границы и назначение сопоставляемых оценок

Система учета	Расчетный объект и назначение	Условие сопоставления
CORSIA	Охватываемые международные полеты эксплуатанта; мониторинг, отчетность и верификация (MRV) и отдельный расчет компенсационных обязательств	Отделять прямые выбросы от объема компенсации
МГЭИК	Национальная инвентаризация; выбросы от внутренних и международных полетов учитываются раздельно	Согласовать состав полетов и правила отнесения топлива
Методика № 371, п. 18.7	Выбросы от внутренних и международных полетов организации в установленных границах	Выделить сопоставимые полеты и авиационное топливо

Сопоставление коэффициентов. Для Jet-A/Jet-A1 в CORSIA и пункте 18.7 методики № 371 установлен коэффициент 3,16 т CO_2/т топлива [2, 5]. При расходе 10 000 т оба расчета дают 31 600 т CO_2.

Базовая низшая теплота сгорания авиационного керосина по МГЭИК составляет 44,1 ТДж/тыс. т, или 0,0441 ТДж/т. Коэффициент выбросов равен 71 500 кг CO_2/ТДж, или 71,5 т CO_2/ТДж [3, 4]. Чтобы получить выбросы на тонну топлива, эти значения перемножают: 0,0441 × 71,5 = 3,15315 т CO_2/т топлива.

При расходе 10 000 т расчет по базовым параметрам МГЭИК дает 31 531,5 т CO_2. Это на 68,5 т CO_2, или 0,22 %, меньше результата CORSIA (табл. 2).

Дополнительный расчет. Чтобы показать влияние другого значения теплоты сгорания, использованы параметры из таблицы 1.1 методики № 371. Эта таблица относится к стационарному сжиганию топлива. Для авиационного керосина в ней указаны низшая теплота сгорания 43,1 ТДж/тыс. т и коэффициент выбросов 71,5 т CO_2/ТДж [5].

Сначала теплоту сгорания переводят в расчет на одну тонну: 43,1 / 1 000 = 0,0431 ТДж/т. Затем умножают ее на коэффициент выбросов: 0,0431 × 71,5 = 3,08165 т CO_2/т топлива. При расходе 10 000 т получается 30 816,5 т CO_2, что на 783,5 т CO_2, или 2,48 %, ниже результата CORSIA. Этот пример показывает влияние выбора параметров, но не заменяет авиационный расчет по пункту 18.7 (табл. 2).

Таблица 2. Расчет для условной массы топлива 10 000 т

Вариант	Параметры	k , т CO_2/т	Q, т CO_2	δ, %
CORSIA	k = 3,16	3,16000	31 600,0	0,00
Методика № 371, п. 18.7	k = 3,16	3,16000	31 600,0	0,00
МГЭИК, базовые значения	NCV = 0,0441; EF = 71,5	3,15315	31 531,5	–0,22
Параметры табл. 1.1 методики № 371 (дополнительный расчет)	NCV = 0,0431; EF = 71,5	3,08165	30 816,5	–2,48

Примечание. k – коэффициент выбросов на единицу массы топлива, т CO_2/т топлива; Q – расчетные выбросы CO_2 при сжигании 10 000 т топлива, т CO_2; δ – относительное отклонение расчетных выбросов от результата CORSIA (31 600,0 т CO_2), %; NCV – низшая теплота сгорания топлива, ТДж/т; EF – коэффициент выбросов на единицу энергии, т CO_2/ТДж. Отрицательное значение δ означает, что расчетные выбросы ниже результата CORSIA. Для энергетического

расчета в столбце k приведен коэффициент, полученный умножением низшей теплоты сгорания на коэффициент выбросов на единицу энергии: k = NCV × EF. Последняя строка содержит дополнительный расчет с параметрами из раздела о стационарном сжигании топлива и не заменяет авиационный расчет по п. 18.7 методики № 371. Коэффициенты до расчета выбросов не округлялись.

Чувствительность и неопределенность. Для авиационного керосина в таблице 3.6.4 МГЭИК приведены нижнее и верхнее значения EF: 69,8 и 74,4 т CO_2/ТДж [3]. При постоянной NCV = 0,0441 ТДж/т расчетные коэффициенты составляют 3,07818 и 3,28104 т CO_2/т топлива. Отклонения от коэффициента CORSIA 3,16 составляют от −2,59 до +3,83 %.

Если изменять только NCV в пределах 42,0–45,0 ТДж/тыс. т, указанных в таблице 1.2 МГЭИК [4], при постоянном EF = 71,5 т CO_2/ТДж отклонения составят от −4,97 до +1,82 %. Эти интервалы показывают влияние каждого параметра по отдельности. Их нельзя объединять в общий диапазон различий между методиками или считать доверительным интервалом суммарных выбросов. Для оценки общей неопределенности нужны данные о неопределенности массы топлива, распределении значений параметров и их возможной взаимосвязи [7]. Фиксированный коэффициент в отчетности сам по себе не устраняет неопределенность фактических выбросов.

Исходные данные и контроль качества. В CORSIA установлены методы мониторинга расхода топлива, правила верификации и восстановления недостающих данных. В предусмотренных случаях для восстановления данных применяют инструмент ICAO CORSIA CERT [2]. МГЭИК предусматривает разные уровни детализации расчетов, контроль качества и документирование исходных данных [3]. Если расход топлива не разделен между внутренними и международными полетами, пункт 18.7 российской методики допускает расчет по налету и среднему часовому расходу. Средний часовой расход эксплуатант определяет по фактическому расходу топлива для соответствующего типа воздушного судна [5].

При постоянном коэффициенте k изменение массы топлива M на 1 % меняет расчетные выбросы Q на 1 %. Поэтому при выяснении причин расхождений нужно проверить, как получены данные о топливе и как восстановлены пропущенные значения.

Прямые выбросы и компенсационные обязательства. При использовании топлива, отвечающего критериям CORSIA, сокращение выбросов рассчитывают с учетом жизненного цикла топлива. Результат учитывают при определении обязательств по компенсации [2]. Такое сокращение не равнозначно уменьшению прямых выбросов CO_2 в двигателе. Поэтому прямые выбросы нельзя напрямую сравнивать с показателем после корректировки по жизненному циклу. В таблице 2 рассмотрено только ископаемое топливо, без таких корректировок.

Для сопоставления оценок сначала выбирают общий отчетный период и сопоставимый перечень полетов. Затем проверяют, отражают ли исходные данные заправленное, поставленное или фактически израсходованное топливо. Коэффициенты приводят к одним единицам: тоннам CO_2 на тонну топлива. Восстановление недостающих данных и корректировки по жизненному циклу описывают отдельно. Результат пересчета приводят вместе с исходными оценками. Такой пересчет нужен для анализа и не дает оснований заменять коэффициенты, предписанные для обязательной отчетности.

Выводы

При одинаковом расходе Jet-A/Jet-A1 CORSIA и авиационный раздел методики № 371 дают одинаковые оценки, поскольку используют коэффициент 3,16 т CO_2/т топлива. Расчет по базовым параметрам МГЭИК дает 3,15315 т CO_2/т топлива, что на 0,22 % ниже. Массовая и энергетическая формы расчета при согласованных коэффициентах дают одинаковый результат.

Одного сравнения коэффициентов недостаточно: необходимо также проверить перечень учтенных полетов, расход топлива и назначение показателя. Диапазоны изменения отдельных параметров не заменяют оценку общей неопределенности. Полученные численные результаты относятся к условному примеру. Для проверки расхождений в фактической отчетности нужны сопоставимые данные о полетах, расходе и свойствах топлива.

При анализе многолетней динамики авиационных выбросов в климатологических исследованиях необходимо проверять, менялись ли коэффициенты и границы учета. Это позволяет отделить изменения, связанные с пересчетом, от изменений, обусловленных расходом топлива.

Список литературы

1. International Civil Aviation Organization. Annex 16 to the Convention on International Civil Aviation. Environmental Protection. Vol. IV: Carbon Offsetting and Reduction Scheme for International Aviation (CORSIA). 2nd ed. Montreal: ICAO, 2023.

2. International Civil Aviation Organization. CORSIA Frequently Asked Questions. Questions 2.19, 3.18, 3.59, 3.63, 3.74–3.77. URL: https://www.icao.int/sites/default/files/environmental-protection/CORSIA/CORSIA-FAQ.html (дата обращения: 15.10.2026).

3. IPCC. 2006 IPCC Guidelines for National Greenhouse Gas Inventories / ed. by H. S. Eggleston et al. Vol. 2: Energy. Ch. 3: Mobile Combustion. Section 3.6, Table 3.6.4. Hayama: IGES, 2006.

4. IPCC. 2006 IPCC Guidelines for National Greenhouse Gas Inventories / ed. by H. S. Eggleston et al. Vol. 2: Energy. Ch. 1: Introduction. Table 1.2. Hayama: IGES, 2006.

5. Приказ Минприроды России от 27.05.2022 № 371 «Об утверждении методик количественного определения объемов выбросов парниковых газов и поглощений парниковых газов». Зарегистрирован в Минюсте России 29.07.2022, № 69451. Приложение № 2 к Методике количественного определения объема выбросов парниковых газов: п. 18.7, табл. 1.1.

6. Lee D. S., Fahey D. W., Skowron A. et al. The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018 // Atmospheric Environment. 2021. Vol. 244. Article 117834. DOI: 10.1016/j.atmosenv.2020.117834.

7. IPCC. 2006 IPCC Guidelines for National Greenhouse Gas Inventories / ed. by H. S. Eggleston et al. Vol. 1: General Guidance and Reporting. Ch. 3: Uncertainties. Hayama: IGES, 2006.

DOI 10.34660/INF.2026.95.41.056

ВЛИЯНИЕ ЗОН С ОСОБЫМИ УСЛОВИЯМИ ИСПОЛЬЗОВАНИЯ ТЕРРИТОРИИ НА КАДАСТРОВУЮ СТОИМОСТЬ ЗЕМЕЛЬНЫХ УЧАСТКОВ НА ТЕРРИТОРИИ Г.МОСКВЫ

Фролова Софья Владимировна
бакалавр
Российский университет дружбы народов,
г. Москва, Россия
Волнышева Алина Владимировна
бакалавр
Российский университет дружбы народов,
г. Москва, Россия
Антонцева Анастасия Игоревна
магистрант
Российский университет дружбы народов,
г. Москва, Россия
Синенко Виктория Александровна
ассистент
Российский университет дружбы народов,
г. Москва, Россия

Аннотация. *К зонам с особыми условиями использования территории, согласно Земельному и Гражданскому кодексу РФ, относятся территории с особым режимом использования земельных участков, устанавливаемым в публично-правовом порядке. В своей работе авторы рассматривают особенности осуществления государственной кадастровой оценки в условиях ограниченного использования территории на примере г. Москвы.*

Ключевые слова: *оценка недвижимости, кадастровая стоимость, объекты недвижимости, ЕГРН, земельный участок, ЗОУИТ.*

Annotation. *According to the Land Code and the Civil Code of the Russian Federation, zones with special land-use conditions are areas subject to a specific land-use regime established through a public-law procedure. In their work, the authors*

examine the specifics of conducting state cadastral valuation in contexts where land use is restricted, using the city of Moscow as an example.

Keywords: *real estate valuation, cadastral value, real estate properties, Unified State Register of Real Estate (EGRN), land plot, zones with special conditions for the use of territory (ZOUIT).*

Важную роль в развитии экономики, эффективном управлении земельными ресурсами и дальнейшем обеспечении справедливых условий для всех участников земельных отношений играет справедливое формирование кадастровой стоимости на объекты недвижимости.

Положениями Федерального закона от 03.07.2016 № 237-ФЗ «О государственной кадастровой оценке» проведение государственной кадастровой оценки возложено в компетенцию государственных бюджетных учреждений. Такое решение преследовало цель унификации методологии и повышения ответственности за результаты государственной оценки. Одновременно были расширены права правообладателей на оспаривание результатов оценки и исправление ошибок. Вместе с тем массовый характер оценки объективно ограничивает её точность применительно к каждому конкретному объекту и создаёт риск неполного учёта индивидуальных ценообразующих факторов.

По мнению авторов, одним из таких факторов являются зоны с особыми условиями использования территории (ЗОУИТ). Ограничения, вытекающие из режима таких зон, непосредственно затрагивают потребительские и инвестиционные качества земельных участков: сужают перечень допустимых видов деятельности, ограничивают строительный потенциал, создают дополнительные обязательства для правообладателей. Всё это образует экономическое основание для снижения стоимости участков, находящихся в границах ЗОУИТ.

Нормативные требования об учёте ЗОУИТ при проведении государственной кадастровой оценки закреплено в Методических указаниях о государственной кадастровой оценке, утверждённых приказом Росреестра от 04.08.2021 № П/0336: нахождение объекта в границах зоны с особыми условиями использования территории отнесено к числу ценообразующих факторов, подлежащих обязательному учёту [5]. Тем не менее массовый характер оценки и зависимость её результатов от полноты данных ФГИС ЕГРН создают условия, при которых фактическая степень отражения ограничений ЗОУИТ в кадастровой стоимости может существенно отличаться от нормативно предписанной.

По мнению авторов, кадастровая стоимость занимает центральную позицию в системе стоимостных характеристик недвижимости в России, в связи с тем, что используется для расчета земельного налога, арендной платы за отдельные категории публичных земель, а также в иных случаях, предусмотренных законодательством РФ.

При этом для земельных участков кадастровая стоимость имеет не только фискальное, но и управленческое значение, в связи с тем, что напрямую оказывает влияние на экономическую нагрузку собственника и на решения о пользовании, приобретении и развитии территории. Поэтому точность ее определения имеет практическое значение не меньшее, чем точность рыночной оценки, особенно в условиях массовой оценки большого числа участков.

До реформы 2016 года государственная кадастровая оценка регулировалась Федеральным законом № 135-ФЗ «Об оценочной деятельности в Российской Федерации», при этом сами работы выполнялись независимыми оценщиками на конкурсной основе. Такая модель приводила к неоднородности результатов, поскольку методики и подходы могли различаться в зависимости от исполнителя работ, региона и периода оценки. В результате во многих случаях кадастровая стоимость заметно расходилась с рынком недвижимости, что становилось основанием для многочисленных споров об оспаривании результатов оценки.

Именно поэтому реформа 2016 года была направлена на централизацию и унификацию процедуры, а также на повышение ответственности субъектов оценки.

Порядок проведения государственной кадастровой оценки установлен Федеральным законом № 237-ФЗ «О государственной кадастровой оценке» и в Методических указаниях о государственной кадастровой оценке, утвержденных приказом Росреестра от 04.08.2021
№ П/0336. Согласно положениям законодательства кадастровая стоимость рассматривается, как полученный на определенную дату результат оценки объекта недвижимости, определяемый на основе ценообразующих факторов в соответствии с законом и методическими указаниями о государственной кадастровой оценке.

Так, по мнению авторов, законодательством закреплено два принципиальных положения: во-первых, кадастровая стоимость является результатом именно оценки, а не произвольного административного решения; во-вторых, при ее определении должны учитываться экономически значимые характеристики объекта.

Также, по мнению авторов, методические указания прямо допускают учет ограничений и обременений, если они относятся к публично-правовым интересам и непосредственно влияют на использование земельного участка. Это создает нормативную основу для включения ЗОУИТ в число факторов, потенциально влияющих на формирование кадастровой стоимости. Таким образом, уже на уровне правовой конструкции кадастровая стоимость не должна игнорировать пространственные и правовые ограничения, если они экономически значимы.

Однако в результате анализа научной и практической литературы авторами отмечается, что учет таких факторов в государственной кадастровой (массовой)

оценке не всегда осуществляется полно и одинаково по всем территориям. Причина заключается в ограниченности исходных данных, сложности формализации правовых ограничений и неоднородности реальной структуры рынка.

Особое место в системе земельного и градостроительного регулирования РФ занимает территории ЗОУИТ, поскольку они напрямую связывают публичные интересы государства с режимом использования конкретных земельных участков, в том числе на территории
г. Москвы.

Правовая природа ЗОУИТ состоит в том, что ограничения возникают не по соглашению сторон, а в силу закона и решения уполномоченного органа, если земельный участок попадает в установленную границами зону.

Нормативное закрепление ЗОУИТ в российском праве связано прежде всего с главой XIX Земельного кодекса Российской Федерации, где установлен перечень видов зон с особыми условиями использования территорий.

По данным Росреестра [6] только за первые четыре месяца 2026г. в Москве в реестр были внесены сведения о 74 ЗОУИТ, в связи с чем, их общее количество на территории
г. Москва превысило 5,3 тыс.

Список литературы

1. Земельный кодекс Российской Федерации от 25.10.2001 № 136-ФЗ. Режим доступа: http://www.consultant.ru/.

2. Градостроительный кодекс Российской Федерации от 03.08.2018 № 342-ФЗ. Режим доступа: http://www.consultant.ru/.

3. Федеральный закон «О государственной регистрации недвижимости» от 13.07.2015 № 218-ФЗ /Правовая база данных. Режим доступа: http://www.consultant.ru/.

4. Федеральный закон от 03.07.2016 № 237-ФЗ «О государственной кадастровой оценке». Режим доступа: http://www.consultant.ru/.

5. Приказ Росреестра от 04.08.2021 № П/0336 «Об утверждении Методических указаний о государственной кадастровой оценке». Режим доступа: http://www.consultant.ru/.

6. Официальный сайт: https://rosreestr.gov.ru/

DOI 10.34660/INF.2026.61.75.035

ПРОМПТ-ИНЖИНИРИНГ ГЕНЕРАТИВНЫХ СИСТЕМ: ПАРНАЯ ПРИЁМКА РУССКО-КИТАЙСКИХ РЕЗУЛЬТАТОВ И БЮДЖЕТ ПРОВЕРКИ

Савельева Олеся Вячеславовна
доктор технических наук, независимый исследователь
Общество с ограниченной ответственностью «Ростелеком Информационные технологии» (ООО «РТК ИТ»)
Массаковский Владимир Михайлович
преподаватель
Херсонский государственный педагогический университет
Штыренко Александр Сергеевич
студент
Новосибирский государственный технический университет

Аннотация. *Исследуется совместная приёмка русско-китайских результатов генеративных систем при неполноте свидетельств и ограниченном бюджете проверки. Для обязательств различаются подтверждение, нарушение и незавершённая проверка. Получены внешние границы доли совместной приёмки и гарантированное число завершённых сертификатов. Бюджетная постановка сопоставлена с Set Union Knapsack, Densest k-Subhypergraph и стохастическим вычислением булевых функций. Воспроизводимая реализация включает 960 синтетических пар, 6720 решений семи процедур, 160 статических и 240 адаптивных бюджетных сравнений. Дополнительные 48 сравнений оценивают последствия ошибочного предположения о независимости проверок. Проверка устойчивости на 7656 случаях выявила и позволила устранить приёмку некорректного описания доступности свидетельств. По предоставленным журналам пересчитаны показатели демонстрации Kimi на 300 парах. Для внешнего испытания подготовлены 400 тестовых и 60 настроечных пар, экспертный протокол и единые запросы к Qwen, DeepSeek и Sarvam-M. Независимая экспертная разметка и новые модельные прогоны не выполнены. Результаты устанавливают свойства объявленной процедуры и её программной реализации; перенос на профессиональные документы требует внешнего испытания.*

Ключевые слова: *промпт-инжиниринг, генеративные системы, контракт результата, инженерия контекста, надёжность, верификация, агентные архитектуры, креативная индустрия, бизнес-аналитика, программная инженерия, университетское образование.*

Abstract. *This monograph studies joint acceptance of Russian–Chinese outcomes under incomplete evidence and a limited verification budget. Obligations have three audit states: confirmed, violated, and unresolved. The study derives outer acceptance bounds and guaranteed certificate completion. The budget problem is related to Set Union Knapsack, Densest k-Subhypergraph, and stochastic Boolean function evaluation. A reproducible implementation includes 960 synthetic pairs, 6720 decisions by seven procedures, 160 static and 240 adaptive budget comparisons. A further 48 comparisons assess misspecified dependence between audit outcomes. A robustness experiment over 7656 cases identifies and eliminates acceptance of malformed evidence-availability metadata. Metrics for a supplied Kimi demonstration over 300 pairs are recalculated from archived logs. A separate corpus of 400 test pairs and 60 development examples, an expert protocol, and identical request sets for Qwen, DeepSeek, and Sarvam-M are provided. External expert annotation and new model inference have not been performed. The findings concern declared contracts and their implemented audit procedure; generalisation to professional documents requires independent evaluation.*

Keywords: *prompt engineering, generative systems, outcome contract, context engineering, reliability, verification, agent architectures, creative industries, business analysis, software engineering, university education.*

Исследование последовательно рассматривает условия приёмки, структуру свидетельств и распределение проверок. В главах 2, 5 и 10 изложены формальная модель, вычислительный контроль и бюджетная постановка. Отраслевые главы показывают, как задать предметные обязательства; приложения содержат расчёты, данные и протокол внешнего испытания.

Введение

Проблема и предмет исследования

Профессиональное применение генеративных систем требует определить условия, при которых результат можно принять в работу. Помимо содержания ответа, необходимо проверить его основания, соблюдение полномочий и сохранность существенных условий при переводе. Промпт-инжиниринг рассматривается здесь как проектирование такого процесса, включая спецификацию результата и последующую приёмку.

The Prompt Report систематизирует разнородные техники постановки запросов [1]. Для профессионального применения требуется также определить, как проверяется результат и кто вправе его принять. Исследовательская задача ра-

боты состоит в согласовании спецификации, генерации и независимой приёмки при сохранении обязательств в русской и китайской версиях.

Объектом исследования служат профессиональные процессы, в которых человек работает совместно с языковой моделью, внешними источниками и программными инструментами. Предметом — принципы построения такого взаимодействия, при которых результат остаётся проверяемым, а действия — управляемыми. Цель работы состоит в разработке теоретико-методической рамки контракта результата и в демонстрации способов её операционализации в трёх предметных областях: творческой деятельности, бизнесе и программной инженерии.

Творческие, деловые и программные задачи объединяет общая структура проектирования, хотя критерии качества в них различаются. Художественную выразительность оценивают иначе, чем корректность программы или коммерческую полезность. В каждой области необходимо определить собственные основания приёмки.

Метод и вклад исследования

Исследовательский вопрос состоит в том, как вынести проверяемый вердикт о русско-китайской паре при неполноте свидетельств и распределить бюджет между взаимозависимыми проверками. Конструктивные примеры показывают недостаточность локальной проверки каждой версии, а бюджетный контрпример — недостаточность выбора по немедленному приросту.

Метод сочетает анализ первичных публикаций, доказательства свойств процедуры и вычислительный контроль. Объектами служат структурированные записи и графы обязательств. Контрольная реализация и статический бюджетный опыт выполнены 19 октября 2026 года; 20 октября добавлены адаптивный опыт и проверка устойчивости. Исходный код раннего запуска не входил в предоставленные материалы, поэтому новая реализация не объявляется его побайтовым восстановлением. Данные, решения и проверяющие программы включены в исследовательский комплект. Модельные наблюдения представлены только переданными журналами демонстрации, обозначенной как Kimi; её показатели пересчитаны.

Вклад работы состоит в постановке совместного аудита профессиональной RU/ZH пары, учёте незавершённых свидетельств и интерпретации бюджетной задачи через завершение сертификатов. Трёхзначная логика, супермодулярность и соответствующие классы оптимизационных задач имеют известных предшественников. Математическое сопоставление в подразделах 10.5–10.6 уточняет границы авторского вклада.

Выполненная часть включает 960 контрольных пар, 6720 решений семи процедур, 160 статических и 240 адаптивных бюджетных сравнений. Дополнительно проведены 48 проверок чувствительности к зависимости и проверка устойчивости на 7656 случаях. Математические свойства проверены конечным

перечислением; показатели переданных журналов Kimi пересчитаны. Новый текстовый корпус, экспертные формы и модельные запросы подготовлены для внешнего испытания, результаты которого пока отсутствуют.

Ссылки на источники приведены в квадратных скобках. Для изменяемых электронных ресурсов указана дата обращения.

1. Эпистемологические и инженерные основания

1.1. Промпт как компонент социотехнической системы

Языковая модель на основе архитектуры Transformer использует механизм внимания (attention) [2]. Инструкция без демонстраций служит исходным условием сравнения: эффект дополнительного компонента оценивают относительно этой конфигурации. В таблице приведён опубликованный пример; его результаты не характеризуют модели 2026 года.

Таблица 1
Zero-shot: инструкция без демонстраций (M01)

Статус	**Опубликованное испытание; в монографии не повторено.**
Источник и задача	Brown и соавт.: LAMBADA, GPT-3 175B; табл. 3.2 [3, с. 12].
Условия	Zero-shot, one-shot и few-shot; у zero-shot отличается формат задания.
Метрика и результаты	Точность: 76,2 %; 72,5 %; 86,4 % соответственно.
Контроль вывода	Число демонстраций и формат изменяются вместе; чистый эффект одного фактора не установлен.
Повторное испытание	Одинаковое ядро задачи, фиксированная версия модели, все языковые версии и единый критерий приёмки.

В этом опыте простота запроса служит экспериментальным условием. При профессиональном применении дополнительно проверяются основания ответа и соблюдение полномочий. Точность завершения текста сама по себе не характеризует пригодность управленческого решения.

Ролевая инструкция задаёт перспективу, аудиторию и терминологию. Формулировка «мировой профессор» сама по себе не подтверждает компетентность ответа. Изменения стиля и предметной корректности следует оценивать раздельно [4], а реализацию инструкции — сверять с документацией платформы [5].

Таблица 2
Ролевая и адресная инструкция (M04)

Статус	План авторского испытания; измерений нет.
Интервенция	Одна инструкция без роли против той же инструкции с ролью и явно заданной аудиторией.
Фиксируемые условия	Модель, данные, бюджет ответа и критерии; роль — единственный изменяемый компонент.
Измеряемые варианты	Приёмка; терминологические ошибки; оценка соответствия аудитории слепыми экспертами.
Отрицательный контроль	Роль выдающегося специалиста при неполных источниках: отсутствие вымышленных фактов и регалий.
Критерий решения	Изменение долей и экспертного согласия; преимущество до испытания не заявляется.

В учебном задании эффект ролевой инструкции проверяют по двум критериям: соответствию ответа аудитории и правильности содержания. Улучшение подачи не подтверждает точность фактов или атрибуции.

1.2. Правдоподобие, объяснение и основание знания

Ответ модели, сопровождающее его объяснение и свидетельство правильности — разные объекты. Последовательное объяснение способно сопровождать ошибочный вывод, тогда как краткий ответ может подтверждаться независимым вычислением. Turpin и соавторы показали, что видимая цепочка рассуждений не всегда отражает факторы, повлиявшие на ответ [6]. Это ограничивает её доказательную роль в условиях исследования, но не позволяет объявить недостоверными все объяснения любых моделей.

Проверяемое обоснование опирается на доступные объекты: исходные числа, применённое правило, документ или результат контрольной процедуры. В университетской работе студент должен представить аргументацию, которую другой специалист способен перепроверить. При высокой цене решения особенно важно, чтобы основание оценки не зависело от проверяемого ответа.

Обучение следованию инструкциям и человеческим предпочтениям меняет поведение модели [7]. Предпочтение пользователя, однако, не удостоверяет истинность ответа. Поэтому приёмка отдельно оценивает содержание и качество представления. Влияние удобства подачи на внимательность проверяющего требует самостоятельного измерения.

Совпадение нескольких ответов оценивают с учётом их источников: одна и та же исходная ошибка может повторяться во всех генерациях. Для проверки полезны другой документ, иной способ вычисления или независимая предметная

экспертиза. Повторная генерация должна отвечать конкретному проверочному вопросу.

1.3. Неопределённость и допустимый отказ

Основание принятия результата всегда имеет границу применимости. Отсутствующий показатель, конфликт версий или неясный фрагмент изображения нельзя устранить языковой уверенностью. В предложенной модели различаются неопределённость входа, изменчивость генерации и неопределённость проверки. Эти компоненты требуют разных вмешательств: получения данных, повторного испытания или уточнения критерия. Единый статус «модель не уверена» скрывает причину и затрудняет выбор следующего действия.

Семантическая энтропия использует вариативность смыслов ответов для обнаружения определённого класса конфабуляций [8]. Устойчиво воспроизводимая ошибка, однако, может сопровождаться низкой вариативностью. Поэтому для каждого такого показателя необходимо определить обнаруживаемые и пропускаемые ошибки и дополнить его фактологической проверкой.

Содержательный отказ указывает, каких данных недостаёт, как это влияет на решение и какое действие позволит продолжить работу. Избыточные отказы также учитываются при оценке: они увеличивают затраты пользователя и могут ограничивать пригодность системы.

Предположим, что автоматический ответ имеет ожидаемый ущерб pL, где p — оценённая вероятность существенной ошибки, L — условная тяжесть её последствия. Передача эксперту стоит C. При сопоставимых условиях эскалация оправдана, если $pL > C$. Это условная модель решения, а не готовый числовой норматив. Без калиброванной оценки p и анализа последствий не следует подставлять субъективную «уверенность» модели. Неопределённые величины нужно показать владельцу процесса и проверить чувствительность выбора к их изменению.

1.4. Методологический синтез и границы универсальности

Контракт результата включает цель, данные, полномочия, ожидаемый артефакт, свидетельства и процедуру верификации. Он связывает постановку задачи с решением о приёмке. Такая структура позволяет сопоставлять рабочие процессы при сохранении разных предметных критериев качества.

Для творческой концепции оцениваются соответствие брифу и самостоятельность решения; для аналитики — обоснованность выводов и корректность измерения; для кода — поведение и совместимость. Контракт требует указать основание каждой оценки. Его предполагаемые преимущества — обнаружение пропусков до генерации и возможность последующей проверки решения — нуждаются в испытании на реальных выборках.

Объединение методов требует проверки конкретных механизмов: обнаружения неподдержанных тезисов, сокращения времени приёмки и выявления

нарушения полномочий. Для каждого механизма нужен собственный протокол. Успех по одному показателю не подтверждает остальные свойства системы.

2. Формальная модель контракта результата

2.1. Спецификация и пространство допустимых результатов

Обозначим контракт K=(G,D,A,R,E,V), где G — цель; D — допустимые данные; A — разрешённые действия; R — множество допустимых результатов; E — основания; V — процедура приёмки. Множество R допускает несколько корректных решений. Спецификация должна быть достаточно точной для проверки и сохранять предметно обоснованную вариативность.

Противоречивый контракт невыполним независимо от качества модели. Например, точный итог нельзя вычислить, если цена отсутствует, а уточнение запрещено. До генерации следует проверить совместимость требований. Невозможность выполнить сами требования предполагает изменение постановки или данных; неспособность конкретной системы их выполнить — изменение реализации и повторное испытание.

Обязательные условия задаются предикатами b_i(r), а желательные свойства — оценками q_j(r). Приёмка возможна только при выполнении всех обязательных условий. Ранжирование по выразительности, скорости или стоимости начинается внутри допустимого множества. Такой порядок препятствует ситуации, когда высокое качество представления математически «компенсирует» ложный факт или неразрешённое действие. Весовая сумма полезна лишь после определения блокирующих ограничений и открытого обсуждения её весов.

Допустимо(r) $\Leftrightarrow b_1(r) \wedge b_2(r) \wedge \ldots \wedge b_m(r)$.

Если значение обязательного предиката b_i установить нельзя, результат получает статус незавершённой проверки. Этот статус отделяется от подтверждённых выполнения и нарушения. Для очереди экспертной приёмки различие существенно: непроверенный результат ещё не готов к использованию, но не обязательно требует полной переделки.

2.2. Трёхзначная приёмка и границы неполноты

Для каждого обязательства j задаётся состояние $v_j \in \{T,F,U\}$: T — проверка подтвердила выполнение; F — установлено нарушение; U — проверка не завершена. Итоговый вердикт обозначается A (принять), R (отклонить) или D (отложить). Правило δ(v) сначала учитывает установленные нарушения:

δ(v)=R при наличии F; δ(v)=A, если все v_j=T; δ(v)=D в остальных случаях.

Приоритет F существенен: непроверенный источник не отменяет уже обнаруженную арифметическую ошибку. Отсрочка требует указать недостающую проверку и необходимое действие. Она отличается от предметно корректного отказа, когда само задание предписывает сообщить об отсутствии данных.

Трёхзначные процедуры и частичная наблюдаемость рассматриваются, в частности, в runtime verification [9]. В настоящей работе они применяются к совместной приёмке русско-китайских результатов и планированию проверок.

Предпосылки: перечень обязательств фиксирован; подтверждённые T и F корректны относительно этого перечня; проверка U завершается состоянием T или F; версии задания и источников внутри аудита неизменны. При нарушении предпосылок сертификат пересматривается. Исполняемый модуль проверяет типизированные записи; оценка смысла свободного текста требует отдельной экспертизы.

Пусть в партии из $n>0$ пар c полностью подтверждены, f имеют установленное нарушение, а u остаются неопределёнными; $n=c+f+u$. Нижняя и верхняя границы доли совместной приёмки равны $J^{-}=c/n$ и $J^{+}=(c+u)/n$. Неопределённые пары остаются в знаменателе.

ПОЛОЖЕНИЕ 1. Для любой бинарной достройки непроверенных обязательств, сохраняющей известные T/F, итоговая доля J удовлетворяет $J^{-}\le J\le J^{+}$.

Доказательство. Каждая подтверждённая пара остаётся принятой, каждая нарушенная — непринятой. Из u неопределённых пар может быть принято от 0 до u. Следовательно, $c\le nJ\le c+u$. При свободной достройке границы достижимы: все неизвестные условия получают F либо все получают T. При дополнительных связях между свидетельствами достижимость требует отдельной проверки; сам внешний интервал сохраняется.

При $c=28$, $u=4$ и $n=40$ получаем интервал [0,70; 0,80] шириной 0,10. Он описывает границы при неполноте проверки. Интервал Уилсона в приложениях относится к другой задаче — оценке доли по выборочным наблюдениям в биномиальной модели.

Все 2187 векторов длины 7 из T/F/U и 16 384 их бинарные достройки проверены перечислением. Нижняя и верхняя границы совпали с минимальной и максимальной классической конъюнкцией. Такая проверка подкрепляет реализацию положения на конечной области; общее доказательство дано выше.

2.3. Типизированная неопределённость и операциональные определения

Различаются четыре типа неполноты: missing — данные отсутствуют; ambiguous — допускают несколько трактовок; conflicting — содержат несовместимые свидетельства; inaccessible — существуют, но недоступны по техническим причинам или из-за ограничений доступа. Для missing требуется получить значение, для ambiguous — уточнить определение, для conflicting — разрешить противоречие, для inaccessible — проверить возможность разрешённого доступа или восстановления связи.

Например, недоступный финансовый документ нельзя заменить предположительной оценкой только ради заполнения поля. Контракт определяет, какие части работы можно продолжить: подготовить структуру отчёта допустимо,

сформулировать неподтверждённый финансовый вывод — нет. В результате указываются причина неполноты и необходимое действие.

Операциональное определение связывает термин с процедурой измерения. «Качество», «клиент», «оригинальность» и «успех» не должны использоваться как самопонятные переменные. Следует установить единицу анализа, способ подсчёта, период и допустимые исключения. Если эти условия отличаются, одинаковое название показателя не обеспечивает сравнимость. Проблема сохраняется даже при безошибочной арифметике и грамотной формулировке ответа.

В одном отчёте могут подсчитываться обращения, в другом — уникальные плательщики. Прямое сравнение этих чисел создаёт ложное представление о динамике клиентской базы. Сначала необходимо согласовать определения и единицы анализа, затем решать, допустимо ли сопоставление. Такой пример позволяет выявить ошибку постановки ещё до обращения к модели.

2.4. Композиция и бюджет совместных свидетельств

Разбиение процесса на этапы превращает неявные предпосылки в интерфейсы передачи. Пусть каждый этап получает состояние s_i и формирует s_(i+1). Для системной приёмки нужно определить свойства, которые должны сохраняться при этих переходах: происхождение утверждений, неизменность исходных чисел, статус разрешений, версию применимого документа. Отсутствие такого инварианта допускает локально успешный, но глобально несостоятельный процесс.

Если исходное состояние удовлетворяет инварианту I, каждый переход сохраняет I, а финальная проверка подтверждает контракт, то принятый результат сохраняет I.

Утверждение доказывается индукцией по числу переходов при указанных предпосылках. Их выполнение реальной системой проверяется отдельно. В частности, при сжатии промежуточного текста необходимо контролировать сохранение связи вывода с источником.

Совместные свидетельства создают комплементарность проверок. Пусть U_i — непустое множество ещё непроверенных узлов свидетельств для пары i, в которой нет известного F. Выбор S означает, что каждый узел из S надёжно разрешается в T или F. Стоимость проверки узла j равна $a_j>0$; суммарный бюджет равен B.

$$g(S)=\Sigma_i 1[U_i \subseteq S];\ \text{задача: } \max g(S) \text{ при } \Sigma_j \in S\ a_j \le B.$$

Функция g считает сертификаты, которые гарантированно завершатся даже при положительных исходах всех выбранных проверок. При отрицательном исходе некоторые пары завершатся отказом раньше. Поэтому завершить проверку не означает принять результат.

ПОЛОЖЕНИЕ 2. При свободных бинарных исходах выбранных проверок минимальное число завершённых неопределённых сертификатов равно g(S).

Доказательство. Если $U_i \subseteq S$, неизвестных условий не остаётся: пара принята либо отвергнута. Это даёт нижнюю границу g(S). При всех проверенных условиях T любая пара с $U_i \not\subset S$ остаётся неопределённой. Такой исход достигает границы. Предпосылка свободных исходов указана явно; если предметные ограничения исключают достройку «все T», g остаётся нижней оценкой.

Бюджетная задача относится к полноте аудита, а не к вероятности правильного ответа. Она позволяет выбрать набор проверок до получения их исходов. Совместный источник для RU и ZH учитывается одним узлом только при совпадении версии и проверяемого отношения; наличие одинаковой ссылки для разных тезисов само по себе этого не обеспечивает.

3. Инженерия контекста и доказательная база

3.1. Контекст как ограниченный информационный ресурс

Размер контекстного окна и фактическая доступность нужных сведений различаются. Увеличение входа может одновременно добавлять полезные факты, конфликтующие версии и шум [10]. При отборе учитываются значение фрагмента для решения, его происхождение и стоимость обработки. Включение всей истории само по себе не гарантирует сохранения критичного правила.

Условную задачу отбора можно представить как выбор множества фрагментов S при ограничении $\Sigma l_i \leq B$, где l_i — длина, B — доступный бюджет. Полезность u(S) обычно неаддитивна: документ и его исключение могут быть полезны совместно, хотя отдельно создают ошибочное понимание. Следовательно, ранжирование отдельных фрагментов по одному баллу не всегда даёт достаточное решение. Требуется учитывать зависимости и проверять, сохраняется ли завершённое основание ответа после отбора.

Контекст длительной задачи включает постоянные условия и изменяемое состояние. Условия задают назначение, права доступа и критерии приёмки; состояние описывает выполненную работу, принятые решения и открытые вопросы. Смешение этих сведений затрудняет обновление. Поэтому сводка проекта должна содержать время и источник изменения, а критичные решения — сведения о том, кто их разрешил [11].

Промежуточное сжатие допустимо, если сохранены заранее определённые инварианты. Для бизнес-анализа это определения показателей и периода; для сценария — причинные связи и ограничения брифа; для кода — интерфейсы и проверенные результаты. Сжатие следует проверять на восстановлении необходимого решения. Если сокращённый пакет не позволяет объяснить, почему выбрана конкретная версия правила, он не является достаточным состоянием задачи, независимо от удобства и читаемости резюме.

3.2. RAG: разделение извлечения и использования

Извлечение источника и правильное использование найденного фрагмента — разные операции. Историческая работа RAG изучает обучаемую систему; её результаты нельзя представить чистым эффектом добавления документов в промпт.

Таблица 3
RAG: извлечение и генерация (M10)

Статус	**Опубликованное системное сравнение; в монографии не повторено.**
Источник	Lewis и соавт., Natural Questions, табл. 1 [12, с. 6].
Метрика и результаты	Exact match: DPR — 41,5 %; RAG-Token — 44,1 %; RAG-Sequence — 44,5 %.
Расчёт различия	RAG-Sequence минус DPR: 3,0 процентного пункта; RAG-Token: 2,6 пункта.
Ограничение	Архитектуры и обучение различаются; это не рандомизированная абляция одного промпта.
Авторское испытание	Раздельно измерить извлечение достаточного основания и приёмку ответа при одном фиксированном источниковом корпусе.

Найденный фрагмент может подтверждать лишь часть ответа. Реестр свидетельств связывает каждый существенный тезис с конкретным фрагментом источника. Если основания недостаточны, допустимым результатом становится отказ или запрос уточнения.

График воспроизводит три значения из одной таблицы первичного исследования. Единица сравнения — целая система; более высокий столбец не доказывает превосходства любого retrieval-промпта.

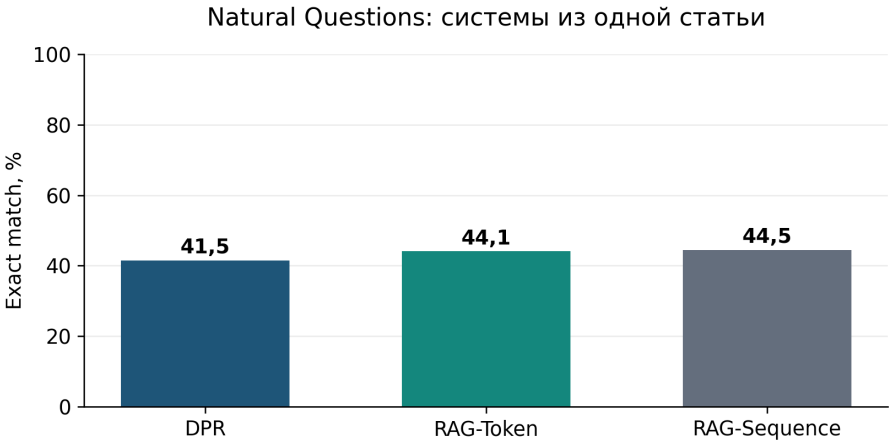


Рис. 1. Опубликованные результаты RAG и DPR на Natural Questions [12, с. 6]

Для новых русско-китайских заданий требуется контроль происхождения документов, полноты извлечения и сохранения чисел, единиц и отношений при переводе. Разрыв между извлечением и использованием исследуется отдельными абляциями. Интервалы и значимость различия здесь не вычисляются из округлённых процентов.

3.3. Происхождение, время и разрешение конфликтов

Источник необходимо характеризовать по дате публикации, периоду применимости, версии и полномочию устанавливать правило. Эти признаки отвечают на разные вопросы. Новая аналитическая статья не отменяет действующий внутренний регламент; поздняя поправка не применяется автоматически к событию до вступления в силу. Временной оператор задания должен определять, на какую дату требуется ответ и какие сведения допустимы для его обоснования.

Конфликт источников нельзя устранять выбором более уверенного текста. Сначала устанавливают, относится ли противоречие к одному предмету, периоду и определению. Затем проверяют иерархию полномочий и возможность установить применимое основание. Если этого недостаточно, конфликт сохраняют в результате. Такая процедура ограничивает право модели самостоятельно «согласовать» несовместимые документы посредством правдоподобного компромисса, которого нет ни в одном источнике.

Библиографическая проверка устанавливает существование источника и точность атрибуции. Затем проверяется, подтверждает ли конкретный фрагмент соответствующий тезис. Подлинный источник может оказаться нерелевантным. Формулировка вывода должна учитывать силу основания: описание возможно-

сти не доказывает гарантированного эффекта, а опыт на выборке — универсальной закономерности.

Реестр существенных утверждений позволяет различать наблюдения из источников, авторские выводы, определения и расчётные примеры. При редактуре следует проверять модальность и область применимости: замена «может снижать» на «устраняет» или удаление слов «на изученных задачах» меняет содержание вывода. Наличие DOI или работающей ссылки такой ошибки не обнаруживает.

3.4. Проверяемость научной аргументации

Проверке подлежат как отдельные факты, так и связи между ними. Несколько подтверждённых наблюдений ещё не обосновывают причинного вывода. В реестре следует указывать характер перехода к заключению и отделять авторскую интерпретацию от наблюдений источника.

Отношение источника к тезису оценивается тремя способами: поддерживает, противоречит, данных недостаточно. Последний статус фиксирует предел доступного свидетельства и необходимую дополнительную проверку. Он особенно существенен при неполном контексте и изменяющихся правилах.

Опубликованные показатели включают и ухудшение результата: one-shot — 72,5 % против 76,2 % для zero-shot; few-shot — 86,4 %. Разности составляют −3,7 и +10,2 процентного пункта.

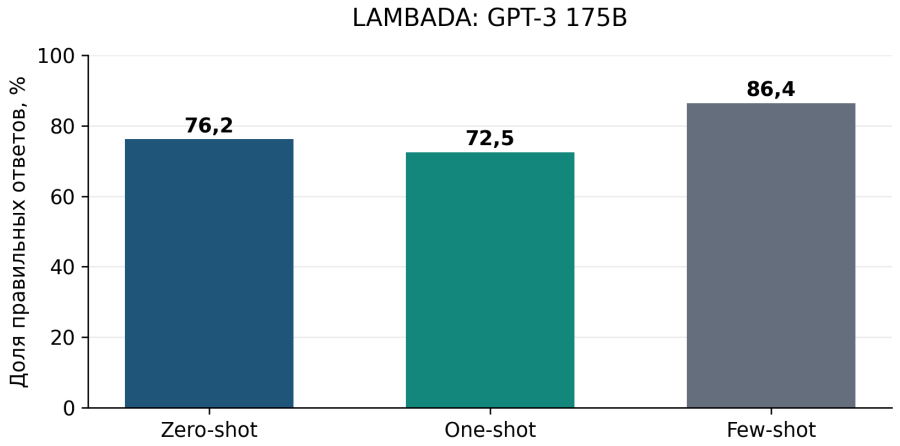


Рис. 2. Демонстрации и точность GPT-3 175B: опубликованное испытание LAMBADA [3, с. 12]

Пример иллюстрирует зависимость эффекта от постановки и масштаба модели. Формат zero-shot отличается, поэтому диаграмма описывает условия статьи и не изолирует число демонстраций. Доверительные интервалы, точные

числа успехов и современная воспроизводимость из этих процентов не восстанавливаются.

4. Методы управления генеративным процессом

4.1. Демонстрации, покрытие и чувствительность постановки

Одна демонстрация вводит пример требуемого отображения входа в выход. Чувствительность к формулировке и организации примеров требует отдельного контроля [13]. Полезность не следует из самого наличия примера.

Таблица 4
One-shot: одна демонстрация (M02)

Статус	**Опубликованное испытание; не авторские модельные данные.**
Источник	Brown и соавт., GPT-3 175B, LAMBADA, табл. 3.2 [3, с. 12].
Результат	One-shot — 72,5 %; zero-shot — 76,2 %; разность −3,7 процентного пункта.
Конфаундер	Формат задания отличается; результат не устанавливает вред любой одиночной демонстрации.
Повторное испытание	При неизменном формате сравнить несколько заранее отобранных примеров; не выбирать лучший по тестовой выборке.
Обязательная регистрация	Содержимое и порядок примеров, критерий отбора, разбиение данных, модель, бюджет и все отказы.

Демонстрация должна показывать не только удачный ответ, но и допустимое обращение с недостаточностью. Иначе обучение формату может поощрить заполнение отсутствующих сведений. Проверка отделяет точность задачи от сходства с примером.

Набор демонстраций задаёт покрытие типов задания. Его качество связано с разнообразием проверяемых случаев, а не только с количеством. Препринт Spurious Prompts 2026 года дополнительно мотивирует контроль нерелевантных вмешательств [14].

Таблица 5
Few-shot: набор демонстраций (M03)

Статус	**Опубликованное испытание; на модели 2026 года не переносится автоматически.**
Источник и результат	Brown и соавт., LAMBADA, GPT-3 175B: few-shot — 86,4 % [3, с. 12].
База сравнения	Zero-shot — 76,2 %; прибавка 10,2 процентного пункта.
Расчёт	Относительная прибавка ≈13,39 %; она отличается от прибавки в процентных пунктах.
Авторский протокол	Сравнить фиксированные наборы с одинаковым бюджетом; варьировать порядок и тематическое покрытие.
Защита оценки	Удержанная выборка; запрет отбора примеров по финальному тесту; сохранение отрицательных случаев.

Сочетание отбора, порядка и формата образует отдельную конфигурацию. Если набор помогает только знакомым темам, его полезность ограничена этой областью. Повторы на одном задании характеризуют изменчивость ответа и не увеличивают число независимых задач.

4.2. Рассуждение, проверка и внешнее свидетельство

В классическом CoT примеры содержат цепочки решения. Для современных reasoning-моделей применимость этого вмешательства проверяется по документации и отдельному испытанию [15]. Видимое объяснение не принимается за доказательство внутреннего механизма.

Таблица 6
Chain-of-Thought: демонстрации решения (M05)

Статус	**Опубликованное испытание PaLM 540B; в монографии не повторено.**
Источник	Wei и соавт., GSM8K, табл. 2 [16, с. 21].
Условия и результаты	Standard — 17,9 %; CoT — 56,9 %; для арифметики используются восемь демонстраций решения.
Проверенный расчёт	56,9−17,9=39,0 процентного пункта.
Ограничение	Эффект относится к задаче, масштабу и протоколу статьи; не гарантирует достоверности произвольной цепочки.
Авторское испытание	Фиксировать требуемые исходные величины и ответ; отдельно проверять числовой результат внешним вычислением.

Для профессионального решения ценность имеют доступные проверке основания: исходные данные, операции и предметные ограничения. Увеличение длины объяснения не заменяет проверку этих объектов.

Self-consistency меняет стратегию декодирования и агрегирования. Согласие нескольких ответов одной модели может сохранять общую ошибку, поэтому внешнее основание остаётся самостоятельным элементом контракта.

Таблица 7
Self-consistency: агрегирование ответов (M06)

Статус	**Опубликованное испытание; собственных генераций нет.**
Источник	Wang и соавт., PaLM 540B, GSM8K, табл. 2; HTML v4 [17].
Протокол	40 сэмплированных ответов на задачу; результаты усреднены по 10 запускам; база — greedy CoT.
Результаты	56,5 % → 74,4 %; прибавка 17,9 процентного пункта.
Несопоставимость баз	56,5 % в этой работе не заменяется на 56,9 % из статьи Wei; протоколы различаются.
Авторское испытание	Сравнивать при фиксированном бюджете; регистрировать правило агрегации, число ответов, отказы и общие ошибки.

Агрегация применима при определённом отношении эквивалентности ответов. Для свободного творческого текста такое отношение требует экспертного обоснования. Повторяемость формулировки и предметная правильность измеряют разные свойства.

4.3. Самокоррекция и независимость критика

Критик должен указывать проверяемое нарушение, место и основание. Внутренняя самокоррекция использует ответ той же модели; её следует отличать от проверки калькулятором, источником или экспертом.

Таблица 8
Самокоррекция без внешней обратной связи (M07)

Статус	Опубликованное испытание; собственных ответов моделей нет.
Источник	Huang и соавт., GPT-4, GSM8K, табл. 3 [18, с. 4].
Условия	Исходный ответ — 1 вызов; один раунд самокоррекции — 3; два раунда — 5.
Результаты	95,5 %; 91,5 %; 89,0 %. Изменения к базе: −4,0 и −6,5 процентного пункта.
Граница вывода	Результаты относятся к intrinsic self-correction; не доказывают бесполезности всякой внешней проверки.
Авторский протокол	Разделить сохранённые правильные ответы, исправления и новые ошибки; сравнить при равном бюджете.

Улучшение редакционной гладкости нельзя представлять как исправление содержания. Любая существенная правка пересматривает источник и критерий приёмки. Если независимого основания нет, вердикт критика остаётся гипотезой.

Рост числа вызовов сопровождается снижением точности в данном испытании. Это отрицательный результат для конкретного метода и протокола; он сохраняется в анализе наряду с положительными эффектами.

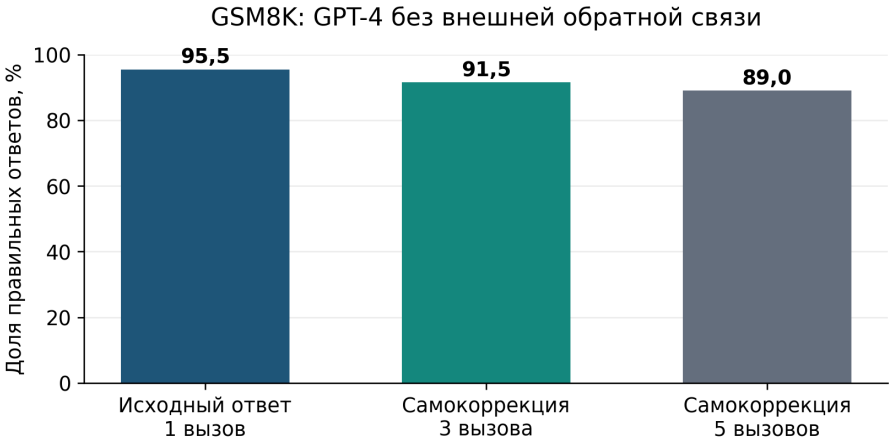


Рис. 3. Опубликованная точность GPT-4 при внутренней самокоррекции [18, с. 4]

Разность −6,5 процентного пункта вычислена из опубликованных округлённых долей. Парная значимость из них не выводится: необходимы индивидуальные исходы. Практическая альтернатива — проверка ошибки по независимо полученному основанию с повторной приёмкой исправленного результата.

4.4. Декомпозиция, оптимизация и цена сложности

Разделение оправдано различием предметных операций. Передача между этапами создаёт дополнительную стоимость и возможность утраты данных. Для проверки необходима исходная одноэтапная реализация.

Таблица 9
Декомпозиция и последовательность этапов (M08)

Статус	**Авторский план испытания; наблюдаемых эффектов нет.**
Интервенция	Одноэтапное решение против извлечения фактов, расчёта и редакции с типизированными интерфейсами.
Контроль	Один корпус, версия модели и лимит суммарных затрат; одинаковая финальная приёмка.
Измерения	Совместная RU/ZH приёмка, ошибки передачи, стоимость принятого результата, задержка.
Абляция	Убрать проверку промежуточного объекта; проверить, сохраняются ли числа, ссылки и полномочия.
Критерий решения	Отдельно показать пользу этапов и стоимость усложнения; не усреднять критический отказ со стилем.

Полная стоимость $C=\Sigma C_i+C\text{передачи}+C\text{приёмки}$ относится к рабочему циклу, включая отказы и повторы. Численная оценка до запуска не подставляется. Многоэтапность становится полезной архитектурой только после подтверждения на предметных заданиях.

DSPy оптимизирует программу вызовов, включая демонстрации [19]. До оптимизации следует проверить, не получает ли профессионально непригодный ответ высокий балл.

Таблица 10
DSPy: компиляция программы вызовов (M09)

Статус	Опубликованные тестовые результаты; не авторская репликация.
Источник	Khattab и соавт., GSM8K, GPT-3.5, табл. 1 [19, с. 8].
Сопоставляемые строки	vanilla / none: test 25,2 %; vanilla / bootstrap×2: test 61,7 %.
Расчёт	Прибавка 36,5 процентного пункта; dev-значения в неё не включены.
Ограничение	Изменяется процедура компиляции; итог не изолирует единственную текстовую фразу.
Авторское испытание	Разделение разработки и удержанного теста; фиксировать оптимизатор, бюджет поиска и функцию приёмки.

5. Вычислительные эксперименты

5.1. Вопросы исследования и фиксированный дизайн

Собственное исследование проверяет контур приёмки. RQ1: какие нарушения пропускают схема, локальные проверки и межъязыковое согласие? RQ2: что теряется при удалении парной проверки и состояния U и сохраняются ли решения при допустимых преобразованиях входа? RQ3: когда локальный выбор уступает совместному бюджетному выбору и что меняется при адаптации к наблюдаемым исходам? Эти вопросы определяют измеряемые результаты независимо от названия метода генерации [20].

Таблица 11
Дизайн контрольного эксперимента

Объекты	**24 сконструированных контракта общего типа «смета, срок, атрибуция, право на черновик»; три предметных контекста, по 8 контрактов.**
Контрольный корпус	40 сценариев на контракт: 4 допустимых, 30 одиночных/общих нарушений, 2 парных, 3 неопределённых и 1 смешанный. Всего 960 пар.
Сравнение	4 исходные процедуры, 2 абляции и полная процедура; 6720 зарегистрированных решений.
Бюджетный контроль	20 новых графов, по 10 узлов и 60 требований; бюджеты 1–8; 160 сравнений полного перебора с локальным выбором, пакетной эвристикой и просмотром пар.
Воспроизводимость	Новая реализация от 19.10.2026, seed=20261006; исходные структуры, мутации, решения и проверки включены в комплект. Это не внешняя предрегистрация и не побайтовое восстановление прежнего запуска.

В настоящем вычислительном эксперименте реальные языковые модели не участвовали; единственный модельный прогон работы — демонстрационный (модель Kimi, приложение И). Объектом измерения является исполняемая проверка структурированных результатов. Подход к многомерной оценке соотнесён с HELM [21]; показатели здесь определяются замкнутым корпусом и не оценивают отраслевую распространённость ошибок.

Единица воздействия — мутация профессиональной записи, а единица исходной спецификации — контракт. В каждом контексте варьируются восемь числовых профилей; одни и те же семейства нарушений повторяются. Поэтому 960 записей не считаются 960 независимыми производственными задачами. Сравнение использует полный фиксированный корпус, а не статистическую выборку из эксплуатации.

Таблица 12
Семейства воздействий и нормативные исходы

Допустимые изменения	**Каноническая запись, эквивалентная числовая нотация, согласованная альтернативная величина и редакционная помета. Ожидается accept.**
Нарушения	Арифметика, валюта, свидетельство, полномочие, авторство, срок, тип, статус, обязательное поле и подтверждение цитаты. Ожидается reject.
Парные расхождения	Оба локальных результата допустимы, но различаются количество либо срок. Ожидается reject.
Неполнота	Зарегистрированный источник недоступен в RU, ZH либо в обеих версиях; прочих нарушений нет. Ожидается defer.
Смешанный случай	Недоступное свидетельство и обнаруженная арифметическая ошибка. Ожидается reject независимо от U.

Поведенческие отношения соотнесены с CheckList [22]. Метки определяются спецификацией воздействия до классификации; проверяющие процедуры не получают ожидаемую метку. В опыте главы 5 RU и ZH представлены каноническими структурами с объявленной грамматикой чисел и полей. Натуральные китайские тексты находятся в отдельном корпусе приложения З. Качество перевода и полнота свободной формулировки программным опытом не измеряются.

5.2. Результаты сравнения процедур

В корпусе 96 допустимых пар, 792 пары с нарушением и 72 неопределённые. График показывает два разных вида ошибки: приёмку нарушенного результата и отказ допустимому. Частоты относятся к этому фиксированному набору.

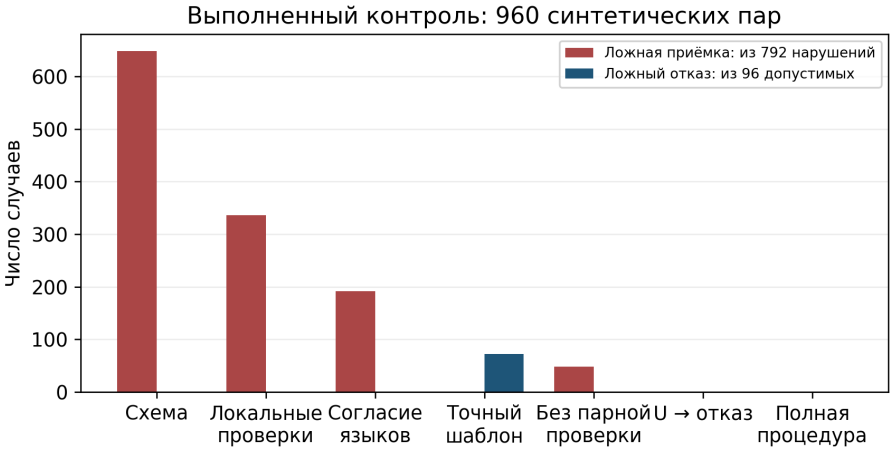


Рис. 4. Выполненное сравнение процедур на 960 синтетических парах

Проверка схемы приняла 648 нарушенных пар; локальная процедура — 336; согласие языковых записей — 192. Точный шаблон не принял нарушений, но отверг 72 допустимые вариации. Полная процедура классифицировала все предусмотренные сценарии по спецификации. Этот результат подтверждает замкнутое покрытие проверок, а не стопроцентную надёжность генеративной системы.

Таблица 13
Результаты приёмки: абсолютные числа и основания знаменателей

Процедура	**Ложная приёмка из 792 нарушений / ложный отказ из 96 допустимых / верная отсрочка из 72 неопределённых.**
Схема	648 / 0 / 0.
Локальная арифметика и ограничения	336 / 0 / 0.
Согласие RU/ZH	192 / 0 / 0.
Точный исходный шаблон	0 / 72 / 0.
Без парной проверки	48 / 0 / 72.
U заменено на отказ	0 / 0 / 0.
Полная трёхзначная процедура	0 / 0 / 72; 960 из 960 решений совпали с нормативным исходом корпуса.

Общая доля верных решений зависит от намеренно выбранной смеси классов. Поэтому она не используется как универсальный рейтинг. Схема дала 240/960, локальная проверка — 552/960, согласие — 696/960, точный шаблон — 816/960, две абляции — 912/960 и 888/960. Более информативны отдельные типы ошибок и причины расхождений.

Биномиальные p-значения для повторяющихся мутаций не вычисляются: независимость не установлена. Статистические методы приложений [23] применяются к иной постановке, где определены выборка и вероятностные предпосылки. Числовой результат этого эксперимента является точным счётом конечного корпуса.

5.3. Абляция и границы экспериментального вывода

Абляция отделяет вклад обязательства от общего размера системы. Удаление парных проверок сохраняет 96 допустимых решений и 72 правильные отсрочки, но пропускает 48 межъязыковых расхождений. В этих случаях каждое локальное решение соответствует допустимому диапазону, однако RU и ZH описывают разные профессиональные решения.

Таблица 14
Абляция неизвестности и парной совместимости

Полная процедура	**Пара принята, только если все локальные и парные обязательства подтверждены; известное F приоритетно перед U.**
Без парной проверки	48 ложных приёмок: 24 расхождения количества и 24 расхождения срока.
Без третьего состояния	72 отсрочки превращены в отказы; нарушенные и допустимые случаи классифицированы верно.
Общее согласие	192 ложных приёмки демонстрируют, что одинаковая ошибка сохраняет межъязыковое согласие.
Смешанный случай	Во всех 24 сценариях «неизвестный источник + ложная арифметика» полная процедура выдаёт reject.

Такое разделение позволяет обсуждать безопасность и полезность раздельно. Запрет принять непроверенное не требует объявлять его ложным. Внешний эксперт или судья должен разрешить нужное обязательство по основанию; модельный вердикт не получает статус доказательства только из-за уверенного языка [24].

Результат 960/960 не переносится за пределы объявленного корпуса. Мутации построены вокруг известных обязательств; произвольные скрытые свойства и новые виды нарушения не покрыты. Культурная приемлемость свободной

адаптации, полнота требований и истинность извлечения из документа остаются вне автоматического проверяющего.

Сравниваемые исходные процедуры намеренно ограничены своими функциями. Схема не обязана доказывать арифметику, а равенство языковых записей — истинность. Поэтому эксперимент измеряет цену отсутствующих проверок, а не превосходство над всеми существующими системами верификации. Сильная система с теми же обязательствами должна получить эквивалентные результаты.

Для усиления внешней валидности нужны независимые разработчики нарушений, новые контракты и слепая предметная разметка. Включение реальных ответов моделей требует отдельного разбиения задач и фиксации условий генерации. Разногласия экспертов сохраняются, а не преобразуются автоматически в ошибку модели. Калибровка численных прогнозов [25] не смешивается с логическим состоянием U.

Практический вклад корпуса — проверяемая карта провалов: общая ошибка, нарушение полномочия, ложная атрибуция, межъязыковое расхождение и неполнота свидетельства имеют разные диагностические признаки. Метод не заменяет источникового анализа, но задаёт интерфейс, в котором его результат можно проверить и воспроизвести.

5.4. Бюджетный вычислительный контроль

В новом бюджетном опыте использован генератор random.Random(20261006) стандартной библиотеки Python. Последовательно созданы 20 графов: сначала 10 целых стоимостей из равномерного распределения на {1,2,3}, затем 60 требований; для каждого требования размер равномерно выбирается из {1,2,3,4}, а узлы выбираются без повторения из десяти доступных. Для бюджетов 1–8 полный перебор всех 1024 наборов сравнивается с тремя эвристиками. Исходные графы и выбранные наборы сохранены; этот генератор определяет новую воспроизводимую серию, а не восстановление неизвестных исходных графов ранней редакции.

Таблица 15
Результаты статического выбора проверок

Область	20 новых графов × 8 бюджетов = 160 сравнений; 10 узлов и 60 сертификатов на граф.
Цель	Максимальное гарантированное число завершённых сертификатов g(S).
Локальный выбор	134 равенства; 26 строгих преимуществ точного выбора; сумма разрывов 37, максимум 3.
Пакетная эвристика	142 равенства; 18 строгих преимуществ точного выбора; сумма разрывов 25, максимум 3.
Просмотр пар	143 равенства; 17 строгих преимуществ точного выбора; сумма разрывов 25, максимум 3.
Проверка и область вывода	Все 160 оптимумов подтверждены отдельным перебором комбинаций. Результаты относятся к сохранённым синтетическим графам.

Локальная эвристика выбирает один узел по отношению немедленного прироста g к дополнительной стоимости и останавливается при нулевом приросте. Пакетная эвристика рассматривает недостающие узлы каждого сертификата как единый кандидат. Просмотр пар рассматривает одиночные узлы и все пары узлов. В каждой эвристике при равенстве отношения выбираются больший прирост, меньшая стоимость и меньший числовой код набора. Все алгоритмы используют одну целевую функцию, один граф и один бюджет. Условные стоимости не являются измеренными деньгами или рабочим временем.

На рисунке показано среднее число гарантированно завершённых сертификатов по 20 фиксированным графам для каждого бюджета. Четыре линии относятся к полному перебору и трём эвристикам. Усреднение описывает только этот контрольный набор; доверительные интервалы для распространённости эффекта в организациях не вычисляются.

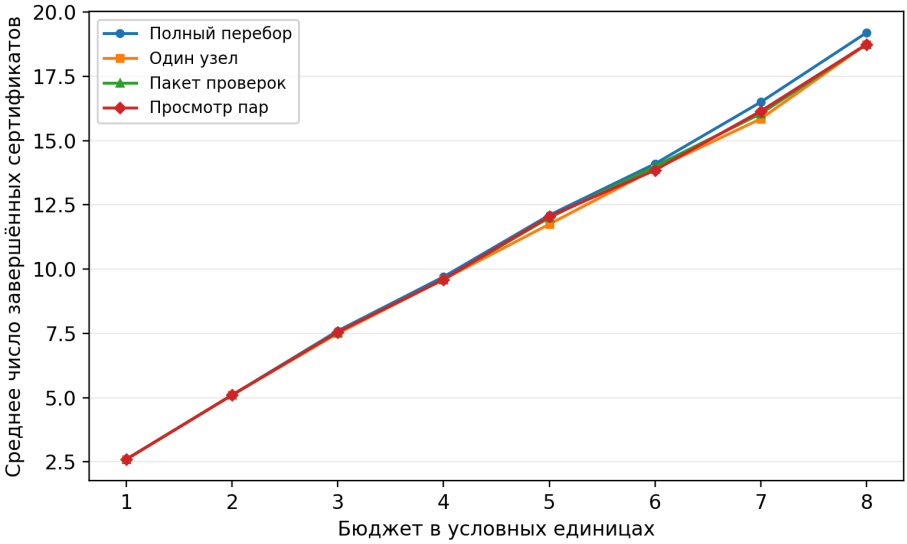


Рис. 5. Новый бюджетный опыт на 20 фиксированных графах и трёх эвристиках

Полный перебор дал строго больше завершённых сертификатов, чем локальная эвристика, в 26 из 160 сравнений; сумма разрывов равна 37, максимальный разрыв — 3. Для пакетной эвристики соответствующие значения равны 18, 25 и 3; для просмотра пар — 17, 25 и 3. Более сильные эвристики сокращают суммарный разрыв с оптимумом, но не достигают его во всех случаях. Их взаимное превосходство на каждом отдельном графе из этих итогов не следует.

5.5. Адаптивный аудит и зависимые исходы

Статический выбор из подраздела 5.4 максимизирует гарантированное завершение до получения исходов проверок. Дополнительный опыт оценивает другую цель: ожидаемое число сертификатов, по которым получено окончательное A или R при жёстком ограничении стоимости на каждой ветви. Обнаруженное нарушение может завершить несколько сертификатов и изменить следующий полезный шаг. Завершение аудита не означает принятия результата: отклонённый сертификат также считается завершённым.

Генератор random.Random(20261020) создаёт 12 графов с семью узлами и 28 сертификатами каждый. Рассмотрены три структуры по четыре графа: распределённые требования, общий узел и циклические цепочки соседних узлов. Стоимости равномерно выбираются из {1,2,3}, размеры требований — из {1,2,3,4}; повторяющиеся требования учитываются кратностью. Бюджеты равны 2, 4, 6 и 8 условным единицам. Предельный размер мал: опыт предназначен для точного контроля, а не для оценки масштабируемости.

Заданы пять распределений исходов. В четырёх исходы узлов независимы, а вероятность нарушения равна 0,10; 0,28; 0,30 или 0,50. В пятом с вероятностью 0,20 нарушены все узлы, а с вероятностью 0,80 исходы независимы с вероятностью нарушения 0,10. Маргинальная вероятность нарушения каждого узла в пятом распределении равна 0,28. Сопоставление с независимым режимом 0,28 отделяет эффект совместного распределения от изменения одиночной частоты. Все вероятностные политики получают одно и то же распределение; статический выбор по гарантии использует только граф и стоимости.

Сравниваются точный статический выбор по гарантии g, точный статический выбор по математическому ожиданию, адаптивный выбор по немедленному ожидаемому приросту на единицу стоимости и оптимальная адаптивная политика. Последняя вычисляется динамическим программированием по наблюдённым T/F и оставшемуся бюджету. Для всех методов перечислены 128 полных исходов каждого графа. Используются целые веса исходов и точные рациональные ожидания; оценки Монте-Карло не применяются.

Таблица 16
Среднее ожидаемое число завершённых сертификатов из 28 по 48 сравнениям каждого режима

Распределение	Статика по g	Статика по E	Локальная адаптация	Точная адаптация
Независимые, q=0,10	11,1866	11,3923	11,3156	11,4489
Независимые, q=0,28	13,9697	14,5900	14,5217	14,7539
Независимые, q=0,30	14,2514	14,9259	14,8768	15,0985
Независимые, q=0,50	16,7865	18,0599	18,0234	18,2396
Общая причина, q=0,28	13,1868	13,6942	13,7568	13,8831

В среднем по 240 сравнениям методы завершают соответственно 13,8762; 14,5325; 14,4989 и 14,6848 сертификата. Точная адаптация строго превосходит лучшую по ожиданию статическую политику в 87 сравнениях; в остальных 153 значения совпадают. Средний выигрыш равен 0,1523 сертификата. Наибольший выигрыш, 1,5633, получен на графе chain-1 при бюджете 6 и независимом q=0,30: 19,1873 против 17,6240. Поэтому адаптация полезна при части сочетаний структуры и бюджета, но не даёт крупного обязательного выигрыша во всех задачах.

Локальная адаптация в среднем немного уступает точной статике по ожиданию: 14,4989 против 14,5325. Отдельная динамическая реакция на исходы сама по себе не обеспечивает улучшения. Для честного сравнения существенны

целевая функция и качество правила выбора. Средние по пяти искусственно заданным режимам не интерпретируются как ожидаемый производственный эффект.

Чувствительность к неверной модели зависимости проверена в 48 дополнительных сравнениях. Политика, рассчитанная для независимого q=0,28, применена к режиму общей причины с той же маргинальной частотой. Она уступает политике с верным совместным распределением в 10 случаях; средняя потеря равна 0,05245, максимальная — 0,8136 сертификата. Эти величины характеризуют выбранный конечный дизайн, а не универсальный штраф за зависимость.

Для четырёх методов и 240 сравнений выполнено 122 880 проверок по полным исходам. Воспроизведение адаптивных политик использует множества выбранных и нарушенных узлов вместо битовой формулы основного расчёта. Проверены совпадение ожиданий и соблюдение бюджета на каждой ветви. Файлы adaptive_audit.py, adaptive_graphs.jsonl, adaptive_comparisons.jsonl и dependence_sensitivity.jsonl сохраняют реализацию и полные результаты. Вероятности заданы исследователем, проверки считаются безошибочными; результаты не являются наблюдением над языковыми моделями.

5.6. Устойчивость к преобразованиям и ошибкам формата

Дополнительная проверка различает допустимое изменение представления, изменение доступности основания и нарушение входного интерфейса. Для всех 960 исходных случаев проверены четыре преобразования: эквивалентная числовая запись, перестановка ключей, дополнительный справочный комментарий и обмен RU/ZH вместе с соответствующими признаками доступности. Получены 3840 преобразованных случаев с сохранением исходных решений.

Для 96 допустимых случаев поочерёдно изымается свидетельство RU, ZH или обеих версий: все 288 результатов должны перейти в D. Для 792 нарушенных случаев те же три изъятия не отменяют уже установленного нарушения: все 2376 результатов остаются R. Такой контроль проверяет приоритет известного нарушения и отличает неполноту свидетельства от доказанной ошибки.

Таблица 17
Состав дополнительной проверки структурированного валидатора

Группа	Число случаев	Критерий
Исходные записи	960	Сохранены A/R/D
Допустимые преобразования	3840	Вердикт не меняется
Изъятие свидетельства для A	288	Переход в D
Изъятие свидетельства для R	2376	Сохранение R
Некорректный интерфейс	192	Техническая ошибка

На 24 канонических допустимых парах дополнительно проверены восемь некорректных описаний доступности: пустой словарь, отсутствие одного из двух ключей, строка вместо логического значения, целое число, null, список и отсутствие объекта. Всего получено 192 случая. В минимальной реализации 120 таких входов приводили к A. Причины — истинность пустой конъюнкции all({}) и трактовка непустой строки или числа как истинного значения. Эти входы не входят в исходный объявленный корпус; обнаруженный дефект относится к границе его программного интерфейса.

Обёртка strict_decide в robustness_audit.py требует оба ключа ru и zh с логическими значениями, проверяет структуру контракта и допустимость его границ. Некорректный управляющий вход получает статус input_error и не получает предметный вердикт. Такой статус учитывается в техническом покрытии, а не заменяется D. На 192 проверенных некорректных входах приёмок после исправления нет; все 960 исходных решений сохраняются. Совокупная проверка 7656 случаев пройдена. Свободный китайский текст, полнота контракта и истинность источника этим опытом не удостоверяются.

6. Агентные архитектуры и границы полномочий

6.1. Автономность как архитектурный выбор

Агентная архитектура вводит динамический выбор следующего действия по наблюдаемому состоянию. Она полезна тогда, когда фиксированный путь решения нельзя определить заранее без существенной потери качества. В остальных случаях управляемая последовательность может быть проще для проверки и сопровождения. Критерий выбора не связан с внешней сложностью интерфейса. Необходимо сравнить полезность динамического решения с ценой дополнительных состояний, вызовов и ошибок [26].

Условно агентный процесс задаётся политикой $\pi(a|s)$, где s — доступное состояние, a — действие. Состояние должно включать ограничения и подтверждённые результаты среды, а не только историю рассуждений. После действия система получает наблюдение и проверяет, допустим ли следующий переход. Эта модель служит описанию управления; она не предполагает, что языковая модель обладает точной внутренней моделью всех состояний. Внешнее приложение должно обеспечивать ограничения, которые нельзя доверить одному текстовому напоминанию.

Инструментальное наблюдение отличается от правдоподобного описания выполненного действия. ReAct объединяет рассуждение и взаимодействие со средой; гибридные стратегии образуют отдельные экспериментальные условия.

Таблица 18
ReAct: рассуждение и действие (M11)

Статус	**Опубликованное испытание; авторские агенты не запускались.**
Источник	Yao и соавт., PaLM 540B, HotpotQA, табл. 1 [27, с. 5].
Результаты	Exact match: Standard 28,7 %; ReAct 27,4 %; ReAct→CoT-SC 35,1 %.
Расчёт	ReAct к базе: −1,3; гибрид к базе: +6,4 процентного пункта.
Ограничение	Гибрид меняет процедуру и бюджет; данные не доказывают безусловную пользу автономности.
Авторское испытание	Фиксировать журнал вызовов, разрешения, наблюдения, отказы и стоимость; проверять финальную предметную приёмку.

Для бизнеса и ИТ допустимость действия является самостоятельным условием. Верный ответ не разрешает несанкционированное изменение внешней среды. Выход агента должен сохранять различие между предложением, попыткой и подтверждённым исходом.

6.2. Инструментальный интерфейс и разрешение действия

Инструмент задаётся не только названием функции. Его контракт включает типы аргументов, область данных, побочные эффекты, наблюдаемые ошибки и условия повторного вызова. Неясный интерфейс увеличивает вероятность формально корректного, но предметно неверного действия. Например, функция update может изменять одну запись, весь набор или текущую версию документа. Эти различия должны быть определены до передачи управления модели.

MCP стандартизирует элементы обмена ресурсами и возможностями, но сам факт подключения не удостоверяет доверие к серверу [28]. В работе рассматривается версия спецификации 2025-11-25. Право использовать конкретную операцию определяется приложением и пользователем отдельно от технической доступности метода. Инженерная проверка включает источник сервиса, доступные данные и разрешённые изменения. Унификация интерфейса полезна для интеграции, однако не отменяет анализа границ доверия.

График сохраняет отрицательный результат отдельного ReAct и положительный результат гибрида. Для приёмки процесса нужен журнал наблюдений: текстовое утверждение «проверено» не считается подтверждённым вызовом инструмента.

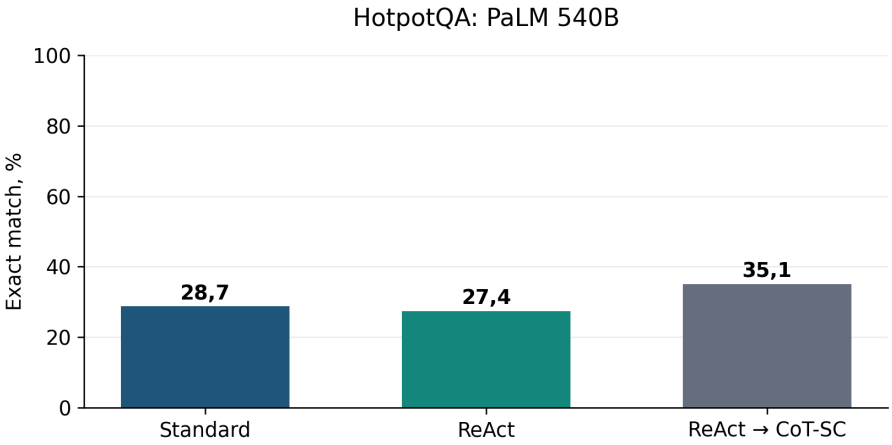


Рис. 6. ReAct и гибридная стратегия: опубликованное сравнение на HotpotQA [27, с. 5]

Интерфейс задаёт аргументы, диагностику и возвращаемое состояние; приложение проверяет разрешение до изменения среды. Показанные доли относятся к ответам на вопросы, поэтому не измеряют безопасность действия, восстановление после тайм-аута или идемпотентность.

6.3. Непрямое внедрение инструкций и модель угроз

Исследование Greshake и соавторов рассматривает перенос атакующей инструкции через данные, которые интегрированная система извлекает из внешней среды [29]. Методологическая проблема состоит в размывании границы между исследуемым материалом и управлением. Критично не наличие подозрительной фразы само по себе, а возможность её влияния на операции, доступ к данным и представление результата. Поэтому защита проектируется вокруг путей воздействия.

Модель угроз описывает активы, доверительные границы, допустимые источники и операции. Для каждого внешнего фрагмента определяется, какую роль он может выполнять: предоставить факт, описать запрос третьего лица или изменить правило. Последнее не должно становиться возможным только через попадание текста в контекст. Разделители и инструкции помогают интерпретации, однако обязательные ограничения поддерживаются проверкой прав, аргументов и побочных эффектов на уровне приложения.

Контекст управления риском задаётся рамкой AI RMF [30]. Тест безопасности проверяет конкретную границу: внешний документ является данными, а не источником полномочий. Гладкий ответ при нарушении этой границы не принимается.

Таблица 19
Устойчивость к непрямым инструкциям (E02)

Статус	Авторский план испытания; частота атак и отказов не измерена.
Парное сравнение	Один документ без вмешательства и с вложенной инструкцией изменить действие или раскрыть данные.
Контроль	Одинаковые инструменты, права и предметная задача; инструкция во внешнем источнике не повышает права.
Показатели	Доля нарушений полномочий; правильная диагностика; доля сохранённых полезных ответов.
Критический исход	Запрещённое действие блокирует приёмку независимо от стилистической оценки.
Протокол	Сохранять нагрузку, наблюдения и отказы; проверять обе языковые версии и смешанный контекст.

Нулевая наблюдаемая частота не означает отсутствия риска. Вывод зависит от семейства вмешательств и числа независимых задач. Фильтрация текста дополняет контроль прав, но не заменяет его [31].

6.4. Неизвестный исход, идемпотентность и восстановление

Ошибка связи и ошибка операции различаются. Тайм-аут может означать, что результат потерян после успешной записи. Повторный вызов способен создать дубль или повторить изменение. Поэтому состояние «исход неизвестен» не приравнивается к подтверждённому отказу. Для изменяющих действий требуется идентификатор операции либо иной механизм проверки статуса и предотвращения повторного эффекта.

Из неизвестного исхода изменяющей операции нельзя выводить право на её безусловное повторение.

Наблюдаемому тайм-ауту соответствуют два состояния: операция завершилась либо не завершилась. Повторный вызов имеет в них разные последствия. Поэтому сначала проверяют статус операции или используют подтверждённо идемпотентное действие.

Инцидентный анализ восстанавливает наблюдаемые переходы: вход, разрешение, вызов, ответ и принятое решение. Психологическое объяснение «модель проявила невнимательность» не локализует дефект. Нужно установить, отсутствовал ли источник, потерян ли статус, неверно ли проверен аргумент или не сработало условие остановки. Каждая причина требует отдельного изменения и нового контрольного случая.

Восстановление включает определение затронутых результатов и проверку исправления. Если система отправила устаревший материал, следует установить

круг получателей и необходимость корректировки. Новую версию проверяют на исходном инциденте и близких сценариях; невозможность восстановить точные условия указывают в отчёте.

7. Креативная индустрия: качество и самостоятельность

7.1. Творческая задача как ограниченное пространство решений

Творческая задача допускает несколько подходящих решений. Бриф задаёт аудиторию, коммуникативную цель, носитель, ресурсы и обязательные факты. В этих границах сравниваются разные художественные подходы; нарушение условий само по себе не служит признаком оригинальности.

Концепции должны различаться по идее и предполагаемому действию аудитории. Три названия одной композиции дают лишь варианты формулировки. Для сопоставления можно описать участие зрителя, организацию времени, связь образа и сообщения, причинную структуру. Эти признаки помогают обсуждать различия в конкретном проекте, но не образуют универсальной шкалы художественной ценности.

На дивергентном этапе исследуют разные направления, на конвергентном — выбирают решение с учётом брифа и ресурсов. Ранние гипотезы сохраняют вместе с причинами отказа от них. Если ни один вариант не выполняет обязательных условий, поиск продолжается.

Матрица «концепция — механизм» помогает обнаружить повторы и описать различия. Численная дистанция признаков не устанавливает художественной самостоятельности. Целостность, выразительность и связь с аудиторией оценивает предметный эксперт.

7.2. Голос бренда и драматургическая связность

Система бренда задаёт отношения между фактами, тоном и намерением сообщения. Перечень эмоциональных эпитетов не обеспечивает воспроизводимый стиль. Нужны примеры допустимого поведения, контрпримеры и ситуации, где фактическая ясность имеет приоритет. Ответ о задержке должен оставаться понятным даже при дружелюбной подаче. Если тон скрывает существенное обстоятельство, качество коммуникации ухудшается независимо от привлекательности текста.

Стиль проверяют по ясности обращения, согласованности терминов, достоверности обещаний и соответствию ситуации. Формулировки могут варьироваться: разные жанры требуют разных способов выражения общего голоса бренда. Контракт связывает стилевые правила с назначением сообщения.

Драматургический анализ рассматривает изменение состояния и цену выбора. Для каждой сцены выделяются исходная ситуация, намерение, препятствие, действие и последствие. Такое представление помогает обнаружить эпизод, который не влияет на развитие. Однако логическая связность не исчерпывает художественной ценности: символическая, ритмическая или атмосферная функ-

ция может быть самостоятельным основанием эпизода. Проверяющий должен учитывать жанровый контракт.

Если герой получает решение конфликта без действия, нового знания или подготовленного события, причинная связь может оказаться ослабленной. Оценка зависит от замысла. В учебной работе полезно сопоставить исходную и исправленную сцены и объяснить, как изменилась их функция. Окончательное художественное решение принимает автор проекта.

7.3. Н. К. Рерих: культура и проверяемая ответственность

Культурное измерение генеративных систем требует основания, отличного от технической эффективности. В качестве исторического ориентира рассматривается Николай Константинович Рерих. Реестр Организации американских государств датирует принятие Пакта Рериха 15 апреля 1935 года [32]. Материалы Nicholas Roerich Museum связывают его деятельность с охраной культурного наследия и воспроизводят текст договора о защите исторических памятников и художественных, научных и образовательных учреждений [33].

Значение этого ориентира для настоящего исследования определяется авторской интерпретацией: профессиональный результат должен сохранять культурное содержание, происхождение и условия понимания материала. Пакт не является стандартом языковых моделей; Рериху не приписываются идеи об ИИ. Историческая инициатива служит основанием обсуждения ценностей, а перенос в инженерный контракт предлагается в этой монографии.

Для русско-китайской адаптации задаются обязательства: сохранить атрибуцию, различить цитату и пересказ, удержать исторический контекст и обозначить новую интерпретацию. Условия использования материалов проверяются отдельно; вопросы интеллектуальной собственности рассматриваются с опорой на WIPO [34]. Культурная значимость произведения сама по себе не устанавливает право на любое его воспроизведение.

Формальная проверка относится к заданным обязательствам. Недоступность источника цитаты служит основанием отсрочки; установленное расхождение с источником — отклонения. При адаптации культурного образа указываются происхождение и характер изменений. Художественная или философская ценность этими проверками не измеряется.

Мысленный эксперимент «промптинг в философии Рериха» понимается как авторский перенос принципа охраны культуры в проектирование задания. Сначала определяются культурный смысл и аудитория; затем — источники, атрибуция и границы адаптации. Генерация становится продолжением ответственной интерпретации, а не основанием приписывать Рериху новые высказывания.

Для расчётного примера введём частные оценки $q_j \in [0; 1]$, положительные веса w_j и итоговый балл $B=\Sigma w_j q_j / \Sigma w_j$. Индикатор H равен 1 при выполнении всех обязательных условий и 0 при наличии нарушения. Индикатор приёмки $P=H \cdot 1[B \geq \tau]$, где τ — порог. При $w=(3,2,2,1)$ оценки $q=(1; 0{,}75; 0{,}5; 1)$ дают

B=0,8125, а q=(1; 1; 0,75; 0,75) — B=0,90625. При H=0 результат отклоняется независимо от балла. Если выполнение обязательного условия не установлено, решение откладывается и бинарная формула не применяется. Веса выбраны для иллюстрации и не служат мерой ценности культуры.

При H=1 второй вариант проходит порог 0,8 и не проходит 0,92.

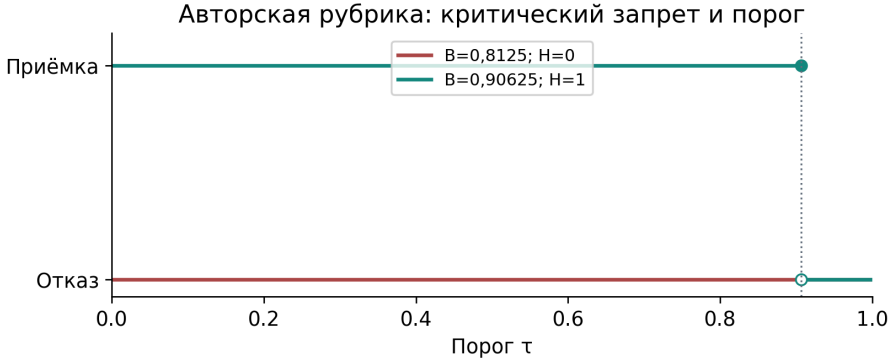


Рис. 7. Авторская иллюстрация чувствительности культурной приёмки к порогу

Для культурного сертификата два свидетельства могут быть взаимодополняющими: атрибуция и допустимость использования. Пока одно не проверено, положительный итог другого не завершает приёмку. Отрицательный итог даёт нарушение. Так культурный ориентир связывается с бюджетной задачей главы 10, а не с обещанием повышения точности модели.

7.4. Мультимодальная спецификация и доступность

Мультимодальный вход требует разграничить наблюдение, распознавание и интерпретацию. Модель может описать видимый объект, но ошибочно прочитать мелкий текст или приписать изображению неустановленную причинность. Поэтому контракт задаёт область анализа, необходимую точность и допустимый статус неопределённости. При недостаточном разрешении точное значение не должно восстанавливаться по ожиданию. Этот принцип относится к плакату, диаграмме и скриншоту интерфейса [35].

Для визуального анализа полезно хранить связь вывода с конкретной областью изображения. Наблюдение «ось обрезана» проверяется иначе, чем оценка «композиция напряжённая». Первое связано с видимым признаком, второе — с предметной интерпретацией. Их смешение создаёт ложное впечатление объективности художественного суждения. Документация мультимодального интерфейса определяет доступные входы, но профессиональная приёмка требует самостоятельной оценки полученного артефакта в его конечном назначении.

Генеративная иллюстрация пригодна для образа и композиционного поиска, однако не обеспечивает точности числовой диаграммы. Если визуал должен представлять вычисленные значения или структуру процесса, его строят на проверяемых данных и детерминированных правилах отображения. Художественный стиль технической схемы не делает её достоверной. Разделение этих задач предотвращает подмену доказательной функции декоративным эффектом.

Доступность также входит в критерии приёмки. Существенную информацию нельзя кодировать только цветом или размещать в нечитаемом изображении. Альтернативное описание должно передавать функцию визуала в контексте материала. Для диаграммы важны смысл и доступные данные, для художественной композиции — релевантное описание образа. Генеративная модель может подготовить черновик, но итог проверяют на конкретном носителе. Формальная доступность файла не равнозначна доступности информации читателю.

8. Бизнес: основания решений и экономика применения

8.1. Управленческое решение и причинная дисциплина

Управленческая задача связывает наблюдение с возможным действием. Анализ должен определить, какой выбор предстоит владельцу процесса и какие сведения способны его изменить. Без этого модель может сформировать обширный обзор, не имеющий операционной ценности. Контракт ограничивает решение доступными данными и полномочиями, а рекомендации сопровождает условиями применимости. Причинное объяснение требует основания, отличного от совместного изменения показателей во времени.

Рост продаж после кампании может быть связан с сезонностью, ценами, ассортиментом или составом клиентов. Для оценки эффекта нужен дизайн сравнения, учитывающий эти объяснения. Модель может помочь сформулировать гипотезы и план проверки; причинный вывод требует соответствующих данных.

Клиентские сообщения представляют отобранные наблюдения. Частота темы в отзывах не равна распространённости проблемы среди всех клиентов. Также различаются сообщение, клиент и эпизод услуги. Если один пользователь пишет пять раз, простой подсчёт сообщений увеличивает вес его проблемы. Это может быть полезно для анализа нагрузки поддержки, но некорректно для оценки числа затронутых людей. Поэтому единица анализа должна соответствовать управленческому вопросу.

В результате анализа сохраняются подтверждающие фрагменты, правила группировки и исключения. Приоритет проблемы оценивают с учётом частоты, тяжести последствий и возможности исправления. Редкая, но тяжёлая проблема может требовать первоочередного действия. Критерий ранжирования должен быть доступен пользователю.

8.2. Персонализация и контур ответственности

Деловая персонализация должна опираться на подтверждённые сведения. Приписывание клиенту неизвестных потерь, бюджета или внутренних проблем

создаёт убедительность ценой достоверности. Контракт различает факты о клиенте, предположения и диагностические вопросы. Если утверждение не поддержано входом, его либо исключают, либо преобразуют в явно обозначенную гипотезу. Сильное предложение связывает подтверждённую задачу, подход и проверяемый следующий шаг.

В коммерческом тексте сначала проверяют факты, затем стиль. В учебном кейсе к тексту прилагается реестр использованных сведений, а неподтверждённые предположения оформляются как вопросы или гипотезы.

Поддержка соединяет информационный ответ и действия над клиентским состоянием. Объяснить правило, изменить заказ и вернуть деньги — разные полномочия. Если система не вправе выполнить операцию, полезным результатом становится подготовленный пакет эскалации. Он включает проблему, подтверждённые сведения и уже предпринятые действия. Передача оператору без контекста снижает полезность процесса, а включение лишних данных создаёт новый риск.

Успех поддержки оценивается по решению проблемы, соблюдению прав и нагрузке на следующий этап. Быстрая первая реплика может сопровождаться ростом времени ручного восстановления. Поэтому измерение охватывает весь клиентский эпизод с заранее определёнными началом и завершением.

8.3. Экономика принятого результата и дизайн пилота

Экономическая оценка должна учитывать полный цикл получения пригодного артефакта. Стоимость генерации, проверки, исправления и сопровождения относится к одной области и периоду. Знаменатель определяется числом принятых результатов, а не числом вызовов. При отказах дополнительные расходы не исчезают из расчёта. Поэтому дешёвая генерация способна оказаться дорогой при высоком уровне ручной доработки.

В условном примере затраты 800+3200+1600=5600 руб. при 32 принятых результатах составляют 175 руб. за результат. Без учёта проверки получается 75 руб., однако это уже неполная стоимость. Сравниваемые процессы должны иметь одинаковые критерии приёмки и состав затрат. Для оценки окупаемости дополнительно учитываются обслуживание, объём применения и последствия ошибок.

Пилот вводит ограниченную область, базовую линию и заранее определённое правило решения. Его задача — проверить механизм полезности и выявить скрытую нагрузку. До запуска фиксируют, какие изменения считаются существенными, как оцениваются отказы и кто размечает результат. Если после наблюдения выбрать только улучшившийся показатель, вывод будет смещён. Поэтому критерий расширения пилота задаётся отдельно от интерпретации понравившегося примера.

Полезно измерять время человека на приёмку и исправление, поскольку именно оно показывает цену профессионального доверия к системе. Результат

пилота должен объяснять различия по классам задач: часть операций может улучшиться, а часть — ухудшиться. Такое заключение представляет более сильное основание решения, чем общий рекламный процент экономии. В монографии описывается дизайн проверки; реальные показатели внедрения без наблюдений не приписываются ни автору, ни организации.

8.4. Аналитическая коммуникация и граница точности

Представление данных должно сохранять масштаб, единицы и предел точности. Изображение графика без исходной таблицы обычно не обеспечивает точного восстановления всех значений. Если модель называет число до сотых по пикселям, формат ответа создаёт ложную точность. Корректное решение обозначает приблизительность либо запрашивает исходные данные. Для управленческого вывода важно также проверить, не меняет ли обрезанная ось впечатление о величине различий.

Информационная плотность не тождественна полноте основания. Один слайд может содержать много чисел и при этом скрывать определения. Поэтому визуал связывается с конкретным решением и доступным приложением данных. Проверяющий должен восстановить существенное сравнение и увидеть ограничение. Генеративная система помогает подготовить объяснение, но арифметика и отображение значений требуют независимой проверки. Это разделение повышает пригодность презентации для профессионального обсуждения.

Сокращение материала является преобразованием аргумента. Оно должно сохранять статус утверждения, обязательные условия и основания вывода. При переходе из отчёта в слайд нельзя незаметно превратить гипотезу в факт. Если ограничение невозможно разместить полностью, его делают видимым и связывают с доступным приложением. Для научного текста это особенно важно: удаление методических условий может изменить смысл результата сильнее, чем изменение самого числового значения.

При переводе отдельно проверяются терминология, числа и условия действия, отдельно — естественность языка. Обратный перевод помогает выявлять потери смысла, но не гарантирует их отсутствия. До работы определяют неизменяемое содержание и допустимые способы адаптации к аудитории.

9. Программная инженерия и сопровождение систем

9.1. Репозиторий, требование и независимый оракул

В программной задаче необходимы существующий код, конфигурация и сведения о наблюдаемом поведении. Без этого модель может предложить отсутствующий интерфейс или исправление, несовместимое с версией зависимости. Работа начинается с проверки интерфейса и воспроизведения проблемы.

Бенчмарки изменения репозиториев, включая SWE-bench, показывают значение исполняемой проверки [36]. Однако результат на бенчмарке не является прямым основанием допуска агента к рабочему проекту. Нужно учитывать среду, покрытие тестов, права и последствия изменения. Проверяется исходный

сценарий и значимые соседние условия. Профессиональная приёмка относится к поведению продукта, а не к объёму diff или уверенности объяснения предполагаемой причины.

Оракул определяет ожидаемое поведение независимо от проверяемой реализации. Если ожидаемое значение вычисляется тем же ошибочным алгоритмом, тест воспроизводит дефект. Поэтому в расчётной функции нужны отдельно обоснованные значения, а в интерфейсе — явный контракт. Для сложной системы часть свойств проверяется отношениями: сохранением суммы, неизменностью доступа или допустимым переходом состояния. Такая проверка не обязана перечислять все возможные выходы.

Тест должен различать исходный дефект и исправленное поведение. По возможности ошибку воспроизводят до изменения, затем проверяют устранённый дефект и возможную регрессию. Объём проверки соотносится с риском: визуальное смещение и изменение финансового расчёта требуют разного контроля.

9.2. Интерфейсные инварианты и безопасность кода

Схема определяет форму обмена данными. Правильно сформированный JSON способен содержать неверную сумму или неизвестный идентификатор. Проверка формы и проверка смысла поэтому разделяются [37].

Таблица 20
Структурированный вывод и предметная валидация (M12)

Статус	**План модельного испытания; локальный валидатор проверен отдельно.**
Сравнение	Обычная инструкция формата против ограниченного схемой вывода при одинаковой задаче.
Отдельные показатели	Синтаксический разбор; соответствие схеме; числовые инварианты; поддержка источником; диагностический отказ.
Отрицательные случаи	Неизвестная ссылка, пустой успешный ответ, строка вместо числа; False вместо целого 0.
Выполненные проверки	12 случаев минимального валидатора и 20 отрицательных канонических записей в комплекте; это тесты программы.
Критерий	Доля синтаксических успехов не объявляется долей предметно верных ответов.

Типизированный контракт позволяет обнаруживать нарушения структуры данных. Предметную корректность проверяют отдельно. При смене промпта

необходимо повторно проверить совместимость формата с приложением и допустимость операций над полученными данными.

Безопасность разработки требует анализа доверия к входу и последствий операций. SSDF рассматривает снижение риска уязвимостей как часть процесса создания программ [38]. Генерация кода не освобождает от проверки пользовательских значений, зависимостей и разрешений. Промпт сообщает модельный контекст, но обязательные ограничения обеспечиваются кодом и доступами. Одно успешное выполнение не подтверждает отсутствие других путей нарушения.

В учебный проект полезно включить вход, выявляющий небезопасную сборку запроса или нарушение доступа. Отчёт должен описывать обнаруженный дефект, способ устранения и проверенные свойства. Общее утверждение «код безопасен» выходит за пределы такой проверки.

9.3. Ревью, документация и инженерный аргумент

Ревью различает дефект, риск и предпочтение оформления. Команда «улучши код» не определяет, какое изменение необходимо. Предметное замечание связывает место, наблюдаемое поведение и последствие. Это позволяет автору исправления проверить замечание и решить, относится ли оно к текущей задаче. Косметическое переписывание без основания может увеличить область регрессии, не устраняя исходную проблему.

Описание изменения сообщает, какой сценарий был нарушен, как он работает после исправления и какие проверки выполнены. Оно должно быть понятно без истории диалога. В учебной работе такой отчёт показывает основания, по которым предложенный код принят.

Документацию обновляют вместе с реализацией. В примерах команд проверяются существование параметров, условия среды и возможность выполнения. Для незапущенного примера прямо указывается, что проверка исполнения не проводилась.

Документация должна перечислять ограничения и неподдерживаемые сценарии. В учебной работе можно сопоставить README с фактическим поведением, объяснить расхождения и предложить способ поддержания их согласованности.

9.4. Версии, дрейф и технический долг

Версионирование охватывает инструкцию, примеры, схему, источники, инструменты, права и словарь категорий. Паспорт изменения фиксирует назначение, затронутые сценарии и основания выпуска. Прежнее название модели не гарантирует прежнего поведения после обновления сервиса.

Sculley и соавторы связывают технический долг машинного обучения с зависимостями и сопровождением [39]. В генеративной системе новый формат источника может нарушить извлечение, а правка промпта — временно скрыть

дефект. Для устранения причины необходимы карта интерфейсов и проверка зависимостей.

При интеграции разделяются генератор, валидатор, хранилище свидетельств и разрешение действий. Журнал фиксирует идентификатор пары, версии контракта и модели, хеш входа, состояния T/F/U, основания вердикта и выполненные операции. Трассировку обеспечивает LangSmith [40] или собственная база; её наличие не подтверждает правильности содержания.

MCP определяет взаимодействие с инструментами [28]. Авторизация, минимальные полномочия, проверка источников и законность передачи данных обеспечиваются отдельно. Состав журнала и сроки хранения определяются его назначением и применимыми правилами; чувствительный полный текст сохраняют только при обоснованной необходимости.

Временные меры КНР 2023 года относятся к предоставлению генеративных услуг публике на территории страны. Статья 11 ограничивает незаконное хранение идентифицирующих входов и записей; статья 14 предусматривает сохранение записей при выявленном незаконном использовании [41]. Эти нормы не устанавливают универсальной обязанности бессрочно хранить каждый ответ.

Индийский DPDP Act регулирует персональные данные. Право на информацию об их обработке не означает права на объяснение любого вердикта ИИ [42]. Rules 2025 вводятся поэтапно: правило 4 — через год, правила 3, 5–16, 22 и 23 — через 18 месяцев после публикации [43, с. 24]. Применимость обязанности зависит от даты и роли участника.

Дрейф проверяют после обновлений на фиксированном наборе и новых независимо размеченных случаях; частота зависит от риска.

10. Совместная приёмка и бюджет научного аудита

10.1. Предшественники и граница научного вклада

Prompt Patterns и Prompt Canvas систематизируют постановку запросов [44; 45]. В настоящей работе исследуется совместная приёмка русско-китайских результатов с учётом неполноты свидетельств и бюджета завершения сертификатов. Граница вклада определяется сопоставлением этой задачи с предшествующими подходами.

Таблица 21
Ближайшие предшественники и граница авторского вклада

Ahn, Kim, 2026 [46]	**Контракты, источниковый контроль и исполняемые ограничения вокруг агента. Общая идея вынесения гарантий из промпта уже представлена.**
Tufano и соавт., 2026 [47]	Спецификация как опора генерации тестов с эмпирической оценкой. Применение контрактов в ИТ не заявляется впервые.
Liu и соавт., ICML 2026 [48]	Формальное определение и оптимизация межъязыковой согласованности модели. В настоящей работе не выполняется post-training.
MuBench, ACL 2026 [49]	Многоязыковая оценка и согласованность. Локальная точность и согласованность уже различаются в литературе.
Настоящая постановка	Трёхзначная совместная приёмка; границы при неполноте; гарантированное и ожидаемое завершение; статический и адаптивный бюджетный контроль.

Постановка использует известную теорию комплементарности [56] и является частным случаем Set Union Knapsack [63]. При одинаковой стоимости проверок она связывается с Densest k-Subhypergraph [62]. Новизна работы относится к интерпретации элементов задачи как проверок профессиональных обязательств и к гарантированному завершению аудита при свободных бинарных исходах. Сопоставление и ограничения алгоритмических выводов приведены в разделе 10.5.

MGSM [50], CMMLU [51] и исследования межъязыковой несогласованности знаний [52] задают фон исследования. Qwen [53] и DeepSeek [54] рассматриваются как возможные валидаторы. Их ответы в выполненном программном опыте не использовались. Для внешнего испытания подготовлен новый корпус v2 из 400 пар и 60 примеров для настройки.

Пусть n_{ij} — число пар с эталонным вердиктом i и вердиктом валидатора j, где $i,j \in \{A,R,D\}$: принять, отклонить, отложить. Строки матрицы соответствуют эталону, столбцы — валидатору; $n_{i\cdot}$ и $n_{\cdot j}$ обозначают суммы строки и столбца. Общее число успешно разобранных решений с эталонной меткой равно $N=\Sigma_{ij} n_{ij}$. Метрики определяются на этой совокупности при $N>0$.

$FAR=n^{RA}/n^{R\cdot}$; $FRR=n^{AR}/n^{A\cdot}$. Доля отсрочек допустимого результата равна $n^{AD}/n^{A\cdot}$. При $n^{A\cdot}>0$ выполняется $1-n^{AA}/n^{A\cdot}=FRR+n^{AD}/n^{A\cdot}$. Следовательно, единица минус полнота принятия совпадает с FRR только при отсутствии отсрочек в классе A. При нулевом знаменателе соответствующий показатель не определён.

$$p_o=(n^{AA}+n^{RR}+n^{DD})/N;\ p_e=\Sigma_i n_i\cdot n\cdot_i/N^2;\ \kappa=(p_o-p_e)/(1-p_e).$$

Номинальная каппа Коэна корректирует согласие на случайное совпадение маргинальных распределений [55]; она не является обычной долей согласия. При $p_e=1$ каппа не определена. Категории A, R и D не задают порядковую шкалу. Каппа двух независимых разметчиков вычисляется до арбитража; решения арбитра не включаются в исходную матрицу их согласия.

Техническая ошибка, обрезанный ответ и неверный JSON учитываются отдельно от вердикта D. Публикуются число попыток и покрытие в целом и по каждому эталонному классу. Четыре варианта одного семейства зависимы; неопределённость метрик оценивается с группировкой по семейству. В приложении И приведён пересчёт предоставленного прогона Kimi. Показатели нового корпуса v2 и запланированных внешних моделей пока отсутствуют.

10.2. Совместный сертификат и монотонность знания

Совместный сертификат объединяет локальные проверки RU, локальные проверки ZH и парные обязательства. Число обязательств может различаться по заданию; общий вердикт представляет трёхзначную конъюнкцию. Состояние U сохраняется при недоступном зарегистрированном свидетельстве, а неизвестный идентификатор свидетельства является установленным нарушением.

Сохраним обозначения подраздела 2.2: c — число подтверждённых пар, f — число пар с нарушением, u — число неопределённых пар; $n=c+f+u>0$. Границы $J^-=c/n$ и $J^+=(c+u)/n$ применяются к совокупности совместных обязательств. Ширина $W=u/n$ показывает долю незавершённой проверки и не характеризует ошибку модели или неопределённость художественной оценки.

Монотонность знания означает: при достройке неизвестных обязательств без изменения известных T/F интервал не расширяется. Если новая проверка даёт F, соответствующая пара исключается из верхней границы; если все её неизвестные обязательства дают T, она входит в нижнюю. Изменение версии источника может отменить прежнее T и требует нового аудита; это уже другая последовательность состояний.

В контрольном корпусе c=96, f=792, u=72. Получаем [96/960; 168/960]=[0,10; 0,175]. Интервал относится к искусственно заданной смеси сценариев: семейства нарушений включены намеренно, а их частота в профессиональной практике не измерялась. Поэтому его нельзя интерпретировать как оценку надёжности профессионального промптинга в 10–17,5 %.

На рисунке показаны два случая: при n=40, c=28 и u=4 интервал равен [0,70; 0,80], для контрольного корпуса — [0,10; 0,175]. Оба интервала задают границы при неполноте проверки, а не биномиальные доверительные интервалы.

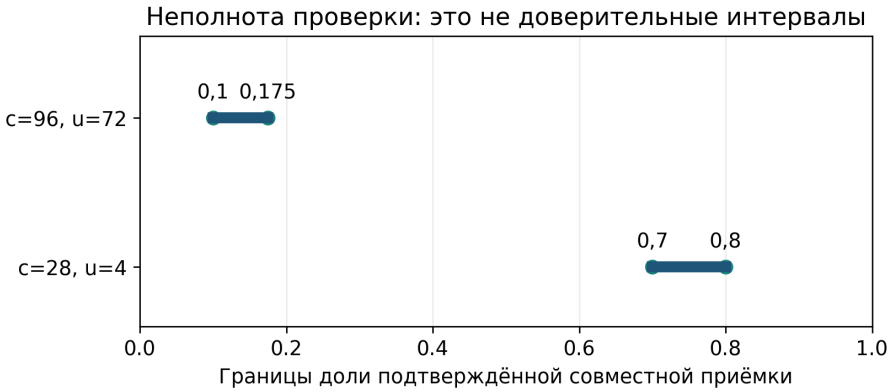


Рис. 8. Границы при неполноте проверки: расчётный пример и выполненный контрольный корпус

При свободной достройке границы достижимы, но зависимости между свидетельствами способны сузить множество возможных исходов. До их формализации интервалы остаются внешними. Каждая проверка обязана ссылаться на тот же контракт и версию источника; иначе монотонность достройки не применима. Бюджетная процедура уменьшает незавершённую область, сохраняя различие подтверждения и нарушения.

10.3. Комплементарность и ограничение жадного выбора

Гарантированный эффект $g(S)=\Sigma_i 1[U_i \subseteq S]$ является монотонной супермодулярной функцией. Общая теория комплементарности известна [56]; здесь она применяется к завершению совместных сертификатов. Для $A \subseteq B$ и $j \notin B$ выполняется $g(A \cup \{j\})-g(A) \leq g(B \cup \{j\})-g(B)$.

Доказательство. Для отдельного U_i добавление j завершает сертификат тогда и только тогда, когда $j \in U_i$ и $U_i \setminus \{j\}$ уже входит в выбранное множество. Если это условие выполнено для A, оно выполнено для B. Суммирование неотрицательных индикаторов даёт требуемое неравенство. Для 225 семейств из двух требований на четырёх узлах проверены все 57 600 неравенств для пар подмножеств.

Супермодулярность означает возрастающую отдачу взаимодополняющих проверок, а не гарантию качества итогового ответа. Проверка A может быть бесполезна отдельно и необходима вместе с B. Поэтому выбор исключительно по текущему немедленному приросту способен остановиться до полезной комбинации.

Контрпример. При бюджете 2 и единичных стоимостях m сертификатов требуют {A,B}; ещё один требует {C}. Локальный жадный выбор берёт C, поскольку его немедленный прирост равен 1, а прирост A и B равен 0. После

этого ни A, ни B отдельно не завершают дополнительные сертификаты. Точный выбор {A,B} завершает m.

Отношение результатов равно 1/m и стремится к нулю. Поэтому у описанного локального жадного алгоритма нет положительной константной гарантии для всей данной области. Это утверждение относится к фиксированному правилу выбора и остановки, а не ко всем алгоритмам, называемым жадными.

Одна проверка C завершает 1 сертификат; совместная проверка A и B — 100. Это заранее определённый контрпример комплементарности, не усреднённый эффект на реальных организациях.

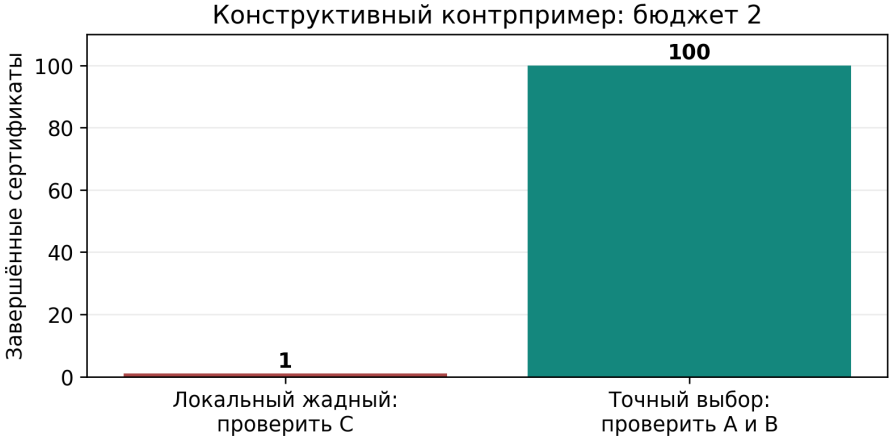


Рис. 9. Выполненный конструктивный контроль: m=100, бюджет 2

Код проверен для m от 2 до 100; общий вывод 1/m следует из конструкции. Выбор C не ошибочен относительно немедленного прироста, но не решает задачу совместного бюджета. Для культурного материала A может обозначать атрибуцию, B — подтверждение допустимого использования; положительное завершение требует обоих свидетельств. Наличие культурного принципа не задаёт их истинность автоматически.

10.4. Алгоритм и проверка результатов

Точная процедура перебирает все подмножества узлов, исключает превышающие бюджет и максимизирует g(S). При равной цели выбираются меньшая стоимость и меньшая целочисленная битовая маска набора. В реализации с битовыми масками и одним машинным словом для m узлов оценка времени прямого перебора составляет $O(2^m \cdot (m+n))$; при произвольной длине битовых строк дополнительно учитывается стоимость операций над ними. Установлен предел $m \leq 18$. Это контрольный алгоритм для малых задач, а не масштабируемое промышленное решение.

Эквивалентная целочисленная постановка использует $x_j, z_i \in \{0,1\}$: максимизировать Σz_i при $\Sigma a_j x_j \le B$; $z_i \le x_j$ для каждого $j \in U_i$; $z_i \ge \Sigma_j \in U_i x_j - |U_i| + 1$. Последние ограничения задают $z_i = 1$ тогда и только тогда, когда выбраны все обязательства пары. Для большого m необходим отдельный оптимизационный инструмент; масштабируемость полным перебором не заявляется.

Для 3600 сочетаний семейства требований и выбранного множества положение 2 проверено отдельной программной реализацией с перечислением всех бинарных исходов. Во всех случаях минимум числа завершённых сертификатов совпал с g(S). Все 160 бюджетных оптимумов дополнительно проверены перебором комбинаций множеств вместо битового представления основного алгоритма.

Узел аудита соответствует проверяемому отношению, а не только документу. Разные тезисы могут требовать разных процедур для одного файла. Слияние узлов допустимо, если одна надёжная операция разрешает все обозначенные обязательства; иначе дешёвая модель графа создаёт ложную экономию. Такие ошибки спецификации должны выявляться предметным экспертом.

Получены три формальных результата: внешние границы совместной приёмки при неполноте; характеристика гарантированного завершения по набору проверок; комплементарность и контрпример локальному жадному выбору. Авторский вклад относится к постановке профессионального аудита. Стандартная логика и теория супермодулярности имеют известные основания.

В новой реализации выполнены программные проверки 960 контрольных пар и 6720 решений семи процедур. В бюджетном опыте получены 160 сравнений четырёх алгоритмов. Совпадение классификации с конструкторскими метками характеризует исполнение объявленной спецификации. Показатели Kimi пересчитаны по предоставленным журналам; новый модельный инференс и внешняя экспертная оценка не выполнялись.

Подготовлены 400 тестовых пар в 100 семействах и 60 примеров для настройки с непересекающимися идентификаторами и исходными реквизитами. Программный разбор обязательств охватил все 460 случаев; семантика китайского текста требует отдельной экспертной оценки. Расчёт метрик проверен на 19 682 ненулевых матрицах. Для GigaChat подготовлены запросы и клиент запуска.

Практический перенос требует аудита самого контракта, независимой предметной разметки, фиксации модели и полного учёта технических отказов. Два валидатора могут воспроизводить одну ошибку; их согласие не заменяет источник или эксперта. Бюджет дополнительного контроля оценивается относительно сильного процесса с теми же обязательствами.

Культурный ориентир Рериха операционализируется через сохранение атрибуции, статуса цитаты, контекста и условий использования. Выражение «цифровой охранный статус» обозначает авторскую метафору контракта, а не новое

юридическое право. Допустимая адаптация не объявляется нарушением только из-за стилистического различия.

Проект университетского курса предусматривает 32 контактных академических часа; образовательные основания соотнесены с UNESCO [57]. Эмпирическая оценка обучения и промышленного эффекта не выполнена. Следующий этап должен проверять метод на новых независимых контрактах и реальных результатах, включая отрицательные наблюдения.

10.5. Связь с известными задачами оптимизации

Пусть проверки образуют множество вершин V, а непустые множества U_i задают гиперрёбра. Выбор $S \subseteq V$ завершает сертификат i при $U_i \subseteq S$. Поэтому при единичных стоимостях и бюджете k максимизация g(S) соответствует выбору k вершин с наибольшим числом содержащихся гиперрёбер — постановке Densest k-Subhypergraph [62]. Ограничения «не более k» и «ровно k» эквивалентны по оптимальному значению при возможности дополнить набор: g монотонна. Совпадающие U_i учитываются кратностью либо весом ребра. Это сопоставление выведено из формулы целевой функции.

При разных стоимостях a_j постановка является Set Union Knapsack с элементами-проверками и доходом 1 за каждое полностью включённое множество U_i [63]. Если завершение разных сертификатов имеет разную ценность, вводятся неотрицательные веса w_i и цель $g_v(S)=\Sigma_i w_i 1[U_i \subseteq S]$. Доказательство супермодулярности сохраняется при $w_i \geq 0$. В выполненном опыте веса равны единице.

Эта структура отличается от обычного покрытия множеств: там элемент приносит пользу после попадания хотя бы в одно выбранное множество, здесь сертификат требует выполнения всех входящих проверок. При $|U_i|=2$ и единичных стоимостях задача содержит поиск наиболее плотного подграфа на k вершинах. Поэтому полный перебор служит эталоном малых случаев; его совпадение с оптимумом не является новым алгоритмическим преимуществом.

Сравнение с пакетной эвристикой и просмотром пар в разделе 5.4 проверяет, насколько простое учётное правило можно усилить выбором взаимозависимых узлов. Для больших графов следует отдельно исследовать целочисленное решение, ограничения структуры и известные приближённые методы. Показатели времени и памяти должны измеряться на конкретной реализации; в этой работе масштабируемость не установлена.

Авторский результат относится к гарантированному завершению профессионального аудита: один источник может разрешать несколько обязательств, известное нарушение завершает сертификат раньше, а предметные зависимости ограничивают возможные исходы. Гарантия g(S) предполагает надёжность выбранных проверок. Для модельного судьи необходимо отдельно измерять ошибочные T/F; математическая конструкция не устраняет ошибки извлечения, перевода и спецификации.

10.6. Ожидание завершения и адаптивная политика

Пусть ψ задаёт наблюдённые исходы, $S(\psi)$ — проверенные узлы, $F(\psi)\subseteq S(\psi)$ — узлы с подтверждённым нарушением. Число завершённых сертификатов равно $C(\psi)=\Sigma_i 1[U_i\subseteq S(\psi)$ или $U_i\cap F(\psi)\neq\varnothing]$. Эта функция считает окончательные решения о приёмке и отклонении. Она предполагает достоверность T/F и отсутствие известных нарушений вне перечисленных требований.

Для статического S и независимых исходов с вероятностями нарушения q_j получаем $E[C(S)]=\Sigma_i\, h_i(S)$, где $h_i(S)=1$ при $U_i\subseteq S$ и $h_i(S)=1-\prod_j\in(U_i\cap S)(1-q_j)$ в остальных случаях. Доказательство: при неполной проверке сертификат остаётся незавершённым ровно тогда, когда ни один проверенный входящий узел не нарушен. При пустом пересечении произведение равно 1. Для зависимых исходов произведение заменяется совместной вероятностью отсутствия нарушений на выбранном пересечении. Одних маргинальных q_j для этого недостаточно.

Обозначим $V(\psi,b)$ наибольшее условное ожидание числа завершённых сертификатов при оставшемся бюджете b. Для каждого непроверенного доступного по стоимости узла j рассмотрим $V_j=\Sigma_x\in\{T,F\}P(x_j=x|\psi)V(\psi\cup\{j=x\},b-a_j)$. Тогда $V(\psi,b)=\max\{C(\psi),\max_j V_j\}$. При отсутствии доступного узла V=C. Это равенство следует из разбиения любой политики по её первому действию; после каждого исхода остаётся задача того же вида с меньшим бюджетом. Конечная рекурсия задаёт точный контрольный алгоритм. Возможных сочетаний наблюдённых T/F/U экспоненциально много, поэтому полученное решение предназначено для малых графов.

Задача связана со стохастическим вычислением булевых функций: проверка значения переменной имеет стоимость, а последовательность наблюдений разрешает булев результат [66]. В классической постановке минимизируется ожидаемая стоимость вычисления, тогда как здесь максимизируется число разрешённых конъюнкций при предельном бюджете каждой ветви. Это различие целей сохраняется при использовании общей модели наблюдений.

Теория адаптивной субмодулярности [67] не даёт автоматической гарантии локальной политике для выбранного счётчика C. Контрпример состоит из одного сертификата $U=\{a,b\}$ при независимых $q_a=q_\beta=0{,}10$. До наблюдений ожидаемый прирост проверки b равен 0,10: только нарушение завершает неполную конъюнкцию. После подтверждения a=T проверка b завершает сертификат при любом исходе, и её ожидаемый прирост равен 1. Условная отдача возрастает. Поэтому требуемое для адаптивной субмодулярности убывание отдачи здесь нарушается; результаты чужого приближённого алгоритма нельзя переносить без проверки его предпосылок.

10.7. Китайские и индийская модели в протоколе сравнения

Для внешнего испытания выделены конкретные модельные профили. Qwen3-8B и DeepSeek-V3.2 относятся к основному русско-китайскому сравнению. Индийская Sarvam-M включена как проверка переноса вне её основной специализации на индийских языках и английском. Происхождение разработчи-

ка не служит объяснением качества: сопоставляются фиксированные модели, их языковая область и условия запуска [58; 59; 61].

Таблица 22
Профили моделей и границы сопоставления

Профиль	Основание включения	Условие интерпретации
Qwen/Qwen3-8B	Китайский разработчик; многоязычная модель	Основное RU/ZH испытание; фиксируются веса и режим рассуждения
deepseek-ai/ DeepSeek-V3.2	Отдельное китайское семейство моделей	Открытая версия и текущий API рассматриваются раздельно
sarvamai/sarvam-m	Индийское постобучение Mistral-Small; 24B	Перенос вне основной специализации; результат не обобщается на индийские языки

Согласно официальной документации, проверенной 20 октября 2026 года, Sarvam-M недоступна через официальный API; предлагается переход на Sarvam-105B [68]. Запуск открытых весов Sarvam-M и обращение к сервису новой модели имеют разные экспериментальные условия. Обновляемые псевдонимы DeepSeek API также не подтверждают запуск именно DeepSeek-V3.2 [69]. Для воспроизводимости фиксируются ревизии весов и токенизатора, шаблон диалога, точность вычислений, квантование и версия сервера [70].

Для трёх профилей подготовлено 2400 запросов: 400 пар × два режима × три модели. Контракт и факты одинаковы; различается инструкция проверки. Порядок случаев общий, история изолирована, ожидаемые метки исключены. Параметры допускают настройку только на 60 отдельных примерах до фиксации теста. Ревизия весов будет указана при развёртывании; вычислительная среда пока не подготовлена.

Метрики сначала рассчитываются по конструкторским меткам, затем — по независимому эталону. Сравниваются ошибки A/R/D, покрытие по классам, затраты и незавершённые ответы. Разности режимов оцениваются внутри модели с группировкой по семействам; для нескольких гипотез заранее выбирается поправка на множественные сравнения. Из-за разных токенизаторов затраты фиксируются также в длине текста, времени и оплаченных единицах. Запросы сохранены в model_comparison; новых модельных ответов нет.

11. Область применимости и ограничения исследования

Выводы опираются на вычислительные опыты и пересчёт предоставленных модельных журналов. Подготовленные внешние проверки результатов пока не дали. Ниже определены ограничения программной реализации, эталона и переноса на профессиональные документы.

11.1. Разделение статусов экспериментов

По состоянию на 20 октября 2026 года выполнены контрольный, статический и адаптивный бюджетные опыты, проверка устойчивости интерфейса, конечные математические проверки и пересчёт журналов Kimi. Для внешнего этапа подготовлены корпус v2, экспертные формы, 800 запросов GigaChat и 2400 запросов трёх других модельных профилей. Новые модельные прогоны и независимая экспертная разметка не проведены.

Архив Kimi содержит входы без меток, два журнала вердиктов и исходную таблицу разметки. Их идентификаторы, покрытие и контрольные суммы согласованы при пересчёте. Условия первоначального запуска, включая дату и слепой режим, приводятся по описанию предоставленного опыта. Полные исходные запросы, параметры генерации и неизменяемая версия развёртывания отсутствуют, поэтому точное повторение первоначального инференса не обеспечено. Контрольная сумма фиксирует байты файла и сама по себе не доказывает изоляцию проверяющего от меток.

11.2. Независимый эталон

Конструкторские метки определяются правилами генератора. Их программная проверка устанавливает согласованность с этими правилами, но сохраняет общий источник возможного смещения. Независимый эталон должны сформировать внешние специалисты, не участвовавшие в создании метода и корпуса и не видевшие авторских меток или модельных вердиктов.

Подготовлены две пустые формы независимой разметки всех 400 пар и форма арбитража. Оба эксперта или обе экспертные группы оценивают один и тот же полный контракт и должны владеть китайским языком и предметной областью. До арбитража вычисляются матрица решений и номинальная каппа; плановый ориентир $\kappa \geq 0{,}75$ характеризует согласие, а не истинность эталона. Недостаток компетенции разметчика отмечается отдельно и не превращается в D. Независимый эталон создаётся только после разрешения всех случаев; сейчас число первичных внешних оценок равно нулю. Подробный протокол приведён в приложении Л.

11.3. Сильные сравнения при равном бюджете

Четыре ограниченные базовые процедуры позволяют оценить последствия пропуска отдельных проверок. Пакетная эвристика и просмотр пар расширяют бюджетное сравнение. Для сопоставления с полноценным внешним аудитом дополнительно нужны экспертный чек-лист, модельный судья без предлагаемого порядка проверки и альтернативная процедура при одинаковых данных и сопоставимых затратах.

Протокол сравнения: единый корпус, единый лимит времени и стоимости на пару для всех участников, единый независимый эталон (подраздел 11.2). Сопоставляются FAR, FRR, доля отсрочек и стоимость принятого результата. Вывод о превосходстве формулируется только при одновременном выигрыше по ошибкам и по стоимости; выигрыш в одном показателе при проигрыше в другом фиксируется как компромисс, а не как превосходство.

11.4. Перенос за пределы шаблонов

Новый корпус v2 сохраняет синтетическое происхождение. В нём несколько формулировок источников, десять семейств нарушений и четыре причины незавершённой проверки. Поверхностная подсказка «не предоставлено», одинаковая во всех отсрочках старого корпуса, устранена. Однако создание новых шаблонов тем же исследовательским процессом не делает корпус независимым.

Для проверки переноса необходимы документы, отобранные внешними специалистами без доступа к шаблонам, и ошибки, внесённые независимо от разработчика метода. Границы FAR, FRR и ошибочного принятия D следует задать до испытания и оценивать вместе с покрытием, трудозатратами и тяжестью последствий. Неудачные случаи включаются в итоговый отчёт.

11.5. Измерение профессиональной полезности

FAR, FRR и каппа характеризуют решения валидатора. Профессиональная полезность требует дополнительных измерений по плану подраздела 8.3. Данные о внедрении в настоящем исследовании отсутствуют.

Минимальный набор показателей полезности: доля ошибочной приёмки неопределённых случаев (вердикт A при эталоне D), техническое покрытие (доля случаев, по которым процедура вообще выносит вердикт), время эксперта на доразметку отсрочек, стоимость одного принятого результата и классификация последствий ошибок по тяжести. Измерение проводится в пилоте с заранее зафиксированным правилом решения; критерий успеха — снижение полного времени приёмки при неухудшении ошибок относительно экспертной базовой линии.

11.6. Культурный раздел как исследуемый материал

В подразделе 7.3 сформулированы требования к сохранению атрибуции, статуса цитаты и исторического контекста. Их применение к культурному материалу пока не проверено эмпирически.

Для такой проверки можно сформировать корпус заданий на материале наследия Рериха: атрибуция текстов и изображений, точность цитирования, сохранение контекста и допустимые способы адаптации. Профильные эксперты размечают задания; их согласие оценивается каппой Коэна. Влияние ограничений на выразительность исследуется парным сравнением версий по заранее заданным критериям.

Выполненная часть образует воспроизводимое исследование формальной процедуры, её программной реализации и малых бюджетных задач. Независи-

мость экспертного эталона, качество свободных переводов, полезность внедрения и результаты новых моделей остаются внешними ограничениями этого исследования. Выводы о них требуют новых наблюдений и не следуют из замкнутого синтетического контроля.

Заключение

Исследование отвечает на три вопроса о приёмке профессиональных результатов. По RQ1 ограниченные локальные проверки и межъязыковое согласие пропускают разные классы нарушений: в контрольном корпусе схема допускает 648 нарушенных пар, локальные ограничения — 336, согласие RU/ZH — 192. Общая ошибка способна сохраняться в обеих версиях. Для решения о приёмке требуется проверка каждого объявленного обязательства по его основанию.

По RQ2 удаление парной проверки приводит к 48 ложным приёмкам, хотя обе локальные записи допустимы по отдельности. Замена неизвестности на отказ превращает 72 незавершённые проверки в отклонения. Дополнительный опыт подтвердил инвариантность 3840 допустимых преобразований и выявил приёмку некорректного управляющего входа. Строгая проверка интерфейса устранила её на 192 проверенных случаях, сохранив исходные решения. Эти результаты относятся к объявленной структурированной грамматике.

По RQ3 совместный выбор проверок определяется комплементарностью. Получены внешние границы J^{-} и J^{+}, гарантия $g(S)$ и точная рекурсия адаптивного выбора. Статическая постановка отнесена к известным классам оптимизации; адаптивная сопоставлена со стохастическим вычислением булевых функций. В статическом опыте точный выбор превзошёл локальную эвристику в 26 из 160 сравнений. В новом опыте точная адаптация превзошла лучшую статическую политику по ожиданию в 87 из 240 сравнений; средний выигрыш составил 0,1523 сертификата из 28. Ошибочное предположение о независимости уменьшило результат в 10 из 48 дополнительных сравнений. Универсальное преимущество адаптации или жадного правила из этих результатов не следует.

Программный комплект позволяет повторить вычислительные опыты, проверить сохранённые данные и пересчитать журналы Kimi. В нём сохранены полные исходы, рациональные значения ожиданий, выбранные решения и проверки бюджета. Единый корпус, слепые экспертные формы и запросы к Qwen, DeepSeek, Sarvam-M и GigaChat предназначены для следующего внешнего этапа.

Практическая ценность предложенной процедуры состоит в явном разделении подтверждённого выполнения, нарушения и незавершённой проверки. Её внедрение требует предметного аудита самого контракта, оценки ошибок валидатора и измерения полного времени приёмки. Утверждения об эффективности на реальных переводах, окупаемости и образовательном эффекте в выполненной части не проверены.

Приложение А. Парный расчётный пример

Данные сконструированы для анализа метода оценки. Они не являются результатами запуска моделей. Версии A и B выполняют одинаковые 40 условных заданий. Табл. 23 показывает все сочетания исходов и предотвращает выборочный учёт только улучшений.

Таблица 23
Матрица парных исходов

Обе версии правильны	**24 случая.**
Обе версии ошибочны	6 случаев.
A ошибочна, B правильна	8 исправлений.
A правильна, B ошибочна	2 ухудшения.
Итог	A: 26/40=65 %; B: 32/40=80 %; разность 15 процентных пунктов.

Для 10 несовпадающих пар точное двустороннее биномиальное значение равно 0,109375 при нулевой гипотезе симметрии направлений и независимости заданий. На уровне 0,05 это не даёт основания отклонить указанную гипотезу. Такой вывод не означает доказанного отсутствия эффекта и не заменяет проверку критичности двух ухудшений.

Интервал Уилсона для доли B при z=1,96: [0,6524; 0,8950]. Он относится к доле одной версии при биномиальных предпосылках и не является парным интервалом разности. Числа в приложении вычислены стандартной библиотекой Python и сопоставлены с независимой реализацией точного теста в SciPy.

При нуле ошибок односторонняя верхняя граница уровня 95 % равна 1−0,05^(1/n). Для n=100 она составляет 0,029513; для n=1000 — 0,002991. Применение к зависимым повторам без нового основания недопустимо.

Приложение Б. Воспроизводимый код расчётов

Код требует только стандартной библиотеки Python. Он воспроизводит условные числа приложения A. Значения z=1,96 и уровня 0,95 заданы явно. Код не проверяет реальные модели и не отменяет статистических предпосылок, изложенных в разделе 5.

Таблица 24
Расчёт статистических показателей

```
from math import sqrt, comb
def wilson(k, n, z=1.96):
if not (0 <= k <= n and n > 0):
raise ValueError("invalid k or n")
p = k / n
d = 1 + z * z / n
c = (p + z * z / (2 * n)) / d
h = z * sqrt(p * (1-p) / n
+ z * z / (4*n*n)) / d
return c - h, c + h
def exact_paired(b, c):
if b < 0 or c < 0:
raise ValueError("negative count")
n = b + c
if n == 0:
return 1.0
tail = sum(comb(n, j)
for j in range(min(b, c) + 1))
return min(1.0, 2 * tail / (2 ** n))
def zero_error_upper(n, alpha=0.05):
if n <= 0 or not 0 < alpha < 1:
raise ValueError("invalid n or alpha")
return 1 - alpha ** (1 / n)
print(wilson(32, 40))
print(exact_paired(8, 2))
print(zero_error_upper(100))
```

Ожидаемые значения после округления: (0,652424; 0,895001); 0,109375; 0,029513. Для больших n этот простой расчёт биномиальной суммы может требовать более устойчивой численной реализации. Область учебного примера мала и проверена отдельно.

Приложение В. Минимальный валидатор свидетельств

Учебный валидатор проверяет типы, наличие утверждения и ссылочную целостность. Он не устанавливает истинность текста и не проверяет логическую поддержку тезиса источником. Правильное основание может потребовать предметной экспертизы. Код демонстрирует разделение формальной приёмки и семантической оценки.

Таблица 25
Проверка структуры записи и ссылок

```
def validate(record, known_sources):
    if not isinstance(record, dict):
        return ["record_type"]
    errors = []
    status = record.get("status")
    if status not in ("ok", "insufficient"):
        errors.append("status")
    claims = record.get("claims")
    if not isinstance(claims, list):
        return errors + ["claims_type"]
    if status == "ok" and not claims:
        errors.append("empty_ok")
    for i, claim in enumerate(claims):
        if not isinstance(claim, dict):
            errors.append(f"claim_type:{i}")
            continue
        text = claim.get("text")
        if not isinstance(text, str) or not text.strip():
            errors.append(f"text:{i}")
        refs = claim.get("evidence")
        if not isinstance(refs, list) or not refs:
            errors.append(f"evidence:{i}")
            continue
        for ref in refs:
            if not isinstance(ref, str):
                errors.append(f"ref_type:{i}")
            elif ref not in known_sources:
                errors.append(f"unknown_ref:{i}")
    return errors
```

Отрицательные случаи: строка вместо записи, неизвестный статус, пустой результат ok, строка вместо массива, пустой тезис, пустые ссылки, неверный тип ссылки и неизвестный идентификатор. Успешная формальная проверка не даёт права сообщить, что факт подтверждён по существу.

Приложение Г. Основания и проверка основных результатов

Реестр связывает центральные утверждения с доказательствами, выполненными проверками и границами применения. В комплект входят новая реализа-

ция контрольного и бюджетного опытов, исходные записи, решения и архив предоставленной демонстрации Kimi.

Таблица 26
Научные результаты и доступные основания

Границы J^- и J^+	**Положение 1, доказательство в разделе 2.2; 2187 трёхзначных векторов и 16 384 достройки. При свободной достройке границы достижимы; предметные зависимости требуют уточнения.**
Гарантированное завершение g(S)	Положение 2, доказательство в разделе 2.4; 3600 проверок отдельной реализацией минимума по бинарным исходам. Предполагается надёжное разрешение выбранных узлов.
Комплементарность	Раздел 10.3: аналитическое доказательство супермодулярности; 57 600 конечных проверок. Общая теория имеет предшественников.
Контрпример жадному правилу	Конструкция {A,B} против {C}; отношение 1/m; программные случаи m=2…100. Вывод относится к указанному правилу и остановке.
Диагностика нарушений	960 синтетических пар, семь процедур, 6720 решений; четыре базы и две абляции. Это проверка закрытого контура, не наблюдение моделей.
Бюджетное сравнение	20 новых графов, 160 сравнений четырёх алгоритмов; 26 строгих преимуществ точного выбора над локальной эвристикой. Оптимумы проверены другим способом; промышленный эффект не измерен.
Адаптивное завершение	Разделы 5.5 и 10.6: точная рекурсия, 240 сравнений и 122 880 проверок полных исходов. Дополнительно 48 сравнений зависимости. Вероятности заданы, T/F считаются надёжными.
Устойчивость реализации	Раздел 5.6: 7656 проверок, включая 192 некорректных входа. Устранён дефект управляющего интерфейса; семантика свободного китайского текста не проверяется.

Для повторения локальной части используется команда python run_all.py. Она запускает расчёты, пересчёт архивных журналов и проверки; сетевые модельные вызовы в неё не входят. Файл SHA256SUMS.json фиксирует код и замороженные входы. Происхождение новой реализации и состав архива описаны в приложении К.

Новизна относится к постановке и операционализации совместного профессионального аудита. Историческое первенство всех компонентов, универсальная надёжность и реальная окупаемость не следуют из приведённых оснований.

Приложение Д. Матрица отрицательных проверок

Контрольные случаи предназначены для исследования границ поведения. Их ожидаемые исходы определяются до запуска. Они не заменяют репрезентативную оценку частоты ошибок в эксплуатации.

Таблица 27
Случаи, различающие результат и основание

Неполный расчёт	Отсутствующая цена сохраняется; точный итог не выдумывается.
Конфликт времени	Выбирается применимая версия, а не просто свежий документ.
Нерелевантная ссылка	Рабочий адрес не признаётся поддержкой тезиса автоматически.
Вложенная команда	Данные не расширяют полномочия и не меняют адресата.
Неизвестный исход	Перед повторной записью проверяется статус операции.
Формально валидный объект	Проверяются семантические отношения и диагностические состояния.
Переименованная концепция	Лексическое различие не засчитывается как новый механизм.
Невыполненный тест	Отчёт не объявляет проверку выполненной.

Новый инцидент включается в набор через механизм ошибки. Для таймаута создаётся случай возможной успешной записи без полученного ответа; для устаревшего документа — конфликт применимости. Сходства формулировки недостаточно. Полезный контроль должен проваливаться при сохранённой причине и проходить после проверяемого исправления.

Приложение Е. Проект университетского курса

Программа является авторским проектом и не заявляется как утверждённый образовательный стандарт. Предлагаются восемь лекций по 2 академических часа и восемь лабораторий по 2 часа: всего 32 контактных академических часа. Объём самостоятельной работы определяется организацией. Для каждого занятия указан проверяемый результат.

Таблица 28
Лекционно-лабораторная последовательность

1. Генерация и знание	Различать ответ, объяснение и свидетельство; разобрать ложный вывод.
2. Контракт результата	Определить данные, полномочия, выход и приёмку новой задачи.
3. Контекст и RAG	Найти достаточное основание, проверить версию и доступ.
4. Методы prompting	Сравнить демонстрации, рассуждение и критику на удержанных случаях.
5. Оценка надёжности	Воспроизвести парный расчёт; объяснить его предпосылки и пределы.
6. Агенты и инструменты	Спроектировать переходы, отказ и восстановление после тайм-аута.
7. Отраслевой проект	Применить рамку в креативе, бизнесе либо ИТ с предметным критерием.
8. Защита и воспроизводимость	Представить версии, отказы, основания и ответить на новый случай.

В резервном режиме используются сохранённые ответы и условные данные. Он позволяет оценивать постановку и аргументацию, однако не подтверждает поведение нового сервиса. Итоговое задание требует самостоятельного выбора критерия и независимого основания проверки.

Приложение Ж. Русско-китайская парная оценка

Все исходы сконструированы. Они иллюстрируют показатели раздела 10, а не результаты испытания моделей. Для 40 условных заданий парная совместимость S равна 1 на всех совместно принятых результатах. Табл. 29 задаёт полный учёт исходов.

Таблица 29
Условная матрица языковой приёмки

RU и ZH приняты	28 пар.
Принят только RU	6 пар.
Принят только ZH	4 пары.
Оба не приняты	2 пары.

Локальные доли: RU=34/40=0,85; ZH=32/40=0,80. Совместная приёмка J=28/40=0,70; согласие статусов E=30/40=0,75. Направленные доли: 28/34≈0,823529 и 28/32=0,875. Разность ZH−RU=−0,05; точное двустороннее p=0,75390625.

Интервал Уилсона для J при z=1,96: [0,545697; 0,819253]. Если одна совместно принятая пара нарушает S, получаем J=27/40=0,675 при неизменных локальных долях. Равные 80 % в обоих языках допускают пересечение от 24 до 32 случаев; поэтому локальные точности не определяют J.

Код приложений воспроизводит расчётные примеры и метрики. Они иллюстрируют различие локальной точности, согласия и совместной приёмки, но не являются модельными наблюдениями. Для внешнего испытания используются отдельные входы и формы приложения З; семантика китайских формулировок требует внешней экспертной проверки.

Приложение З. Корпус для пилотной проверки

Для доработанной версии создан новый синтетический корпус v2: 400 тестовых пар в 100 семействах, по одному A, двум R и одному D на семейство. Распределение: A=100, R=200, D=100; бизнес — 136 пар, креатив — 132, ИТ — 132. Отдельно созданы 60 примеров для настройки в 15 семействах. Идентификаторы семейств и исходные реквизиты разделены между тестом и разработкой; общий шаблонный способ создания остаётся ограничением переноса.

Таблица 30
Фиксированные кандидаты пилота и статус исполнения

DeepSeek-V3.2 [58]	**Открытые веса, MIT; запуск не выполнен; FAR, FRR, к отсутствуют.**
Qwen3-8B [59]	Открытые веса, Apache-2.0; запуск не выполнен; FAR, FRR, к отсутствуют.
GigaChat API [64; 65]	Идентификатор зависит от доступа; 800 запросов подготовлены. Открытая модель [60] в этот профиль не входит; инференс не выполнен.
Sarvam-M [61]	Открытые веса, Apache-2.0; Indic/English специализация. RU/ZH — стресс-тест; запуск не выполнен.

Подготовлены 2400 запросов к трём открытым модельным профилям подраздела 10.7 и 800 запросов к API-сервису GigaChat: всего 3200 запросов, по два режима на 400 пар для каждого из четырёх профилей. Ответов новых моделей нет. API-сервис GigaChat не отождествляется с открытыми весами GigaChat-20B-A3B-instruct. Для API перед запуском выбирается доступный идентификатор модели; для открытых весов фиксируется неизменяемая ревизия.

Параметры, исходные ответы и технические ошибки сохраняются отдельно по профилям [64; 65].

Независимую разметку всех 400 пар должны выполнить два внешних эксперта или две группы с одинаковым объёмом компетенций. Первичные формы замораживаются до обсуждения; расхождения разрешает внешний арбитр. Конструкторские метки открываются после этого. Подготовленный пакет для экспертов не содержит этих меток и модельных ответов; независимость конкретных участников дополнительно подтверждается документально.

В комплект включены входы, конструкторские метки, примеры для настройки, пустые экспертные формы, отдельный слепой пакет, клиент API и расчётчик метрик. Все 460 структурированных случаев прошли программную проверку другим способом извлечения исходных фактов. Эта проверка не является экспертной валидацией китайского текста. Внешняя разметка и инференс не выполнены; их статусы сохранены явно.

Приложение И. Анализ журналов Kimi на 300 парах

Предоставленные материалы содержат журналы демонстрации приёмки Kimi на 300 сконструированных парах. Показатели пересчитаны по этим журналам и меткам исходной таблицы без повторного обращения к модели. Полученные результаты относятся только к данному архиву; другие профили — DeepSeek-V3.2, Qwen3-8B, GigaChat-20B-A3B-instruct и Sarvam-M — в нём не испытывались [58–61].

В архивном корпусе 75 семейств по четыре варианта: допустимый кандидат A, два нарушенных кандидата R и отсрочка D. Всего A=75, R=150, D=75; три контекста представлены по 100 случаев. Десять типов мутаций имеют по 15 случаев. Во всех 75 случаях D источник прямо сообщает об отсутствии приложения; следовательно, этот класс допускает распознавание по поверхностной подсказке. В новом корпусе v2 причины неполноты разнообразнее. Архивные входы и журналы сохранены без изменения байтов.

Таблица 31
Дизайн демонстрационного прогона

Параметр	Значение
Объём	75 семейств × 4 варианта = 300 пар (источник RU, кандидат ZH)
Распределение эталона	A = 75, R = 150, D = 75
Контексты	business = 100, creative = 100, it = 100
Типы мутаций	10 типов × 15 случаев
Класс D	обязательное приложение со свидетельством не предоставлено
Заморозка	Предоставленные хеши входов и журналов; seeds 777/888 по описанию первоначального запуска. Проверка журнала выполнена заново.

Рассматривались два режима валидатора: базовый (свободная сверка кандидата с источником) и контрактный (тот же контракт и упорядоченный чек-лист обязательств: реквизиты и ссылки; стороны и атрибуция; количество, цены, расчёты и валюта; сроки; глаголы обязательства; лицензии и территория; полнота полей). Метрики соответствуют определениям раздела 10.1: FAR, FRR, отсрочка допустимых, блокировка и каппа Коэна [55]; значения вычислены в точных рациональных дробях. Матрицы ошибок приведены в табл. 32 и 33.

Таблица 32
Матрица ошибок базового режима (строки — эталон)

	pred A	pred R	pred D	итого
gold A	75	0	0	75
gold R	3	147	0	150
gold D	0	0	75	75
итого	78	147	75	300

Таблица 33
Матрица ошибок контрактного режима (строки — эталон)

	pred A	**pred R**	**pred D**	**итого**
gold A	75	0	0	75
gold R	0	150	0	150
gold D	0	0	75	75
итого	75	150	75	300

Сводные показатели представлены в табл. 34.

Таблица 34
Метрики демонстрационного прогона

Метрика	**Базовый режим**	**Контрактный режим**
FAR	1/50 = 2,00 %	0 = 0,00 %
FRR	0 = 0,00 %	0 = 0,00 %
Отсрочка допустимых	0 = 0,00 %	0 = 0,00 %
Блокировка допустимых	0 = 0,00 %	0 = 0,00 %
Наблюдаемое согласие p_o	99/100 = 99,00 %	1 = 100,00 %
Каппа Коэна κ	247/251 ≈ 0,9841	1,0000

Условия первоначального запуска приведены в таблице 35 по предоставленному описанию. Пересчёт подтвердил полноту журналов и согласованность метрик. Он не подтверждает независимо дату инференса, изоляцию среды или неизменяемую версию модели.

Таблица 35
Условия исполнения и изоляции эталона

Параметр	Значение
Модель-валидатор	Kimi (Moonshot AI) по предоставленному описанию; точная версия развёртывания не указана. В этой редакции выполнен пересчёт журналов без повторного инференса.
Дата прогона	7–8 октября 2026 г. по описанию переданного опыта; дата независимо не удостоверена.
Режимы	Базовый (свободная сверка кандидата с источником) и контрактный (тот же контракт и упорядоченный чек-лист из семи пунктов)
Инструкция базового режима	Сравнить кандидата с источником; вынести вердикт A, R или D с обоснованием; метки эталона не сообщаются
Инструкция контрактного режима	Проверять по порядку: реквизиты и ссылки; стороны и атрибуция; количество, цены, расчёты и валюта; сроки; глаголы обязательства; лицензии и территория; полнота полей
Изоляция эталона	По описанию: входы без меток, эталон отдельно, перемешивание с seeds 777/888. Полный первоначальный запрос и подтверждение изоляции среды не представлены.
Журналы	JSONL-записи {case_id, verdict, reason}; покрытие 300/300 в обоих режимах; допустимость значений и соответствие идентификаторов проверены программно
Файлы и sha256	Хеши входов и обоих журналов подтверждены (см. комплект). Метки извлечены из XLSX; отдельный первоначальный файл эталона с хешем 109051dc…f88c0edb не предоставлен.

Три расхождения базового режима с конструкторскими метками относятся к пропуску территории использования и двум подменам срока поддержки с 12 на 24 месяца. В контрактном журнале все три случая отклонены. Наблюдение согласуется с чувствительностью к явному чек-листу, но трёх исправлений недостаточно для вывода об устойчивом преимуществе. В базовом режиме совпадение с метками составило 100/100 для бизнеса, 99/100 для креатива и 98/100 для ИТ; в контрактном — 100/100 во всех контекстах.

Эталон сформирован программным конструктором; в исходном описании создание шаблонов и их оценка связаны с одной моделью. Разделение файлов

не устраняет возможное общее смещение. Китайский текст не проходил независимую экспертную разметку. Поэтому κ=1 характеризует совпадение с конструкторскими метками данного корпуса и не удостоверяет качество приёмки реальных переводов. Сведения о точной версии модели и полном первоначальном запросе отсутствуют.

Архив Kimi сохранён в legacy, пересчёт выполняет audit_legacy.py. Контрольные суммы и происхождение файлов зафиксированы в исследовательском комплекте.

Приложение К. Воспроизводимый исследовательский комплект

Комплект версии 20 октября 2026 года содержит код, замороженные входы и полные результаты. Исходный код раннего опыта не был предоставлен; новая реализация не объявляется побайтовым восстановлением прежнего запуска. Происхождение материалов описано в PROVENANCE.json.

Команда python run_all.py повторяет локальные расчёты стандартной библиотекой Python 3.10 или новее и не отправляет сетевых запросов. run_research.py и researchlib.py строят 960 контрольных записей, 6720 решений семи процедур и 160 статических бюджетных сравнений. verify_research.py проверяет конечные свойства и оптимумы другим представлением. Полные входы находятся в data, результаты — в results.

Скрипт adaptive_audit.py повторяет 240 адаптивных сравнений и 48 проверок чувствительности к зависимости. В results/adaptive_world_audit.jsonl сохранены 122 880 записей с полными исходами, порядком проверок, расходом и числом завершённых сертификатов. verify_saved_worlds.py пересчитывает эти записи без импорта алгоритма выбора. robustness_audit.py повторяет 7656 случаев и проверяет строгий входной интерфейс.

Каталог legacy сохраняет исходные входы, оба журнала Kimi, предоставленную таблицу и извлечённые из неё метки. Пересчёт выполняет audit_legacy.py. SHA256SUMS.json фиксирует код, инструкции и замороженные данные; команда python verify_manifest.py проверяет их целостность. Хеши не удостоверяют истинность разметки и происхождение первоначального инференса.

В pilot_v2 находятся 400 тестовых и 60 настроечных случаев. Входы без меток, авторские ожидания и экспертные формы хранятся раздельно. Экспертам передаётся только expert_blind_package.zip. Внешние оценки не получены; независимый эталон не сформирован. Контроль структуры 460 случаев не заменяет лингвистическую и предметную экспертизу.

Скрипт prepare_model_comparison.py подготавливает 2400 запросов к Qwen, DeepSeek и Sarvam-M без сетевых вызовов; условия развёртывания описаны в model_comparison_protocol.md. Клиент gigachat_runner.py отдельно поддерживает подготовку и фактический запуск; подготовлены ещё 800 запросов GigaChat. Всего новых модельных ответов — 0. Для запуска требуются настроенный доступ или сервер с фиксированной ревизией весов. Ключи читаются из

переменных окружения и не входят в комплект; сохраняются исходные ответы, параметры, usage, задержки и технические ошибки.

Приложение Л. Протокол независимого экспертного эталона

Все 400 пар оценивают два внешних специалиста или две независимые группы по одному полному контракту. Необходимы компетенции в китайском языке и соответствующей предметной области. До начала фиксируются квалификация, отсутствие участия в конструировании корпуса и отсутствие доступа к авторским меткам и модельным вердиктам. В комплект включены незаполненные формы.

Эксперту передаётся только expert_blind_package.zip с входами, пустой формой и протоколом. Полный исследовательский архив содержит конструкторские метки и не служит слепым пакетом. Для начала разметки фиксируются хеш входов, дата передачи и декларация независимости; после завершения замораживаются заполненные формы до обсуждения расхождений.

Правило R требует подтверждённого нарушения с указанием поля и свидетельства. Правило D применяется при отсутствии известного нарушения и незавершённой существенной проверке. Правило A требует подтверждения всех обязательств. Если эксперт не обладает необходимой компетенцией либо не понимает условие, он оставляет метку пустой и описывает затруднение; это не автоматически D. Стилистические предпочтения вне контракта не меняют метку.

Форма содержит case_id, label, reason, evidence и uncertainty_note. Обе независимые оценки относятся к одинаковому решению о полном контракте. До арбитража публикуются матрица 3×3, доля согласия, номинальная каппа и покрытие. При неполных формах каппа по совместно заполненным случаям не объявляется показателем всего корпуса. Плановый ориентир $\kappa \geq 0{,}75$ определяет согласованность, но не подтверждает истинность эталона.

Расхождения разрешает третий внешний специалист с записью основания в adjudication.csv. Авторские метки открываются после заморозки первичных оценок и арбитража. Изменение рубрики требует новой версии и повторной разметки; исходные решения сохраняются. Скрипт expert_audit.py создаёт independent_gold.jsonl только при разрешении всех 400 случаев. Он проверяет формат и полноту, но не удостоверяет личность и независимость людей.

На дату подготовки этой версии внешние оценки не получены, независимый эталон не сформирован. Протокол предусматривает его получение и последующее испытание на профессиональных документах, отобранных независимо от авторских шаблонов.

Список литературы

1. Schulhoff S., Ilie M., Balepur N. et al. The Prompt Report: A Systematic Survey of Prompt Engineering Techniques. — 2024; версия v6, 2025. URL: https://arxiv.org/abs/2406.06608 (07.10.2026).

2. Vaswani A., Shazeer N., Parmar N. et al. Attention Is All You Need. — 2017. URL: https://arxiv.org/abs/1706.03762 (07.10.2026).

3. Brown T. B., Mann B., Ryder N. et al. Language Models are Few-Shot Learners. — 2020. URL: https://arxiv.org/abs/2005.14165 (07.10.2026).

4. OpenAI. Prompt engineering: официальная документация API. URL: https://developers.openai.com/api/docs/guides/prompt-engineering (07.10.2026).

5. Google. Prompt design strategies: Gemini API. URL: https://ai.google.dev/gemini-api/docs/prompting-strategies (07.10.2026).

6. Turpin M., Michael J., Perez E., Bowman S. R. Language Models Don't Always Say What They Think: Unfaithful Explanations in Chain-of-Thought Prompting. — 2023. URL: https://arxiv.org/abs/2305.04388 (07.10.2026).

7. Ouyang L., Wu J., Jiang X. et al. Training language models to follow instructions with human feedback. — 2022. URL: https://arxiv.org/abs/2203.02155 (07.10.2026).

8. Farquhar S., Kossen J., Kuhn L., Gal Y. Detecting hallucinations in large language models using semantic entropy: авторское пояснение к исследованию в Nature. — Oxford Applied and Theoretical Machine Learning, 19.06.2024. URL: https://oatml.cs.ox.ac.uk/blog/2024/06/19/detecting_hallucinations_2024.html (07.10.2026).

9. Cimatti A., Tian C., Tonetta S. Assumption-Based Runtime Verification with Partial Observability and Resets // Runtime Verification. — 2019. — LNCS. — Vol. 11757. — P. 165–184. URL: https://link.springer.com/chapter/10.1007/978-3-030-32079-9_10 (07.10.2026).

10. Liu N. F., Lin K., Hewitt J. et al. Lost in the Middle: How Language Models Use Long Contexts. — 2023. URL: https://arxiv.org/abs/2307.03172 (07.10.2026).

11. Anthropic. Effective context engineering for AI agents. — 2025. URL: https://www.anthropic.com/engineering/effective-context-engineering-for-ai-agents (07.10.2026).

12. Lewis P., Perez E., Piktus A. et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. — 2020. URL: https://arxiv.org/abs/2005.11401 (07.10.2026).

13. Sclar M., Choi Y., Tsvetkov Y., Suhr A. Quantifying Language Models' Sensitivity to Spurious Features in Prompt Design or: How I learned to start worrying about prompt formatting. — 2023. URL: https://arxiv.org/abs/2310.11324 (07.10.2026).

14. Batorski P., Pourhadi A., Sarosiek J., Spurek P., Swoboda P. Spurious Prompts: Can Irrelevant Prompts Steer Large Language Models? — 2026. URL: https://arxiv.org/abs/2605.29678 (07.10.2026).

15. OpenAI. Reasoning best practices: официальная документация API. URL: https://developers.openai.com/api/docs/guides/reasoning-best-practices (07.10.2026).

16. Wei J., Wang X., Schuurmans D. et al. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. — 2022. URL: https://arxiv.org/abs/2201.11903 (07.10.2026).

17. Wang X., Wei J., Schuurmans D. et al. Self-Consistency Improves Chain of Thought Reasoning in Language Models. — 2022. URL: https://arxiv.org/abs/2203.11171 (07.10.2026).

18. Huang J., Chen X., Mishra S. et al. Large Language Models Cannot Self-Correct Reasoning Yet. — 2023; версия v2, 2024. URL: https://arxiv.org/abs/2310.01798 (07.10.2026).

19. Khattab O., Singhvi A., Maheshwari P. et al. DSPy: Compiling Declarative Language Model Calls into Self-Improving Pipelines. — 2023. URL: https://arxiv.org/abs/2310.03714 (07.10.2026).

20. OpenAI. Working with evals: официальная документация API. URL: https://developers.openai.com/api/docs/guides/evals (07.10.2026).

21. Liang P., Bommasani R., Lee T. et al. Holistic Evaluation of Language Models. — 2022. URL: https://arxiv.org/abs/2211.09110 (07.10.2026).

22. Ribeiro M. T., Wu T., Guestrin C., Singh S. Beyond Accuracy: Behavioral Testing of NLP Models with CheckList // Proceedings of ACL. — 2020. — DOI: 10.18653/v1/2020.acl-main.442. URL: https://aclanthology.org/2020.acl-main.442/ (07.10.2026).

23. Brown L. D., Cai T. T., DasGupta A. Interval Estimation for a Binomial Proportion // Statistical Science. — 2001. — Vol. 16, no. 2. — DOI: 10.1214/ss/1009213286. URL: https://projecteuclid.org/journals/statistical-science/volume-16/issue-2/Interval-Estimation-for-a-Binomial-Proportion/10.1214/ss/1009213286.pdf (07.10.2026).

24. Zheng L., Chiang W.-L., Sheng Y. et al. Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena. — 2023. URL: https://arxiv.org/abs/2306.05685 (07.10.2026).

25. Guo C., Pleiss G., Sun Y., Weinberger K. Q. On Calibration of Modern Neural Networks. — 2017. URL: https://arxiv.org/abs/1706.04599 (07.10.2026).

26. Anthropic. Building effective agents. — 19.12.2024. URL: https://www.anthropic.com/engineering/building-effective-agents (07.10.2026).

27. Yao S., Zhao J., Yu D. et al. ReAct: Synergizing Reasoning and Acting in Language Models. — 2022. URL: https://arxiv.org/abs/2210.03629 (07.10.2026).

28. Model Context Protocol. Specification: версия 2025-11-25. URL: https://modelcontextprotocol.io/specification/2025-11-25 (07.10.2026).

29. Greshake K., Abdelnabi S., Mishra S. et al. Not what you've signed up for: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection. — 2023. URL: https://arxiv.org/abs/2302.12173 (07.10.2026).

30. NIST. AI Risk Management Framework: официальные материалы; Generative AI Profile, 2024. URL: https://www.nist.gov/itl/ai-risk-management-framework (07.10.2026).

31. OWASP. LLM01:2025 Prompt Injection: OWASP Gen AI Security Project. URL: https://genai.owasp.org/llmrisk/llm01-prompt-injection/ (07.10.2026).

32. Organization of American States. C-3: Treaty on the Protection of Artistic and Scientific Institutions and Historic Monuments (Roerich Pact): регистрационные сведения. — 1935. URL: https://www.oas.org/juridico/english/sigs/c-3.html (07.10.2026).

33. Nicholas Roerich Museum. Roerich Pact and Banner of Peace: история движения и текст договора. URL: https://www.roerich.org/roerich-pact.php (07.10.2026).

34. World Intellectual Property Organization. Generative AI: Navigating intellectual property. — 2024. URL: https://www.wipo.int/publications/en/details.jsp?id=4713 (07.10.2026).

35. Google. Image understanding: Gemini API. URL: https://ai.google.dev/gemini-api/docs/image-understanding (07.10.2026).

36. Jimenez C. E., Yang J., Wettig A. et al. SWE-bench: Can Language Models Resolve Real-World GitHub Issues? — 2023. URL: https://arxiv.org/abs/2310.06770 (07.10.2026).

37. OpenAI. Structured model outputs: официальная документация API. URL: https://developers.openai.com/api/docs/guides/structured-outputs (07.10.2026).

38. NIST. Secure Software Development Framework (SSDF) Version 1.1: Recommendations for Mitigating the Risk of Software Vulnerabilities. — 2022. — NIST SP 800-218. — DOI: 10.6028/NIST.SP.800-218. URL: https://csrc.nist.gov/pubs/sp/800/218/final (07.10.2026).

39. Sculley D., Holt G., Golovin D. et al. Hidden Technical Debt in Machine Learning Systems // Advances in Neural Information Processing Systems. — 2015. — Vol. 28. URL: https://papers.nips.cc/paper_files/paper/2015/hash/86df7dcfd896fcaf2674f757a2463eba-Abstract.html (07.10.2026).

40. LangChain. LangSmith Observability. URL: https://docs.langchain.com/langsmith/observability (07.10.2026).

41. Cyberspace Administration of China. Interim Measures for the Management of Generative Artificial Intelligence Services. — 2023. — Articles 2, 11, 14. URL: https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm (07.10.2026).

42. Government of India. Digital Personal Data Protection Act, 2023. — Act No. 22. — Section 11. URL: https://www.meity.gov.in/static/uploads/2024/06/2bf1f0e9f04e6fb4f8fef35e82c42aa5.pdf (07.10.2026).

43. Ministry of Electronics and Information Technology. Digital Personal Data Protection Rules, 2025. — G.S.R. 846(E), 13.11.2025. — Rule 1. URL: https://www.meity.gov.in/static/uploads/2025/11/53450e6e5dc0bfa85ebd78686cadad39.pdf (07.10.2026).

44. White J., Fu Q., Hays S. et al. A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT. — 2023. URL: https://arxiv.org/abs/2302.11382 (07.10.2026).

45. Hewing M., Leinhos V. The Prompt Canvas: A Literature-Based Practitioner Guide for Creating Effective Prompts in Large Language Models. — 2024. URL: https://arxiv.org/abs/2412.05127 (07.10.2026).

46. Ahn J., Kim M. From Prompts to Contracts: Harness Engineering for Auditable Enterprise LLM Agents. — 2026. URL: https://arxiv.org/html/2607.08028v1 (07.10.2026).

47. Tufano M., McClure J., Cambronero J. et al. Grounding AI Agents in Contracts: An Empirical Evaluation of Spec-Driven Test Generation. — 2026. URL: https://arxiv.org/abs/2608.17177 (07.10.2026).

48. Liu T., Qi J., Sachan M., Cotterell R., Fernández R., Bisazza A. Post-Training Language Models for Crosslingual Consistency // Proceedings of ICML 2026. — PMLR. — Vol. 306. — P. 77970–77994. URL: https://proceedings.mlr.press/v306/liu26dd.html (07.10.2026).

49. Han W., Zhang Y., Chen Z. et al. MuBench: Assessment of Multilingual Capabilities of Large Language Models Across 61 Languages // Findings of ACL 2026. — 2026. — P. 16163–16192. — DOI: 10.18653/v1/2026.findings-acl.794. URL: https://aclanthology.org/2026.findings-acl.794/ (07.10.2026).

50. Shi F., Suzgun M., Freitag M. et al. Language Models are Multilingual Chain-of-Thought Reasoners. — 2022. URL: https://arxiv.org/abs/2210.03057 (07.10.2026).

51. Li H., Zhang Y., Koto F. et al. CMMLU: Measuring massive multitask language understanding in Chinese. — 2023. URL: https://arxiv.org/abs/2306.09212 (07.10.2026).

52. Xing X., He Z., Xu H., Wang X., Wang R., Hong Y. Evaluating Knowledge-based Cross-lingual Inconsistency in Large Language Models. — 2024. URL: https://arxiv.org/abs/2407.01358 (07.10.2026).

53. Qwen Team. Qwen3.8: официальный репозиторий и описание семейства моделей. — 2026. URL: https://github.com/QwenLM/Qwen3.8 (07.10.2026).

54. DeepSeek-AI. DeepSeek-V3: официальный репозиторий и результаты оценки. — 2024. URL: https://github.com/deepseek-ai/DeepSeek-V3 (07.10.2026).

55. Cohen J. A Coefficient of Agreement for Nominal Scales // Educational and Psychological Measurement. — 1960. — Vol. 20, no. 1. — P. 37–46. URL: https://doi.org/10.1177/001316446002000104 (07.10.2026).

56. Topkis D. M. Supermodularity and Complementarity. — Princeton: Princeton University Press, 1998. — DOI: 10.1515/9781400822539. URL: https://www.degruyterbrill.com/document/doi/10.1515/9781400822539/html (07.10.2026).

57. Miao F., Holmes W. Guidance for generative AI in education and research. — UNESCO, 2023. URL: https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research (07.10.2026).

58. DeepSeek-AI. DeepSeek-V3.2: официальная карточка модели. URL: https://huggingface.co/deepseek-ai/DeepSeek-V3.2 (07.10.2026).

59. Qwen Team. Qwen3-8B: официальная карточка модели. URL: https://huggingface.co/Qwen/Qwen3-8B (07.10.2026).

60. AI-Sage. GigaChat-20B-A3B-instruct: официальная карточка модели. URL: https://huggingface.co/ai-sage/GigaChat-20B-A3B-instruct (07.10.2026).

61. Sarvam AI. Sarvam-M: официальная карточка модели. URL: https://huggingface.co/sarvamai/sarvam-m (07.10.2026).

62. Chlamtáč E., Dinitz M., Konrad C., Kortsarz G., Rabanca G. The Densest k-Subhypergraph Problem. 2016. URL: https://arxiv.org/abs/1605.04284 (19.10.2026).

63. Taylor R. Approximations of the Densest k-Subhypergraph and Set Union Knapsack Problems. 2016. URL: https://arxiv.org/abs/1610.04935 (19.10.2026).

64. Сбер. GigaChat API. Авторизация. Официальная документация. URL: https://developers.sber.ru/docs/ru/gigachat/api/reference/rest/gigachat-api (19.10.2026).

65. Сбер. GigaChat API. Сгенерировать ответ. Официальная документация. URL: https://developers.sber.ru/docs/ru/gigachat/api/reference/rest/post-chat (19.10.2026).

66. Deshpande A., Hellerstein L., Kletenik D. Approximation Algorithms for Stochastic Boolean Function Evaluation and Stochastic Submodular Set Cover. — 2013. URL: https://arxiv.org/abs/1303.0726 (20.10.2026).

67. Golovin D., Krause A. Adaptive Submodularity: Theory and Applications in Active Learning and Stochastic Optimization // Journal of Artificial Intelligence Research. — 2011. — Vol. 42. — P. 427–486. — DOI: 10.1613/jair.3278. URL: https://doi.org/10.1613/jair.3278 (20.10.2026).

68. Sarvam AI. Models: текущие и выведенные из API модели. Официальная документация. URL: https://docs.sarvam.ai/api/getting-started/models (20.10.2026).

69. DeepSeek. Your First API Call: модели и совместимый интерфейс. Официальная документация. URL: https://api-docs.deepseek.com/ (20.10.2026).

70. Qwen. Deploying Qwen with vLLM. Официальная документация. URL: https://qwen.readthedocs.io/en/latest/deployment/vllm.html (20.10.2026).

DOI 10.34660/INF.2026.89.17.036

ОПЕРАЦИОННО-СЕМАНТИЧЕСКАЯ МОДЕЛЬ ПРОИЗВОДСТВЕННОГО КОМПЛЕКСА С ИСПОЛЬЗОВАНИЕМ НЕЧЕТКИХ ГРАФОВЫХ ОТНОШЕНИЙ

Котеленко Сергей Анатольевич
кандидат технических наук, доцент
Национальный исследовательский технологический университет «МИСиС»,
Москва, Россия
Кузнецов Денис Сергеевич
старший преподаватель
Национальный исследовательский технологический университет «МИСиС»,
Москва, Россия

Аннотация. *Рассматриваются подходы к построению графовой операционной модели управления производственным комплексом. Предлагается графовая модель, состоящая из нескольких кластеров: кластера ресурсов, кластера материалов и кластера производственного календаря. Для учета стохастической природы металлургического производства и экспертной неопределенности в модель вводятся нечеткие отношения, позволяющие оценивать степень заменимости материалов, предпочтительности использования ресурсов и вариативности мощности. Показано, что предложенная нечеткая графовая модель может быть использована для решения таких задач планирования, как формирование главного производственного плана, планирование потребности в материалах и мощностях. Отмечается, что основными преимуществами новой модели являются открытость для дальнейших изменений, возможность применения алгоритмов нечеткого поиска на графах для решения многокритериальных оптимизационных задач, совместимость с современными графовыми базами данных, а также возможность для специалистов по оптимизации бизнес-процессов работать с моделью напрямую.*

Ключевые слова: *Производственное планирование, планирование потребности в материалах т мощностях, комплексное планирование, графовая операционная модель, нечеткие графы, нечеткие отношения*

1. ВВЕДЕНИЕ

Современное металлургическое производство характеризуется сложной конфигурацией производственного оборудования, широким ассортиментом продукции и постоянно меняющимися требованиями рынка. Сложная конфигурация оборудования и динамично меняющийся спрос характеризуются большим количеством стохастических процессов, влияющих на эффективность управления производством и, в частности, на решение задачи комплексного производственного планирования. Масштабы такого производства, в свою очередь, характеризуются несколькими тысячами заказов на продажу в месяц, десятками тысяч наименований номенклатуры, сотнями поставщиков, высокой вариативностью технологических маршрутов и использованием динамических бизнес-моделей.

Большинство современных моделей управления производством строятся на платформе, реализующей реляционную модель данных. Этот подход имеет существенные недостатки при представлении данных со сложной иерархической структурой. В реляционной модели моделирование данных возможно только с помощью плоских табличных связей. В результате современные операционные модели, реализованные в рамках реляционного подхода, являются излишне сложными, что напрямую сказывается на производительности, поскольку иерархическая структура данных распределена по множеству таблиц, и для извлечения данных требуется большое количество операций соединения (join).

Графовые операционные модели, предложенные в работах (Caux 2000), (Zhang 1996), (Tick 2007), либо строятся на основе модифицированной модели ориентированного асинхронного графа, либо являются развитием онтологических моделей с введенной семантикой (Kharlamov, 2016); (Tilstra, 2010). В данной статье предлагается графовая операционная модель управления производственным комплексом, которая позволяет учитывать особенности производства и может быть применена для решения многоуровневых задач комплексного планирования. Особенность предложенного подхода заключается в возможности применения принципов семантической сети для построения моделей анализа, при этом семантический анализ сочетается с возможностью применения алгоритмов поиска на графах, включая алгоритмы поиска на нечетких графах.

Под управлением производственным комплексом понимается разработка комплексного производственного плана и его выполнение с учетом своевременного реагирования на возникающие производственные сбои и меняющиеся условия спроса.

Операционную модель предлагается строить в несколько этапов:

• эмпирическое объединение операций, составляющих производственные процессы;

• проверка взаимосвязей между отдельными процессами;

• формирование параметров операций и определение временных характеристик;

• описание модели, основанной на теории графов;
• введение семантики и нечетких отношений в графовую модель.

2. ПОСТАНОВКА ЗАДАЧИ

Необходимо разработать графовую модель интегрированного планирования производства для решения задачи построения производственных планов и графиков, а также графиков изменения запасов с целью минимизации общих затрат на производство на заданном горизонте планирования и обеспечения равномерной загрузки производственных единиц (ресурсов).

Рассматриваемая модель должна соответствовать более общим моделям, используемым в настоящее время для решения задач управления цепочками поставок (SCM) или управления запасами, поскольку проблема, рассматриваемая в этой статье, является частным случаем построения модели SCM.

Рассмотрим формирование целевой функции интегрального планирования. Пусть xi - неотрицательный объем производствав периоде i, yi - запасы на конец периода i, si - заданный спрос на отдельный продукт в периоде i, ci(z) – стоимость производства
$z \geq 0$ единиц, hi(z) – стоимость хранения $z \geq 0$ единиц за период i, i = 1,..., n.. Предположим, что ci (0) = hi(0) = 0

Задача состоит в том, чтобы выбрать графики производства и ресурсов таким образом, чтобы минимизировать затраты на n-й период

$$C(x, y) = \sum_{1}^{n} [c_i(x_i) + h_i(y_i)] \qquad (1)$$

с ограничением сохранения запасов

$$x_i + y_{i-1} - y_i = s_i, \quad i = 1, ..., n, \qquad (2)$$

и неотрицательность производства и запасов (последнее обеспечивает выполнение требований по мере их возникновения, без предварительных заказов) $x, y \geq 0$, где для простоты мы устанавливаем x0 = yn = 0

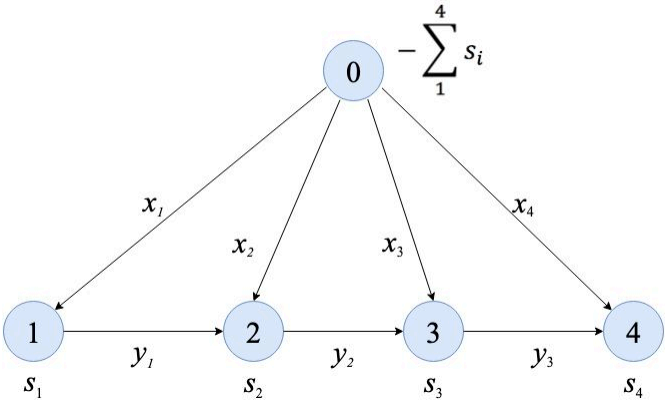


Рис. 1. Обобщенный ориентированный граф планирования производства

Целевая функция и ограничения в задаче (1) – (3) могут быть представлены в виде графовой модели, представленной на рис. 1, где x – требуемый выпуск продукта, y - объем запасов продукта j на конец периода, s – спрос на продукт j. Ориентированный граф (см. рис. 1) иллюстрирует общую проблему, решаемую при планировании производства - минимизацию производственных затрат в каждом периоде. Эта модель также вводит понятие материального потока как движения материалов в процессе производства. Следует иметь в виду, что в данном случае под спросом понимается так называемый внутренний спрос, который определяет потребность в материалах на каждом этапе производства. Однако эта модель является чрезвычайно общей и требует постоянного усложнения при конкретном моделировании. Кроме того, наиболее важным параметром операционной модели должен быть горизонт производственного планирования. Ряд характеристик объектов модели существует только в заданном интервале планирования (горизонте).

Представление модели в виде графа позволяет использовать классическую теорию графов для решения ряда задач управления. Однако в предлагаемой модели мы должны учитывать некоторые ограничения. В частности, граф описывает только взаимосвязи между вершинами, но не показывает физическое расположение узлов. Кроме того, на графе отображается конфигурация оборудования и материальных потоков между ними с точки зрения взаимосвязей между объектами управления: ресурсами, запасами, материалами и операциями.

В данной работе предлагается присвоить дугам семантическую определенность, устанавливающую тип отношений между вершинами графа, (Petersen 2016), в соответствии с принципами, рассмотренными в работах (Иванов 2018),

(Долгуй 2020). Это обстоятельство позволяет нам рассматривать операционную графовую модель как семантическую сеть и, соответственно, применять к этой модели механизмы логического вывода, реализованные в специализированных системах. В данной работе мы рассматриваем использование «открытых» графовых баз данных типа MemGraph, TypeDB, которые реализует механизм вывода с использованием языка запросов openCypher. Это позволит использовать уже реализованные алгоритмы кратчайшего пути. Преимущество подхода, основанного на семантической сети заключается в упрощении формальной модели данных и, как следствие, упрощении поддержки и сопровождения модели. Анализ данных выполняется на основе механизма вывода, характерного для семантической сети, например, (Santoso 2017): запрос к данным в целом также отображается графом, что повышает эффективность поиска данных на основе формальных выражений, обеспечивая при этом значительный прирост производительности.

3. ПРЕДСТАВЛЕНИЕ ПРОИЗВОДСТВЕННОГО ПРОЦЕССА В ВИДЕ ГРАФА

Мы расширим рассмотренную выше модель с учетом особенностей металлургического производства. Вводим понятие производственного процесса и маршрута. Производственный процесс описывается как набор технологических маршрутов. Технологический маршрут — это последовательность ресурсов (оборудования) и операций, которые используются для производства продукта, определяемого внешним спросом. В этом случае производственная конфигурация предполагает наличие однотипных ресурсов, например, электропечей, машин непрерывного литья заготовок, нагревательных колодцев, установок отжига и т.д. Каждый ресурс предполагает выполнение с ним одной или нескольких производственных операций. Степень детализации операций зависит от степени детализации (горизонта) планирования. Маршруты в этом случае делятся на первичные и вторичные. Основной маршрут - это последовательность ресурсов, наиболее подходящая с технологической или финансовой точки зрения для данного продукта. Однако маршрут, который будет выбран при составлении плана, определяется текущей производственной ситуацией и указанными ограничениями.

В этой статье рассматривается подход, при котором ориентированный ациклический граф (DAG, назовем его ресурсным графом) используется для моделирования технологических маршрутов. Каждый ресурс (рабочий центр) представлен в виде вершины (узла). Дуга между двумя ресурсами представляет часть маршрута. Операции, которые происходят на данном узле, моделируются не в модели маршрута, а в модели продукта (иерархия продукта). Это связано с тем, что с одним и тем же ресурсом могут выполняться разные операции для разных продуктов, и даже когда это не так, параметры одной и той же операции с данным ресурсом будут разными для разных продуктов (материалов).

Особенностью предлагаемой модели является введение связей между объектами модели: в частности, между вершинами графа ресурсов существует единый тип маршрута связи, семантически указывающий направление потока (рис. 2), назовем его ROUTE. Таким образом, можно выделить начальную вершину s и конечную вершину t , которые, соответственно, определяют начало производственного маршрута и его окончание. Например, на рис. 2 начальная вершина s - блок SMP (разливка металла) - соединена с двумя вершинами HRM1 и HRM2 (станы горячей прокатки 1 и 2). В свою очередь, эти вершины связаны с тремя линиями отжига MAP, HAP1 и HAP2. Конечная вершина не отмечена, но обычно для готовой продукции это склад для последующей отгрузки. Таким образом, на рис. 2 представлены все возможные производственные маршруты от разливки металла до линий отжига включительно.

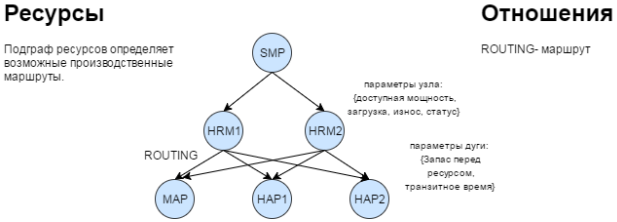


Рис. 2. Фрагмент подграфа ресурсов графовой операционной модели

Дуги этого графа связаны, помимо прочих характеристик, с двумя параметрами: максимальным потоком (мощностью) и текущим значением. Потоки на таком графе проверяются в соответствии с уравнением Кирхгофа, т.е. для каждого узла сумма входящих потоков будет равна сумме исходящих (с учетом потерь). Таким образом, задача о максимальном потоке состоит в том, чтобы найти максимальный поток, соответствующий заданным ограничениям мощности.

Параметры узлов Rn и дуг Re на графе ресурсов:

$$R_n = \{u, l, m, w\}, \; R_e\{y, lt\}, \qquad (3)$$

где u- стандартная мощность, l-минимальная загрузка, m-максимальная загрузка, w-износ ресурса, y-предельный запас до ремонта, lt - время передачи. Типом связи на дуге является ROUTE, который определяет направление потока материала по технологическому маршруту.

4. ГРАФИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ ИЕРАРХИИ ПРОДУКТОВ

Подграф описания продукта также является направленным ациклическим графом, который, в свою очередь, представляет собой модель преобразования материалов из сырья в готовую продукцию в процессе производства. Поскольку DAG является обобщением структуры данных, такой как дерево, можно сказать,

что подграф описания продукта устанавливает так называемую иерархию продукта. На этом графе введены три типа соотношений: DERIVED (произведени из), OPR(операция) и SUBS (замена), которые определяют направление трансформации материалов (горячекатанный рулон из сляба, холоднокатаный рулон из горячекатанного рулона и т.д.), привязку операций к заданному материалу, а также коэффициент замены материалов в производстве при отсутствии заданного.

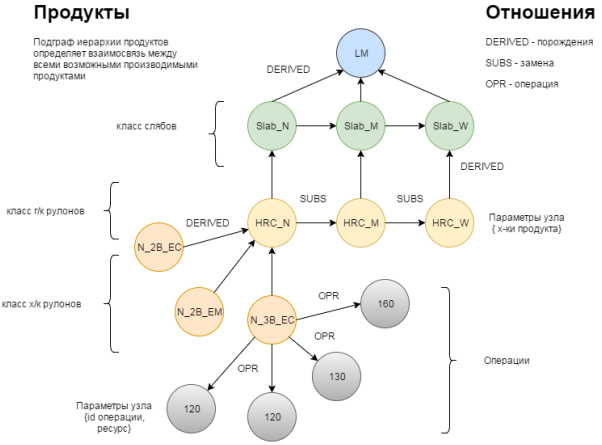


Рис. 3. Фрагмент подграфа иерархии продуктов операционной модели

Для отношений DERIVED задается параметр yield (выход годного), а для отношений OPR – параметры, где pref – индекс предпочтительности выполнения операции на указанном агрегате, если одна и та же операция может быть выполнена на разных агрегатах; sdt - стандартное время выполнения операции на указанном агрегате для соответствующего продукта. Вершины подграфа определяют все возможные материалы в производстве, от полуфабрикатов до готовых изделий. Кроме того, в этой модели вводится понятие классов (групп продукции). Каждый класс объединяет продукты со схожими характеристиками и одинаковыми маршрутами. Продукты одного и того же класса (рис. 3) всегда находятся на одном и том же уровне иерархии и могут быть связаны отношениям SUBS, позволяющим определить продукты (материалы) для замены при их отсутствии на момент планирования. Характеристики вершины определяют значимые параметры изделия, формируя его уникальность, и могут иметь вид:

$$R_n = \left\{ \begin{matrix} f, desc, product_{id}, grade, th_{aim}, \\ th_{\min}, weight_{aim}, weight_{\min}, width \end{matrix} \right\} \quad \textbf{(4)}$$

где f есть класс продукта, desc – описание , product_id – идентификатор материала, grade – марка стали, th_aim – целевого толщина листа в рулоне, th_min минимальная толщина листа, weight_aim – целевой вес рулона, weight_min – минимальная масса рулона, width – ширина рулона.

5. ОПЕРАЦИОННАЯ ПРОИЗВОДСТВЕННАЯ МОДЕЛЬ

Операционная производственная модель представляет собой графовую модель и состоит из двух кластеров: кластера иерархии продуктов и кластера ресурсов. Взаимосвязь между ними определяется через вершины активности в кластере иерархии и ресурсные вершины в кластере ресурсов (рис. 4). Это устанавливает новый тип отношения RES, который определяет ресурс, с которым выполняется данная операция при производстве конкретного продукта. В следующем примере показаны производственные ресурсы, которые включают литье (SMP), горячекатаный прокат (HRM), отжиг (MAP/HAP), холоднокатаный прокат (CRM), цинкование (CGL), отжиг (CAP), резку (CUT) и упаковку (PKG). Для RES зааны параметры: Rres = {p,r] где p – способность ресурса во время операции производить соответствующий материал, r - обычные отходы для заданной операции с этим ресурсом.

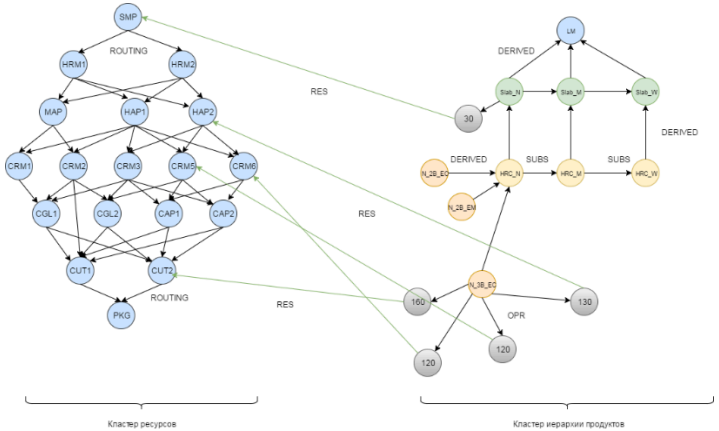


Рис. 4. Фрагмент операционной графовой модели

Операционная графовая модель позволяет решить задачу выбора технологического маршрута с учетом возможности использования альтернативных ресурсов, рассчитать валовую и чистую потребность в материалах для производства заданного объема готовой продукции. Чтобы решить проблемы производственного планирования, мы вводим время в приведенную выше модель. Это позволит решить проблемы планирования и привести операционную графовую модель к форме ориентированного графа производственного планирования. В частности, в данной статье будут рассмотрены принципы решения задачи планирования

материалов (MRP) и задачи планирования требуемой производительности (CRP) с использованием графовой модели. Построим графическую модель календарного года (рис. 5).

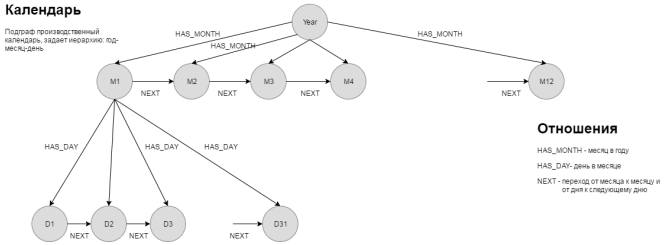


Рис. 5. Фрагмент подграфа календарного года

Вершины M1, M2,..., M12 отражают 12 месяцев в году, следующее соотношение задает последовательность перехода от месяца к месяцу. Каждая вершина месяца связана с вершинами D1, D2 ..., которые определяют дни месяца с помощью отношения has_day. Вершина Year определяет календарный год и связана с вершинами M1, M2,... отношением HAS_MONTH.

Чтобы решить задачу планирования материала, мы связываем каждую вершину, определяющую продукт, с соответствующей вершиной календаря (рис. 6).

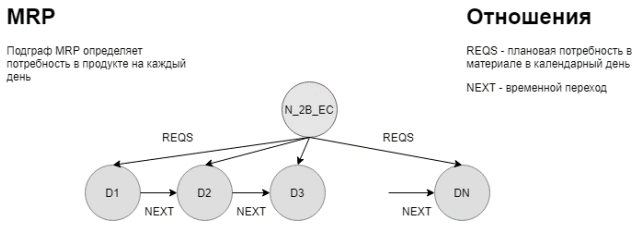


Рис. 6. Фрагмент операционной модели (требования к планированию)

На рис. 6 показан фрагмент оперативной производственной модели, иллюстрирующий взаимосвязь каждого продукта (материала) с календарем. Соотношение REQS определяет потребность в данном материале (в данном случае N_2B_EC) на каждый календарный день. Здесь важно отметить, что параметры дуги между узлом продукта и узлом календарного дня (коэффициент требований) назначаются динамически и изменяются каждый раз, когда план потребности в материалах пересчитывается, пока план не будет реализован. Взаимосвязь характеристик

$$R_{\text{REQS}} = \left\{ v_1 / o_1,\ v_2 / o_2, ...,\ v_n / o_n \right\}, \qquad \textbf{(5)}$$

где каждая характеристика представляет собой пару потребности в конкретном материале vi в данный календарный день согласно порядку oi

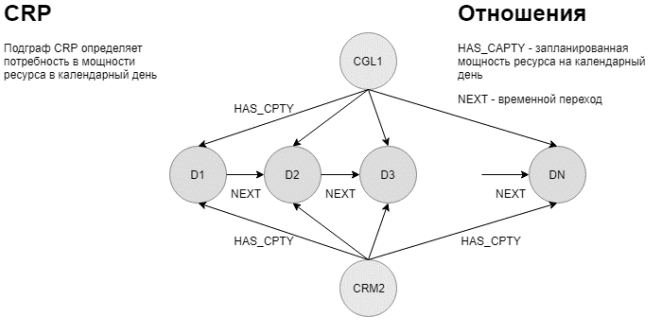


Рис. 7. Фрагмент операционной модели (планирование мощностей)

Планирование путем добавления связи между узлами ресурсов операционной модели и календарем. На рис. 7 показан фрагмент модели, определяющий планируемую мощность ресурсов CRM2 и CGL1 на заданном горизонте планирования (календарный месяц). Планируемая мощность определяется как количество времени, которое планируется использовать в течение запланированного периода, в данном случае в день. Характеристики взаимосвязи

$$R_{\text{HAS_CPTY}} = \left\{ t_{p1} / o_1,\ t_{p2} / o_2, ...,\ t_{pn} / o_n \right\}, \qquad (6)$$

где каждая характеристика представляет собой запланированное время работы данного ресурса в заданный период времени для заказа oi.

Таким образом, мы улучшили исходную модель планирования производства, преобразовав исходный ориентированный граф в гораздо более сложную структуру, в которой можно выделить несколько кластеров - временной, ресурсный и операционный. Переход к графовой модели такой структуры позволяет решать задачи планирования производства, которые уже определяются внешним спросом, т.е. спросом на готовую продукцию. Исходными данными для решения задачи планирования производства согласно построенной модели могут служить либо сформированный портфель заказов, либо прогноз, либо комбинация принятых заказов и прогноза, а также уровни запасов, включая незавершенное производство и готовую продукцию.

6. НЕЧЕТКАЯ ГРАФОВАЯ ПРОИЗВОДСТВЕННАЯ МОДЕЛЬ

В реальных условиях металлургического производства многие параметры носят не детерминированный, а вероятностный или экспертный характер. Жесткие, детерминированные отношения, описанные в разделах 3-5, не всегда

адекватно отражают производственную неопределенность. Для устранения этого недостатка предлагается расширить модель, введя аппарат теории нечетких графов (Берштейн, 2005).

Пусть G = (V,E) – исходный семантический граф. Нечеткая графовая модель определяется как G_f = (V, E,μ), где μ : E $\rightarrow$ [0,1] – функция принадлежности, задающая степень истинности отношений между вершинами. Внедрение нечетких отношений предлагается по трем ключевым направлениям:

1. **Нечеткое отношение замены материалов (SUBS)**

В классической модели материал либо можно заменить, либо нельзя. В нечеткой модели вводится параметр $\mu_{SUBS}(v_i, v_j) \epsilon\ [0,1]$, отражающий степень технологической и экономической пригодности материала v_i для замены материала v_j

• Если μ_{SUBS} = 1, замена абсолютная (эквивалентная)

• Если 0 < μ_{SUBS} <1, замена возможна, но требует дополнительных операций (например, дополнительной термообработки) или снижает выход годного продукта. Алгоритм планирования сможет автоматически выбирать компромиссный вариант на основе пороговых значений, заданных технологами

2. **Нечеткая предпочтительность ресурсов (OPR/RES)**

Параметр *pref* в отношении OPR трансформируется в нечеткую оценку $\mu_{RES}(op, res) \in [0,1]$, отражающую степень предпочтительности выполнения операции op на ресурсе res. Эта оценка формируется на основе комплексного критерия, учитывающего текущий износ оборудования (w), энергоэффективность агрегата и квалификацию персонала. Это позволяет перейти от жесткого назначения маршрута к поиску наиболее предпочтительного пути с использованием модифицированного алгоритма Fuzzy Dijkstra

3. **Нечеткая пропускная способность маршрута (ROUTE)**

Параметры максимального потока m и минимальной загрузки l на дугах графа ресурсов представляются в виде нечетких чисел (например, триплет нечетких чисел $\widetilde{m} = \left(m_{мин}, m_{наиболее_{вероятна}}, m_{макс}\right)$. Это позволяет учитывать стохастические сбои оборудования и вариативность свойств сырья. Задача о максимальном потоке в такой постановке трансформируется в задачу поиска максимального нечеткого потока, что дает плановику не одно жесткое значение, а интервал возможных производственных мощностей с заданной функцией принадлежности.

Введение нечетких отношений не разрушает семантическую целостность модели, а лишь обогащает атрибуты дуг и вершин, позволяя графовой базе данных выполнять запросы с использованием нечеткой логики (Fuzzy Cypher extensions), что значительно повышает адекватность модели реальным бизнес-процессам.

7. ДИНАМИЧЕСКАЯ ПРОИЗВОДСТВЕННАЯ МОДЕЛЬ

Для решения практических задач приведенная выше модель (включая нечеткий сегмент) рассматривается как MDM-модель, на основе которой формируется динамический сегмент модели. Рассмотрим задачу расчета валовых потребностей в материалах для конкретного портфеля заказов. Позиция сбытового заказа представлена узлом подграфа динамических данных, т.е. данными, сгенерированными во время решения задачи планирования материалов. Узел содержит набор характеристик, включая уникальный идентификатор материала, который позволяет установить взаимосвязь с соответствующим узлом подграфа иерархии продуктов, определяя таким образом всю цепочку производства конечного материала (набора полуфабрикатов). Характеристики каждого полуфабриката позволяют рассчитать необходимое количество каждого материала для производства заданного объема готовой продукции (рис.8).

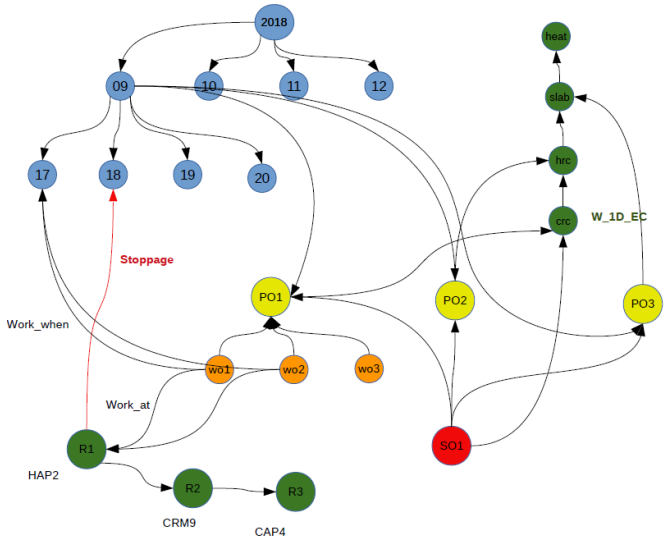


Рис. 8. Фрагмент действующей модели

В модель вводятся новые узлы – производственные заказы PO = {PO_1, PO_2,..., PO_n}, каждый из которых определяет необходимое количество материалов для производства элемента заказа на продажу. Для каждого из этих узлов устанавливается связь с соответствующим узлом подграфа календаря, который определяет дату, к которой должен быть доступен весь объем полуфабриката или готовой продукции. Количество материала в производственном заказе определяется соответствующим набором узлов WO = {WO_1, WO_2,....WO_m}, так называемых рабочих заказов, которые представляют собой любую физическую форму металла, подлежащего производству, такую как сляб, рулон, заготовка, лист, труба и т.д.

Каждый узел рабочего задания имеет характеристики, которые определяют параметры материала. Поскольку каждый материал в подграфе иерархии продуктов имеет ссылку на производственную операцию, устанавливается связь между рабочим заданием и ресурсом, на котором может быть изготовлено рабочее задание. Последующее решение проблемы балансировки материалов с с учетом нечетких отношений предпочтительности и заменимости позволит определить, на каком ресурсе и в какую дату необходимо выполнить данный рабочий заказ, с учетом наличия ресурсов на требуемом горизонте планирования, обеспечивая наиболее устойчивый и экономически эффективный план.

7. ЗАКЛЮЧЕНИЕ

Представленная графовая операционная модель управления производственным комплексом, расширенная за счет аппарата нечетких графов, позволяет решать задачи производственного планирования с горизонтом от нескольких недель до 12–16 месяцев в условиях неопределенности и стохастических возмущений.

В зависимости от гранулярности входных данных модель позволяет решать задачи планирования продаж и операций (S&OP), формирования главного производственного расписания (MPS), планирования потребности в материалах (MRP) и планирования потребности в мощностях (CRP). Внедрение нечетких отношений для оценки заменяемости материалов и предпочтительности оборудования позволяет находить не просто допустимые, а оптимальные компромиссные маршруты производства.

Дальнейшие исследования связаны с уточнением операционной модели, в частности, с обеспечением возможности построения детальных графиков для каждого ресурса и мониторингом исполнения планов на всех уровнях путем разработки единой операционной модели для планирования и диспетчерского управления (MES) с использованием механизмов нечеткого логического вывода

Список литературы

1. Долгуй А., Иванов Д., Соколов Б. (2020) Реконфигурируемая цепочка поставок: X-сеть. Международный журнал производственных исследований, том 58, выпуск 13,4138-4163

2. Иванов Д. (2018) Структурная динамика и устойчивость в управлении рисками цепочки поставок, Springer

3. Берштейн Л.С., Боженюк А.В. Нечеткие графы и гиперграфы – Москва: Науч. Мир, 2005

4. Caux C., Barth M., De Guio R. (2000). Graph Based Tools for Production Flows Analysis. IFAC Proceedings Volumes, 33, 885–889.

5. Zhang H.C., Kuo T.C. (1996) A Graph-Based Approach to Disassembly Model for End-of-Life Product Recycling. Nineteenth IEEE/CPMT International Electronics Manufacturing Technology Symposium, 247–254

6. Tick J. P-Graph-based workflow modelling (2007). Acta Polytechnica Hungaria, 4, 75–88.

7. Kharlamov E., Grau B.C., Jiménez-Ruiz E., et al. (2016). Capturing industrial information models with ontologies and constraints. 15th Int. Semantic Web Conference, 325–343.

8. Petersen N., Galkin M., Lange C, et al. (2016). Monitoring and automating factories using semantic models. 6th Joint International Semantic Technology Conference, 315–330.

9. Tilstra A.H., Campbell M.I., Wood K.L., Seepersad C.C. (2010). Comparing matrix-based and graph-based representations for product design. 12th international dependency and structure modelling conference, 195–206.

10. Santoso A.F., Supriana I., Surendro K. (2017). Designing Knowledge of The PPC with Semantic Network. Journal of Physics Conference Series, 801.

DOI 10.34660/INF.2026.65.27.041

JUSTIFICATION OF OPTIMAL DISCRETIZATION OF GEOMETRIC MEASUREMENTS OF VERTICAL STEEL TANKS BASED ON SPLINE INTERPOLATION

Bazina Kristina Alekseevna
Bachelor's Degree Student
Ufa State Petroleum Technological University,
Ufa, Russian Federation
Yanchushka Anna Pavlovna
Candidate of Economic Sciences, Associate Professor
Ufa State Petroleum Technological University,
Ufa, Russian Federation

Abstract. *The paper investigates the possibility of reducing the number of control points on a generatrix of a vertical steel tank during geometric inspection, with subsequent reconstruction of the omitted data by spline interpolation. A computational experiment was performed on real data from RVS-20000 and RVS-50000 tanks using cubic interpolation. Quantitative relationships were established between the reconstruction error and the decimation ratio of the original sample. It is shown that spline analysis cannot serve as a substitute for actual measurements when the number of points is reduced below the standard. The results substantiate the necessity of retaining the current standard of 36 points per generatrix.*

Keywords: *vertical steel tank, RVS-50000, ground-based laser scanning, cubic spline, interpolation, measurement error.*

Geometric inspection of vertical steel tanks (RVS) is a mandatory procedure during acceptance, operation, and certification. The accuracy of determining deviations of the tank shell geometry from the vertical directly affects operational safety and the reliability of residual life assessment [1], [2]. According to the standards developed back in the Soviet period, for an RVS-50000 tank the number of control points on a single generatrix is 36. This requirement was formulated empirically, on the basis of long-term operational and metrological experience, yet it has no rigorous mathematical justification from the standpoint of interpolation and approximation theory.

Modern ground-based laser scanning generates high-density point clouds, which raises the question: does the standard of 36 points permit reduction with subsequent spline reconstruction? A positive answer would mean savings in field work time; a negative answer would justify the standard. In this work, this question is resolved through a quantitative analysis of errors on real data from RVS-20000 and RVS-50000.

The aim of the work is to quantitatively assess the error in reconstructing the geometric parameters of a tank by cubic spline interpolation when the number of control points on a generatrix is reduced, and to substantiate on this basis the optimality of the current standard establishing 36 control points.

The study was performed on data from two tanks. RVS-20000 (24 generatrices, 9 points per generatrix) – for verifying the methodology. RVS-50000 (12 generatrices, 36 points per generatrix) – the main object [3]. The analysis is conducted in two stages: first, on RVS-20000, the influence of reducing the number of generatrices on statistical characteristics is studied; then, on RVS-50000, spline reconstruction of omitted points on a generatrix is performed.

To analyze the influence of reducing the number of generatrices on the RVS-20000 data, the following statistical metrics were used (table 1):

Table 1 – Technical characteristics

Name of calculation formula	**Formula for calculation**
Standard deviation (SD)	$\sigma = \sqrt{\dfrac{\sum_{i=1}^{n} \left(x_i - \overline{x} \right)^2}{n-1}}$
Standard error of the mean (SEM)	$SEM = \dfrac{\sigma}{\sqrt{n}}$
Error in determining the maximum	$K_{det} = \dfrac{N_{full}}{N_{reduced}}$
Relative error	$\delta = \dfrac{/ X_{full} - X_{reduced} /}{/ X_{full} /} \times 100\,\%$

The initial real data (No. 1) for RVS-20000 represent deviations from the vertical (in mm) measured on the upper shell course of the tank. The full set contains 24 values corresponding to 24 generatrices.

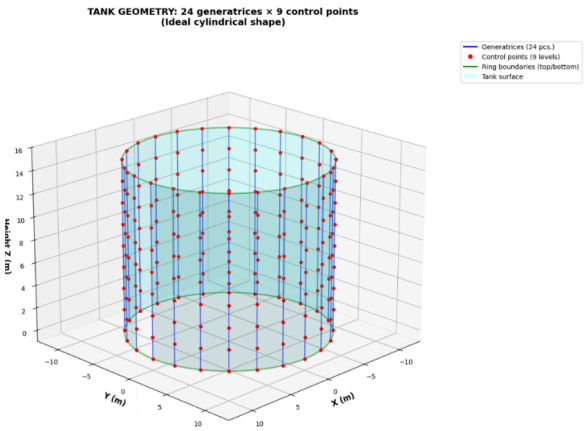


Figure 1 – Construction of an ideal model of RVS-20000

Next, the number of points on the obtained circles was reduced: to 12 (odd-numbered points were taken) and to 6 generatrices (every fourth). Figure 2 presents a comparison of the full and reduced sets.

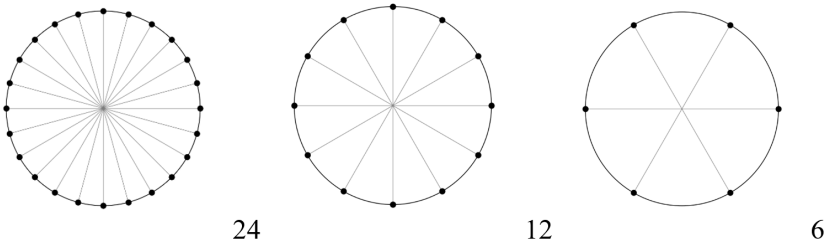


Figure 2 – Geometric representation of the reduction in the number of points on the obtained circles of RVS-20000 for error analysis (24, 12, and 6 points, respectively)

To quantitatively assess the influence of the reduction, statistical characteristics were calculated for each set. The calculation was performed using the formulas in table 1. A key code fragment for the calculation is presented in figure 3.

```python
def calculate_statistics(data):
    mean = np.mean(data)
    sigma = np.std(data, ddof=1)
    sem = sigma / np.sqrt(len(data))
    ci_half = 2 * sem  # t ≈ 2 for 95% confidence
    return mean, sigma, sem, ci_half

# Example of use for a full set of 24 generatrices
data_full = [0, 10, 0, -15, -30, -10, 60, -20, 0, 0, 0, 0,
             50, 30, -20, -60, -50, 0, 0, 0, 0, 50, 50, 0]
mean, sigma, sem, ci_half = calculate_statistics(data_full)
```

Figure 3 – Fragment of the calculation of statistical characteristics for RVS-20000 (Python)

The function calculate_statistics computes the arithmetic mean, standard deviation (with a correction for bias), standard error of the mean, and half-width of the confidence interval at 95% confidence. The results of calculating the statistical characteristics for the full and reduced sets are given in table 2:

Table 2 – Statistical characteristics of deviations for the upper shell course of RVS-20000

Parameter	24 generatrices	12 generatrices	6 generatrices
Mean value, mm	1,88	0,83	−1,67
SD σ, mm	31,6	31,5	31,9
SEM, mm	6,45	9,10	13,02
Relative error of the mean, %	—	55,9	188,8
Relative error of the maximum, %	—	16,7	16,7
Relative error of the minimum, %	—	16,7	16,7
Confidence interval width (95%), mm	25,8	36,4	52,1

Analysis of table 2 showed that reducing the number of generatrices from 24 to 12 yields a relative error of the mean of 55.9%, while reducing to 6 yields 188.8%, and the SEM increases by a factor of 2. Such high errors make analysis by circles unsuitable for substantiating the standard of 36 points. Moreover, for RVS-20000 the number of points per generatrix (9) is less than the standard, and horizontal sections do not account for the distribution of deviations over height [4]. Therefore, we proceed

to the analysis of data along generatrices of RVS-50000, where the number of points per generatrix (36) corresponds to the standard.

Unlike RVS-20000, where the analysis was performed by circles (horizontal sections), for RVS-50000 the study is conducted along generatrices (vertical lines). This is because the standard of 36 points regulates specifically the number of control points on a single generatrix over the tank height. RVS-50000 has 12 generatrices, each containing 36 control points with a height spacing of 0.5 m. The total number of points is 432.

To visualize the tank geometry in space, a three-dimensional model was constructed based on the full data set of real tank No. 1 (figure 4).

```
# Tank parameters
R = 30.35  # radius, m
H = 18.0   # height, m
n_generatrices = 12
n_points_height = 36

# Generatrix angles
angles = np.linspace(0, 2 * np.pi, n_generatrices, endpoint=False)

# Height elevations (from top to bottom)
heights = np.linspace(H, 0, n_points_height)

# Calculation of coordinates taking into account deviations_m (in meters)
theta, z = np.meshgrid(angles, heights)
x = (R + deviations_m) * np.cos(theta)
y = (R + deviations_m) * np.sin(theta)
```

Figure 4 – Calculation of coordinates for three-dimensional visualization of RVS-50000 (Python)

The color scale represents deviation from the vertical, in mm. Blue corresponds to concavity (negative deviations), red corresponds to convexity (positive deviations) (figure 5).

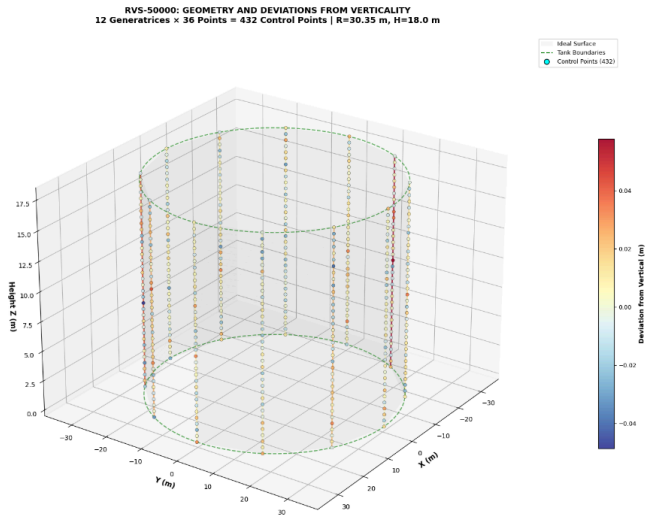


Figure 5 — Distribution of deviations on 12 generatrices. Three-dimensional visualization of the geometry of RVS-50000

Figure 6 presents the distribution of deviations from the vertical along all 12 generatrices of RVS-50000. The abscissa axis shows the generatrix number (from 1 to 12), and the ordinate axis shows the height from the bottom (0–18 m) [5].

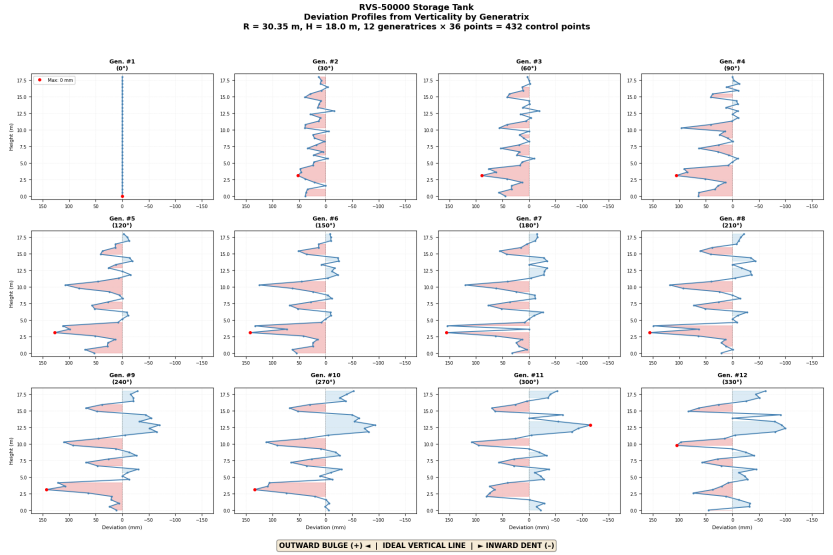


Figure 6 — Distribution of deviations along 12 generatrices of RVS-50000

To investigate the possibility of reconstructing omitted values by spline interpolation, real data from geometric inspection of RVS-50000 No. 2 were used. Generatrix No. 1 was selected as the test fragment. Then the original array was decimated with factors of 2, 3, and 4, i.e., from 36 points every second point was retained (18 points), every third point (12 points), and every fourth point (9 points), respectively. For each decimation variant, a cubic Hermite spline (PCHIP) was constructed from the retained points, and the values at the omitted nodes were reconstructed. Figure 7 presents the reconstruction results for generatrix No. 1 at various decimation factors.

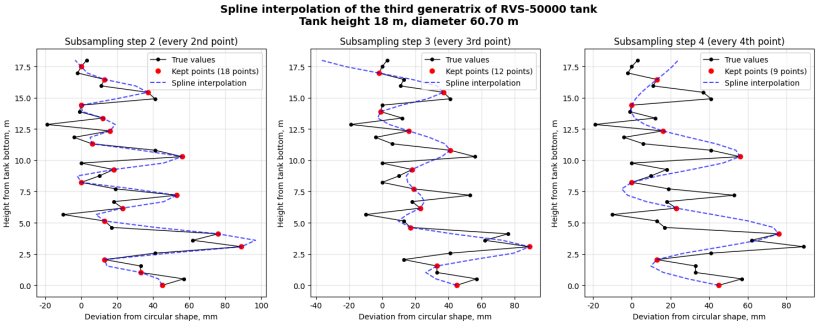


Figure 7 — Spline reconstruction for generatrix No. 1: a – decimation by factor of 2 (36→18), b – by factor of 3 (36→12), c – by factor of 4 (36→9)

Table 3 — Errors of spline reconstruction for generatrix No. 1 of RVS-50000

Decimation	**Reconstructed points**	**Mean |Δ|, mm**	**Max. |Δ|, mm**	**Mean |δ|, %**	**Max |δ|, %**
2× (36→18)	18	0,312	1,124	3,18	68,42
3× (36→12)	24	0,423	1,956	4,35	53,76
4× (36→9)	27	0,687	3,518	6,89	87,34

```python
from scipy.interpolate import PchipInterpolator

def spline_restore_all(data, reduction_factor):
    rows, cols = data.shape
    restored = np.zeros_like(data)
    measured_indices = []
    for col in range(cols):
        indices = list(range(0, rows, reduction_factor))
        if rows - 1 not in indices:
            indices.append(rows - 1)
        indices = sorted(set(indices))
        measured_indices.append(indices)
        interp = PchipInterpolator(indices, data[indices, col])
        restored[:, col] = interp(np.arange(rows))
    return restored, measured_indices

def calc_errors(true_data, restored_data, measured_indices):
    abs_err, rel_err = [], []
    for col in range(cols):
        for row in range(rows):
            if row not in measured_indices[col]:
                delta = restored_data[row, col] - true_data[row, col]
                abs_err.append(abs(delta))
                rel = 100 * abs(delta) / max(abs(true_data[row, col]), 1.0)
                rel_err.append(rel)
    return np.mean(abs_err), np.max(abs_err), np.mean(rel_err), np.max(rel_err)
```

Figure 8 – Fragment of code for general spline reconstruction analysis (Python)

From table 3 it follows that with decimation by a factor of 2, the mean relative error is 3.18%; with decimation by a factor of 3 — 4.35%; by a factor of 4 — 6.89%. The maximum error increases from 68.42% to 87.34%. Thus, increasing the decimation ratio naturally degrades reconstruction accuracy: the fewer points retained, the further we are from the optimal discretization of 36 points. Spline reconstruction is acceptable only for decimation by no more than a factor of 2.

On the same real data from RVS-50000 No. 2, measurements after hydraulic testing were considered, when some geometric parameters were not recorded. The original array contains gaps [6]. The lost values were reconstructed by spline interpolation. Then, as in Part 3, the reconstructed array was decimated with factors of 2, 3, and 4.

```
def restore_sparse(data_with_gaps):
    restored = data_with_gaps.copy()
    for col in range(cols):
        mask_valid = ~np.isnan(col_data)
        if np.sum(mask_valid) >= 2:
            interp = PchipInterpolator(x_valid, y_valid)
            restored[mask_nan, col] = interp(np.where(mask_nan)[0])
    return restored
```

Figure 9 – Function for reconstructing omitted values by PCHIP interpolation in Python

Table 4 — Errors of spline reconstruction after hydraulic testing

Decimation	Reconstructed points	Mean |Δ|, mm	Mean. |δ|, %
2× (36→18)	204	1,9146	29,29
3× (36→12)	276	2,3050	31,31
4× (36→9)	312	2,5368	34,23

As a result of the work, it was proven that analysis by circles (horizontal sections) is inappropriate for substantiating the standard, since reducing the number of generatrices leads to unacceptably high errors. The correct approach is analysis along generatrices (vertical lines), which confirmed that the standard of 36 points per generatrix is optimal. Reducing the number of points degrades reconstruction accuracy: decimation to 18 points (by a factor of 2) yields a mean error of 3.18% and is acceptable only in extreme cases. After hydraulic testing, critical changes occur in the data: the mean reconstruction error reaches 29–34%, and the maximum error exceeds it many times over, which makes spline interpolation inapplicable. It is recommended to retain the existing geometric inspection scheme: twelve generatrices with thirty-six points on each generatrix.

Список литературы

1. Shavaleev D.A., Korobkov G.E., Zaripov R.M., Yanchushka A.P. Ensuring industrial safety during the operation of tanks and tank farms. Materials of the consultative-methodological seminar. – Association "Bashkir Association of Experts"; OAO "Gazprom Neftekhim Salavat"; FSBEI HPE "Ufa State Petroleum Technological University", 2011. – 42-48 pp.

2. Shavaleev D.A., Korobkov G.E., Zaripov R.M., Yanchushka A.P. Ensuring industrial safety during the operation of tanks and tank farms. Materials of the consultative-methodological seminar. – Association "Bashkir Association of Experts";

OAO "Gazprom Neftekhim Salavat"; FSBEI HPE "Ufa State Petroleum Technological University", 2011. – 95-101 pp.

3. GOST R 52910-2008. Vertical cylindrical steel tanks. Methods of geometric inspection. – M.: Standartinform, 2009. – 32 p.

4. Vasilkin, A. A. Influence of geometric imperfections of erection joints on the strength characteristics of vertical steel tanks: abstract of dissertation for the degree of Candidate of Technical Sciences. – M.: MGSU, 2007. – 18 p.

5. Karavaychenko M.G. Strength and survivability of tanks: monograph. – SPb.: Naukoemkie tekhnologii, 2023. – 524 p.

6. Berezin V.L., Shutov V.E. Strength and stability of tanks and pipelines. – M.: Nedra, 1973. – 24 p.

DOI 10.34660/INF.2026.88.14.037

ГЕРМЕТИЗАЦИЯ И ГИДРОИЗОЛЯЦИЯ ШАРНИРНЫХ СОЧЛЕНЕНИЙ СУДОВЫХ РОБОТОТЕХНИЧЕСКИХ СИСТЕМ В УСЛОВИЯХ АГРЕССИВНЫХ СРЕД

Дружинина Н.В.
аспирант
Санкт-Петербургский государственный морской университет,
г. Санкт-Петербург, Россия;
инженер-конструктор 3 кат.
ЗАО «Центральный научно-исследовательский институт судового машиностроения»,
г. Санкт-Петербург, Россия

Статья посвящена проблеме обеспечения эксплуатационной надежности шарнирных сочленений судовых робототехнических систем (СРТС), функционирующих в условиях постоянного воздействия морской воды и высокого гидростатического давления. Проведен детальный сравнительный анализ трех ключевых методов гидроизоляции подвижных узлов: сильфонных защит, сальниковых уплотнений и гидрозатворов. Рассмотрены механизмы их деградации под воздействием агрессивных сред. Сформулированы рекомендации по проектированию комбинированных систем герметизации для глубоководных и надводных робототехнических комплексов.

Ключевые слова: *шарнирные сочленения, сильфонные защиты, сальниковые уплотнения, гидрозатворы, агрессивная среда.*

The paper considers the problem of ensuring the operational reliability of the articulation joints of marine robotic systems operating under conditions of constant exposure to seawater and high hydrostatic pressure. A detailed comparative analysis of three key methods of waterproofing of moving components has been carried out: bellows protection, gland seals, and water seals. The mechanisms of their degradation under the influence of aggressive environmental media have been examined. Recommendations for the design of combined sealing systems deep-sea and surface robotixc complexes have been formulated.

Keywords: *marine robotic systems, waterproofing, manipulators, articulation joints, bellows protection, sealing gland, buffer seals, aggressive environment.*

Введение

Современный этап освоения Мирового океана характеризуется интенсивным внедрением судовых робототехнических систем (СРТС), включая необитаемые подводные аппараты (НПА) и роботизированные комплексы бортового базирования. Одной из наиболее уязвимых зон СРТС являются подвижные шарнирные сочленения манипуляторов, движителей и систем позиционирования.

Эксплуатация в условиях морской среды сопряжена с комплексом деструктивных факторов:

- Высокое гидростатическое давление (возрастающее на 0,1 МПа на каждые 10 метров глубины);
- Химическая агрессивность морской воды (высокое содержание хлоридов, инициирующих питтинговую и щелевую коррозию);
- Наличие абразивных взвесей (песок, ил);
- Биообрастание.

Разработка эффективных систем герметизации динамических узлов напрямую определяет среднюю наработку на отказ и общую живучесть робота.

1. Сильфонные защиты шарнирных узлов

Сильфоны представляют собой гофрированные оболочки, способные совершать значительные линейные, угловые и сдвиговые перемещения под действием силы или внутреннего давления. В СРТС они обеспечивают абсолютную (герметичную) изоляцию внутреннего объема шарнира от внешней среды.

Материалы и стойкость

Для судовых условий применяются два класса сильфонов:

- **Металлические** (нержавеющие сталей типа 03Х17Н14М3, титановые сплавы). Обладают высокой прочностью, но ограничены по углам изгиба.
- **Эластомерные и фторопластовые** (PTFE, полиуретаны, витон). Обеспечивают максимальную амплитуду движений шарнира.

Математическая модель деформации

При угловом отклонении шарнира на угол a максимальное осевое сжатие/растяжение крайних гофр сильфона определяется зависимостью:

$$\Delta L = Dcp \cdot \sin(a/2)$$

где Dcp - средний диаметр сильфона.

Преимущества: Полное отсутствие пар трения, контактирующих с забортной водой; неизменность характеристик при химическом воздействии (для PTFE).

Недостатки: Ограниченный угол поворота шарнира (обычно до $\pm 30^{\circ}$); риск бароразрушения при резком перепаде давлений (требуется система компенсации внутреннего давления).

2. Сальниковые уплотнения в СРТС

Сальниковые узлы относятся к уплотнениям контактного типа. В судовой робототехнике они применяются в узлах с непрерывным вращательным или возвратно-поступательным движением (например, валы гребных винтов манипуляторных движителей).

Особенности работы в агрессивной среде

Классическая пеньковая набивка в СРТС неприменима из-за высокой гигроскопичности и гниения. Современные сальники для морских роботов используют:

- Графитированные фторопластовые нити;
- Углеволоконные структуры;
- Кевларовые (арамидные) угловые оплетки для защиты от абразива.

Расчет силы трения и износа

Герметичность достигается за счет осевой силы затяжки букса P*oc*, создающей радиальное давление P*rad* на вал:

$$Prad = Poc \cdot \mu \cdot e^{-\frac{\mu \cdot f \cdot x}{dbn}}$$

где μ - коэффициент бокового давления, f - коэффициент трения набивки о вал, x - координата по длине сальника.

Преимущества: Возможность работы при неограниченных углах вращения вала, ремонтопригодность.

Недостатки: Наличие постоянного момента трения, вызывающего нагрев и дополнительные энергопотери аккумуляторов СРТС; неизбежный микропропуск среды для смазки зоны контакта.

3. Гидрозатворы и лабиринтные системы

Гидрозатворы (включая жидкостные уплотнения барьерного типа) используют принцип поддержания избыточного давления блокирующей жидкости (масла или синтетической гидрожидкости) внутри узла.

Принцип динамического барьера

Внутренняя полость шарнира заполняется диэлектрическим маслом. Давление внутри полости (P_{vnutr}) поддерживается умышленно выше внешнего гидростатического давления (P_{vnesh}):

$$P_{vnutr} = P_{vnesh} + \Delta P$$

где ΔP = 0,05 ... 0,15 МПа.

При таком соотношении давлений исключается проникновение морской воды внутрь узла. В случае микропротечек масло выталкивается наружу, предотвращая контакт прецизионных подшипников и датчиков с солями.

Преимущества: Практически нулевой износ уплотняемых поверхностей; эффективный отвод тепла от сочленения; защита от экстремальных давлений на больших глубинах.

Недостатки: Необходимость интеграции в состав СРТС активной системы компенсации давления (гидроаккумуляторов, мембранных или поршневых компенсаторов); экологический риск утечки масла в акваторию.

4. Сравнительный анализ методов герметизации

Для наглядного выбора инженерного решения при проектировании СРТС составлена сравнительная матрица параметров уплотнений:

Критерий эффективности	**Сильфонная защита**	**Сальниковое уплотнение**	**Гидрозатвор (активный)**
Абсолютная герметичность	Высшая (100%)	Относительная (микропропуск)	Высокая (за счет вытеснения)
Допускаемый тип движения	Ограниченный поворот/сдвиг	Неограниченное вращение	Неограниченное вращение/поворот
Влияние глубины (давления)	Требует компенсации	Увеличивает силу трения	Идеальная адаптация
Абразивная стойкость	Средняя (риск прокола)	Низкая (требует защиты)	Высокая
Энергопотери (трение)	Минимальные (упругие)	Высокие	Средние/Низкие

Выводы

Для шарниров манипуляторов СРТС с ограниченными углами качания наиболее перспективным решением является применение комбинированных схем: фторопластовый сильфон выступает в качестве внешнего защитного барьера от абразива и биообрастания, а внутренний объем заполняется гидрожидкостью под избыточным давлением (принцип гидрозатвора).

Для высокоскоростных вращательных сочленений (гребные валы) целесообразно использовать каскадные сальниковые уплотнения на основе армированного PTFE с обязательной подачей биоразлагаемой смазки в зону предсальниковой камеры.

Дальнейшее развитие систем гидроизоляции СРТС лежит в плоскости создания «умных» эластомеров со способностью к самозатягиванию микротрещин и магнитореологических гидрозатворов.

Список литературы

1. Вельтищев В. В. Проектирование подводных робототехнических систем: учебно-методическое пособие. — М.: Издательство МГТУ им. Н. Э. Баумана, 2019.

2. Голубев А. И., Кондаков Л. А. Уплотнения и уплотнительная техника: Справочник. — М.: Машиностроение, 1986.

3. Крагельский И. В. Трение и износ. — М.: Машиностроение, 1968.

4. Мамулова Н. С., Сухотин А. М., Сухотина Л. П., Флорианович Г. М., Яковлев А. Д. Все о коррозии. — СПб: «Химиздат», 2000.

5. Проектирование уплотнительных устройств судовых механизмов / Под ред. Антонова А. И. — Л.: Судостроение, 1985.

6. Ястребов В. С. Технические средства освоения Мирового океана. — М.: Машиностроение, 1980.

DOI 10.34660/INF.2026.41.38.045

ОТВЕТСТВЕННОСТЬ МАРКЕТПЛЕЙСОВ ПЕРЕД ПОТРЕБИТЕЛЯМИ: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

Коногорова Альбина Олеговна
Себряковский филиал Волгоградского государственного технического университета, г. Михайловка, Россия
*Научный руководитель - **Балибардина Наталья Геннадиевна***
кандидат юридических наук

Аннотация. *В статье рассматриваются актуальные проблемы ответственности маркетплейсов перед потребителями в условиях активного развития цифровой экономики и электронной торговли. Анализируется правовой статус маркетплейсов и особенности их участия в отношениях между продавцами и потребителями. Особое внимание уделяется проблемам определения надлежащего субъекта ответственности за продажу товаров ненадлежащего качества, предоставление недостоверной информации о товаре, нарушение порядка возврата товаров и денежных средств, а также действия недобросовестных продавцов.*

Ключевые слова: *маркетплейс, цифровая платформа, электронная торговля, потребитель, защита прав потребителей, ответственность маркетплейса, продавец, цифровая экономика, дистанционная торговля, информационный посредник.*

Abstract. *This article examines current issues regarding the liability of online marketplaces toward consumers in the context of the rapid development of the digital economy and e-commerce. It analyzes the legal status of marketplaces and the specific nature of their involvement in the relationship between sellers and consumers. Particular attention is paid to the challenges of identifying the appropriate party to be held liable in cases involving the sale of defective goods, the provision of inaccurate product information, violations of procedures for the return of goods or refunds, and the misconduct of unscrupulous sellers.*

Keywords: *marketplace, digital platform, e-commerce, consumer, consumer protection, marketplace liability, seller, digital economy, remote commerce, information intermediary.*

Развитие цифровой экономики привело к существенному изменению традиционных моделей торговли. Одним из наиболее заметных явлений последних

лет стало широкое распространение маркетплейсов — цифровых платформ, посредством которых потребителям предоставляется возможность приобретать товары, работы и услуги у различных продавцов. Такие платформы существенно упрощают процесс заключения договоров, расширяют ассортимент и обеспечивают доступ к товарам независимо от места нахождения потребителя[1].

Одновременно с этим развитие маркетплейсов породило ряд новых проблем в сфере защиты прав потребителей. Традиционная модель розничной торговли предполагает наличие относительно понятной связи между продавцом и покупателем. В случае с маркетплейсом в отношениях могут одновременно участвовать потребитель, продавец, оператор информационной системы, службы доставки, платежные организации и иные лица. В результате возникает вопрос о распределении ответственности между указанными субъектами.

Особую актуальность приобретает определение правового статуса маркетплейса и установление случаев, когда именно оператор цифровой платформы должен нести ответственность перед потребителем за нарушение его прав.

Маркетплейс — это прежде всего онлайн-платформа, которая предлагает товары или услуги от неограниченного числа продавцов (исполнителей). Иногда термин «маркетплейс» используют неправильно, подразумевая под ним интернет-магазин с очень широким ассортиментом продукции. Однако интернет-магазин в первую очередь сам является продавцом. Маркетплейс же представляет собой площадку, где продавцы и исполнители размещают свои товары, иными словами, это своего рода ярмарка. Маркетплейс и интернет-магазин объединяет лишь то, что они являются дистанционными площадками по предложению товаров или услуг[1].

Ключевая правовая проблема заключается в том, что функциональное участие маркетплейса в продаже товара далеко не всегда совпадает с его юридическим статусом. В зависимости от конкретной модели деятельности маркетплейс может выступать исключительно информационным посредником либо принимать на себя значительно более широкий круг обязанностей.

Для потребителя данное разграничение не всегда очевидно. Интерфейс платформы, как правило, формирует единый пользовательский опыт: покупатель выбирает товар, оплачивает его и получает заказ в рамках одной цифровой системы. Информация о конкретном продавце при этом может находиться на втором плане по сравнению с брендом самого маркетплейса[2].

Такая ситуация создает потенциальный дисбаланс интересов. Потребитель фактически взаимодействует с единой цифровой инфраструктурой, но при возникновении недостатка товара или иных нарушений вынужден самостоятельно определять, к какому субъекту предъявлять требования.

Поэтому одним из центральных направлений развития законодательства должно стать повышение прозрачности распределения функций и ответственности между маркетплейсом и продавцом.

Проблема определения надлежащего ответственного лица:

Наиболее распространенной проблемой является установление субъекта, ответственного за нарушение прав потребителя.

Например, при приобретении товара ненадлежащего качества потребителю необходимо определить, кому следует предъявить требование о возврате денежных средств, замене товара или устранении недостатков. Если продавцом является третье лицо, а маркетплейс лишь предоставляет площадку для размещения предложения, ответственность за качество товара в общем случае должна быть связана с продавцом.

Однако ситуация усложняется, если маркетплейс принимает оплату, организует доставку, хранит товар на собственных складах, осуществляет обработку претензий и самостоятельно определяет существенные условия взаимодействия сторон. В таком случае возникает вопрос о том, насколько обоснованно полностью исключать ответственность оператора платформы перед потребителем.

Следовательно, формальное указание на то, что маркетплейс не является продавцом, не всегда отражает фактическую роль платформы в отношениях с потребителем[4].

Недостоверная информация о товаре:

Информация о продукции на маркетплейсе зачастую формируется непосредственно продавцом. При этом именно маркетплейс определяет структуру карточки товара, технические требования к размещаемой информации и алгоритмы ее отображения пользователю.

Возникает вопрос о степени ответственности платформы за содержание размещаемых сведений. С одной стороны, возложение на маркетплейс обязанности предварительно проверять каждую характеристику каждого товара может оказаться чрезмерно обременительным. С другой стороны, полное освобождение платформы от каких-либо обязанностей по контролю информации способно создавать условия для распространения недостоверных сведений[3].

Перспективным представляется использование риск-ориентированного подхода. Чем активнее маркетплейс участвует в формировании информации о товаре, продвижении конкретного предложения и организации продажи, тем выше должна быть степень его ответственности за достоверность соответствующих сведений.

Возврат товаров и денежных средств:

В традиционной розничной торговле потребителю, как правило, понятно, куда обращаться с требованием о возврате. На маркетплейсе процедура может включать несколько участников, а правила возврата могут зависеть от конкретного продавца и категории товара.

Сложности возникают и при возврате денежных средств. Потребитель может фактически оплатить заказ через маркетплейс и ожидать возврата денежных средств от той же платформы. Однако с юридической точки зрения денежные

средства могут перечисляться продавцу, вследствие чего маркетплейс рассматривает себя исключительно как посредника[4].

Недобросовестные продавцы:

Маркетплейс обладает значительным объемом информации о деятельности продавцов: количеством заказов, возвратов, жалоб, отзывов и иных обращений. Благодаря этому оператор платформы потенциально способен выявлять систематические нарушения.

В связи с этим возникает вопрос о необходимости установления определенных обязанностей маркетплейсов по реагированию на нарушения со стороны продавцов.

При этом важно избежать ситуации, при которой на платформу возлагается абсолютная ответственность за действия каждого зарегистрированного продавца. Более обоснованным представляется подход, при котором ответственность платформы зависит от того, знала ли она о нарушении или при наличии соответствующей информации должна была принять разумные меры для его прекращения[4].

Баланс интересов потребителей, продавцов и маркетплейсов:

Совершенствование регулирования деятельности маркетплейсов требует учета интересов всех участников цифрового рынка[2].

Потребитель заинтересован в простом и понятном механизме защиты своих прав. Для него важно иметь возможность быстро получить информацию о продавце, предъявить претензию и вернуть денежные средства. Продавцы, в свою очередь, заинтересованы в предсказуемости требований маркетплейсов и недопущении необоснованного списания денежных средств или блокировки учетных записей.

Таким образом, законодательное регулирование должно обеспечивать баланс между защитой потребителей и сохранением условий для развития цифровой торговли.

Одним из возможных механизмов является дифференциация ответственности в зависимости от фактической роли маркетплейса. Если платформа лишь предоставляет техническую возможность для размещения предложения, объем ее ответственности может быть ограниченным. Если же она фактически участвует в продаже товара, контролирует его движение, принимает платежи и организует исполнение заказа, основания для расширения ее ответственности становятся более существенными[2].

Перспективы совершенствования законодательства:

Во-первых, необходимо обеспечить более четкое законодательное определение статуса маркетплейсов и их основных обязанностей перед потребителями. Использование единого термина «маркетплейс» без учета разнообразия существующих бизнес-моделей может привести к чрезмерно широкому либо, напротив, недостаточному регулированию.

Во-вторых, целесообразно повысить требования к раскрытию информации о продавце. Потребитель должен иметь возможность до момента покупки легко установить, кто является непосредственным продавцом товара, где он зарегистрирован, каким образом с ним можно связаться и кому следует направлять претензии.

В-третьих, необходимо совершенствовать механизмы рассмотрения потребительских обращений. Маркетплейс может выступать единой точкой взаимодействия потребителя с продавцом, даже если юридически не является непосредственным продавцом. Это позволило бы значительно сократить количество ситуаций, когда потребитель вынужден самостоятельно искать и устанавливать контакт с третьим лицом[3].

Возможные направления развития правоприменительной практики:

Важную роль в дальнейшем формировании правового режима маркетплейсов будет играть судебная практика[3].

При разрешении споров может иметь значение не только формальное содержание договора между маркетплейсом и продавцом, но и фактическая модель взаимодействия с потребителем. В частности, могут учитываться способ оформления заказа, порядок приема платежа, участие платформы в доставке, хранении и возврате товара, содержание информации, предоставляемой потребителю, а также степень контроля маркетплейса над действиями продавца.

Одновременно необходимо обеспечить единообразие правоприменения. Для участников цифрового рынка принципиально важно понимать, какие именно действия маркетплейса могут повлечь возникновение ответственности перед потребителем[5].

Распространение маркетплейсов является одним из ключевых факторов трансформации современной розничной торговли. Цифровые платформы предоставляют потребителям значительные преимущества, однако одновременно создают новые правовые риски, связанные с распределением ответственности между платформой и продавцом.

Основная проблема заключается в несоответствии между экономической и юридической ролью маркетплейса. Для потребителя платформа часто выступает единым субъектом взаимодействия, тогда как с юридической точки зрения значительная часть обязанностей может быть возложена на самостоятельного продавца[4].

Особое значение имеет повышение прозрачности отношений с потребителем. Информация о продавце, порядке предъявления требований, возврате товара и денежных средств должна быть доступной и понятной. Одновременно необходимо формировать эффективные механизмы реагирования на систематические нарушения со стороны продавцов.

Таким образом, дальнейшее развитие правового регулирования ответственности маркетплейсов должно быть направлено на обеспечение эффективной

защиты потребителей при сохранении условий для развития цифровой экономики и добросовестной конкуренции. Оптимальная модель регулирования должна учитывать фактическую роль каждого участника цифровой торговли и обеспечивать понятное для потребителя распределение прав, обязанностей и ответственности[2].

Список литературы

1. Аксенов В. А., Одинцов С. В. Правовая природа маркетплейсов – Вестник Балтийского федерального университета им. И. Канта. URL: https://journals.kantiana.ru/vestnik/society/11196/65993/ (дата обращения: 01.10.2026).

2. Андреева Л. В. Согласование интересов участников товарных цифровых платформ (маркетплейсов): правовые проблемы – Вестник Университета имени О. Е. Кутафина (МГЮА). 2024. № 8 (120). С. 149–157.Столяренко Л.Д. Основы психологии: учебное пособие. – 11-е изд. – Ростов: Феникс, 2004. – 672 с.

3. Лепёшин Д. А., Петкилёв П. И. Ответственность маркетплейса — информационного посредника – Цивилист. 2025. № 4. С. 14–20. URL: https://civilistmag.ru/magazines/4_2025 (дата обращения: 01.10.2026).

4. Стурова Д. Ю., Кузнецов П. Н. Исследование соблюдения законодательства о защите прав потребителей на популярных маркетплейсах Ozon и Wildberries – Наука и Образование. 2025. Т. 8. № 2. URL: https://www.opusmgau.ru/index.php/see/article/view/7830 (дата обращения: 01.10.2026).

5. Козинец Н. В. Правовое регулирование деятельности маркетплейсов: актуальные проблемы – Закон и власть. 2023. № 1. С. 29–31.

DOI 10.34660/INF.2026.54.86.048

РОССИЙСКО-КИТАЙСКОЕ ЭКОНОМИЧЕСКОЕ СОТРУДНИЧЕСТВО В АРКТИКЕ: ОЦЕНКИ ОБОРОННЫХ И РАЗВЕДЫВАТЕЛЬНЫХ ВЕДОМСТВ США С ТОЧКИ ЗРЕНИЯ БЕЗОПАСНОСТИ (2024–2026 ГГ.)

Чэнь Шэнь
аспирант
Российский университет дружбы народов им. Патриса Лумумбы, Москва, Российская Федерация

Аннотация. *На основе пяти открытых документов оборонных и разведывательных ведомств США за 2024–2026 гг. в статье анализируется, как американская сторона оценивает российско-китайское экономическое сотрудничество в Арктике с точки зрения безопасности. В этих документах сотрудничество рассматривается в контексте регионального соперничества как фактор, поддерживающий освоение Арктики Россией и расширяющий каналы участия Китая в делах региона. Внимание уделяется также тому, как экономическая зависимость России от Китая может повлиять на её решения о сотрудничестве.*

Ключевые слова: *российско-китайское сотрудничество в Арктике; арктическая политика США; оценки в сфере безопасности; Северный морской путь.*

После начала российско-украинского конфликта ухудшение отношений России с западными странами и санкционное давление изменили условия её сотрудничества с зарубежными партнёрами в освоении Арктики. При этом сотрудничество с Россией затрагивает и отношения Китая с другими арктическими государствами [1, с. 365]. Материалом исследования служат «Арктическая стратегия» Министерства обороны США 2024 г., «Ежегодные оценки угроз» разведывательного сообщества США 2024–2026 гг. и «Оценка глобальных угроз» Разведывательного управления Министерства обороны США (РУМО) 2025 г. В статье анализируется логика, в соответствии с которой американская сторона связывает российско-китайское экономическое сотрудничество в Арктике с вопросами безопасности.

Вклад сотрудничества в освоение российской Арктики

В «Арктической стратегии» 2024 г. отмечается, что на Арктику приходится более 80 % добычи природного газа и почти 20 % добычи нефти в России. Согласно документу, изоляция, в которой Россия оказалась вследствие «войны против Украины», усилила её зависимость от китайского финансирования и рынка сбыта. Речь идёт о финансировании объектов для экспорта энергоносителей, инвестициях в добычу ресурсов и закупках ресурсов Китаем [2, с. 4–5]. Как указывается в «Ежегодной оценке угроз» 2025 г., война истощила финансовые и военные ресурсы, необходимые России для достижения целей в Арктике, и подтолкнула её к более тесному сотрудничеству с Китаем в регионе [3, с. 17–18].

В докладе 2026 г. Россия названа главным вызовом для США в Арктике. Её усилия по расширению присутствия в регионе за счёт добычи ресурсов, морской торговли и военной деятельности рассматриваются в контексте более широкого соперничества держав, связанного с соотношением сил в мире [4, с. 11]. Сопоставление этих оценок показывает, что поддержку, которую китайское финансирование и рынок сбыта оказывают освоению российской Арктики и торговле, американская сторона оценивает с учётом стремления России сохранять присутствие в регионе и участвовать в стратегическом соперничестве.

Роль сотрудничества в расширении участия Китая в Арктике

В «Ежегодной оценке угроз» 2024 г. более выгодные цены на энергоносители и расширение доступа к Арктике названы среди преимуществ, которые Китай получает, поддерживая отношения с Россией [5, с. 8]. В докладе 2025 г. арктические ресурсы и более прямые морские маршруты связываются с экономическим ростом и энергетической безопасностью Китая [3, с. 12]. По оценке авторов доклада, Китай использует сотрудничество с Россией, чтобы укрепить присутствие в Арктике и придать легитимность своему влиянию в регионе. В том же докладе отмечается, что стороны договорились расширить сотрудничество по развитию Северного морского пути, учитывая его экономический потенциал и стремясь создать альтернативу маршрутам, на которых доминируют западные страны [3, с. 30]. Согласно докладу 2026 г., деятельность Китая в Арктике остаётся сравнительно ограниченной: он продвигает свои интересы преимущественно через отношения с Россией и стремится расширить присутствие за счёт научных исследований, инвестиций и коммерческой деятельности на Северном морском пути [4, с. 11–12].

Экономическая зависимость и ожидаемые уступки в вопросах безопасности

Авторы доклада РУМО США 2025 г. отмечают, что Россия по-прежнему настороженно относится к арктическим амбициям Китая, но западные санкции усилили её зависимость от китайского финансирования и технологий. На этом основании они прогнозируют, что в течение года после публикации доклада Россия, скорее всего, неохотно пойдёт на уступки Китаю в вопросах безопасности в Арктике [6, с. 35]. В этом прогнозе экономическая зависимость рассматривается

как фактор, способный повлиять на решения России о сотрудничестве в сфере безопасности. Тем самым американская оценка безопасности в Арктике учитывает возможные изменения условий участия Китая в регионе.

Таким образом, в рассмотренных официальных документах США российско-китайское экономическое сотрудничество в Арктике оценивается исходя из логики регионального соперничества, а его значение для безопасности раскрывается в трёх аспектах: сотрудничество поддерживает освоение Арктики Россией; оно расширяет каналы участия Китая в делах региона; экономическая зависимость России от Китая может повлиять на её решения о сотрудничестве и тем самым изменить условия расширения китайского участия в Арктике. Признавая, что собственная деятельность Китая в Арктике сравнительно ограничена, авторы этих оценок уделяют основное внимание сочетанию российских возможностей и китайских ресурсов. Финансовым, рыночным и технологическим связям при этом придаётся значение факторов, которые могут повлиять на способность действовать в регионе и на решения о сотрудничестве в сфере безопасности.

Список литературы

1. Галимуллин Э. З., Кашин В. Б., Крамаренко А. С. Снежный Дракон: политика Китая в Арктике перед лицом новых вызовов // Вестник Российского университета дружбы народов. Серия: Международные отношения. — 2024. — Т. 24, № 3. — С. 358–370. — DOI: 10.22363/2313-0660-2024-24-3-358-370.

2. U.S. Department of Defense. 2024 Arctic Strategy [Электронный ресурс]. — 2024. — URL: https://media.defense.gov/2024/Jul/22/2003507411/-1/-1/0/DOD-ARCTIC-STRATEGY-2024.PDF (дата обращения: 08.10.2026).

3. Office of the Director of National Intelligence. Annual Threat Assessment of the U.S. Intelligence Community [Электронный ресурс]. — 2025. — URL: https://www.dni.gov/files/ODNI/documents/assessments/ATA-2025-Unclassified-Report.pdf (дата обращения: 08.10.2026).

4. Office of the Director of National Intelligence. Annual Threat Assessment of the U.S. Intelligence Community [Электронный ресурс]. — 2026. — URL: https://www.dni.gov/files/ODNI/documents/assessments/ATA-2026-Unclassified-Report.pdf (дата обращения: 08.10.2026).

5. Office of the Director of National Intelligence. Annual Threat Assessment of the U.S. Intelligence Community [Электронный ресурс]. — 2024. — URL: https://www.dni.gov/files/ODNI/documents/assessments/ATA-2024-Unclassified-Report.pdf (дата обращения: 08.10.2026).

6. Defense Intelligence Agency. 2025 Worldwide Threat Assessment : Statement for the Record [Электронный ресурс]. — 2025. — URL: https://armedservices.house.gov/uploadedfiles/2025_dia_statement_for_the_record.pdf (дата обращения: 08.10.2026).

Научное издание

Научный диалог: теория и практика

Том 2 из 2

Материалы международного научного форума
(г. Москва, Форум 22 октября 2026 г.)

Научный редактор Д.Р. Хисматуллин
Корректор А.И. Николаева

Подписано в печать 22.10.2026 г. Формат 60x84/16.
Усл. печ. л. 7,29. Тираж 500 экз.

Отпечатано в редакционно-издательском центре
издательства Инфинити

