

Риск-ориентированная оценка экономической эффективности LLM в сервисном секторе Республики Беларусь



Екатерина Пискун,
доцент кафедры проектирования информационно-компьютерных систем Белорусского государственного университета информатики и радиоэлектроники, кандидат экономических наук;
espiskun@gmail.com



Егор Крячев,
магистрант кафедры проектирования информационно-компьютерных систем БГУИР



Акбаржон Азизов,
магистрант кафедры проектирования информационно-компьютерных систем БГУИР

УДК 338.24:004.8

Аннотация. Представлен анализ трансформации национального бизнеса под воздействием больших языковых моделей (LLM), являющихся внутренним катализатором интенсификации труда в сервисном секторе экономики Республики Беларусь. Основываясь на синтезе системного анализа и экономико-математического моделирования, адаптирована концепция производственной функции с постоянной эластичностью замещения для оценки потенциала вытеснения рутинных операций интеллектуальными агентами. Эмпирически подтверждена высокая детерминация между интеграцией LLM-решений и операционной продуктивностью. Установлено, что в условиях белорусского рынка внедрение больших языковых моделей позволяет минимизировать время при обработке клиентских данных с нескольких часов до полутора минут. Количественная оценка демонстрирует возможность замещения трудовых ресурсов, сопоставимую по масштабам с деятельностью сотен штатных специалистов. Сформулированы оригинальные модели расчета рисков от внедрения LLM, учитывающие не только прямые выгоды, но и вероятностные распределения ущерба от специфических киберрисков. Полученные выводы могут быть использованы топ-менеджментом при переходе к AI-ориентированным моделям управления в условиях лимитированного доступа к аппаратным мощностям и жестких требований национального регулятора.

Ключевые слова: генеративный интеллект, экономический эффект, производительность, Республика Беларусь, технологический суверенитет, эластичность замещения, риск-менеджмент.

Для цитирования: Пискун Е., Крячев Е., Азизов А. Риск-ориентированная оценка экономической эффективности LLM в сервисном секторе Республики Беларусь // Наука и инновации. 2026. №3. С. 36–44.

<https://doi.org/10.29235/1818-9857-2026-03-36-44>

Цифровизация мировой хозяйственной системы посредством внедрения больших языковых моделей позволила выйти за рамки сугубо технологических инноваций, превратившись в стабильный фактор повышения операционной эффективности бизнеса. Интенсивный переход к решениям на базе LLM, способных к автономному синтезу семантически связанного контента, характеризуется замещением рутинных когнитивных функций цифровыми решениями, что в свою очередь обеспечивает изменение себестоимости сервиса и приводит к трансформации взаимодействия с потребителями.

Динамика инвестиций в ИИ-ориентированную инфраструктуру подтверждает необратимый характер данных изменений. Согласно прогнозам американской исследовательской и консалтинговой компании IDC, к 2028 г. мировые расходы в данном сегменте могут превысить 631 млрд долл. [1]. При этом стратегический потенциал технологии оценивается экспертами Bloomberg Intelligence в масштабе 1,3 трлн долл. к 2032 г. [2].

С позиции экономики наибольший интерес представляют оценки McKinsey, фиксирующие ежегодный вклад генеративного ИИ в глобальный ВВП в диапазоне 2,6–4,4 трлн долл. [3]. Высокая скорость адаптации технологии бизнесом подтверждается тем, что к 2025 г. уже более 70% организаций интегрировали решения на базе искусственного интеллекта в свои ключевые бизнес-функции [4].

Цель работы состоит в систематизации и сравнительном анализе экономических результатов от внедрения больших языковых моделей на примере международного опыта и практики ведущих предприятий Республики Беларусь. В рамках статьи решаются задачи по выявлению ключевых метрик эффективности операционных процессов, оценке потенциала высвобождения рабочего времени персонала и анализу системных рисков, сопровождающих проникновение данных инноваций в национальное экономическое пространство.

Методика и инструментарий исследования

В основу данной работы заложен принцип количественно-качественного анализа, направленный на подтверждение гипотезы о влиянии внедрения LLM на экономическую эффективность компаний. При этом технология больших языковых моделей рассматривается не как изолированное IT-решение, а как экономический фактор, способствующий сдвигу границы производственных возможностей в организациях сервисного сектора. Под таким понимается совокупность отраслей, нацеленных на удовлетворение потребностей юридических и физических лиц посредством предоставления интеллектуальных, финансовых и информационных услуг (банковский сектор, розничная торговля, телекоммуникации).

Методический аппарат работы включает следующие этапы:

- **анализ, базирующийся на сопоставлении международной практики (бенчмарков) и локальных кейсов Республики Беларусь.** В качестве эталонных значений используются данные о росте производительности труда,

зафиксированные Национальным бюро экономических исследований (NBER), что позволяет оценить «теоретический потолок» эффективности для отечественных условий [5];

- **изучение выбранных объектов (ОАО «АСБ Беларусбанк» [6], ООО «МТС» [7] и др.),** обусловленное их высокой сервисной активностью. Методика анализа этих кейсов предполагает декомпозицию их бизнес-процессов на автоматизируемые и неавтоматизируемые составляющие с последующей оценкой уровня замещения человеческого труда;
- **экономико-статистическая экстраполяция,** основанная на оценке через систему метрик потенциального вклада внедрения LLM в организации Республики Беларусь, позволяющая спрогнозировать эффективность всей отрасли:
 - коэффициент автоматизации: доля запросов, обработанных без участия человека;
 - эффект высвобождения времени: сокращение длительности операционного цикла;

- эквивалент полной занятости: расчетный объем человеческих ресурсов, замещаемый внедрением одной единицы LLM-решения;
- **сценарное моделирование рисков, где каждый дестабилизирующий фактор** (правовой, технологический, кадровый и т.д.) ранжирован по степени влияния развитие отрасли.

Принцип работы и анализ международного опыта внедрения LLM в деятельность организаций

Типовой цикл работы LLM-ассистента в компании представлен на *рис. 1*.

Первичная стадия «*Send Query*» связана с получением входящего обращения. В классической схеме это момент активации дорогих переменных издержек вследствие дорогостоящего рабочего времени персонала, так как любая информация требует ручной классификации и анализа.

Этап «*RAG System searches a database*» исключает необходимость самостоятельного изучения инструкций сотрудником. Технология автоматического поиска мгновенно извлекает нужные сведения из архивов, что приводит к замещению трудоемких операций капиталоемкими, а также снижается риск использования неактуальной информации (минимизация убытков от асимметрии данных) и кардинально сокращается период поиска, преобразуя «время ожидания» во «время полезного действия».

На этапе обработки «*LLM Brain icon processes data*» нейросеть синтезирует итоговое решение. Здесь происходит ключевое экономическое замещение: сложная умственная деятельность человека подменяется

машинными вычислениями. При этом затраты на подготовку каждого последующего документа становятся ничтожно малыми.

Контрольный этап «*Human-in-the-Loop shield icon checks quality*» предусмотрен для наиболее ответственных направлений. Здесь специалист или защитные алгоритмы проводят верификацию результата, что позволяет избежать имиджевых потерь и ошибок.

Финальная стадия «*Final output document delivered to user*» завершает процесс предоставлением готового документа адресату. Автоматическая фиксация результата в базе данных ускоряет сервис, что напрямую влияет на удовлетворенность потребителя.

В настоящее время LLM-системы внедряются практически во все сферы деятельности человека.

Шведская финтех-компания Klarna Bank AB, которая предоставляет финансовые услуги в сети Интернет, активно использует AI-ассистента, который за первый месяц провел 2,3 млн диалогов, обработав около двух третей обращений в чате; среднее время решения сократилось с 11 до 2 мин., количество повторных обращений снизились на 25%, а совокупный эффект компания оценила как эквивалент работы 700 сотрудников поддержки [8]. Также отмечалось ожидание улучшения финансового результата (рост прибыли) на 40 млн долл. в 2024 г. за счет масштабирования ассистента и снижения затрат на обслуживание [8].

Американская транспортная компания C.H. Robinson до внедрения генеративного ИИ занималась обработкой входящих электронных версий тендерной заявки до 4 часов, после внедрения – около 90 сек. [9]. Компания сообщает о масштабе автоматизации: до 5500 заказов на перевозку в день обраба-



Рис. 1. Типовой цикл работы LLM-ассистента в организации

тываются за 90 сек.; около 2600 ценовых предложений в день формируются за 32 сек.; около 3 тыс. заявок на согласование слотов в день обрабатываются в пределах 60 сек. по более чем 26 тыс. локаций [9].

По сообщению Reuters со ссылкой на Bloomberg News, Microsoft оценила экономию от применения ИИ в своих контакт-центрах более чем в 500 млн долл. в 2024 г.; также сообщалось, что ИИ генерирует 35% кода в новых продуктах компании [10].

Один из крупнейших розничных и коммерческих банков Великобритании NatWest в пресс-релизе сообщил, что функциональность GenAI в цифровом ассистенте Soqa+ дала улучшение удовлетворенности клиентов на 150% и одновременно снизила долю случаев, когда требуется вмешательство сотрудника [11].

Британская телекоммуникационная компания BT Group отмечает, что в кейсе внедрения AI-ассистента Aimee приводятся следующие метрики: до 60 тыс. диалогов в неделю проходит через чат-бот, в то время как двумя годами ранее эта цифра была вдвое меньше; по клиентским сценариям показатели успешной автоматизации приближаются к 50%; использование канала растет на 51% от года к году [12].

Крупнейший австралийский Commonwealth Bank сообщил, что ИИ-ориентированная коммуникация в мобильном приложении помогла сократить время ожидания в колл-центре на 40% за последний финансовый год [13].

Австралийская телекоммуникационная компания Telstra отметила, что по итогам тестовых кейсов в 2023 г. 90% персонала, использовавшего инструмент краткого резюмирования истории клиента, сэкономили время; в результате повторных контактов стало 20% меньше. В другом пилоте более 80% участников отметили положительное влияние генеративного ИИ на взаимодействие с пользователями; первоначально решение тестировалось на 200 фронтлайн-сотрудниках [14].

Для понимания механизмов трансформации сервисных отраслей недостаточно констатации факта автоматизации. Необходимо формализовать процесс замещения живого труда технологическим фактором. В связи с этим предлагается задействовать модифицированную производственную функцию с постоянной эластичностью замещения, где выпуск сервисного продукта (Q) определяется как:

$$Q = A \times \left[\alpha \times (L_r + \beta \times LLM)^\rho + (1 - \alpha) \times L_c^\rho \right]^{\frac{1}{\rho}},$$

где A – коэффициент, отражающий уровень технологического прогресса и эффективности организации процессов в компании, который не зависит напрямую от объема затраченных ресурсов (труда или ИИ);

α – коэффициент, который показывает долю задач в отрасли, поддающихся алгоритмизации (в банках α выше, чем в творческих индустриях), безразмерный;

LLM – объем инвестиций в вычислительные мощности или количество активных ИИ-агентов, денежная единица;

L_r – затраты рутинного труда (фронт-офис, первая линия поддержки), денежная единица к единице времени;

L_c – когнитивный (не рутинный) труд экспертов, денежная единица к единице времени;

β – коэффициент технологического превосходства LLM в обработке структурированных данных, безразмерный;

ρ – параметр замещения, определяющий эластичность замещения (σ) персонала нейросетью, безразмерный.

Предложенная математическая модель демонстрирует, что высокая эластичность смены рутинных операций подтверждает гипотезу об эндогенном характере влияния ИИ на рост производительности. Математически случай совершенного замещения факторов достигается при $\rho = 1$ (модель становится линейной), что соответствует полной автоматизации типовых операций. При $\rho \rightarrow 0$ модель стремится к функции Кобба-Дугласа, а при $\rho \rightarrow -\infty$ факторы становятся взаимодополняющими. Анализ параметра α показывает, что сервисный сектор (финансы, ритейл) обладает высокой чувствительностью к внедрению LLM в силу доминирования рутинных когнитивных операций. При этом чем выше параметр A за счет совершенствования качества менеджмента, скорости интеграции LLM в существующие CRM и ERP-системы и общей цифровой зрелости отрасли, тем быстрее каждый вложенный в ИИ рубль даст кратный эффект.

Белорусские кейсы внедрения LLM и близких NLP/диалоговых решений

В основу исследования заложен принцип количественно-качественного анализа, направленный на верификацию гипотезы о влиянии внедрения LLM и диалоговых AI-платформ на трансформацию операционной эффективности организаций. В работе данные технологии рассматриваются не как

изолированное IT-решение, а как фундаментальный экономический фактор, способствующий сдвигу границы производственных возможностей.

Проведенный анализ источников [15–22] позволяет систематизировать экономические эффекты по ключевым отраслевым кластерам.

1. Финансовый сектор: переход к парадигме Intelligent Automation. Банковский сектор демонстрирует высокую эластичность операционных процессов по отношению к внедрению генеративных моделей:

- ЗАО «Альфа-Банк» реализует стратегию, где ИИ (RAG-архитектура) обеспечивает дивергенцию роста бизнеса и операционных затрат. Рост клиентской базы (+27% по сравнению с аналогичным периодом прошлого года) сопровождается удержанием лидерских позиций по NPS (44 процентных пункта), что свидетельствует об эффективности гибридных моделей обслуживания;
- ЗАО «Сбер Банк» подтверждает гипотезу о высокой мобильности технологий: миграция на платформу SaluteRPA с API GigaChat была реализована менее чем за месяц. Ключевой актив – способность алгоритмов к семантической саммаризации и извлечению сущностей из неструктурированных данных;
- ЗАО «МТБанк» фиксирует эффект масштаба через внедрение платформы Digital2Speech. Обработка 455 тыс. звонков с уровнем полной автоматизации порядка 9–10% свидетельствует о высвобождении значительного объема человеко-часов для задач высокой когнитивной сложности.

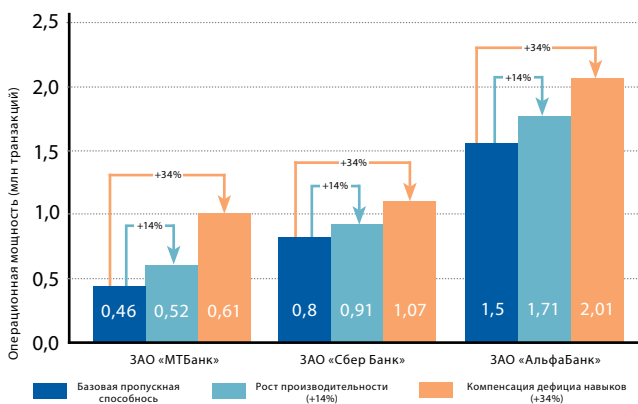


Рис. 2. Прогнозная модель масштабирования пропускной способности сервисных систем на основе коэффициентов эффективности внедрения LLM

Результаты сценарного моделирования (рис. 2) базируются на экстраполяции коэффициентов NBER к верифицированным метрикам белорусских банков. В расчетах использованы подтвержденные данные пилотного внедрения в ЗАО «МТБанк» (455 тыс. транзакций), модель реконструкции потока заявок ОАО «Сбер Банк» (уровень автоматизации >40%) и нормативная оценка пропускной способности ЗАО «Альфа-Банк» на основе базы в 1 млн активных клиентов. Моделирование демонстрирует, что интеграция LLM позволяет компенсировать дефицит квалифицированных кадров (+34%) и обеспечить чистый прирост производительности труда (+14%), трансформируя структуру операционных издержек без экстенсивного расширения штата.

2. Телекоммуникации и экосистемы: коммерциализация LLM. ООО «МТС» запустило интеллектуального помощника RomoAI, ориентированного на автоматизацию клиентской поддержки и создание интеллектуального контента. Сервис уже привлек более 1 тыс. активных пользователей с доминирующей долей текстовых запросов (+80%), а пользователи, в свою очередь, сгенерировали около 1 тыс. изображений через мультимодальные функции. Инструмент доступен массовой аудитории за фиксированную подписку, что демонстрирует коммерческую готовность LLM-решений к повседневному использованию и снижению операционных затрат без расширения контакт-центра.

3. Ритейл и E-commerce: сквозная автоматизация цепочки стоимости. В данном секторе зафиксирован синергетический эффект от внедрения ИИ на этапе операционного цикла.

К примеру, работа омниканальной платформы CraftTalk (интегратор – «НЬЮЛЭНД технолоджи») в ЗАО «Интернет-магазин Евроопт» подтверждает эффективность ИИ в управлении волатильностью спроса. Достигнутый уровень автоматизации составляет 30–85% в зависимости от типа запроса. Экономический эффект выражается в обеспечении круглосуточного цикла обслуживания (24/7) и сглаживании пиковых нагрузок без экстенсивного расширения штата операторов.

4. Сектор услуг, малый и средний бизнес: эластичность конверсии. Эмпирические данные проектов-победителей конкурса HostFly.by и кейсы интегратора JS Global демонстрируют прямую корреляцию между внедрением AI-агентов и коммерческими показателями:

- проект AI Sapiens показывает рост конверсии до 50% при сокращении времени отклика до 15–30 сек.;

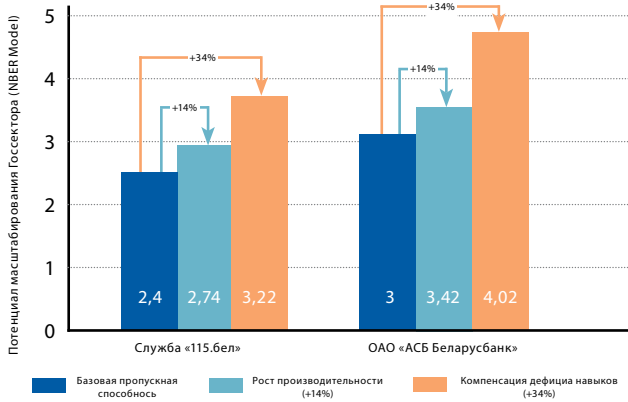


Рис. 3. Прогнозная модель масштабирования пропускной способности сервисных систем ОАО «АСБ Беларусбанк» и «115.бел» на основе коэффициентов эффективности внедрения LLM

- кейс ADVAGA (BBRent) иллюстрирует возможность полного замещения линейного персонала: AI-бот сократил время обработки заявки на 65%, увеличив конверсию на 32%.

5. Государственный сектор: масштабирование пропускной способности. ОАО «АСБ Беларусбанк» и единая диспетчерская служба «115.бел» эксплуатируют NLP-системы для обработки типовых обращений, при этом находятся на стадии активного внедрения LLM-решений. Сценарное моделирование с использованием коэффициентов NBER демонстрирует потенциал экстенсивного роста пропускной способности сервисных систем без увеличения штатной численности персонала, что указывает на расширение производственных возможностей (рис. 3).

Риск-ориентированные модели внедрения LLM и методика расчета совокупного экономического эффекта

Риски внедрения LLM в белорусских организациях во многом совпадают с международной практикой, но имеют и локальные особенности: ограниченный доступ к высокопроизводительным вычислениям и зарубежным облачным сервисам, высокие требования к защите информации (банковская тайна, хранение и обработка персональных данных), а также необходимость надежной работы в критичных для населения сервисах (банки, связь, ЖКХ, транспорт).

Ниже приведены наиболее часто встречающиеся в практике риски и меры их снижения (таблица), сформированные с учетом международного стандарта NIST AI RMF 1.0 (Artificial Intelligence Risk Management Framework) и профилем для генеративного ИИ (NIST AI 600-1) [23].

Для нашей страны наиболее применима модель, где внедрение LLM опирается на уже действующие требования к защите персональных данных, информационной безопасности и отраслевому комплаенсу (банковская тайна, требования к критическим информационным системам), а также дополняется внутренними корпоративными политиками использования ИИ.

На уровне организаций это означает необходимость формализовать перечень допустимых кейсов, правила работы с данными, ответственность за решения, порядок тестирования и мониторинга, процедуры реагирования на инциденты и обучение персонала.

Заключение

Экономический анализ подтверждает, что в таких отраслях, как банковское дело, телекоммуникации и ритейл, большие языковые модели выступают в роли эндогенного катализатора производительности. Применение модифицированной оценки выпуска сервисного продукта позволило математически доказать, что современный интеллектуальный капитал способен эффективно замещать рутинные когнитивные операции, обеспечивая при этом нелинейный рост пропускной способности систем без привлечения дополнительных штатных единиц.

Результаты деятельности ведущих организаций Республики Беларусь демонстрируют высокую эластичность сервисных процессов к внедрению AI-решений. Установлено, что автоматизация типовых сценариев взаимодействия с потребителями достигает 85%, что сопровождается существенным сокращением временных издержек и повышением лояльности аудитории за счет мгновенности реакции.

Показано, что масштабное проникновение LLM в национальное экономическое пространство сопряжено с комплексом инфраструктурных, регуляторных и операционных барьеров. Устойчивое масштабирование технологии в Беларуси требует исключения угроз, связанных с информационной безопасностью и потенциальной деградацией качества ответов.

Таблица. Реестр рисков внедрения больших языковых моделей в организациях Республики Беларусь и меры по их минимизации

Риск	Как проявляется в условиях Республики Беларусь	Меры по минимизации
Инфраструктурно-ресурсные риски		
Доступность вычислительных ресурсов и вендорная зависимость	Дефицит GPU/ускорителей, ограничения на поставки и использование зарубежных облаков, риск резкого роста стоимости API	Гибридная архитектура (on-prem + локальные облака), резервные модели, план миграции, закупки/пулы вычислений, контрактные SLA

Математическая модель совокупных инфраструктурных издержек (C_{infra}) можно представить как функцию от способа развертывания:

$$C_{infra} = (P_{cloud} \times V \times R_{FX}) + (CapEx_{local} + OpEx_{local}),$$

где $P_{cloud} \times V$ – стоимость использования зарубежных API. В условиях Республики Беларусь этот параметр несет риск резкого роста стоимости или блокировки доступа, ден. ед. \times число токенов (шт.);

R_{FX} – валютный коэффициент (риск девальвации национальной валюты, увеличение стоимости импортных облачных услуг);

$CapEx_{local}$ – капитальные затраты на закупку собственных GPU-серверов для локального развертывания, что необходимо для соблюдения требований банковской тайны и защиты данных, ден. ед.;

$OpEx_{local}$ – операционные расходы на обслуживание локальной фермы серверов и оплату электроэнергии для высокопроизводительных вычислений, ден. ед.

Информационно-безопасностные и регуляторные риски		
Утечки данных и нарушение конфиденциальности	Риск передачи клиентских данных во внешние сервисы; ошибки сотрудников при вводе конфиденциальной информации в чат	Классификация данных, запрет/фильтрация чувствительных данных, анонимизация данных, уходящих во внешние сервисы, локальное развертывание для критичных процессов, журналирование и DLP
Косвенная инъекция промптов	Риск, при котором злоумышленник размещает вредоносные инструкции на внешнем сайте, который LLM считывает через Интернет или RAG-систему, что заставляет ассистента выполнять несанкционированные действия (например, пересылать данные)	Пентест и red teaming, контроль входов/выходов, изоляция инструментов, управление привилегиями, рекомендации OWASP для LLM-приложений [24]
Чрезмерная агентность	Риск предоставления ИИ-ассистенту избыточных полномочий (доступ к API, удаление файлов, проведение транзакций) без промежуточного подтверждения человеком	Принцип минимальных привилегий: отдельные сервисные учетные записи, RBAC, строго ограниченные права, разделение сред (dev/test/prod), список разрешенных инструментов и команд (что агенту можно вызывать), запрет произвольных запросов, шаблоны действий

Для оценки рисков, связанных с утечкой данных или инъекцией промптов, целесообразно использовать следующую формулу:

$$ALE = SLE \times ARO,$$

где ALE (Annualized Loss Expectancy) – ожидаемые годовые потери, (ден. ед./год); SLE (Single Loss Expectancy) – ожидаемый ущерб от одного успешного инцидента, ден. ед./инцидент; ARO (Annualized Rate of Occurrence) – ожидаемая годовая частота инцидентов, инцидентов/год.

В свою очередь, данные показатели имеют следующий математический вид:

$$SLE = AV \times EF,$$

где AV (Asset Value) – стоимость (ценность) актива, ден. ед.; EF (Exposure Factor) – доля потерь при реализации события ($0 \leq EF \leq 1$);

$$ARO = \frac{N_{incidents}}{T_{years}},$$

где $N_{incidents}$ – количество случившихся угроз или негативных событий (например, инцидент безопасности, технический сбой или критическая ошибка LLM), шт.;

T_{years} – длительность периода наблюдения, лет.

Регуляторный риск при внедрении LLM в финансовом секторе Республики Беларусь можно описать как сумму вероятностных санкций за несоблюдение нормативов (банковская тайна, защита данных и т.д.):

$$R_{reg} = \sum (P_i \times S_i) + C_{comp},$$

где P_i – вероятность нарушения i -го регуляторного требования (например, закона о персональных данных) ($0 \leq P_i \leq 1$); S_i – финансовая санкция (штраф/потери) при нарушении i , ден. ед.; C_{comp} – затраты на обеспечение соответствия (аудиты, разработка внутренних политик и т.д.), ден. ед. в год.

Утечка конфиденциальной информации может произойти через обучающие выборки для LLM, риск которой можно оценить следующей вероятностной моделью:

$$D_{risk} = V \times T \times I,$$

где V (Vulnerability) – вероятность того, что сотрудник введет чувствительные данные в публичный чат-бот ($0 \leq V \leq 1$);

T (Threat) – вероятность того, что эти данные будут извлечены злоумышленником или использованы моделью для ответа другому пользователю ($0 \leq T \leq 1$);

I (Impact) – стоимость потери данных (например, рыночная стоимость клиентской базы и т.д.), ден. ед.

Математическая эффективность комплекса мер по управлению рисками на основе NIST AI RMF [24] выражается через снижение совокупного риска:

$$Risk_{total} = (Risk_{base} - Risk_{mitigated}) + Cost_{protection},$$

где $Risk_{base}$ – исходный риск (например, при использовании зарубежных облачных LLM без фильтрации), ден. ед. за период;

$Risk_{mitigated}$ – величина риска, которую удалось снизить/нивелировать ($0 \leq Risk_{mitigated} \leq Risk_{base}$), ден. ед.;

$Cost_{protection}$ – прямые расходы на внедрение мер защиты (гибридная архитектура, DLP, «песочницы» и т.п.), ден. ед. за тот же период.

Операционно-технологические риски

Ошибки генерации (hallucinations) и качество ответов	Риск генерации уверенных, но ложных утверждений. В экономическом контексте это может привести к ошибочным финансовым прогнозам или неверному толкованию законодательства	Использование RAG-архитектуры на основе верифицированной базы знаний, шаблоны ответов, обязательная эскалация человеку в риск-сценариях, мониторинг качества, тестовые наборы, агент-ревьюер для проверки ответов
--	--	---

Математическая модель совокупного ущерба от операционно-технологических рисков строится как сумма вероятностных потерь от сбоев в работе модели, некорректной генерации данных и деградации системы со временем:

$$L_{op_tech} = L_{hallucination} + L_{degradation} + L_{recovery}$$

где $L_{hallucination}$ – ущерб от генерации ложных или «уверенных, но неверных» утверждений, ден. ед.;
 $L_{degradation}$ – потери из-за снижения качества ответов по мере изменения внешних данных или «отравления» модели, ден. ед.;
 $L_{recovery}$ – прямые затраты на восстановление работоспособности системы и исправление бизнес-данных, ден. ед.

Рассмотрим входящие в математическую модель параметры более детально.

Ущерб от «галлюцинаций» ($L_{hallucination}$) является наиболее критичным для рассматриваемых экономических сфер. Он рассчитывается как:

$$L_{hallucination} = N \times P_{err} \times (1 - P_{detect}) \times SLE_{avg}$$

где N – общее количество запросов, обрабатываемых LLM, шт.;
 P_{err} – вероятность критической ошибки (галлюцинации), не выявленной системой фильтрации;
 P_{detect} – вероятность того, что оператор или фильтр заметит ошибку до того, как она попадет клиенту;
 SLE_{avg} – средний размер финансового убытка от одной неверной консультации (например, ошибочное одобрение льготного кредита), ден. ед.

Ущерб от деградации модели ($L_{degradation}$) рассчитывается как:

$$L_{degradation} = \int_{t_0}^{t_1} (E_{large} - E_{actual}(t)) dt + C_{unit}$$

где E_{large} – целевой уровень автоматизации; E_{actual} – фактический уровень автоматизации, снижающийся из-за неактуальности базы знаний (RAG);
 C_{unit} – стоимость ручной обработки одного запроса сотрудником (стоимость замещения), ден. ед. на единицу времени.

Для комплексного анализа рисков интеграции LLM предлагается ввести базовую формулу совокупного ущерба (L_{total}), включив в нее компоненты нефинансовых и регуляторных потерь:

$$L_{total} = L_{direct} + L_{rep} + L_{reg}$$

где L_{direct} – прямые финансовые потери (ошибки транзакций, затраты на восстановление инфраструктуры), ден. ед.;
 L_{rep} – репутационный ущерб, ден. ед.; L_{reg} – регуляторные и штрафные санкции, ден. ед.;

Для Республики Беларусь это включает штрафы за нарушение законодательства о персональных данных, например Закона №99-3 [25], и несоблюдение нормативов операционной надежности:

$$L_{reg} = \sum (S_i \times P_i) + C_{lic}$$

где S_i – размер штрафа, ден. ед.; P_i – вероятность наложения санкции, %; C_{lic} – стоимость риска отзыва лицензии или приостановки операций, ден. ед.

Социально-кадровые и репутационные риски

Правовые и комплаенс-риски	Неопределенность авторства и ответственности за контент; соблюдение требований по персональным данным и отраслевым ограничениям	Внутренние политики использования ИИ, маршруты согласования, маркировка AI-контента, хранение доказательности (audit trail), обучение персонала
Кадровые и организационные риски	Недостаток специалистов по ML/LLM и инженеров по интеграции; сопротивление изменениям, «теневое» использование публичных чат-ботов	Центр компетенций, обучение пользователей, регламент внедрения, KPI по качеству и экономии, управление изменениями
Репутационные риски	Негативные ответы, неправильные рекомендации, непредсказуемые реакции ассистента в публичных каналах поддержки	Политики контента, безопасные шаблоны, ограничение тем, обязательная эскалация, контроль тональности, публичные дисклеймеры

Социально-кадровые риски связаны с изменением роли человека в производственном процессе и внутренней организационной культурой.

Математическую модель кадрового риска (R_{HR}) можно представить в следующем виде:

$$R_{HR} = (C_{train} + C_{recruit}) \times (1 + k_{recruit})$$

где C_{train} и $C_{recruit}$ – затраты на обучение текущего персонала и найм новых экспертов, ден. ед.;
 $k_{recruit}$ – коэффициент снижения производительности из-за организационного сопротивления изменениям.

Также стоит учитывать риски, связанные с внешним восприятием бренда при взаимодействии клиентов с LLM-системами (токсичность и некорректность ответов, снижение эмпатии вследствие падения лояльности клиентов и отток клиентов из-за утечек).

Модель ущерба деловой репутации (L_{rep}) можно оценить как:

$$L_{rep} = \Delta NPS \times LTV_{avg} \times N_{cust}$$

где ΔNPS – прогнозируемое снижение индекса лояльности, пункты NPS (NPS в диапазоне -100...+100); LTV_{avg} – средняя пожизненная ценность клиента в сервисном секторе экономики, ден. ед. к клиенту; N_{cust} – количество клиентов, затронутых негативным опытом взаимодействия с LLM-системами, чел.
 В условиях высокой концентрации банковского сектора Республики Беларусь утечка данных через LLM-интерфейс может привести к потере доверия к цифровым каналам обслуживания в целом.

Предложенный алгоритм управления рисками на базе международных стандартов (NIST AI RMF) [24] позволяет нивелировать негативные последствия и обеспечить технологическую независимость бизнеса.

На основании проведенных исследований видно, что в условиях текущей трансформации мировой хозяйственной системы большие языковые модели перестают быть объектом экспериментов, становясь базисным инструментом повышения конкурентоспособности и операционной продуктивности сервисного сектора Республики Беларусь. ■

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. IDC's Worldwide AI and Generative AI Spending – Industry Outlook // <https://www.idc.com/resource-center/blog/ids-worldwide-ai-and-generative-ai-spending-industry-outlook/>.
2. Generative AI to Become a \$1.3 Trillion Market by 2032, Research Finds // <https://www.bloomberg.com/company/press/generative-ai-to-become-a-1-3-trillion-market-by-2032-research-finds/>.
3. The economic potential of generative AI: The next productivity frontier // <https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier>.
4. The state of AI in early 2024 // <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-2024>.
5. Brynjolfsson, E. Generative AI at Work // <https://www.nber.org/papers/w31161>.
6. Почти 3 млн обращений поступило в контакт-центр Беларусбанка за прошлый год // <https://belta.by/special/society/view/pochti-3-mln-obraschenij-postupilo-v-kontakt-tsentr-belarusbanka-za-proshlyj-god-699756-2025/>.
7. МТС: виртуальный консультант обработал 4 млн вопросов по телефону и 600 тыс в чатах // <https://www.mts.by/news/91225/>.
8. Klarna: Scaling customer support with an AI assistant // <https://openai.com/index/klarna/>.
9. C.H. Robinson Automates Freight Lifecycle with Generative AI (Press release) // <https://www.chrobinson.it/en-us/chrglobal/about-us/newsroom/press-releases/2024/generative-ai-for-freight-shipment-lifecycle/>.
10. Microsoft racks up over \$500 million in AI savings while slashing jobs, Bloomberg News reports // <https://www.reuters.com/business/microsoft-racks-up-over-500-million-ai-savings-while-slashing-jobs-bloomberg-2025-07-09/>.
11. NatWest & OpenAI collaborate to accelerate cutting-edge AI transformation // <https://www.natwestgroup.com/news-and-insights/news-room/press-releases/ai-and-data/2025/mar/natwest-open-ai-collaborate-to-accelerate-cutting-edge-ai-transf.html>.
12. BT Group leans on AI to transform customer service experience // <https://newsroom.bt.com/bt-group-leans-on-ai-to-transform-customer-service-experience/>.
13. Customer safety, convenience and recognition boosted by early implementation of GenAI // <https://www.commbank.com.au/articles/newsroom/2024/11/reimagining-banking-nov24.html>.
14. Telstra scales up AI adoption following promising pilots of generative AI solutions // <https://www.telstra.com.au/aboutus/media/media-releases/telstra-scales-up-ai-adoption>.
15. Будущее наступило: IT-гуру Альфы рассказали, что уже делает искусственный интеллект для клиентов // <https://myfin.by/article/banki/stali-izvestny-finansovye-pokazateli-alfa-banka-smotrim-pribyl-i-rentabelnost-37312>.

■ **Summary.** The paper presents an analysis of the transformation of national business under the influence of Large Language Models (LLMs), which act as an internal catalyst for labor intensification in the service sector of the economy of the Republic of Belarus. Based on a synthesis of system analysis and economic-mathematical modeling, the concept of a production function with constant elasticity of substitution has been adapted to assess the potential for displacing routine operations with intelligent agents. A high determination between the integration of LLM solutions and operational productivity is empirically confirmed. It has been established that in the conditions of the Belarusian market, the implementation of large language models allows minimizing the time for processing client data from several hours to one and a half minutes. Quantitative assessment demonstrates the possibility of substituting labor resources comparable in scale to the activities of hundreds of full-time specialists. Original models for calculating risks from LLM implementation have been formulated, taking into account not only direct benefits but also probabilistic distributions of damage from specific cyber risks. The findings can be used by top management during the transition to AI-oriented management models under conditions of limited access to hardware capacities and strict requirements of the national regulator.

■ **Keywords:** Generative Intelligence, economic effect, productivity, Republic of Belarus.

■ <https://doi.org/10.29235/1818-9857-2026-03-36-44>

16. СберБанк (Беларусь) внедрил платформу SaluteRPA для роботизации бизнес-процессов // <https://www.sber-bank.by/new/business-process-10042025>.
17. Проект: МТБанк (Минский транзитный банк) – BSS Digital2Speech // [https://www.tadviser.ru/index.php/Проект:МТБанк_\(Минский_транзитный_банк\)_BSS_Digital2Speech](https://www.tadviser.ru/index.php/Проект:МТБанк_(Минский_транзитный_банк)_BSS_Digital2Speech).
18. МТС представила ИИ-помощника PomogAI // <https://tech.onliner.by/2025/10/21/mts-predstavila-ii-pomoshhnika-pomogai>.
19. Евроторг внедрил интеллектуального чат-бота в свои онлайн-сервисы // <https://belretail.by/news/evrotorg-vnedril-intellektualnogo-chat-bota-v-svoi-onlayn-servisyi>.
20. ИИ меняет белорусский бизнес: результаты конкурса «Искусственный интеллект в деле» // <https://www.hostfly.by/about/news/ii-menyayet-belorusskiy-biznes-rezultaty-konkursa-iskusstvennyy-intellekt-v-dele-ot-hostfly-by/>.
21. Кейс: ИИ-регулирование. Вебинары для сотрудников [Электронный ресурс] // <https://belarus.revera.legal/proekty/451-kejs-ii-regulirovanie-vebinarny-dlya-sotrudnikov/>.
22. Контакт-центр 115 за сутки обрабатывает 6,5 тыс. звонков // <https://www.belta.by/society/view/kontakt-tsentr-115-za-sutki-obrabatyvaet-65-tys-zvonkov-538647-2022/>.
23. Artificial Intelligence Risk Management Framework (AI RMF 1.0) // <https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-ai-rmf-10>.
24. OWASP Top 10 for Large Language Model Applications // <https://owasp.org/www-project-top-10-for-large-language-model-applications/>.
25. О защите персональных данных: Закон Респ. Беларусь, 07.05.2021 г., №99-3 / Эталон-Online // <https://etalonline.by/document/?regnum=h12100099>.

Статья поступила в редакцию
29.12.2025 г.