

ОБЗОРЫ

УПРАВЛЕНИЕ ЗНАНИЯМИ ОРГАНИЗАЦИИ И БОЛЬШИЕ ЯЗЫКОВЫЕ МОДЕЛИ

Ю. А. ЗЕЛЕНКОВ

Высшая школа бизнеса, Национальный исследовательский университет «Высшая школа экономики», Россия

Цель исследования: обобщение, классификация и анализ современных научных публикаций, посвященных использованию больших языковых моделей (large language models) в управлении знаниями организации. **Методология исследования:** систематический обзор литературы на основе модели PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Для анализа отобрано 75 публикаций — научные статьи и отчеты консалтинговых компаний, вышедшие начиная с 2020 г. **Результаты исследования:** выделены четыре основные направления исследований: (1) проблемы внедрения больших языковых моделей; (2) влияние больших языковых моделей на эффективность управления знаниями; а также применение больших языковых моделей в процессах (3) использования и (4) создания знаний. В рамках каждого направления рассмотрены ключевые публикации и открытые вопросы. **Оригинальность и значимость результатов:** представлен систематический обзор современных зарубежных публикаций, предложена классификация исследовательских тем, определены потенциальные направления новых исследований. Рассмотрены ограничения, препятствующие внедрению больших языковых моделей в практику управления знаниями организации.

Ключевые слова: генеративный AI, большие языковые модели, управление знаниями, систематический обзор литературы.

JEL: M10, M15

ВВЕДЕНИЕ

Современные организации делают ставку на знания для поддержания долгосрочного преимущества. Это требует точного понимания процессов управления знаниями (knowledge management — КМ) и, в частности, того, как знания оцениваются, при-

обретаются, распространяются, хранятся и применяются в рамках организационной системы [Zelenkov, 2018]. Однако с начала нового тысячелетия на процессы КМ сильно влияет развитие искусственного интеллекта (artificial intelligence — AI) [Manesh et al., 2021], который повышает способности организации идентифицировать, ин-

Адрес организации: Высшая школа бизнеса, Национальный исследовательский университет «Высшая школа экономики», ул. Шаболовка, 28, Москва, 119049, Россия.

© Ю. А. Зеленков, 2024

<https://doi.org/10.21638/spbu18.2024.309>

терпретировать, делать выводы и учиться на основе данных [Mikalef, Gupta, 2021].

Классические модели AI (машинное обучение с учителем) позволяют автоматизировать процедуры принятия решений. Однако такие системы прежде всего ориентированы на поддержку повторяющихся рутинных процессов, основанных на структурированных данных (например, кредитный скоринг в банках). Их преимущество заключается в нахождении и использовании статистических зависимостей, которые не может обнаружить человек.

Генеративный AI (generative artificial intelligence — GenAI) представляет собой новую парадигму, поскольку такие модели могут имитировать творческие способности человека [Shanahan, 2024], что делает их еще более полезным в процессах, подразумевающих интенсивное использование знаний [Berg, Raj, Seamans, 2023]. Модели GenAI решают задачу создания нового контента, преобразуя сложные данные в удобные для человека формы — текст, изображение, звук или видео [Mohr, Yunus, 2023].

Наибольший интерес в контексте КМ представляют большие языковые модели (large language models — LLM), ключевой особенностью которых является способность отвечать на вопросы, сформулированные на естественном языке [Lan et al., 2022]. Хотя эта способность опирается лишь на предсказание наиболее вероятного слова в последовательности, генерируемой в соответствии с запросом пользователя, многие исследования показывают, что в LLM создается модель внешней среды, которая служит для интерпретации, обобщения и выделения абстрактных категорий [Roberts, Raffel, Shazeer, 2020; Wang et al., 2022; Grosse et al., 2023].

Некоторые авторы (например, [Eloundou et al., 2023]) указывают, что GenAI можно рассматривать как технологию общего назначения, поскольку она удовлетворяет трем необходимым критериям, таким как: 1) совершенствование с течением времени; 2) повсеместное распространение

в экономике; 3) способность порождать дополнительные инновации [Lipsey, Carlaw, Bekar, 2005]. С использованием LLM связаны потенциальные риски [Weidinger et al., 2022], однако наблюдается взрывной рост применения этой технологии [Teubner et al., 2023].

На основании анализа рынка США отмечается, что уже 15 % всех видов работ выполняются при использовании GenAI быстрее, но с тем же качеством, расширение его использования потенциально будет способствовать тому, что доля таких работ составит 47–53 % [Eloundou et al., 2023]. Все это привело к тому, что инвестиции в GenAI увеличились за год почти в 10 раз с 2,85 млрд долл. в 2022 г. до 25,23 млрд долл. в 2023 г., при этом суммарные инвестиции в AI всех направлений снизились за этот же период со 103,4 до 96 млрд долл. [Maslej et al., 2024, p. 243–244].

Однако только в 12 % российских компаний присутствует понимание того, для каких задач будут использованы приложения GenAI, а приоритетами являются коммуникации с клиентами, маркетинг и продажи [Яков и Партнеры, 2023; Долотов, 2024]. Это содействует тому, что менее 10 % отечественных компаний планируют потратить на приложения GenAI более 3 % средств из бюджета на развитие информационных технологий. В то же время опрос компаний из списка Fortune 500 показывает, что зарубежные лидеры использования GenAI увеличили соответствующие бюджеты в 2,5 раза по сравнению с 2023 г., и среди приоритетных направлений КМ и автоматическое реферирование текстов (60 и 62 % из опрошенных компаний соответственно) опережают клиентский сервис и маркетинг (59 и 53 % соответственно) [Wang, Xu, 2024].

Такая ситуация типична для освоения новой технологии, когда на первом этапе тестируется несколько вариантов ее использования и интеграции в существующие системы и еще отсутствуют доминирующие решения, гарантирующие эффек-

тивность. Тем не менее очевиден сдвиг приоритетов российских компаний относительно зарубежных, что дополнительно осложняется ограничением доступа к наиболее производительным LLM.

Цель данного исследования — обобщение современных зарубежных научных публикаций, исследующих проблемы и перспективы использования LLM в управлении знаниями организации.

Вопрос исследования формулируется достаточно узко: как появление и быстрое развитие LLM влияют на управление корпоративными знаниями? Поэтому в качестве методологии выбран систематический обзор литературы.

Статья имеет следующую структуру. В первом разделе представлена методология подготовки обзора и четыре выделенных направления исследований. Во втором кратко рассматриваются основные технические вопросы построения LLM. Последующие разделы посвящены обсуждению направлений исследований: в третьем рассматриваются общие вопросы внедрения LLM, в четвертом — влияние LLM на эффективность управления знаниями в организации в целом, в пятом и шестом — применение LLM в процессах использования и создания знаний соответственно. В заключении сформулированы наиболее важные выводы, которые, на взгляд автора, описывают текущее состояние дел.

МЕТОДОЛОГИЯ ИССЛЕДОВАНИЯ

В исследовании представлен систематический обзор литературы [Snyder, 2019], т. е. отбор, анализ и обобщение научных публикаций, соответствующих критериям, установленным согласно заданной теме. Для обеспечения требований прозрачности, повторяемости и полноты использован фреймворк PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [Rethlefsen et al., 2021], который

включает следующие шаги: 1) поиск публикаций в базах данных; 2) отбор релевантных статей; 3) анализ полных текстов статей; 4) качественный анализ и синтез.

Поиск публикаций в базах данных. Критерий поиска сформулирован в соответствии с вопросом исследования, т. е. в первую очередь отбирались статьи на английском языке, заголовки, аннотации и полные тексты которых включали термины “knowledge management” и “large language model / LLM”.

Кроме того, в критерии отбора вошли показатели качества публикации — это квартиль источника, в котором опубликована статья (не ниже Q2 по данным сайта SJR¹), либо рейтинг соответствующей конференции (не ниже A по данным CORE Rankings Portal²). Для публикаций, не прошедших рецензирование (препринты arXiv.org, SSRN, консалтинговые обзоры и т. д.), для оценки качества использовалось количество цитирований по данным системы Google Scholar³. Рассматривались только те публикации, которые вышли в свет начиная с 2020 г.

Поиск осуществлялся в базах данных ключевых издательств Science Direct, IEEE Xplore, SpringerLink, ACM, а также в системе Google Scholar. Последняя покрывает публикации издательств, не вошедших в основной список, а также труды конференций, препринты и обзоры консалтинговых компаний.

Отбор релевантных статей. Отбирались публикации, посвященные вопросам влияния LLM на эффективность фирмы в целом и на процессы КМ (прежде всего поиск, использование и создание знаний) в частности, интеграции LLM в деятельность фирмы, а также практические кей-

¹ Scimago Journal & Country Rank. URL: <https://www.scimagojr.com> (дата обращения: 05.04.2024).

² CORE Rankings Portal. URL: <https://www.core.edu.au/conference-portal> (дата обращения: 05.04.2024).

³ Google Scholar. URL: <https://scholar.google.com> (дата обращения: 05.04.2024).

Таблица 1

Распределение публикаций: источники и направления исследований

База данных	Направление исследований				Всего
	Внедрение LLM	Влияние LLM на КМ	Использование знаний	Создание знаний	
ACM	5	1	0	0	6
Google Scholar	17	8	4	10	39
IEEE Xplore	3	0	5	2	10
ScienceDirect	1	2	9	1	13
SpringerLink	1	1	2	3	7
Всего	27	12	20	16	75

сы. Однако кейсы использования LLM в медицине и образовании исключены из рассмотрения, поскольку фокус настоящего исследования — применение LLM в бизнесе.

Анализ и синтез. Поскольку количество отобранных источников относительно невелико (75 документов, табл. 1), применение количественных методов связано с риском систематической ошибки [Carrera-Rivera et al., 2022], поэтому использован качественный анализ.

Ряд методов качественного анализа предполагает кодирование статей на основе заранее определенных конструктов (направлений исследований). Однако здесь выбран способ идентификации этих конструктов на основе данных [Sauer, Seuring, 2023], поскольку исследуемая область является новой и отсутствует общепринятое разделение тем. Для этого была реализована следующая процедура: для каждой статьи после анализа ее полного текста кратко формулировалась исследуемая проблема. Затем выделенные проблемы группировались по степени близости в более широкие темы, а темы — в направления исследований. Этот подход базируется на технике *metaplan*, которая используется для сбора идей при групповой работе, но с успехом применяется в междисциплинарных исследованиях для структуриро-

вания предметной области [Andersen et al., 2009]. В результате выделены четыре основные направления исследований: (1) проблемы внедрения LLM; (2) их влияние на эффективность КМ; а также применение LLM в процессах (3) использования и (4) создания знаний.

Кроме того, необходимо убедиться в том, что используемые для анализа публикаций конструкты достаточно определены, взаимоисключающи и обеспечивают полноту классификации [Sauer, Seuring, 2023]. Соответствующие данные (выделенные направления и темы, а также количество статей, отобранных после оценки релевантности и анализа полных текстов) представлены в табл. 1. Более детальная классификация исследований приведена на рис. 1.

КЛЮЧЕВЫЕ ТЕХНОЛОГИИ LLM

Успех LLM базируется на нескольких ключевых технологиях — векторном представлении слов (*embedding*), механизме внимания и самообучении. Поскольку компьютеры манипулируют числами, а не символами, первым этапом любой задачи обработки естественного языка является трансформация входных данных в числовое представление.

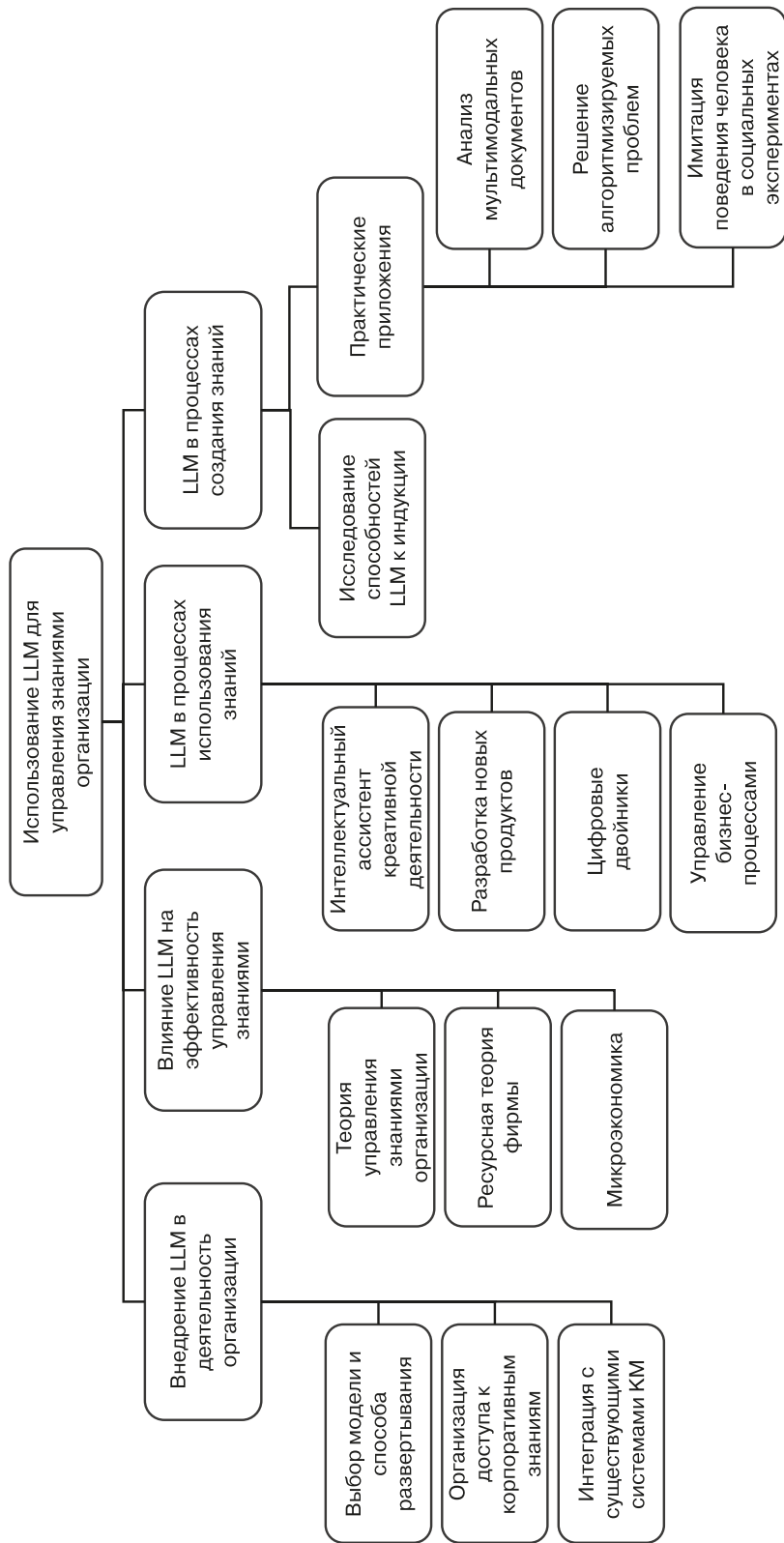


Рис. 1. Использование LLM для управления знаниями в организации: классификация направлений исследований

Векторное представление кодирует значение токена⁴ вектором в n -мерном пространстве таким образом, что термины, близкие по смыслу, будут находиться ближе и в векторном пространстве [Asudani, Nagwani, Singh, 2023]. Однако семантика слова может меняться в зависимости от контекста, поэтому в LLM используется механизм внимания — алгоритм, который позволяет модели фокусироваться на определенных частях входных данных [Vaswani et al., 2017]⁵. Эта технология дает возможность понимать контекст предложения и генерировать более точные ответы. Самообучение — техника, при которой модель учится предсказывать недостающие части входных данных. Указанный подход используется для обучения LLM предсказывать следующее слово или заполнять недостающие слова в предложении [Radford et al., 2018].

Однако разработчики LLM учитывают, что обученная таким образом модель может продуцировать этические и экзистенциальные риски, возникающие, если машины нарушают ценности и цели своих создателей и пользователей [Weidinger et al., 2022; Harrer, 2023]. Поэтому на финальной стадии LLM осуществляется обучение с подкреплением, когда пользователь явно указывает модели, какой из сгенерированных вариантов ответа является предпочтительным. В [Chang et al., 2024] представлен обзор методов оценки LLM.

Известно несколько архитектур моделей глубокого обучения, предназначенных для решения задач, связанных с обработкой естественного языка [Yang et al., 2024b]. В частности, предшественник LLM — трансформер [Vaswani et al., 2017] — является комбинацией двух нейронных сетей (энкодера и декодера) и предназначен для решения задачи

машинного перевода. Энкодер трансформирует входную последовательность слов на одном языке во внутреннее представление модели, а декодер превращает это представление в последовательность на другом языке. Обе сети аналогичны, и очень упрощенно каждую из них можно описать как слой, создающий векторные представления, за которым следует несколько слоев внимания и сетей прямого распространения.

LLM предсказывает следующее слово, поэтому используется только декодер. Предсказание делается на основе внутреннего представления модели, созданного на основе данных, на которых она была обучена. Поэтому объем и качество данных играют решающую роль. Например, для обучения модели GPT-3 компания OpenAI использовала около 400 млрд токенов из Интернета, книг и Wikipedia.

Ключевыми характеристиками LLM являются число параметров (весов отдельных нейронов), которое зависит от числа блоков внимания, и размер контекста. Число параметров определяет сложность модели и ее способность выделять абстрактные концепции в данных и устанавливать связи между ними. Экспериментально подтверждено, что при достижении некоторого порога сложности модель приобретает эмерджентные способности, которым ее специально не обучали и которые требуют рассуждений на лету или адаптации к домену — выполнение арифметических действий, объяснение шуток или дополнение программного кода [Brown et al., 2020]. Если модели с небольшим числом параметров могут просто копировать те утверждения, которые встречались в обучающих данных, то с ростом размера LLM создает все более абстрактные обобщения [Grosse et al., 2023; Kandpal et al., 2023].

Необходимо отметить, что существуют определенные сомнения в доказанности эмерджентного поведения LLM. Например, в [Schaeffer, Miranda, Koyejo, 2023] показано, что скачкообразное возрастание некоторых метрик при достижении опре-

⁴ Токен — минимальная часть слова, передающая его смысл.

⁵ Для учета контекста в LLM также применяется позиционное кодирование, когда кодируется не только значение слова, но и его положение в тексте.

Таблица 2

Основные параметры ведущих LLM

Компания	LLM	Год выпуска	Число параметров, млрд	Размер контекста, токен	Доступ для корпоративных заказчиков из РФ
Google	Gemini 1.5 Ultra	2024	1800*	32 тыс. — 10 млн	Нет
OpenAI	GPT-4 Turbo	2023	1760*	128 000	Нет
Anthropic AI	Claude 3.0 Opus	2024	500*	200 000	Нет
Сбер	ruGPT-3.5	2023	13.0	2048	Да
Яндекс	YaGPT-2	2023	1.3	8000	Да

Примечание: * — точные данные разработчиком не опубликованы, приведены оценки независимых экспертов.

Составлено по: [Maslej et al., 2024].

деленного уровня сложности моделей связано с нелинейностью самих метрик. Тем не менее авторы этой работы подчеркивают, что их результаты вовсе не означают, что LLM не обладают эмерджентными свойствами. По их мнению, некоторые ранее заявленные способности оказываются миражами, вызванными некорректным анализом.

Размер контекста определяет длину запроса, который может ввести пользователь в диалоге с LLM. Длинное контекстное окно позволяет передать LLM новые данные без изменения весов модели. Можно сказать, что таким образом производится дополнительное обучение на лету (контекстное обучение) [Agarwal et al., 2024]. Параметры наиболее сложных моделей, доступных в мире и в России на момент подготовки настоящей статьи, представлены в табл. 2.

Обучение сложных моделей требует наличия гигантской специализированной вычислительной инфраструктуры, а также штата высококвалифицированных специалистов. Затраты на обучение моделей GPT-4 и Gemini Ultra оцениваются в 78 и 191 млн долл. соответственно [Maslej et al., 2024, p. 64].

Важно отметить, что, помимо наиболее сложных моделей, также доступны систе-

мы меньшего масштаба (до 400 млрд параметров), которые, хотя и уступают по производительности [Maslej et al., 2024, p. 146], тем не менее могут быть интегрированы в процессы КМ. Кроме того, часть из них поставляется по открытой лицензии, их параметры (т. е. веса, полученные при обучении) могут быть загружены из репозитория Hugging Face⁶. Соответственно, такие системы заказчик может развернуть самостоятельно.

Взаимодействие с LLM происходит в диалоговом режиме. Пользователь вводит данные, которые модель использует для генерации ответа. В английском языке для обозначения вводимых данных применяется слово “prompt” («подсказка») в отличие от слова “query” («запрос») для обращения к структурированным данным. В настоящей работе используется калька «промт», поскольку в русскоязычной профессиональной среде этот термин считается устоявшимся.

Промт может представлять собой вопрос, команду, незавершенное предложение или любые другие данные в зависимости от сценария использования LLM. Как минимум промт ставит задачу, кото-

⁶ Hugging Face. URL: <https://huggingface.co/models> (дата обращения: 05.04.2024).

рую нужно решить, и контекст, в соответствии с которыми модель должна активировать те области внутреннего представления, которые кодируют соответствующие знания.

Поскольку LLM является вероятностной моделью, постольку возникает задача манипулирования промптом таким образом, чтобы распределение генерируемых токенов соответствовало ожиданиям пользователя [Jiang et al., 2020]. Этот процесс называется промпт-инжиниринг. Часто применяются следующие варианты промпта [Sahoo et al., 2024]:

- без примеров (zero-shot prompt) — простейший вариант, когда в промпте не приводятся примеры того, как должна реагировать модель;
- с примерами (few shots prompt) — в запросе приводится несколько примеров (тексты удачных пресс-релизов и т. д.), которые позволяют модели обучаться, используя предложенный контекст. При этом качество ответа модели возрастает с увеличением числа примеров [Agarwal et al., 2024];
- цепочка рассуждений (chain of thought) — вместе с примерами дается цепочка рассуждений, которые надо применять при выводе;
- размышления шаг за шагом (let's think step by step) — модели предлагается продемонстрировать цепочку рассуждений, которые она выполнила при генерации ответа.

Важно отметить, что добавление эмоциональной окраски в промпт (например, «это очень важно для моей карьеры») часто улучшает результат [Li et al., 2023].

ВНЕДРЕНИЕ LLM В ДЕЯТЕЛЬНОСТЬ ОРГАНИЗАЦИИ

При внедрении LLM необходимо принять решения по нескольким вопросам: 1) какую модель выбрать и на какой инфраструктуре развертывать систему — соб-

ственной или внешней; 2) как обеспечить доступ модели к корпоративным знаниям; 3) как интегрировать модель с существующими системами КМ? Рассмотрим сценарии их реализации.

Выбор модели и способа развертывания системы. Возможные варианты (рис. 2) определяются выбором LLM, это могут быть предварительно обученные модели с закрытыми или открытыми весами или обучение собственной сети. Каждый из этих сценариев имеет свои преимущества и недостатки.

Наиболее сложные и производительные модели с триллионами параметров доступны только как удаленные сервисы через сайт разработчика, причем оплата взимается за количество обработанных токенов. Большинство разработчиков предоставляет интерфейс прикладного программирования (application program interface — API), с помощью которого заказчики могут интегрировать LLM в свои приложения. Важно подчеркнуть, что доступ к таким системам закрыт для российских компаний. Среди зарубежных компаний — лидеров использования AI 27 % выбирают этот сценарий [Wang, Xu, 2024].

При использовании предварительно обученных LLM с открытыми весами заказчик получает больший контроль над системой, в частности возможно дообучение моделей под его специфические задачи. В этом случае требуется создание команды квалифицированных инженеров, способных сформировать необходимую инфраструктуру и произвести дообучение. Модель может быть развернута как на собственной инфраструктуре, так и на мощностях облачных провайдеров. Важно подчеркнуть, что самые сложные открытые модели имеют меньшее число параметров, чем ведущие проприетарные LLM, но адаптация моделей к специфике заказчика позволяет добиться высокой эффективности в решении его задач. Поэтому 73 % зарубежных компаний выбирают указанный сценарий, из них примерно треть (28 %) разворачивают такие модели на собственных мощностях,

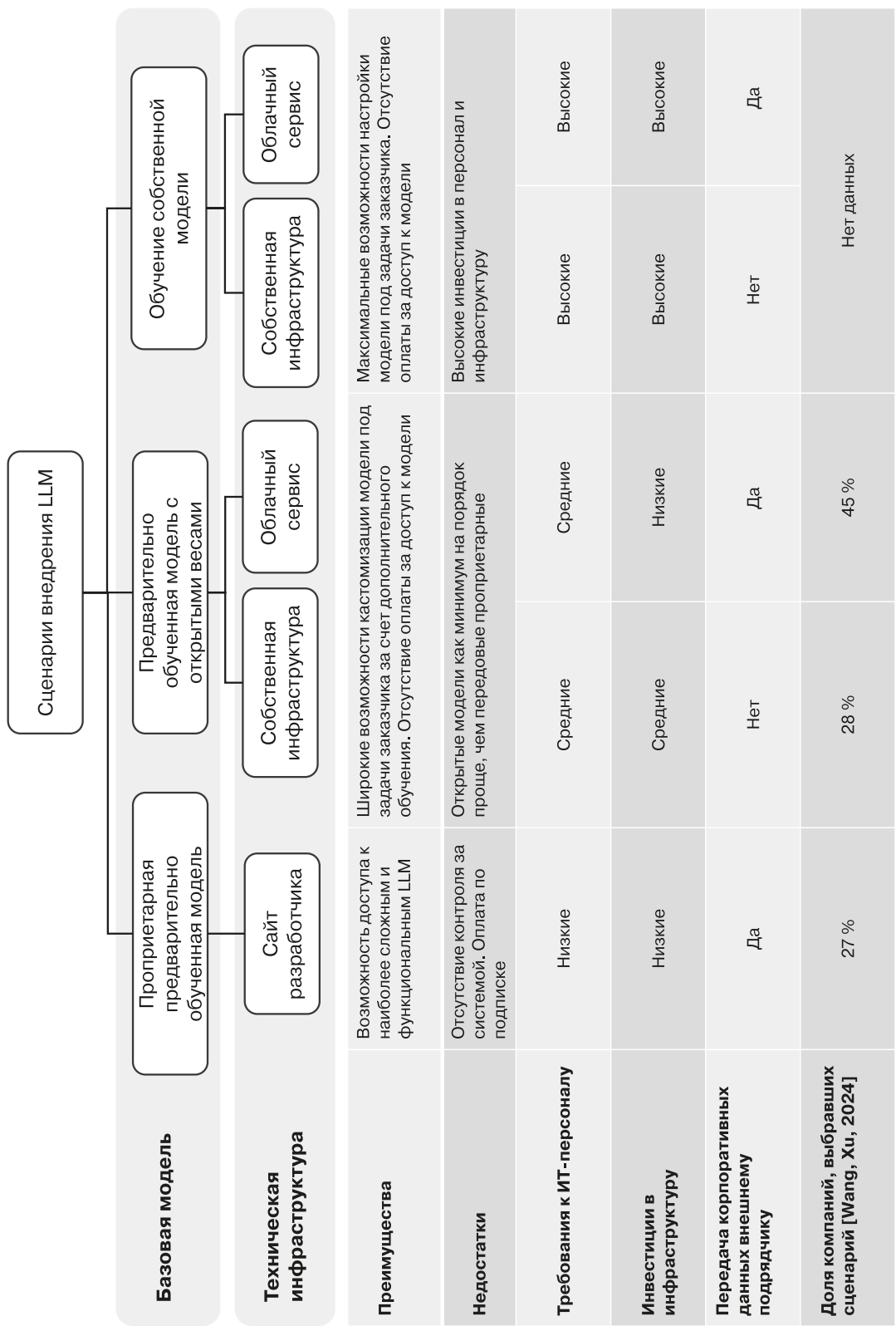


Рис. 2. Сценарии внедрения LLM в организации

а остальные применяют облачные сервисы [Wang, Xu, 2024].

Обучение собственной модели с нуля требует гигантских затрат, и в основном по этому пути идут компании, которые планируют продавать свои системы как сервисы другим фирмам. В качестве примера можно привести BloomerGPT [Wu et al., 2023] и модели российских компаний «Сбер» и «Яндекс». Данный сценарий требует привлечения квалифицированной команды исследователей, а также создания серьезной инфраструктуры — затраты, на которые способны только крупные компании.

Кроме того, доступ к системам, развернутым на инфраструктуре внешнего поставщика — разработчика LLM или облачного провайдера, — подразумевает, что данные потребителя будут передаваться за пределы компании. Это может рассматриваться как серьезная проблема с точки зрения информационной безопасности.

Доступ модели к корпоративным знаниям. Обеспечение доступа LLM к корпоративным ресурсам — важнейший фактор интеграции моделей в процессы КМ [O’Leary, 2024]. Несмотря на то что LLM аккумулируют гигантские знания, полученные во время обучения, необходим учет специфики деятельности конкретной организации. При решении сложных задач в конкретных областях возникает множество проблем, вызванных неоднородностью и сложностью данных, уникальностью целей и разнообразными ограничениями (культурными и социальными нормами, религиозными убеждениями, этическими стандартами и т. д.). Кроме того, необходимо поддерживать актуальность знаний LLM. Предложено несколько способов устранения указанных проблем [Ling et al., 2023; Yu et al., 2023b], основными из которых, помимо передачи дополнительного контекста через промпт, являются тонкая настройка моделей (fine tuning — FT) и генерация с расширенным поиском (retrieval augmented generation — RAG).

RAG представляет собой расширение процесса передачи дополнительной информации через контекст. В этом случае в системе появляется дополнительный модуль ретривер (retriever), функция которого — кодирование запроса на дополнительную информацию и поиск релевантных данных (рис. 3).

Наиболее простой сценарий генерации ответа следующий (более сложные варианты рассмотрены в [Gao et al., 2023]):

- пользователь вводит промпт и запрос на поиск информации для создания расширенного контекста;
- ретривер создает векторное представление запроса и обращается к дополнительным источникам информации (Интернету, документам и базам данных организации и т. д.). Эти данные также должны быть предварительно преобразованы в векторные представления, например, с помощью соответствующих слоев LLM;
- найденная информация (k векторов, наиболее близких по какой-либо метрике к вектору запроса) добавляется в расширенный контекст;
- пользовательский промпт, запрос на дополнительную информацию и расширенный контекст передаются LLM;
- LLM генерирует ответ.

Если методы расширения контекста, включая RAG, не изменяют параметры LLM, то FT предполагает их адаптацию с помощью дополнительного обучения на относительно небольшом объеме специально подготовленных данных. Можно выделить два базовых подхода [Ling et al., 2023]:

- на основе адаптеров — модулей, которые вставляются между слоями предварительно обученной модели так, чтобы изменить поведение частей LLM, критичных для решаемой задачи [Hu et al., 2021]. Преимуществом этого подхода является то, что при использовании нескольких адаптеров можно настроить одну базовую LLM для решения разных задач;



Рис. 3. Процесс генерации с расширенным поиском
Составлено по: [Lewis et al., 2020].

- с использованием инструкций, т. е. предварительно размеченных данных. В этом случае применяется обучение с учителем — модель дообучается выполнению специфических задач (классификация, реферирование, анализ эмоциональной окраски и т. д.) на коллекции промптов, с которыми сопоставлены ожидаемые ответы (рис. 4). В результате корректируются веса всех либо нескольких (обычно выходных) слоев. При этом необходимо отметить, что: 1) глобальное обновление весов LLM может разрушить ее способность к генерации ответов, 2) данный процесс требует значительных вычислительных затрат. Однако при учете этих ограничений удастся добиться возрастания качества (см., например, кейс по классификации правовых норм [Liga, Robaldo, 2023]).

Согласно исследованию лидеров использования GenAI, большинство организаций предпочитает не обучать свою собственную LLM с нуля, а применять RAG

(22 %) или дообучать модель с открытыми весами под свои конкретные нужды (72 %); прочие компании выбирают другие способы кастомизации [Wang, Xu, 2024].

Интеграция модели с существующими системами КМ. До появления LLM одним из основных подходов к созданию систем поддержки КМ были графы знаний, основанные на выделении в корпоративных данных именованных сущностей и связей между ними [Hogan et al., 2022]. Эта технология, известная также как онтологии и семантические сети, обеспечивает формализацию, обмен и повторное использование знаний, в результате чего такие системы также могут отвечать на вопросы с обоснованием ответа.

Сравнительный анализ LLM и графов знаний представлен в [Pan et al., 2024]. В отличие от графа знаний, который является структурированной интерпретируемой моделью, LLM можно рассматривать как черный ящик, аккумулирующий неявные знания. При этом для графа знаний всегда можно оценить недостаток знаний,

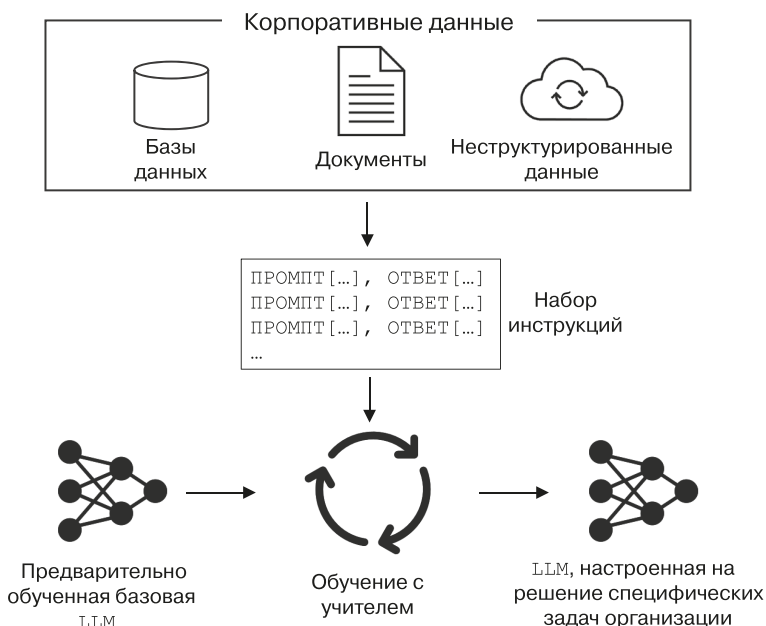


Рис. 4. Процесс тонкой настройки LLM на основе инструкций

в то время как LLM испытывают трудности, если требования не совпадают с усвоенными ими знаниями (т. е. необходимые знания не содержались в обучающей выборке) [Kandpal et al., 2023; Sun et al., 2023]. В этом случае может наблюдаться явление галлюцинации, когда LLM генерирует текст, не основанный на фактах.

Вместе с тем, хотя граф знаний, безусловно, является моделью внешней среды, нет оснований утверждать, что он способен обобщать факты и понятия, т. е. переходить от частного к общему на более высокой степени абстракции. Если эти абстракции отсутствуют в обучающих данных, они не будут представлены в модели, хотя в ряде специфических случаев возможны некоторые индуктивные неточные обобщения [Hogan et al., 2022, p. 67]. Основным преимуществом LLM является то, что она также извлекает паттерны из данных, и чем больше модель, тем более абстрактными становятся эти паттерны [Grosse et al., 2023]. В этом смысле можно утверждать, что граф знаний основан на данных, а LLM — на знаниях [Pan et al., 2024].

Анализируя свойства обоих подходов, авторы работы [Pan et al., 2024] отмечают, что недостатки одной модели соответствуют преимуществам другой, и наоборот, поэтому целесообразно объединять указанные технологии. При этом выделяются несколько возможных схем, которые сводятся к двум основным сценариям — использованию LLM для построения графа знаний и применению графа знаний как источника данных для LLM. Подробно эти схемы и их варианты рассмотрены в [Yang, Liu, Song, 2020; Yu et al., 2022a; Yang et al., 2024a].

В настоящее время входной барьер для реализации инициатив по использованию LLM остается достаточно высоким. Базовые сценарии сводятся либо к применению проприетарной модели как внешнего сервиса, либо к адаптации модели с открытыми весами (предельный случай второго сценария — обучение собственной модели с нуля). Первый вариант требует передачи корпоративных данных третьей стороне, что связано с рисками информационной безопасности. Кроме того, доступ

к таким сервисам закрыт для российских компаний. Второй сценарий предполагает создание собственных компетенций, что сопряжено с большими затратами, причем модели с открытыми весами уступают по производительности ведущим проприетарным LLM.

Несмотря на значительный прогресс, практически каждая из рассмотренных тем содержит множество открытых вопросов, требующих дальнейших исследований. Прежде всего отсутствуют стандартные решения, предлагающие оптимальный вариант интеграции LLM в бизнес. Это порождает две проблемы: 1) пользователи не могут сформулировать ожидания от системы и соответствующие требования; 2) технические специалисты при отсутствии стандартных решений вынуждены прибегать к методу проб и ошибок. Также актуальным является исследование механизмов обеспечения безопасности при передаче корпоративных данных внешнему поставщику. Поэтому стоит ожидать, что уже в ближайшее время на рынке должно появиться предложение услуг по внедрению LLM, которые будут включать идентификацию бизнес-проблемы, выбор оптимальной модели и сценария ее развертывания, создание необходимой инфраструктуры, дообучение и т. д.

ВЛИЯНИЕ LLM НА УПРАВЛЕНИЕ ЗНАНИЯМИ

В настоящее время решающим фактором, определяющим эффективность фирмы, является наличие знаний и возможности их использования и развития [Inkinen, 2016]. Многие авторы (см., напр., [Gaviria-Marin, et al., 2019]) выделяют четыре поколения в развитии КМ как академической дисциплины. На первом этапе (до 1980 г.) сформировалась концепция знаний как ресурса, влияющего на эффективность. На втором этапе (1990-е гг.) основные исследования фокусировались на процессах приобретения, распространения и исполь-

зования знаний. В центре внимания исследований третьего поколения (2000-е гг.) было проектирование корпоративных систем КМ. Однако после 2010 г. КМ в организации рассматривается скорее как социальный процесс со своими стимулами и ограничениями, чем как управленческая система, которую можно спроектировать и внедрить.

В современном понимании КМ базируется на корпоративной культуре, поддерживающей коллаборацию сотрудников и представителей внешних организаций в структурах, аналогичных социальным сетям [Agostini et al., 2020]. Результат этой деятельности — увеличение интеллектуального капитала, что в итоге повышает способности организации к созданию знаний и их реализации в виде инноваций [Azeem et al., 2021]. Современные технологии и алгоритмы являются как минимум существенным компонентом поддержки такого социального взаимодействия [Zelenkov, 2022], а в перспективе могут стать его полноправными участниками [Manesh et al., 2021; Eloundou et al., 2023].

К наиболее популярным теориям, рассматривающим создание ценности за счет знаний, относятся управление знаниями и ресурсная теория фирмы, включающая также анализ динамических и абсорбционных способностей [Durst, Foli, Edvardsoon, 2024]. На наш взгляд, к данному списку необходимо добавить микроэкономику, поскольку эта дисциплина непосредственно изучает поведение отдельных экономических агентов.

Согласно *теории управления знаниями*, КМ генерирует ценность только тогда, когда знания порождают инновации, т. е. воплощаются в жизнь [Dalkir, 2023]. Этот постулат положен в основу новой версии спиральной модели SECI [Nonaka, Takeuchi, 2019], которая, однако, не учитывает AI в качестве актора в процессе создания знания. Чтобы преодолеть этот разрыв, в [Harfouche et al., 2023] предложена теория, которая вводит новые режимы создания знаний: дополнение AI знаниями экс-

пертов и обогащение знаний сотрудников на основе взаимодействия с AI-моделями. Еще один вариант — использование LLM как интеллектуальных ассистентов (дополненная креативность — *augmented creativity*) [Vinchon et al., 2023]) — на сегодняшний день является самым распространенным.

В связи с этим возникает вопрос о том, насколько много LLM «знают», чтобы выступать в роли ассистентов при решении креативных задач? Способность решать такие задачи в значительной степени зависит от бэкграунда, опыта в конкретной области и общих знаний о реальном мире. Отличительное свойство LLM — обширные знания о реальном мире [Yang et al., 2024b].

Доказательством выдающихся способностей LLM служат результаты теста MMLU (Massive Multitask Language Understanding) [Hendrycks et al., 2021], который включает вопросы по 57 категориям из областей технических, социальных и гуманитарных наук и математики. Для ответа на каждый вопрос необходимо выбрать правильный ответ из четырех предложенных вариантов. Неподготовленные люди, проходившие этот тест, дают в среднем чуть менее 35 % правильных ответов, однако, по оценке авторов публикации, только 5 % экспертов в определенной области (в медицине и других), готовившихся к профессиональным экзаменам, могут достигать точности 89,8 % [Hendrycks et al., 2021, p. 3].

На момент подготовки настоящей статьи результаты лучшей LLM (Google Gemini Ultra с 1,76 трлн параметров) составили 90 %⁷. Важно, что этот результат достигнут по всем дисциплинам, представленным в тесте MMLU, т. е. данная LLM выступает на уровне топ-экспертов во всех областях знаний одновременно.

⁷ Multi-task Language Understanding on MMLU. URL: <https://paperswithcode.com/sota/multi-task-language-understanding-on-mmlu> (дата обращения: 28.06.2024).

Однако важно отметить, что на мультидисциплинарных тестах, построенных на мультимодальных вопросах (включающих помимо текста еще и изображения, графики, формулы, электрические схемы, нотные записи и т. д.), таких как MMMU⁸, LLM уступают экспертам [Yue et al., 2023]. Средняя доля правильных ответов экспертов по одной теме — 88,6 %, лучший результат языковых моделей (Gemini Ultra) — 59,4 %. Тем не менее, если учитывать, что модели, ориентированные на работу с объектами одной модальности (текст, изображения), уже превосходят людей [Maslej et al., 2024.], следует ожидать, что и в мультимодальных приложениях уже в скором будущем будет достигнут серьезный прогресс.

В настоящее время большинство систем, декларирующих мультимодальные способности, на самом деле «под капотом» объединяют несколько моделей, каждая из них отвечает за свой тип данных. LLM, которая сможет интегрировать функции трех-четырёх таких специализированных моделей, должна значительно выигрывать при масштабировании, что положительно отразится на ее способностях к выделению абстракций [Sutton, 2019].

Согласно *ресурсной теории фирмы* эффективность использования AI опирается на способности фирмы выбирать, комбинировать и использовать соответствующие ресурсы — материальные (данные и технологии), человеческие (бизнес и технические компетенции) и нематериальные (координация, способность к изменениям, управление рисками) [Mikalef, Gupta, 2021; Sadiq et al., 2021]. Ряд публикаций эмпирически доказывают, что внедрение систем AI положительно влияет на производительность организации [Mikalef, Gupta, 2021].

Авторы одной из самых цитируемых работ в области систем KM [Alavi, Leidner,

⁸ Massive Multi-discipline Multimodal Understanding and Reasoning. URL: <https://mmlu-benchmark.github.io> (дата обращения: 28.06.2024).

2001] считают, что основной задачей является передача знаний туда, где они нужны и могут быть использованы. Однако это не простой процесс, поскольку организации часто не располагают пониманием того, что они знают, и имеют слабые системы для поиска и извлечения знаний, которые находятся в них. Классическая ИТ-поддержка КМ в основном сводится к созданию систем для неформального обмена (форумов, дискуссионных площадок) и формальных хранилищ (графов или карт знаний, корпоративных каталогов). В более поздней работе этих авторов [Alavi, Leidner, Mousavi, 2024] отмечается, что способность GenAI быстро получать доступ к обширным базам знаний существенно меняет взаимодействие сотрудников с системами КМ и способствует индивидуальному и организационному обучению.

С позиции *микрoэкономики* работники и GenAI представляют собой две разные, но взаимозависимые стороны, которые совместно генерируют сетевой эффект в обучении, способствующий росту производительности [Walkowiak, 2023]. Наибольший эффект достигается в сложных креативных рабочих процессах, требующих учета контекста, где LLM выступает как ассистент [Jarrahi et al., 2023; Ritala, Ruokonen, Ramaul, 2024]. К процессам такого рода можно отнести, например, генерацию текстов, реферирование большого корпуса документов и дополнение ранних фаз инновационной деятельности — исследования, разработки идей и создания цифровых прототипов [Bilgram, Laarmann, 2023]. Так, данные работы [Noy, Zhang, 2023] показывают, что для специалистов с профессиональным образованием (окончивших колледж) время написания текстов (пресс-релизов, отчетов, планов) с использованием ChatGPT-3.5 сократилось на 40 %, при этом качество, оцениваемое тремя независимыми экспертами, возросло на 18 %. Поэтому максимальный эффект от применения GenAI ожидается для высокооплачиваемых работников, решающих нестандартные и уникальные задачи

[Eloundou et al., 2023]. В [Felten et al., 2023] предложен индекс АЮЕ (AI Occupational Exposure), измеряющий степень, в которой определенная профессия подвержена воздействию AI, и представлен анализ более 750 специальностей.

Исследователи из компании Forrester отмечают, что LLM очень быстро могут обеспечить следующие преимущества в КМ [Mohr, Yunus, 2023]: 1) трансформацию сложных данных в удобные для человека представления; 2) самообслуживание конечных пользователей; 3) более высокую автономность работников, использующих знания, поскольку с AI-ассистентом они становятся менее зависимы от коллег и различных информационных систем. Кроме того, некоторые кейсы показывают, что LLM, демократизируя доступ к данным, повышают эффективность менее компетентных сотрудников. В результате их производительность в креативных процессах достигает уровня опытных профессионалов [Dell'Acqua et al., 2023].

В завершение обзора механизмов создания ценности на основе LLM следует отметить, что, поскольку GenAI является новой технологией, ставшей доступной широкому кругу организаций лишь в конце 2022 г., еще нет значительного числа эмпирических данных о том, как он влияет на эффективность фирмы в целом. Один из первых результатов представлен в [Eisfeldt, Schubert, Zhang, 2023]. Авторы показывают, что прибыль компаний, начавших использование ChatGPT сразу после ее выхода в свет в конце 2022 г., увеличилась на 0,4 %. Тем не менее, по оценке McKinsey & Company, максимальный рост прибыли (4 % и более) за счет GenAI ожидается в высокотехнологических отраслях, банках, фармацевтике и образовании, а в других индустриях возможный рост ожидается в пределах 2 % [Chui et al., 2023, p. 25].

Таким образом, LLM демонстрируют значительный положительный эффект, выступая в роли интеллектуального ассистента, поскольку они удерживают

в памяти гораздо больше материала, относящегося к различным областям знаний, чем люди. При этом особенно сильно повышается производительность менее квалифицированных работников, использующих знания. Ограничивающим фактором остается способность LLM обрабатывать мультимодальные корпоративные данные, однако многие авторы прогнозируют, что эти способности будут радикально улучшены в сетях следующего поколения.

Тем не менее вопрос об эмпирическом подтверждении положительного влияния LLM на бизнес в целом и КМ в частности изучен недостаточно. Еще одно актуальное направление исследований связано с тем, что потенциальная роль LLM может не ограничиваться функциями интеллектуального ассистента. Помимо процессного сценария, когда LLM следует инструкциям человека, возможны и другие варианты использования — автономные исследования, коллаборация нескольких агентов и т. д. [Xi et al., 2023].

ПРОЦЕСС ИСПОЛЬЗОВАНИЯ ЗНАНИЙ

Как отмечалось, LLM могут существенно повлиять на КМ, облегчая сотрудникам доступ к знаниям, а также обмен и интеграцию знаний между различными подразделениями и направлениями бизнеса. Многие отраслевые применения базируются на стандартной архитектуре: предобученная LLM имеет доступ к корпоративным хранилищам знаний (неструктурированным документам, KG и т. д.) и используется в качестве интеллектуального ассистента, отвечающего на вопросы. Такие кейсы представлены для сельского хозяйства [Yang et al., 2024c], аэрокосмической промышленности [Zhou et al., 2024], строительства [Zhong, Goodfellow, 2024; Saka et al., 2024] и других отраслей. Также необходимо отметить приложения, в которых LLM решает задачи, более ха-

рактерные для классического AI: кредитный скоринг [Babaei, Giudici, 2024]; обнаружение ошибок при сборке авиационной техники [Peifeng et al., 2024]; предсказание изменения стоимости финансовых активов [Lopez-Lira, Tang, 2023]. Фактически все эти задачи сводятся к бинарной классификации, преимущество LLM заключается в том, что, помимо размеченных данных, эти модели могут использовать и неструктурированные данные.

Эмпирическое исследование влияния LLM на эффективность *креативной деятельности* представлено в [Dell'Acqua et al., 2023]. В исследовании участвовали 758 консультантов Boston Consulting Group, которые были случайно разделены на три группы: 1) контрольную группу без доступа к LLM⁹; 2) с доступом к LLM, но без обучения ее использованию; 3) с доступом к LLM и обучением на обзорном курсе по промпт-инжинирингу. Члены всех групп выполняли задачи, относящиеся к консалтинговой деятельности: разработку идей и концепций новых продуктов; интервьюирование заказчиков; решение бизнес-проблем с помощью количественных данных. В эксперименте оценивались качество выполнения задач, для чего привлекались эксперты компании, а также время их выполнения. Участники эксперимента, применявшие LLM, выполнили на 12 % больше задач, и время выполнения 25 % задач сократилось.

При этом наибольший эффект использование LLM принесло участникам, чья квалификация до начала эксперимента оценивалась ниже среднего, — качество их работы возросло на 43 %. Качество работы консультантов с оценкой выше среднего увеличилось на 17 % [Dell'Acqua et al., 2023]. В результате качество более и менее квалифицированных консультантов практически выровнялось. Однако другие исследователи отмечают, что использование LLM для повышения креа-

⁹ В эксперименте использовалась модель GPT-4 от компании OpenAI.

тивности начинающих сотрудников имеет и негативную сторону. GenAI создает такие риски, как предвзятость и недостаточная социализация человека, что потенциально может привести к снижению контактов между экспертами и молодыми консультантами и маргинализации менее квалифицированных сотрудников [Alavi, Leidner, Mousavi, 2024].

Важный вывод указанной работы состоит также в том, что возможности LLM создают «неровную технологическую границу», где одни задачи легко решаются с помощью GenAI, а другие, хотя и кажутся схожими по уровню сложности, находятся за пределами текущих возможностей эти систем [Dell'Acqua et al., 2023, p. 17].

Целесообразно рассмотреть примеры интеграции LLM в деятельность, которая базируется на интенсивном использовании знаний, такие как: 1) разработка новой продукции; 2) цифровые двойники; 3) управление бизнес-процессами.

Разработка новых продуктов опирается на интенсивное применение трех видов знаний: 1) специфичных для конкретной области; 2) универсальных инженерных и технических; 3) общих [Han et al., 2022]. Существует пять категорий представления явных знаний в контексте инженерного проектирования [Hu et al., 2023]: 1) образительные (эскизы, чертежи); 2) символические (диаграммы, онтологии); 3) лингвистические (требования, вербальные коммуникации); 4) виртуальные (CAD-модели, цифровые двойники); 5) алгоритмические (уравнения, алгоритмы). Эти знания содержатся в мультимодальных технических документах [Jiang et al., 2024], которые могут находиться в различных системах и хранилищах.

Традиционным подходом к управлению такими знаниями является создание семантических сетей [Han et al., 2022], однако появление LLM открывает новые перспективы. Как отмечалось, основное преимущество LLM заключается в способности интегрировать данные из источни-

ков различной природы и представлять их человеку в удобном для восприятия виде. Эта способность может быть использована на различных этапах процесса разработки, таких как анализ интервью с потребителями и выделение сегментов с разными требованиями, определение релевантных параметров, генерация вариантов геометрии и выбор оптимального и т. д. [Brossar et al., 2022; Bilgram, Laarmann, 2023]. Хотя в настоящее время LLM еще уступают людям при обработке мультимодальных запросов, в [Hu et al., 2023] представлен ряд сценариев интеграции таких моделей в процессы управления инженерными знаниями. Отдельное направление исследований — дополнение компетенций сотрудников [Beheshti et al., 2023], например использование LLM для создания прототипов программного кода.

LLM также имеют значительный потенциал в повышении эффективности *цифровых двойников*. Эта технология базируется на взаимодействии физического объекта и его виртуальной копии, что усиливает возможности диагностики, анализа, принятия решений, планирования ресурсов и т. д. Одной из проблем при этом выступает обеспечение соответствия цифрового двойника постоянно изменяющемуся объекту, который он моделирует. Поэтому в [Sun et al., 2024] представлена мультиагентная архитектура, управляемая LLM. Агенты воспринимают динамические характеристики физических систем путем интеграции данных различных модальностей, LLM используется для обобщения этих данных и генерации удобного для человека представления. Другая проблема — проведение симуляций на цифровом двойнике, чтобы предсказать поведение физического объекта. Часто полный перебор всех вариантов воздействия на объект невозможен, поэтому многие критические режимы его функционирования остаются неизвестными. В [Galera-Zarco, Floros, 2023] рассматриваются перспективы интеграции LLM и системы проектирования зданий BIM (Building

Information Modelling) для предсказания последствий экстремальных событий (землетрясений, наводнений и т. д.). Согласно предложенному фреймворку, в BIM создается модель здания, затем проводится ряд симуляций, чтобы оценить повреждения его различных элементов. На полученном наборе данных дообучается LLM, которая обобщает полученные данные и может предсказывать поведение различных элементов здания при экстремальных событиях.

Применение LLM для *управления бизнес-процессами* (business process management — BPM) обсуждается в [Kampik et al., 2023; Fahland et al., 2024]. BPM-специалисты должны обладать высоким уровнем технических, социальных и профессиональных навыков, а также значительным опытом работы в конкретной организации; входной барьер для успешной реализации BPM-инициативы очень высок. Поэтому группа исследователей провозгласила инициативу, направленную на создание новых BPM-систем, дополненных AI [Dumas et al., 2023]. Цель данной инициативы — сделать процессы более адаптируемыми, проактивными, объяснимыми и контекстно-зависимыми. LLM может снизить барьер входа, например, ее можно использовать для трансформации текстовых описаний в процессные модели, непрерывной оптимизации процессов, оценки рисков предполагаемых изменений [Kampik et al., 2023], генерации обоснованных объяснений состояния процесса [Fahland et al., 2024].

Таким образом, LLM имеет значительные перспективы в поддержке креативной деятельности, основанной на интенсивном использовании знаний. Однако пока еще отсутствует понимание того, какие задачи могут, а какие не могут быть решены с помощью GenAI; накопление соответствующего опыта происходит методом проб и ошибок. Также следует учитывать, что, хотя использование LLM направлено на упрощение обмена знаниями, обратной стороной этого процесса может стать снижение контактов между сотрудниками.

Рассмотренные варианты являются лишь первыми попытками адаптации способностей LLM и открывают новые направления для практико-ориентированных решений. При этом следует подчеркнуть, что три рассмотренных сценария интеграции LLM в креативные процессы не исчерпывают все возможные варианты. Особый интерес, например, представляет расширение возможностей систем инженерного проектирования (Computer Aided Design / Computer Aided Manufacturing / Computer Aided Engineering — CAD/CAM/CAE). Для этого прежде всего необходимо добиться эффективной работы с мультимодальными данными. Еще одна актуальная задача — разработка формальных методов идентификации технологической границы, разделяющей задачи, сравнимые по сложности, но решаемые или не решаемые с помощью LLM.

СОЗДАНИЕ ЗНАНИЙ

Создание знаний, т. е. обоснованных убеждений, соответствующих реальному миру, во многом опирается на формирование теорий, которые затем подтверждаются или опровергаются на практике [Dalkir, 2023]. Теории формируются в процессе рассуждений, которые разделяются на дедукцию (вывод по правилам логики от общих положений к частным случаям), индукцию (вывод от частных фактов к общей гипотезе) и абдукцию (максимально правдоподобная интерпретация того, что считается истинным). Выше были приведены примеры того, что LLM успешно справляются с дедукцией и абдукцией. В контексте создания знаний особый интерес представляет способность моделей к индукции.

Индуктивные рассуждения — фундаментальная когнитивная задача получения правдоподобных выводов в условиях неопределенности. При этом заключение выводится из предпосылок не строго через законы логики, а через некоторые представления о связи явлений. Поскольку при

решении большинства проблем используются редкие, зашумленные или неточные данные, большинство из них требуют индуктивных, а не дедуктивных рассуждений [Han et al., 2024].

Несколько публикаций последнего времени представляют эмпирические исследования способностей LLM к индукции свойств (property induction). Это один из методов, который применяется для исследования когнитивных способностей детей, взрослых с различным бэкграундом и т. д. Задача испытуемого — определить, какие общие свойства имеют разные категории объектов. Согласно [Han et al., 2024], ранние LLM с малым числом параметров с трудом повторяют многие аспекты человеческого поведения, но модель последнего поколения GPT-4 гораздо успешнее. Решения GPT-4 в основном качественно совпадают с человеческими, единственным заметным исключением является ее неспособность отразить явление немонотонности предпосылок.

В ряде публикаций проблема индуктивного вывода рассматривается в более широкой постановке. Авторы отмечают, что LLM хорошо работают при решении простых индуктивных задач, но в более сложных случаях качество продуцируемых рассуждений снижается [Chen, Saetre, Miya, 2024]. Однако включение в процесс отбора гипотез человека позволяет повысить производительность [Wang et al., 2023]. Повышение качества рассуждений в LLM — активно развивающаяся область исследований. В этой связи представляют интерес такие темы, как оценка надежности рассуждений, самокоррекция [Huang et al., 2023], новые методы промпт-инжиниринга [Yao et al., 2023] и т. д.

С практической точки зрения способы применения LLM для создания знаний, описываемые в литературе, можно разделить на несколько классов (рис. 1). Однако необходимо отметить, что эта классификация составлена с учетом того, что технология LLM только осваивается исследователями и практиками, и в ближайшем бу-

дущем стоит ожидать появления новых вариантов ее использования.

Первый сценарий наиболее очевиден и сводится к *отбору источников знаний* для последующего анализа, например для подготовки обзоров литературы по какой-либо теме [Taylor et al., 2022; Platt, Platt, 2023]. Здесь LLM выступает как интеллектуальный ассистент и фактически не участвует в создании нового знания, используются ее способности к поиску с учетом контекста, что гораздо эффективнее поиска по ключевым словам.

Второй сценарий имеет дело с *проблемами, поддающимися алгоритмизации*, и сводится к поиску в функциональном пространстве решений, удовлетворяющих определенным условиям. Их можно разделить на два направления, различающиеся способом представления входных данных.

В рамках первого известна некая модель функции, связывающей входные параметры, операции над ними и ожидаемый результат. Проблема заключается в том, что пространство решений часто имеет очень высокую размерность и его невозможно исследовать полностью. В качестве примера можно привести химические элементы, уравнения реакции и вещество с определенными параметрами на выходе. В этом случае модель дообучается на известных примерах (набор элементов, химическая реакция и результат), а затем используется для поиска решений, обеспечивающих необходимые параметры вещества на выходе. Такие приложения характерны для поиска новых материалов [Jablonka et al., 2023], лекарств [Sarkar et al., 2023] и в целом для химических задач [Guo et al., 2023]. Однако данный подход может применяться не только в химии или материаловедении — рассмотренная проблема поиска экстремальных режимов с помощью цифрового двойника [Galera-Zarco, Floros, 2023] также может быть отнесена к указанному направлению.

Второе направление в рамках алгоритмируемых проблем связано с поиском

эвристик для случая, когда вид функциональной связи неизвестен. К этому типу относятся большинство комбинаторных задач, например задача об упаковке рюкзака. Такую задачу можно рассматривать как проблему оптимизации, где качество эвристики оценивается внешним по отношению к модели агентом [Yang et al., 2023]. В данном случае LLM используется как генератор эвристик. Например, она создает программный код, решающий заданную проблему, внешний агент оценивает качество решения. Оценка агента возвращается модели, на ее основании LLM улучшает сгенерированный код. Подобный подход в сочетании с эволюционными процедурами использован в [Romera-Paredes et al., 2023], где авторам удалось получить новые более эффективные решения для нескольких хорошо известных комбинаторных проблем.

Третий перспективный сценарий использования LLM для генерации знаний предложен в [Horton, 2023]. В этой работе LLM используется как *заменитель человека в серии экспериментов*, повторяющих классические работы поведенческой экономики. Автор работы показал, что агент на основе LLM повторяет результаты, полученные с участием реальных людей. Этот результат не только открывает новые перспективы в науках, изучающих поведение человека, поскольку снимает многие этические ограничения и значительно удешевляет эксперименты, но и предлагает новые пути к обучению предметно специализированных моделей.

Так, в [Li et al., 2024] представлен виртуальный госпиталь, в котором больные и сотрудники госпиталя (доктора различных специальностей, медсестры и т. д.) являются агентами на основе LLM. Агент, имитирующий больного, генерирует набор симптомов, которые передаются в промпт агента-доктора. Агент-доктор ставит диагноз и отвечает на другие вопросы, затем его вывод валидируется и, если он корректен, сохраняется в памяти. Таким образом, доктор накапливает знания. Кроме того,

агенты-доктора имеют доступ к корпусу медицинских текстов, по которым они сами генерируют вопросы и ответы, что тоже служит для накопления знаний.

Рассмотренные способы генерации знаний при помощи LLM, скорее всего, не исчерпывают весь спектр возможностей, что открывает перспективы для новых исследований. Важно подчеркнуть, что индуктивные способности известных моделей пока ограничены. Указанные сценарии генерации знаний основаны на использовании главного на сегодняшний день преимущества LLM — трансформации сложных данных в удобную для восприятия форму. Следующий прорыв, очевидно, должен быть связан с увеличением способностей LLM по обобщению частных фактов и выделению абстракций. Поиск в функциональном пространстве — актуальная задача, но гораздо важнее идеи, на основе которых такие задачи возникают [Castelvecchi, 2023].

ЗАКЛЮЧЕНИЕ

В ходе исследования проведен анализ 75 публикаций, относящихся к теме применения больших языковых моделей в управлении знаниями организации. Предложена классификация исследований (рис. 1), рассмотрены новые направления изучения данной проблематики.

На основании представленного анализа можно сформулировать следующие основные выводы.

1. LLM обещают и уже демонстрируют значительный положительный эффект в процессах КМ. Эти модели удерживают в памяти гораздо больше материала, относящегося к различным областям знаний, чем люди. Основным сценарием является использование LLM как интеллектуального ассистента. Однако при этом возникает «технологическая граница», разделяющая задачи, легко решаемые с помощью LLM, и задачи, схожие по уровню сложности, но находящиеся за преде-

лами текущих возможностей [Dell'Acqua et al., 2023].

2. Наибольший эффект проявляется при поддержке менее квалифицированных сотрудников, их производительность в креативных процессах возрастает практически до уровня экспертов [Dell'Acqua et al., 2023].

3. LLM повышает автономность сотрудников, использующих знания [Mohr, Yunus, 2023], но также создает такие риски, как предвзятость и недостаточная социализация, что может привести к маргинализации менее квалифицированных сотрудников [Alavi, Leidner, Mousavi, 2024].

4. С позиции использования знаний ограничивающим фактором сейчас является способность LLM обрабатывать мультимодальные корпоративные данные [Yue et al., 2023], однако многие исследователи прогнозируют, что эти способности будут радикально повышены в сетях следующего поколения¹⁰. Кроме того, предполагается, что такие модели должны проявлять еще больше эмерджентных свойств при масштабировании.

5. Создание знаний пока ограничивается способностями LLM к индуктивному выводу [Chen, Saetre, Miyao, 2024]. Эти ограничения частично преодолеваются включением человека в процесс оценки и отбора гипотез, но необходимы дальнейшие исследования прежде всего в направлении повышения способностей LLM к самооценке и самокоррекции [Huang et al., 2023].

6. Входной барьер для реализации инициатив по использованию LLM пока остается достаточно высоким. Базовые сценарии сводятся либо к использованию проприетарной модели как внешнего сервиса, либо к адаптации модели с открытыми весами (предельный случай второго сценария — обучение собственной модели с нуля). Первый вариант требует передачи корпоративных данных третьей стороне, что связано с рисками информационной безопасности¹¹. Кроме того, доступ к таким сервисам закрыт для российских компаний. Второй вариант предполагает создание собственных компетенций в AI, что связано с большими затратами, и модели с открытыми весами уступают по производительности ведущим проприетарным LLM.

Потенциальные ограничения данной работы связаны с большой скоростью развития технологии GenAI, публикации второй половины 2024 г. не попали в этот обзор. Тем не менее радикальных прорывов ни в одном из выделенных направлений исследований пока не наблюдается.

В заключение отметим, что за пределами данного обзора остались многие аспекты, связанные с использованием LLM в бизнесе, в частности безопасность, этика, доверие и т. д. Эти вопросы специфичны как для LLM / КМ, так и для применения AI в целом, причем не только в корпоративных приложениях. Поэтому данные направления требуют отдельных исследований.

¹⁰ 13 мая 2024 г. компания OpenAI представила модель GPT-4o, которая способна обрабатывать мультимодальные данные. Ее результат на тесте MMMU — 69,1 %.

¹¹ 10 июня 2024 г. компания Apple анонсировала подход, интегрированный в персональную систему Apple Intelligence, который решает проблему защиты пользовательских данных.

ЛИТЕРАТУРА НА РУССКОМ ЯЗЫКЕ

Долотов А. 2024. Сложности интеграции: что поможет ускорить GPT-трансформацию бизнеса. *Forbes*. [Электронный ресурс]. <https://www.forbes.ru/mneniya/507501-sloznosti-integracii-cto-pomozet-uskorit-gpt-transformaciu-biznesa> (дата обращения: 05.04.2024).

Яков и Партнеры. 2023. *Искусственный интеллект в России — 2023: тренды и перспективы*. [Электронный ресурс]. <https://yakovpartners.ru/publications/ai-future/> (дата обращения: 05.04.2024).

REFERENCES IN LATIN ALPHABET

- Agarwal R., Singh A., Zhang Lei M., Bohnet B., Rosias L., Chan S., Zhang B., Anand A., Abbas Z., Nova A., Co-Reyes J.D., Chu E., Behbahani F., Faust A., Larochelle H. 2024. Many-shot in-context learning. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2404.11018>
- Agostini, L., Nosella, A., Sarala, R., Spender, J. C., Wegner, D. 2020. Tracing the evolution of the literature on knowledge management in inter-organizational contexts: a bibliometric analysis. *Journal of Knowledge Management* **24** (2): 463–490. <http://dx.doi.org/10.1108/JKM-07-2019-0382>
- Alavi M., Leidner D. 2001. Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly* **25** (1): 107–136. <https://doi.org/10.2307/3250961>
- Alavi M., Leidner D., Mousavi R. 2024. A knowledge management perspective of generative artificial intelligence. *Journal of the Association for Information Systems* **25** (1): 1–12. <https://doi.org/10.17705/1jais.00859>
- Anderson G., Dunlap K., Rademacher J., McMenamy N. 2009. Metaplan: A mutual methodology for interdisciplinary research. *SoTL Commons Conference*: 104. [Electronic resource]. <https://digitalcommons.georgia-southern.edu/sotlcommons/SoTL/2009/104> (accessed: 28.06.2024).
- Asudani D.S., Nagwani N.K., Singh P. 2023. Impact of word embedding models on text analytics in deep learning environment: A review. *Artificial Intelligence Review* **56** (9): 10345–10425. <https://doi.org/10.1007/s10462-023-10419-1>
- Azeem, M., Ahmed, M., Haider, S., Sajjad, M. 2021. Expanding competitive advantage through organizational culture, knowledge sharing and organizational innovation. *Technology in Society* **66**: 101635. <https://doi.org/10.1016/j.techsoc.2021.101635>
- Babaei G., Giudici P. 2024. GPT classifications, with application to credit lending. *Machine Learning with Applications* **16**: 100534. <https://doi.org/10.1016/j.mlwa.2024.100534>
- Beheshti A., Yang J., Sheng Q., Benatallah B., Casati F., Dustdar Sh., Nezhad H.R., Zhang X., Xue Sh. 2023. ProcessGPT: Transforming business process management with generative artificial intelligence. In: *2023 IEEE International Conference on Web Services (ICWS)*; 731–739. <https://doi.org/10.1109/ICWS60048.2023.00099>
- Berg J.M., Raj M., Seamans R. 2023. Capturing value from artificial intelligence. *Academy of Management Discoveries* **9** (4): 424–428. <https://doi.org/10.5465/amd.2023.0106>
- Bilgram V., Laarmann F. 2023. Accelerating innovation with generative AI: AI-augmented digital prototyping and innovation methods. *IEEE Engineering Management Review* **51** (2): 18–25. <https://doi.org/10.1109/EMR.2023.3272799>
- Brossar M., Corbo J., Klaeyle M., Wiseman B. 2022. Deep learning in product design. *McKinsey & Company*. [Electronic resource]. <https://www.mckinsey.com/capabilities/operations/our-insights/deep-learning-in-product-design> (accessed: 02.05.2024).

- Brown T., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., Neelakantan A., Shyam P., Sastry G., Askell A., Agarwal S., Herbert-Voss A., Krueger G., Henighan T., Child R., Ramesh A., Ziegler D., Wu J., Winter C., Hesse C., Chen M., Sigler E., Litwin M., Gray S., Chess B., Clark J., Berner C., McCandlish S., Radford A., Sutskever I., Amodei D. 2020. Language models are few-shot learners. In: *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*. [Electronic resource]. https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf (accessed: 09.04.2024).
- Carrera-Rivera A., Ochoa W., Larrinaga F., Lasa G. 2022. How-to conduct a systematic literature review: A quick guide for computer science research. *MethodsX* **9**: 101895. <https://doi.org/10.1016/j.mex.2022.101895>
- Castelvecchi D. 2023. How will AI change mathematics? Rise of chatbots highlights discussion. *Nature* **615**: 15–16. [Electronic resource]. <https://www.nature.com/articles/d41586-023-00487-2> (accessed: 28.06.2024).
- Chang Y., Wang, X., Wang, J., Wu Y., Yang L. et al. 2024. A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology* **15** (3): 1–45. <https://doi.org/10.1145/3641289>
- Chen B., Sætre R., Miyao Y. 2024. A comprehensive evaluation of inductive reasoning capabilities and problem solving in large language models. In: *Findings of the Association for Computational Linguistics: EACL 2024*; 323–339. [Electronic resource]. <https://aclanthology.org/2024.findings-eacl.22> (accessed: 29.04.2024).
- Chui M., Hazan E., Roberts R., Singla A., Smaje K. et al. 2023. The economic potential of generative AI. The next productivity frontier. *McKinsey & Company*. [Electronic resource]. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-AI-the-next-productivity-frontier#/> (accessed: 23.04.2024).
- Dalkir K. 2023. *Knowledge Management in Theory and Practice*. 4th ed. The MIT Press.
- Dell'Acqua F., McFowland E., Mollick E., Lifshitz-Assaf H., Kellogg K. et al. 2023. Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 24-013*. [Electronic resource]. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4573321 (accessed: 22.04.2024).
- Dumas M., Fournier F., Limonad L., Marrella A., Montali M. et al. 2023 AI-augmented business process management systems: A research manifesto. *ACM Transactions on Management Information Systems* **14** (1): 1–19. <https://doi.org/10.1145/3576047>
- Durst S., Foli S., Edvardsson I.R. 2024. A systematic literature review on knowledge management in SMEs: Current trends and future directions. *Management Review Quarterly* **74**: 263–288. <https://doi.org/10.1007/s11301-022-00299-0>
- Eisfeldt A., Shubert G., Zhang M. 2023. Generative AI and firm value. *NBER Working Paper No. 31222*. [Electronic resource]. <http://www.nber.org/papers/w31222> (accessed: 22.04.2024).
- Eloundou T., Manning S., Mishkin P., Rock D. 2023. GPTs are GPTs: An early look at the labor market impact potential of large language models. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2303.10130>
- Fahland D., Fournier F., Limonad L., Skarbovsky I., Swevels A.J. 2024. How well can large language models explain business processes? *arXiv preprint*. <https://doi.org/10.48550/arXiv.2401.12846>
- Felten E. W., Raj M., Seamans R. 2023. Occupational heterogeneity in exposure to generative AI. *SSRN*: 4414065. <https://dx.doi.org/10.2139/ssrn.4414065>
- Galera-Zarco C., Floros G. 2023. A deep learning approach to improve built asset operations and disaster management in critical events: An integrative simulation model for quicker decision making. *Annals of Operations Research* (Early Access). <https://doi.org/10.1007/s10479-023-05247-z>
- Gao Y., Xiong Y., Gao X., Jia K., Pan J. et al. 2023. Retrieval-augmented generation for large language models: A survey. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2312.10997>

- Gaviria-Marin, M., Merigó, J.M., Baier-Fuentes, H. 2019. Knowledge management: a global examination based on bibliometric analysis. *Technological Forecasting and Social Change* **140**: 194–220. <https://doi.org/10.1016/j.techfore.2018.07.006>
- Grosse R., Bae J., Anil C., Elhage N., Tamkin A. et al. 2023. Studying large language model generalization with influence functions. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2308.03296>
- Guo T., Nan B., Liang Z., Guo Z., Chawla N. et al. 2023. What can large language models do in chemistry? A comprehensive benchmark on eight tasks. In: *Advances in Neural Information Processing Systems (NeurIPS 2023)*; 59662–59688. [Electronic resource]. https://proceedings.neurips.cc/paper_files/paper/2023/hash/bbb330189ce02be00cf7346167028ab1-Abstract-Datasets_and_Benchmarks.html (accessed: 29.04.2024).
- Han J., Sarica S., Shi F., Luo J. 2022. Semantic networks for engineering design: State of the art and future directions. *Journal of Mechanical Design* **144** (2): 020802. <https://doi.org/10.1115/1.4052148>
- Han S.J., Ransom K.J., Perfors A., Kemp C. 2024. Inductive reasoning in humans and large language models. *Cognitive Systems Research* **83**: 101155. <https://doi.org/10.1016/j.cogsys.2023.101155>
- Harfouche A., Quinio B., Saba M., Saba P. 2023. The recursive theory of knowledge augmentation: Integrating human intuition and knowledge in artificial intelligence to augment organizational knowledge. *Information Systems Frontiers* **25**: 55–70. <https://doi.org/10.1007/s10796-022-10352-8>
- Harrer S. 2023. Attention is not all you need: The complicated case of ethically using large language models in healthcare and medicine. *eBioMedicine* **90**: 104512. <https://doi.org/10.1016/j.ebiom.2023.104512>
- Hendrycks D., Burns C., Basart S., Zou A., Mazeika M. et al. 2021. Measuring massive multitask language understanding. In: *Proceedings of the International Conference on Learning Representations*. <https://doi.org/10.48550/arXiv.2009.03300>
- Hogan A., Gutierrez C., Cochez M., de Melo G., Kirrane S. et al. 2022. *Knowledge Graphs*. Springer. <https://doi.org/10.1007/978-3-031-01918-0>
- Horton J. 2023. Large language models as simulated economic agents: What can we learn from homo silicus? *NBER Working Paper No. 31122*. [Electronic resource]: <http://www.nber.org/papers/w31122> (accessed: 18.04.2024).
- Hu E., Shen Y., Wallis P., Allen-Zhu Z., Li Y. et al. 2021. LORA: Low-rank adaptation of large language models. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2106.09685>
- Hu X., Tian Y., Nagato K., Nakao M., Liu A. 2023. Opportunities and challenges of ChatGPT for design knowledge management. *Procedia CIRP* **119**: 21–28. <https://doi.org/10.1016/j.procir.2023.05.001>
- Huang J., Chen X., Mishra S., Zheng H.S., Yu A.W. et al. 2023. Large language models cannot self-correct reasoning yet. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2310.01798>
- Inkinen H. 2016. Review of empirical research on knowledge management practices and firm performance. *Journal of Knowledge Management* **20** (2): 230–257. <https://doi.org/10.1108/JKM-09-2015-0336>
- Jablonka K., Ai Q., Al-Feghali A., Badhwar S., Bocarsly J. et al. 2023. 14 examples of how LLMs can transform materials science and chemistry: A reflection on a large language model hackathon. *Digital Discovery* **2** (5): 1233–1250. <https://doi.org/10.1039/D3DD00113J>
- Jarrahi M., Askay D., Eshraghi A., Smith P. 2023. Artificial intelligence and knowledge management: A partnership between human and AI. *Business Horizons* **66** (1): 87–99. <https://doi.org/10.1016/j.bushor.2022.03.002>
- Jiang S., Hu J., Magee C.L., Luo J. 2024. Deep learning for technical document classification. *IEEE Transactions on Engineering Management* **71**: 1163–1179. <https://doi.org/10.1109/TEM.2022.3152216>
- Jiang Z., Xu F.F., Araki J., Neubig G. 2020. How can we know what language models know? *Transactions of the Association for*

- Computational Linguistics* **8**: 423–438. https://doi.org/10.1162/tacl_a_00324
- Kampik T., Warmuth C., Rebmann A., Agam R., Egger L. et al. 2023. Large process models: business process management in the age of generative AI. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2309.00900>
- Kandpal N., Deng H., Roberts A., Wallace E., Raffel C. 2023. Large language models struggle to learn long-tail knowledge. In: *Proceedings of the 40th International Conference on Machine Learning*, PMLR **202**: 15696–15707. [Electronic resource]. <https://proceedings.mlr.press/v202/kandpal23a.html> (accessed: 09.04.2024).
- Lan Y., He G., Jiang J., Jiang J., Zhao W.X., Wen J.R. 2022. Complex knowledge base question answering: A survey. *IEEE Transactions on Knowledge and Data Engineering* **35** (11): 11196–11215. <https://doi.org/10.1109/TKDE.2022.3223858>
- Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V. et al. 2020. Retrieval-augmented generation for knowledge-intensive NLP tasks. In: *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*. <https://doi.org/10.48550/arXiv.2005.11401>
- Li C., Wang J., Zhang Y., Zhu K., Hou W. et al. 2023. Large language models understand and can be enhanced by emotional stimuli. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2307.11760>
- Li J., Wang S., Zhang M., Li W., Lai Yu. et al. 2024. Agent hospital: A simulacrum of hospital with evolvable medical agents. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2405.02957>
- Liga D., Robaldo L. 2023. Fine-tuning GPT-3 for legal rule classification. *Computer Law & Security Review* **51**: 105864. <https://doi.org/10.1016/j.clsr.2023.105864>
- Ling C., Zhao X., Lu J., Deng C., Zheng C. et al. 2023. Domain specialization as the key to make large language models disruptive: A comprehensive survey. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2305.18703>
- Lipsey R.G., Carlaw K.I., Bekar C.T. 2005. *Economic Transformations: General Purpose Technologies and Long-Term Economic Growth*. Oxford University Press.
- Lopez-Lira A., Tang Y. 2023. Can ChatGPT forecast stock price movements? Return predictability and large language models. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2304.07619>
- Manesh M., Pelligrini M., Marzi G., Dabic M. 2021. Knowledge management in the fourth industrial revolution: Mapping the literature and scoping future avenues. *IEEE Transactions on Engineering Management* **68** (1): 289–300. <https://doi.org/10.1109/TEM.2019.2963489>
- Maslej N., Fattorini L., Perrault R., Parli V., Reuel A. et al. 2024. *The AI index 2024 annual report*, Institute for Human-Centered AI, Stanford University. [Electronic resource]. <https://aiindex.stanford.edu/report/> (accessed: 23.04.2024).
- Mikalef P., Gupta M. 2021. Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Information & Management* **58** (3): 103434. <https://doi.org/10.1016/j.im.2021.103434>
- Mohr J., Yunus A. 2023. Generative AI: What it means for knowledge management. Forrester Report. [Electronic resource]. <https://www.forrester.com/report/generative-ai-what-it-means-for-knowledge-management/RES179524> (accessed: 19.04.2024).
- Nonaka I., Takeuchi H. 2019. *The Wise Company: How Companies Create Continuous Innovation*. Oxford University Press.
- Noy S., Zhang W. 2023. Experimental evidence on the productivity effects of generative artificial intelligence. *Science* **381** (6654): 187–192. <https://doi.org/10.1126/science.adh2586>
- O’Leary D. E. 2024. The rise and design of enterprise large language models. *IEEE Intelligent Systems* **39** (1): 60–63. <https://doi.org/10.1109/MIS.2023.3345591>
- Pan S., Luo L., Wang Y., Chen C., Wang J., Wu X. 2024. Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering* **36** (7): 3580–3599. <https://doi.org/10.1109/TKDE.2024.3352100>
- Peifeng L.I. U., Qian L., Zhao X., Tao B. 2024. Joint Knowledge Graph and Large Language

- Model for Fault Diagnosis and Its Application in Aviation Assembly. *IEEE Transactions on Industrial Informatics* **20** (6): 8160–8169. <https://doi.org/10.1109/TII.2024.3366977>
- Platt M., Platt D. 2023. Effectiveness of generative artificial intelligence for scientific content analysis. In: *IEEE 17th International Conference on Application of Information and Communication Technologies (AICT)*; 1–4. <https://doi.org/10.1109/AICT59525.2023.10313167>
- Radford A., Narasimhan K., Salimans T., Sutskever I. 2018. *Improving language understanding by generative pre-training*. [Electronic resource]. <https://www.mikecaptain.com/resources/pdf/GPT-1.pdf> (accessed: 09.04.2024).
- Rethlefsen M.L., Kirtley Sh., Waffenschmidt S., Ayala A.P., Moher D. et al. 2021. PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews* **10**: 39. <https://doi.org/10.1186/s13643-020-01542-z>
- Ritala P., Ruokonen M., Ramaul L. 2024. Transforming boundaries: How does ChatGPT change knowledge work? *Journal of Business Strategy* **45** (3): 214–220. <https://doi.org/10.1108/JBS-05-2023-0094>
- Roberts A., Raffel C., Shazeer N. 2020. How much knowledge can you pack into the parameters of a language model? *arXiv preprint*. <https://doi.org/10.48550/arXiv.2002.08910>
- Romera-Paredes B., Barekatin M., Novikov A., Balog M., Kumar M.P. et al. 2024. Mathematical discoveries from program search with large language models. *Nature* **625** (7995): 468–475. <https://doi.org/10.1038/s41586-023-06924-6>
- Sadiq R.B., Safie N., Abd Rahman A.H., Goudarzi S. 2021. Artificial intelligence maturity model: A systematic literature review. *PeerJ Computer Science* **7**: e661. <https://doi.org/10.7717/peerj-cs.661>
- Sahoo P., Singh A., Saha S., Jain V., Mondal S., Chadha A. 2024. A systematic survey of prompt engineering in large language models: Techniques and applications. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2402.07927>
- Saka A., Taiwo R., Saka N., Salami B.A., Ajayi S. et al. 2024. GPT models in construction industry: Opportunities, limitations, and a use case validation. *Developments in the Built Environment* **17**: 100300. <https://doi.org/10.1016/j.dibe.2023.100300>
- Sarkar C., Das B., Rawat V.S., Wahlang J.B., Nongpiur A. et al. 2023. Artificial intelligence and machine learning technology driven modern drug discovery and development. *International Journal of Molecular Sciences* **24** (3): 2026. <https://doi.org/10.3390/ijms24032026>
- Sauer P., Seuring S. 2023. How to conduct systematic literature reviews in management research: A guide in 6 steps and 14 decisions. *Review of Managerial Science* **17**: 1899–1933. <https://doi.org/10.1007/s11846-023-00668-3>
- Schaeffer R., Miranda B., Koyejo S. 2023. Are emergent abilities of large language models a mirage? *Advances in Neural Information Processing Systems (NeurIPS 2023)* **36**. [Electronic resource]. https://proceedings.neurips.cc/paper_files/paper/2023/hash/adc98a266f45005c403b8311ca7e8bd7-Abstract-Conference.html (accessed: 18.04.2024).
- Shanahan M. 2024. Talking about large language models. *Communications of the ACM* **67** (2): 68–79. <https://doi.org/10.1145/3624724>
- Snyder H. 2019. Literature review as a research methodology: An overview and guidelines. *Journal of Business Research* **104**: 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Sun K., Xu Y.E., Zha H., Liu Y., Dong X.L. 2023. Head-to-tail: How knowledgeable are large language models (LLM)? AKA will LLMs replace knowledge graphs? *arXiv preprint*. <https://doi.org/10.48550/arXiv.2308.10168>
- Sun Y., Zhang Q., Bao J., Lu Y., Liu S. 2024. Empowering digital twins with large language models for global temporal feature learning. *Journal of Manufacturing Systems* **74**: 83–99. <https://doi.org/10.1016/j.jmsy.2024.02.015>
- Sutton R. 2019. The Bitter Lesson. *Incomplete Ideas (blog)*. [Electronic resource]. <http://>

- www.incompleteideas.net/IncIdeas/Bitter-Lesson.html (accessed: 11.05.2024).
- Taylor R., Kardas M., Cucurull G., Scialom T., Hartshorn A. et al. 2022. Galactica: A large language model for science. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2211.09085>
- Teubner T., Flath C.M., Weinhardt C., van der Aalst W., Hinz O. 2023. Welcome to the era of ChatGPT et al. The prospects of large language models. *Business & Information Systems Engineering* **65** (2): 95–101. <https://doi.org/10.1007/s12599-023-00795-x>
- Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L. et al. 2017. Attention is all you need. In: *31st Conference on Neural Information Processing Systems (NIPS 2017)*. [Electronic resource]. https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf (accessed: 09.04.2024).
- Vinchon F., Bartolotta S., Gironnay V., Botella M., Bourgeois-Bougrine S. et al. 2023. Artificial Intelligence & Creativity: A manifesto for collaboration. *Journal of Creative Behavior* **57** (4): 472–484. <https://doi.org/10.1002/jocb.597>
- Walkowiak E. 2023. Task-interdependencies between generative AI and workers. *Economic Letters* **231**: 111315. <https://doi.org/10.1016/j.econlet.2023.111315>
- Wang C., Liu X., Song D. 2020. Language models are open knowledge graphs. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2010.11967>
- Wang K., Variengien A., Conmy A., Shlegeris B., Steinhardt J. 2022. Interpretability in the wild: A circuit for indirect object identification in GPT-2 small. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2211.00593>
- Wang R., Zelikman E., Poesia G., Pu Y., Haber N., Goodman N.D. 2023. Hypothesis search: Inductive reasoning with language models. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2309.05660>
- Wang S., Xu S. 2024. 16 Changes to the way enterprises are building and buying generative AI. *a16*. [Electronic resource]. <https://a16z.com/generative-ai-enterprise-2024/> (accessed: 05.04.2024).
- Weidinger L., Uesato J., Rauh M., Griffin C., Huang P.-S. et al. 2022. Taxonomy of risks posed by language models. In: *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT'22)*; 214–229. <https://doi.org/10.1145/3531146.3533088>
- Wu S., Irsoy O., Dabravolski V., Dredze M., Gehrmann S. 2023. BloombergGPT: A large language model for finance. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2303.17564>
- Xi Z., Chen W., Guo X., He W., Ding Y. et al. 2023. The rise and potential of large language model based agents: A survey. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2309.07864>
- Yang C., Wang X., Lu Y., Liu H., Le Q.V., Zhou D., Chen X. 2023. Large language models as optimizers. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2309.03409>
- Yang L., Chen H., Li Z., Ding X., Wu X. 2024a. Give us the facts: Enhancing large language models with knowledge graphs for fact-aware language modeling. *IEEE Transactions on Knowledge and Data Engineering*. Early Access. <https://doi.org/10.1109/TKDE.2024.3360454>
- Yang J., Jin H., Tang R., Han X., Feng Q. et al. 2024b. Harnessing the power of LLMs in practice: A survey on ChatGPT and beyond. *ACM Transactions on Knowledge Discovery from Data*. <https://doi.org/10.1145/3649506>
- Yang T., Mei Y., Xu L., Yu H., Chen Y. 2024c. Application of question answering systems for intelligent agriculture production and sustainable management: A review. *Resources, Conservation and Recycling* **204**: 107497. <https://doi.org/10.1016/j.resconrec.2024.107497>
- Yao S., Yu D., Zhao J., Shafran I., Griffiths T. et al. 2023. Tree of thoughts: Deliberate problem solving with large language models. In: *Advances in Neural Information Processing Systems* **36 (NeurIPS 2023)**. [Electronic resource]. https://proceedings.neurips.cc/paper_files/paper/2023/hash/271db9922b8d1f4dd7aaef84ed5ac703-Abstract-Conference.html (accessed: 02.05.2024).

- Yu D., Zhu C., Yang Y., Zeng M. 2022a. JAKET: Joint pre-training of knowledge graph and language understanding. In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI-22)* **36** (10): 11630–11638. <https://doi.org/10.1609/aaai.v36i10.21417>
- Yu W., Zhu C., Li Z., Hu Z., Wang Q. et al. 2022b. A survey of knowledge-enhanced text generation. *ACM Computing Surveys* **54** (11s): 1–38. <https://doi.org/10.1145/3512467>
- Yue X., Ni Y., Zhang K., Zheng T., Liu R. et al. 2023. MMMU: A massive multi-discipline multimodal understanding and reasoning benchmark for expert AGI. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2311.16502>
- Zelenkov Y. 2018. The effectiveness of Russian organizations: The role of knowledge management and change readiness. *Russian Management Journal* **16** (4): 513–536. <https://doi.org/10.21638/spbu18.2018.403>
- Zelenkov Y. 2022. Explaining the IT value through the information support of decision-making. In: *Digitalization of Society, Economics and Management: A Digital Strategy Based on Post-pandemic Developments*. Springer; 29–48. https://doi.org/10.1007/978-3-030-94252-6_3
- Zhong Y., Goodfellow S.D. 2024. Domain-specific language models pre-trained on construction management systems corpora. *Automation in Construction* **160**: 105316. <https://doi.org/10.1016/j.autcon.2024.105316>
- Zhou B., Li X., Liu T., Xu K., Liu W., Bao J. 2024. CausalKGPT: Industrial structure causal knowledge-enhanced large language model for cause analysis of quality problems in aerospace product manufacturing. *Advanced Engineering Informatics* **59**: 102333. <https://doi.org/10.1016/j.aei.2023.102333>

TRANSLATION OF REFERENCES IN RUSSIAN INTO ENGLISH

- Dolotov A. 2024. Integration difficulties: What will help accelerate GPT business transformation. *Forbes*. [Electronic resource]. <https://www.forbes.ru/mneniya/507501-sloznosti-integracii-cto-pomozet-uskorit-gpt-transformaciu-biznesa> (accessed: 05.04.2024). (In Russian)
- Yakov and Partners. 2023. *Artificial intelligence in Russia — 2023: Trends and prospects*. [Electronic resource]. <https://yakovpartners.ru/publications/ai-future/> (accessed: 05.04.2024). (In Russian)

*Статья поступила в редакцию
4 мая 2024 г.
Принята к публикации
27 июля 2024 г.*

Knowledge management in organization and the large language models

Yu. A. Zelenkov

Graduate School of Business, HSE University, Moscow, Russia

Purpose: to summarize, classify and analyze current academic papers on the use of large language models (LLM) in knowledge management in organization. **Methodology:** systematic literature review was conducted. It was based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework. 75 papers were selected for the analysis, including academic papers and reports of consulting companies published since 2020. **Findings:**

four main research areas have been identified: (1) LLM implementation issues; (2) the impact of LLM on knowledge management efficiency; the application of LLM in the processes of (3) knowledge usage and (4) knowledge creation. Within each area, the key papers and open questions have been reviewed. **Originality and contribution:** the paper presents a systematic review of current publications, proposes a classification of research topics, and identifies potential directions for new research. The study also considers limitations hindering the implementation of LLM in the organization's knowledge management practice.

Keywords: generative AI, large language models, knowledge management, systematic literature review.

For citation: Zelenkov Yu. A. 2024. Knowledge management in organization and the large language models. *Russian Management Journal* **22** (3): 573–601.

<https://doi.org/10.21638/spbu18.2024.309> (In Russian)

Для цитирования: Зеленков Ю.А. 2024. Управление знаниями организации и большие языковые модели. *Российский журнал менеджмента* **22** (3): 573–601.

<https://doi.org/10.21638/spbu18.2024.309>

Initial Submission: May 4, 2024

Final Version Accepted: July 27, 2024