DOI: http://dx.doi.org/10.21686/2413-2829-2026-4-48-57

ПРИМЕНЕНИЕ ПОДХОДОВ ПОВЫШЕНИЯ КАЧЕСТВА ГЕНЕРИРУЕМЫХ ОТВЕТОВ LLM-СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ПРИ ИСПОЛЬЗОВАНИИ НАКОПЛЕННОЙ ИНФОРМАЦИИ ПРЕДПРИЯТИЯ

А. Л. Куренков
Российский экономический университет имени Г. В. Плеханова, Москва, Россия
Д. А. Куренкова
ООО «Платформеко», Москва, Россия

Развитие больших языковых моделей (LLM) расширило возможности обработки неструктурированных данных в корпоративных средах. Однако фундаментальное ограничение LLM заключается в их зависимости от данных предварительного обучения, что уменьшает их способность надежно работать с информацией предприятий, неопубликованной в других источниках. В статье представлен анализ современных подходов использования корпоративных данных для повышения качества ответов, генерируемых LLM. Рассматриваются системы поисковой дополненной генерации (Retrieval-Augmented Generation – RAG) как в классической векторной реализации, так и в графовом расширении (GraphRAG). Обсуждаются агентные системы и интеграция в их архитектуру RAG- и GraphRAG-подходов, что позволяет строить сложные системы искусственного интеллекта и решения, требующие многоэтапного анализа корпоративных неструктурированных данных. Описывается архитектура классического векторного RAG-подхода, которая состоит из трех основных этапов: получение векторных представлений индексируемых фрагментов документации; поиск фрагментов, релевантных запросу пользователя; формирование ответа по извлеченному контексту, объединенному с исходным запросом. В качестве основного ограничения выделяется отсутствие механизма целостного анализа загруженных проприетарных данных. В свою очередь GraphRAG использует механизм создания графа знаний по контексту, что позволяет работать не с семантически похожими фрагментами данных, а с иерархически кластеризованной информацией посредством накопительного анализа текстовых сводок. Сформулированы недостатки и преимущества использования подходов, определены критерии целесообразности практического применения RAG и GraphRAG, приведены примеры систем, для которых они могут применяться.
Ключевые слова: большие языковые модели, поисковая дополненная генерация, RAG, GraphRAG, цифровая трансформация, интеллектуальный анализ данных.

EMPLOYING APPROACHES OF RAISING QUALITY OF GENERATED ANSWERS OF AI LLM-SYSTEMS WHEN ACCUMULATED ENTERPRISE INFORMATION IS USED

Aleksandr L. Kurenkov
Plekhanov Russian University of Economics, Moscow, Russia
Daria A. Kurenkova
Platformeco LLC, Moscow, Russia

The development of large language models (LLM) has extended opportunities of processing non-structured data in corporate environment. However, LLM fundamental restriction lies in their dependence on data of preliminary learning, which can decrease their ability to work safely with enterprise information non-published in other sources. The article provides analysis of current approach to using corporate data to upgrade quality of answers generated

by LLM. Systems of Retrieval-Augmented Generation (RAG) are studied both in classical vector realization and in GraphRAG. Agent systems and integration of RAG- and GraphRAG approaches in their architecture were discussed, as it can make it possible to build complicated systems of AI and solutions requiring multi-stage analysis of corporate non-structured data. Architecture of classical vector RAG-approach is described, which consists of three key stages: getting vector presentation of indexed fragments of papers; searching for fragments relevant to user requirement; forming the answer on extracted context combined with the initial request. As a key restriction the author showed the absence of mechanism necessary for integral analysis of loaded proprietary data. In its turn GraphRAG uses mechanism of plotting knowledge graph on context, which can help work not with semantically similar data fragments but with hierarchically clustered information by accumulated analysis of text summaries. Shortcomings and benefits of using approaches were formulated, criteria of expediency of RAG and GraphRAG practical application were identified and examples of systems for their use were provided.
Keywords: large language models, search generation, RAG, GraphRAG, digital transformation, intellectual analysis of data.

Введение

Во всем мире и в России в частности наблюдается всплеск спроса на технологии искусственного интеллекта (ИИ). По оценкам проектного офиса «Цифровая экономика», объем российского рынка ИИ в 2023 г. увеличился на 18% по отношению к предыдущему году, 35% российских компаний разных отраслей разработали и внедрили стратегии развития и использования ИИ (в основном в финансовой сфере).

По данным Comindware и PEX, искусственный интеллект, генеративный ИИ, ИИ-агенты будут наиболее востребованными инструментами для инвестиций в 2026 г.

В части широкого понятия искусственного интеллекта понимается «комплекс технических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека»[1].

Искусственный интеллект как концепция появился еще в середине XX в. [11]. Кибернетики прошлого века [1] уже стали описывать перспективное направление вычислительной техники в области моделирования мыслительных процессов естественного интеллекта и их автоматизации. Продолжительное время концепция искусственного интеллекта не выходила за рамки теоретических изысканий, пока два важнейших условия его работы были невыполнимы: доступ к большому объему данных и большим вычислительным мощностям. В настоящее время эти ограничения технологически сняты, что привело к созданию больших ИИ-платформ, моделей, сервисов в США, Китае, России и других странах. Примерами таких систем являются ChatGPT, OpenAI o3, Claude, DeepSeek AI, Jasper, Copy.ai, Cursor, Obviously AI, GigaChat, YandexGPT, Sarvam AI и ряд других.

Стремительное развитие больших языковых моделей (Large Language Models – LLM) [8] трансформировало парадигму взаимодействия человека с вычислительными системами, существенно повысив возможности обработки неструктурированных данных в масштабах, ранее недоступных традиционным методам компьютерной обработки. Вместе с тем фундаментальным ограничением LLM является их замкнутость на обучающих данных. Модель не способна достоверно оперировать информацией, которая отсутствовала в обучающей ее выборке. Это обстоятельство критично для приложений, работающих с проприетарными данными предприятий (документами, регламентами и т. п.), которые не выложены в публичный доступ и не попали в обучающую выборку. Технология RAG [12] была предложена как механизм преодоления этого ограничения. Она предусматривает встраивание данных

[1] URL: https://www.garant.ru/products/ipo/prime/doc/72738946/#1000 (дата обращения: 01.12.2019).

компании в запросы к LLM, обеспечивая большую точность и качество генерируемого ответа. Эта технология используется в двух вариантах: в классической векторной реализации (RAG) и в ее графовом расширении (GraphRAG).

Параллельно с технологиями извлечения и обработки информации активно развивается парадигма агентных систем (AI agents) – автономных программных сущностей, способных планировать порядок действий, вызывать внешние инструменты и итеративно уточнять результаты [14; 18]. Интеграция таких систем с технологиями RAG открывает возможности для построения сложных систем искусственного интеллекта, способных к многоэтапному анализу с использованием корпоративных неструктурированных данных.

Анализ подходов RAG и GraphRAG, используемых для повышения качества генерируемых ответов LLM систем искусственного интеллекта

Большие языковые модели представляют собой нейронные сети, построенные на основе архитектуры Transformer с использованием механизма self-attention (самовнимания) [16]. Взаимодействие с LLM осуществляется посредством промптов (prompts) – текстовых инструкций, определяющих задачу и контекст для генерации ответа. Промпт-инжиниринг выделился в отдельную практику, включающую техники цепочки рассуждений [17] и итеративной самокоррекции [14].

Агентная система на основе LLM представляет собой программный продукт, в котором языковая модель выступает ядром рассуждения, взаимодействующим с внешними инструментами [18].

Агент, работающий в соответствии с классической векторной технологией RAG [12], предусматривает следующие этапы работы (рис. 1):

1. *Индексация*. Каждый проприетарный документ, материал предприятия разбивается на текстовые фрагменты (чанки) фиксированного размера с заданным перекрытием. Для каждого такого фрагмента вычисляется его эмбеддинг (отображение текста в m-мерное векторное пространство), который заносится в векторную базу данных.

2. *Извлечение*. По запросу пользователя агентной системы вычисляется эмбеддинг текста этого запроса по тем же правилам, что и при подготовке проприетарных материалов на первом этапе. По нему в векторной базе, сформированной на первом этапе, выполняется поиск (как правило, по косинусной близости) некоторого количества k ближайших соседей.

3. *Генерация*. Извлеченные текстовые фрагменты, соответствующие найденным k векторам, добавляются к запросу и формируют с ним вместе промпт, который подается на вход в LLM. Ответ LLM передается пользователю агента или используется в дальнейшем в бизнес-процессе.

Рис. 1. Общая архитектура классической векторной технологии RAG

Таким образом, добавляя в первоначальный запрос дополнительную, специфичную для предприятия информацию, на основе которой языковая модель может дать пользователю более полный и точный ответ, агент может работать более качественно. Индексация новых материалов предприятия может проводиться параллельно с работой агента. Механизм эмбеддинга используется с целью повышения скорости и качества подбора релевантных данных для генерации составного промпта. В качестве источников данных можно использовать любую неструктурированную информацию, например, текстовые документы, материалы wiki-корпоративных и проектных баз знаний, данные в базах данных различных систем. В качестве векторных СУБД можно использовать, например, Qdrant, Weaviate, Milvus, PostgreSQL с расширением pgvector.

При работе агента в зависимости от прикладной задачи необходимо уделять внимание размерам текстовых фрагментов (чанков) при индексации: чем он меньше, тем точнее будет буквальный поиск; чем больше, тем больше поиск приближается к смысловому. Начало и конец чанка должны быть осмысленными, в идеале должны совпадать с началом и концом предложения или абзаца. Если по запросу пользователя найдено очень много информации и она вся не помещается в контекст промпта, то ее можно предварительно сжать с помощью LLM. Для уточнения промпта (особенно при работе с узкоспециализированными областями, сложной терминологией), кроме данных, извлеченных по технике эмбеддинга, можно также добавлять материалы, подобранные другими методами, и/или давать агенту или непосредственно пользователю возможность корректировать агента с использованием больших языковых моделей (LLM). Также следует вводить алгоритмы, позволяющие не передавать во внешние LLM чувствительные данные, например, персональные, коммерческие и т. п.

Несмотря на широкое промышленное распространение, классическая векторная технология RAG имеет ряд существенных ограничений [9; 13]:

– неспособность к глобальному осмыслению, потеря связи между разрозненными фактами (поиск извлекает фрагменты текста, локально релевантные запросу, и не способен агрегировать информацию, распределенную по разным документам; векторный поиск не гарантирует одновременное нахождение всех необходимых фрагментов для установления связей между сущностями из различных документов);

– зашумленность контекстного окна (включение фрагментов текста, семантически близких к запросу, но фактологически нерелевантных, что может дезориентировать LLM);

– LLM уделяют непропорционально мало внимания информации, расположенной в середине контекстного окна, что снижает утилизацию извлеченных фрагментов.

Для преодоления этих ограничений был предложен GraphRAG-подход [9], который представляет собой продвинутый метод, расширяющий RAG и объединяющий графы знаний с большими языковыми моделями для повышения точности и контекстного понимания при работе с данными. Ключевая идея состоит в использовании графа знаний, автоматически извлеченного LLM из исходного корпуса документов предприятия и подаваемого в качестве контекста к запросу в LLM.

При использовании GraphRAG изменяется стадия индексации проприетарных документов компании. Вместо их векторизации используются механизмы построения графа знаний по их данным. Построение такого графового индекса – многоэтапный процесс, в котором неструктурированный текст последовательно преобразуется в иерархическое хранилище знаний (рис. 2):

1. *Этап сегментации*. Каждый проприетарный документ, материал предприятия разбивается на текстовые фрагменты фик-

сированного, достаточно большого размера с заданным перекрытием (например, по 600 символов с перекрытием в 100 символов) для сохранения сущностей в перекрывающихся фрагментах текста.

2. *Этап выделения сущностей и их отношений*. Каждый фрагмент текста, полученный на этапе сегментации, обрабатывается LLM, которые извлекают из него именованные сущности (конкретные объекты предметной области, такие как организации, локации, технологии и т. п.) и отношения – связи между ними. Каждая сущность и каждое отношение сопровождаются текстовым описанием (дескриптором), объясняющим контекст упоминания. Для повышения полноты результата такого процесса используется техника саморефлексии [14], при которой после первичной экстракции модель перепроверяет результат и доизвлекает пропущенное.

3. *Этап построения графа знаний*. Извлеченные сущности становятся вершинами графа, а отношения – его ребрами. Если одна и та же сущность упоминается несколько раз в разных отрывках текста, то ее дескрипторы объединяются через LLM. Вес ребра графа определяется числом независимых фрагментов, в которых данное отношение было обнаружено: чем чаще упоминается связь, тем она значимее.

4. *Этап обнаружения тематических кластеров*. На этом этапе используется алгоритм Leiden [15], один из наиболее эффективных методов кластеризации графов [10], позволяющий выявить в нем тематически связанные группы сущностей. Leiden применяется рекурсивно: сначала к исходному графу, затем к подграфам внутри каждого найденного сообщества. В результате формируется несколько уровней вложенных графов: корневой (крупнейшие тематические блоки), промежуточные (подтемы внутри крупных блоков) и листовой – максимально детальные подгруппы, соответствующие, например, конкретным рабочим ситуациям или процессам.

5. *Этап генерации текстовых отчетов (сводок) для каждого кластера*. Это финальный и ключевой этап индексации. Для каждого кластера LLM составляет текстовый отчет – сводку: связное описание группы сущностей. Для этого LLM получает на вход список всех сущностей кластера, их отношения и дескрипторы. На этой основе модель генерирует структурированный текст, содержащий перечень ключевых сущностей кластера, описание их взаимосвязей, основные факты и выводы, которые можно извлечь из данной группы. На листовых уровнях сводки строятся непосредственно из дескрипторов сущностей и ребер. На более верхних уровнях сводки создаются путем обобщения сводок подчиненных кластеров. Таким образом, на корневом уровне сводка содержит обобщенное описание всего документарного корпуса – максимально сжатое описание всех основных тем, акторов и взаимосвязей.

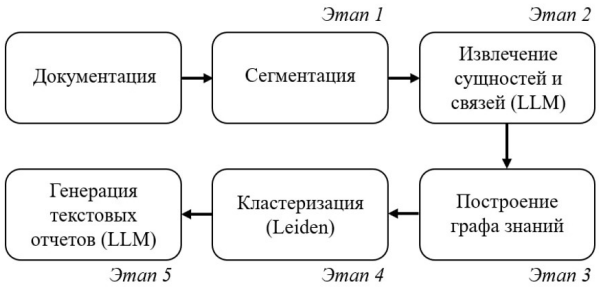

Рис. 2. Цепочка индексации GraphRAG

Именно сводки делают возможным главное преимущество GraphRAG: при ответе на общие вопросы система не ищет отдельные фрагменты текста, а использует подготовленные в сводках обобщения.

После индексации в распоряжении агента GraphRAG находятся три структуры данных: граф знаний с вершинами-сущностями и ребрами-отношениями, иерархия кластеров и текстовые сводки на каждом уровне иерархии.

В случае, когда пользователь задает вопрос, система выбирает одну из стратегий работы в зависимости от характера запроса:

– *Global Search*. Применяется для обзорных и аналитических запросов, для кото-

рых наиболее эффективен GraphRAG. Система берет заранее подготовленные сводки кластеров нужного уровня и обрабатывает их по схеме, когда каждая сводка независимо передается в LLM вместе с вопросом пользователя. Модель генерирует частичный ответ и оценивает его релевантность по шкале от 0 до 100. Частичные ответы с ненулевой релевантностью ранжируются по убыванию релевантности и объединяются в финальный связный ответ;

– *Local Search*. Применяется для вопросов о конкретных сущностях. Система находит сущность в графе и подбирает контекст о ней по графу знаний, который прикладывается к запросу пользователя в LLM;

– *Drift Search*. Комбинированная стратегия для вопросов, которые начинаются с конкретной сущности, но требуют более широкого контекста. Объединяет навигацию по графу от заданной сущности (как в Local Search) с привлечением сводок (как в Global Search), обеспечивая баланс между точностью адресного поиска и широтой охвата.

Общая архитектура GraphRAG приведена на рис. 3.

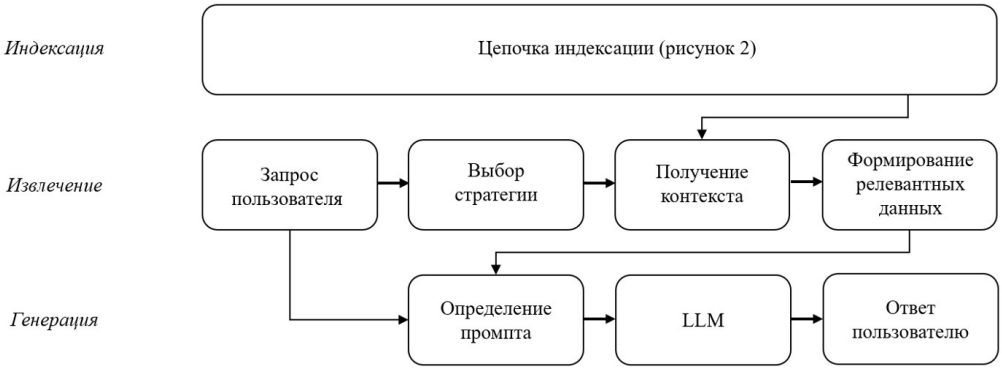

Рис. 3. Общая архитектура GraphRAG

Объективными ограничениями GraphRAG являются:

– высокая стоимость и время индексации, так как построение графа знаний требует многократного обращения к LLM и происходит в несколько трудоемких этапов (для уменьшения этого недостатка на практике возможно использование для ряда задач более экономичных LLM или локально развернутых open-source-моделей);

– отсутствие преимущества перед векторной классической технологией RAG при локальных запросах, направленных на извлечение конкретных фактов (в этом случае векторный поиск предпочтителен);

– зависимость от качества экстракции (точность графа знаний определяется способностью LLM корректно идентифицировать сущности и отношения, что варьируется в зависимости от модели LLM и языка).

Области применения RAG и GraphRAG

В связи с вышеизложенным можно сформулировать критерии целесообразности практического применения RAG и GraphRAG.

Классическую векторную технологию RAG целесообразно рассматривать к применению в агентах искусственного интеллекта:

– направленных на извлечение из LLM ответов на вопросы по конкретным фактам;

– с приоритетом скорости работы;

– с приоритетом стоимости работы перед качеством и полнотой ответов;

– при работе с нестабильным, часто пополняющимся документарным корпусом.

GraphRAG целесообразно рассматривать к применению в агентах искусственного интеллекта, когда:

– документарный корпус содержит сложные взаимосвязи между сущностями, требующие установления неявных зависимостей;

– информационные потребности пользователей включают глобальные аналитические запросы (тематический анализ, выявление трендов, сравнительный анализ);

– корпус стабилен или обновляется нечасто, что амортизирует стоимость индексации;

– качество и полнота ответов имеют приоритет перед скоростью отклика и лаконичностью.

Следует отметить, что подходы RAG и GraphRAG могут использоваться совместно для построения комплексных решений.

Примеры практического применения GraphRAG

В качестве примеров применения GraphRAG можно привести:

– поиск в литературе полных ответов по выбранной научной предметной области. При динамически развивающемся, меняющемся корпусе документов целесообразно рассмотреть использование векторной технологии RAG вместо GraphRAG. При необходимости учитывать связи между статьями, принадлежность одному научному направлению, исследованиям и соавторству рекомендован к применению GraphRAG;

– построение систем подбора туристских предложений, построения маршрутов логистики, требующих учета предпочтений клиента, визовых ограничений, сезонности, связей экскурсионных и развлекательных программ, возрастных ограничений, стоимости и других параметров, приоритетных для конкретного пользователя;

– анализ записанных разговоров коллцентров, когда часть информации отсутствует в CRM системах и содержится только в аудиозаписях. При обработке учитывается дополнительная информация о незафиксированных возражениях, пожеланиях клиентов;

– поиск маркетинговых стратегий, формирование рекламных предложений для продвижения товаров по материалам CRM, рекламным описаниям и документации на продуктовую линейку;

– анализ недостающих компетенций сотрудников, подбор материалов внутренней программы обучения, документации для самостоятельного изучения, ментора, экспертов для консультаций;

– подбор параметров изменения бизнес-среды и прогностика их поведения и влияния на реализацию цифровой трансформации, проводимой в условиях постоянно меняющейся бизнес-среды, ускорения жизненного цикла продуктов, все большей ориентации продуктовой линейки на массового разнородного потребителя с перманентно изменяющимися предпочтениями. Условия работы предприятий и проведения цифровой трансформации изменились и продолжают меняться, что привело к тому, что традиционные методы управления и оценки эффективности становятся все менее актуальными и менее чувствительными к динамике и характеру изменений в бизнес-среде [2–4; 6; 7]. Сложившиеся условия подталкивают развитие новых подходов к управлению всем жизненным циклом цифровой трансформации, базирующимся на принципах Методологии управления цифровой трансформацией (МУЦТ) коммерческих компаний [5], работающих в современных условиях. Оценка эффективности цифровой трансформации и предприятия в целом, корректировки траектории цифровой трансформации, согласно МУЦТ, производятся в зависимости от изменений бизнес-среды и прогностики ее поведения. Параметры бизнес-среды могут быть подобраны и оценены с использованием запросов в LLM, подходов GraphRAG и документации по продуктам предприятия, проводимым им проектам цифровой трансформации, материалам по архитектуре компании. С использованием этих же подходов воз-

можно определение направлений изменений бизнес-среды.

Таким образом, в результате анализа подходов RAG и GraphRAG, используемых для повышения качества генерируемых ответов в системах на базе больших языковых моделей (LLM), сформулированы их недостатки и преимущества, определены критерии целесообразности практического применения RAG и GraphRAG, приведены примеры систем, где они могут использоваться. Практическая ценность исследования состоит в возможности применения данных подходов при реализации современных корпоративных систем искусственного интеллекта.

Список литературы

1. *Винер Н.* Кибернетика, или Управление и связь в животном и машине / пер. с англ. И. В. Соловьева и Г. Н. Поварова; под ред. Г. Н. Поварова. – 2-е изд. – М. : «Наука» Главная редакция изданий для зарубежных стран, 1983.

2. *Денисенко В. Ю.* Мониторинг эффектов цифровых продуктов в условиях цифровой трансформации промышленных предприятий // Креативная экономика. – 2021. – Т. 15. – № 5. – С. 1715–1724.

3. *Кокуйцева Т. В., Овчинникова О. П.* Методические подходы к оценке эффективности цифровой трансформации предприятий высокотехнологичных отраслей промышленности // Креативная экономика. – 2021. – Т. 15. – № 6. – С. 2413–2430.

4. *Кочетков Е. П., Забавина А. А., Гафаров М. Г.* Цифровая трансформация компаний как инструмент антикризисного управления: эмпирическая оценка влияния на эффективность // Стратегические решения и риск-менеджмент. – 2021. – Т. 12. – № 1. – С. 68–81.

5. *Куренков А. Л.* Управление цифровой трансформацией коммерческих предприятий : монография. – М. : Инфра-М, 2025.

6. *Уколов В. Ф., Афанасьев В. Я., Черкасов В. В.* Ключевые эффекты цифровизации и возможные потери // Вестник университета. – 2019. – № 8. – С. 55–58.

7. *Ценжарик М. К., Крылова Ю. В., Стешенко В. И.* Цифровая трансформация компаний: стратегический анализ, факторы влияния и модели // Вестник Санкт-Петербургского университета. Экономика. – 2020. – Т. 36. – Вып. 3. – С. 390–420.

8. *Brown T., Mann B., Ryder N.* et al. Language Models are Few-Shot Learners // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 1877–1901.

9. *Edge D., Trinh H., Cheng N.* et al. From Local to Global: A GraphRAG Approach to Query-Focused Summarization. – URL: https://arxiv.org/abs/2404.16130 (дата обращения: 19.02.2026).

10. *Fortunato S.* Community Detection in Graphs // Physics Reports. – 2010. – Vol. 486. – N 35. – P. 75–174.

11. *Jeffcock P.* What's the Difference between AI, Machine Learning, and Deep Learning? Peter Jeffcock. – URL: https://blogs.oracle.com/bigdata/difference-ai-machine-learning-deep-learning (дата обращения: 09.03.2020).

12. *Lewis P., Perez E., Piktus A.* et al. Retrieval-Augmented Generation for Knowledge Intensive NLP Tasks // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 9459–9474.

13. *Liu N. F., Lin K., Hewitt J.* et al. Lost in the Middle: How Language Models Use Long Contexts. – URL: https://arxiv.org/abs/2307.03172 (дата обращения: 19.02.2026).

14. *Shinn N., Cassano F., Gopinath A.* et al. Reflexion: Language Agents with Verbal Reinforcement Learning. – URL: https://arxiv.org/abs/2303.11366 (дата обращения: 19.02.2026).

15. *Traag V. A., Waltman L., Van Eck N. J.* From Louvain to Leiden: Guaranteeing Well Connected Communities // Scientific Reports. – 2019. – Vol. 9. – N 1. – Article 5233.

16. *Vaswani A., Shazeer N., Parmar N.* et al. Attention is All You Need. – URL: https://arxiv.org/abs/1706.03762 (дата обращения: 19.02.2026).

17. *Wei J., Wang X., Schuurmans D.* et al. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. – URL: https://arxiv.org/abs/2201.11903 (дата обращения: 19.02.2026).

18. *Yao S., Zhao J., Yu D.* et al. ReAct: Synergizing Reasoning and Acting in Language Models // International Conference on Learning Representations (ICLR), 2023. – URL: https://collaborate.princeton.edu/en/publications/react-synergizing-reasoning-and-acting-in-language-models/ (дата обращения: 19.02.2026).

References

1. Winer N. Kibernetika, ili Upravlenie i svyaz v zhivotnom i mashine [Cybernetics or Management and Communication in Living-Being and Machine], translated from English by I. V. Solovev and G. N. Povarov; edited by G. N. Povarov. 2nd edition. Moscow, 'Nauka' the General Editors for Foreign Countries, 1983. (In Russ.).

2. Denisenko V. Yu. Monitoring effektov tsifrovykh produktov v usloviyakh tsifrovoy transformatsii promyshlennykh predpriyatiy [Monitoring Effects of Digital Products in Conditions of Digital Transformation in Industrial Enterprises]. *Kreativnaya ekonomika* [Creative Economy], 2021, Vol. 15, No. 5, pp. 1715–1724. (In Russ.).

3. Kokuytseva T. V., Ovchinnikova O. P. Metodicheskie podkhody k otsenke effektivnosti tsifrovoy transformatsii predpriyatiy vysokotekhnologichnykh otrasley promyshlennosti [Methodological Approaches to Estimating Efficiency of Digital Transformation in Enterprises of Highly-Technological Industries]. *Kreativnaya ekonomika* [Creative Economy], 2021, Vol. 15, No. 6, pp. 2413–2430. (In Russ.).

4. Kochetkov E. P., Zabavina A. A., Gafarov M. G. Tsifrovaya transformatsiya kompaniy kak instrument antikrizisnogo upravleniya: empiricheskaya otsenka vliyaniya na effektivnost [Digital Transformation of Companies as Tool of Crisis Management: Empiric Assessment of Impact on Efficiency]. *Strategicheskie resheniya i risk-menedzhment* [Strategic Solutions and Risk-Management], 2021, Vol. 12, No. 1, pp. 68–81. (In Russ.).

5. Kurenkov A. L. Upravlenie tsifrovoy transformatsiey kommercheskikh predpriyatiy: monografiya [Managing Digital Transformation of Commercial Enterprises: monograph]. Moscow, Infra-M, 2025. (In Russ.).

6. Ukolov V. F., Afanasev V. Ya., Cherkasov V. V. Klyuchevye effekty tsifrovizatsii i vozmozhnye poteri [Key Effects of Digitalization and Possible Losses]. *Vestnik universiteta* [Bulletin of the University], 2019, No. 8, pp. 55–58. (In Russ.).

7. Tsenzharik M. K., Krylova Yu. V., Steshenko V. I. Tsifrovaya transformatsiya kompaniy: strategicheskiy analiz, faktory vliyaniya i modeli [Digital Transformation of Companies: Strategic Analysis, Impact Factors and Models]. *Vestnik Sankt-Peterburgskogo universiteta. Ekonomika* [Bulletin of the St. Petersburg University. Economics], 2020, Vol. 36, Issue 3, pp. 390–420. (In Russ.).

8. Brown T., Mann B., Ryder N. et al. Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, 2020, Vol. 33, pp. 1877–1901.

9. Edge D., Trinh H., Cheng N. et al. From Local to Global: A GraphRAG Approach to Query-Focused Summarization. Available at: https://arxiv.org/abs/2404.16130 (accessed 19.02.2026).

10. Fortunato S. Community Detection in Graphs. *Physics Reports*, 2010, Vol. 486, No. 35, pp. 75–174.

11. Jeffcock P. What's the Difference between AI, Machine Learning, and Deep Learning? Peter Jeffcock. Available at: https://blogs.oracle.com/bigdata/difference-ai-machine-learning-deep-learning (accessed 09.03.2020).

12. Lewis P., Perez E., Piktus A. et al. Retrieval-Augmented Generation for Knowledge Intensive NLP Tasks. *Advances in Neural Information Processing Systems*, 2020, Vol. 33, pp. 9459–9474.

13. Liu N. F., Lin K., Hewitt J. et al. Lost in the Middle: How Language Models Use Long Contexts. Available at: https://arxiv.org/abs/2307.03172 (accessed 19.02.2026).

14. Shinn N., Cassano F., Gopinath A. et al. Reflexion: Language Agents with Verbal Reinforcement Learning. Available at: https://arxiv.org/abs/2303.11366 (accessed 19.02.2026).

15. Traag V. A., Waltman L., Van Eck N. J. From Louvain to Leiden: Guaranteeing Well Connected Communities. *Scientific Reports*, 2019, Vol. 9, No. 1, Article 5233.

16. Vaswani A., Shazeer N., Parmar N. et al. Attention is All You Need. Available at: https://arxiv.org/abs/1706.03762 (accessed 19.02.2026).

17. Wei J., Wang X., Schuurmans D. et al. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. Available at: https://arxiv.org/abs/2201.11903 (accessed 19.02.2026).

18. Yao S., Zhao J., Yu D. et al. ReAct: Synergizing Reasoning and Acting in Language Models // International Conference on Learning Representations (ICLR), 2023. Available at: https://collaborate.princeton.edu/en/publications/react-synergizing-reasoning-and-acting-in-language-models/ (accessed 19.02.2026).

Поступила: 04.03.2026 Принята к печати: 29.04.2026

Сведения об авторах

Александр Львович Куренков
кандидат технических наук, доцент
базовой кафедры цифровой экономики
института развития информационного
общества РЭУ им. Г. В. Плеханова.
Адрес: ФГБОУ ВО «Российский
экономический университет
имени Г. В. Плеханова», 109992,
Москва, Стремянный пер., д. 36.
E-mail: kurenkov.al@rea.ru
ORCID: 0009-0009-2998-0288

Дарья Александровна Куренкова
аналитик ООО «Платформеко».
Адрес: ООО «Платформеко», 117630,
Москва, ул. Воронцовские пруды,
д. 3, офис 48/2.
E-mail: s02240475@gse.cs.msu.ru

Information about the author

Aleksandr L. Kurenkov
PhD, Assistant Professor
of the Basic Department of Digital Economy
of the Institute of the Information Society
of the PRUE.
Address: Plekhanov Russian University
of Economics, 36 Stremyanny Lane,
Moscow, 109992,
Russian Federation.
E-mail: kurenkov.al@rea.ru
ORCID: 0009-0009-2998-0288

Daria A. Kurenkova
Analyst of Platformeco LLC.
Address: Platformeco LLC,
office 48/2, 3 Vorontsovskiye prudy Str.,
Moscow, 117630, Russian Federation.
E-mail: s02240475@gse.cs.msu.ru