ISSN 1999-6640 (print)
ISSN 1999-6659 (online)

http://umj.ru

DOI 10.15826/umpa.2025.01.007

ВНЕДРЕНИЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОБРАЗОВАТЕЛЬНЫЙ ПРОЦЕСС: УПРАВЛЕНЧЕСКИЕ ВЫЗОВЫ

Т. А. Орешкина, А. Ю. Долганов, Е. А. Маяцкая, О. Ю. Артюгин

Уральский федеральный университет им. первого Президента России Б. Н. Ельцина
Россия, 620002, Екатеринбург, ул. Мира, 19;
t.a.oreshkina@urfu.ru

Аннотация. В статье рассмотрены фундаментальные междисциплинарные вопросы и ключевые управленческие вызовы, возникающие в ситуации необходимости принятия решений о легализации и внедрении технологий и сервисов искусственного интеллекта в образовательный процесс. Цель исследования – оценить возможные эффекты, преимущества и риски внедрения технологий искусственного интеллекта на основе больших языковых моделей в образовательный процесс (на уровне учебной дисциплины). Оригинальным теоретическим подходом авторов является использование теории ассамбляжей, разработанной Мануэлем Деланда. Подход позволяет помещать в модель коммуникации всех акторов независимо от их материального носителя, что необходимо в исследуемой ситуации, когда коммуникация становится гетерархичной и не только человеческой. На основе данной теории разработаны новые методические подходы для систематизации профессиональных задач преподавателя, который действует в гибридной (phygital) реальности совместно с технологиями ИИ. Проводится анализ соответствия функциональных возможностей технологий искусственного интеллекта задачам, стоящим перед преподавателем, а также предлагаются методы оценки эффективности использования ТИИ. На примере авторской разработки структуры учебного курса показано, как именно трансформируются задачи преподавателя при разработке учебного контента и реализации синхронного учебного курса совместно с ТИИ. Приводится классификация подходов, позволяющих более эффективно использовать большие языковые модели для решения образовательных задач: промпт-инжиниринг, RAG, LoRA, мультиагентный подход. Анализируются процессы цифровой трансформации высшего образования, обусловленные внедрением технологий искусственного интеллекта (ИИ). Основное внимание уделяется управленческим аспектам интеграции ИИ на различных уровнях образовательной организации. Публикация будет интересна менеджерам системы высшего образования, ученым и педагогам, занимающимся вопросами цифровизации обучения и цифровой трансформации вузов.
Ключевые слова: цифровая трансформация, искусственный интеллект в образовании, педагогические компетенции, большие языковые модели, цифровизация обучения, управление образованием, LLM
Для цитирования: Орешкина Т. А., Долганов А. Ю., Маяцкая Е. А., Артюгин О. Ю. Внедрение технологий искусственного интеллекта в образовательный процесс: управленческие вызовы // Университетское управление: практика и анализ. 2025. Т. 29, № 1. С. 92–105. DOI: 10.15826/umpa.2025.01.007.

DOI 10.15826/umpa.2025.01.007

IMPLEMENTATION OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES IN EDUCATION: MANAGERIAL CHALLENGES

T. A. Oreshkina, A. Yu. Dolganov, E. A. Mayatskaya, O. Yu. Artyugin

Ural Federal University named after the first President of Russia B. N. Yeltsin
Mira st., 19, Yekaterinburg, 620002, Russian Federation;
t.a.oreshkina@urfu.ru

Abstract. This study examines fundamental interdisciplinary issues and key managerial challenges associated with decision-making processes regarding the legalization and implementation of artificial intelligence (AI) technologies in educational settings. The research aims to assess potential effects, advantages, and risks of integrating large language model (LLM)-based AI technologies into educational processes at the discipline level. The authors propose an original theoretical framework utilizing Manuel DeLanda's assemblage theory. This approach enables the incorporation of all

actors into communication models regardless of their material substrates—a crucial consideration in contexts where communication becomes heterarchical and extends beyond human participants. Building on this theoretical foundation, novel methodological approaches have been developed to systematize professional teaching tasks in hybrid (phygital) environments incorporating AI technologies. The study includes analysis of AI functionality alignment with pedagogical requirements, development of AI effectiveness evaluation methodologies, demonstration of task transformation through a case study of course structure design, classification of enhanced LLM utilization approaches (industrial engineering, RAG, LoRA, multi-agent systems). The paper analyzes digital transformation processes in higher education driven by AI adoption, with particular emphasis on managerial considerations at various organizational levels. This research will benefit higher education administrators, researchers, and educators engaged in educational digitalization and institutional transformation.

Keywords: digital transformation, artificial intelligence in education, pedagogical competencies, large language models, learning digitalization, education management, LLM

For citation: Oreshkina T. A., Dolganov A. Yu., Mayatskaya E. A., Artyugin O. Yu. Implementation of Artificial Intelligence Technologies in Education: Managerial Challenges. *University Management: Practice and Analysis*, 2025, vol. 29, nr 1, pp. 92-105. doi 10.15826/umpa.2025.01.007. (In Russ.).

Введение

В современных образовательных системах цифровые технологии играют ключевую роль в трансформации процесса обучения и воспитания. В эпоху цифровизации большинство российских вузов имеют возможность использовать электронную образовательную среду. Обучение в онлайн-формате становится неотъемлемой частью образовательного процесса[1]. Сервисы, основанные на технологиях искусственного интеллекта (ИИ), могут стать инструментом повышения эффективности образовательного процесса. Например, ИИ может использоваться абитуриентами при выборе будущей профессии[2], студентами при подготовке к экзаменам, преподавателями для перевода иностранных текстов и исследователями при подготовке научных работ.

В 2023–2025 гг. был проведен ряд зарубежных и российских исследований влияния ИИ на образование, в которых проанализированы цели и специфика применения ИИ в образовательном процессе, а также выявлены ключевые вызовы, с которыми сталкиваются образовательные учреждения при внедрении ИИ-технологий.

Основные вызовы:

1. *Трансформация компетенций сотрудников вузов*: требуется освоение навыков работы с различными технологиями искусственного интеллекта (ТИИ), обучение применению data-driven подходов и повышение квалификации преподавателей для внедрения ТИИ в образовании[3]; необходима трансформация должностных инструкций и регламентов обучения студентов, поскольку преподаватель теряет роль единственного источника информации, требуется интеграция человеческого опыта с ИИ для повышения качества образовательных материалов[4].

2. *Сложность разработки комплексной системы*: каждый элемент образовательной цифровой экосистемы обеспечивается целым набором технологий ИИ, необходимо обеспечение согласованности всех элементов системы, и, если стратегия внедрения ИИ отсутствует, менеджменту вуза становится сложно обеспечить согласованность и оптимизировать этапы интеграции технологий[5].

3. *Неравенство в образовании*: следует разрабатывать общедоступную систему, технологии ИИ не должны усиливать разрыв между регионами с разным уровнем доходов и образования[6], актуальны также вопросы этики и доверия к технологиям ИИ[7].

4. *Нехватка вычислительных ресурсов и структурированных данных*: для обучения некоторых передовых моделей, таких как большие языковые модели, требуется доступ к высокопроизводительным

[1] Образование в цифрах: 2024: краткий статистический сборник / Т. А. Варламова, Л. М. Гохберг, О. А. Зорина и др.; Нац. исслед. ун-т «Высшая школа экономики». М.: НИУ ВШЭ, 2023.

[2] См. авторский анализ эффективности применения ИИ для работы с абитуриентами. Забокрицкая Л. Д. и др. Применение алгоритма машинного обучения для профориентации абитуриентов высшего учебного заведения // Вестник Томского государственного университета. 2022. № 485. С. 217–225.

[3] Искусственный интеллект и образование. Коротко о том, что происходит // Санкт-Петербург, 2024. URL: https://method.gsom.spbu.ru/white-book-ai#rec721568900 (дата обращения: 09.02.2025).

[4] Elhussein G., Hasselaar E., Lutsyshyn O. Shaping the future of learning: The role of ai in education 4.0 // World Economic Forum, 2024. URL: https://www3.weforum.org/docs/WEF_Shaping_the_Future_of_Learning_2024.pdf (дата обращения: 09.02.2025).

[5] Искусственный интеллект. Глобальные и образовательные тенденции, направления применения // СберУниверситет, 2024. URL: https://sberuniversity.ru/upload/iblock/9eb/8g9zq62hzjajlcqqvv9o4gag45ztiiym/iskusstvennyjintellekt.pdf С. 15 (дата обращения: 09.02.2025).

[6] Влияние ИИ на образование // АНО Цифровая экономика, 2024. URL: https://d-economy.ru/analitic/vlijanie-ii-na-obrazovanie/ (дата обращения: 09.02.2025).

[7] Баранников К.А. и др. Искусственный интеллект и высшее образование: возможности, практики и будущее // Яндекс. Образование. URL: https://education.yandex.ru/aihighreport (дата обращения: 09.02.2025).

вычислительным кластерам и большим объемам структурированных данных организации. Для эффективного применения ИИ-моделей необходимо не только предоставить пользователям доступ к самим моделям, но и наладить процессы сбора, хранения и обработки данных, которые будут использоваться этими моделями[8].

Помимо этого, перед Россией стоит задача, обусловленная экономическими ограничениями и санкциями, а именно – обеспечение технологического суверенитета. Это подразумевает разработку собственных технологий для защиты данных и обеспечения независимости, безопасности и непрерывности процессов обучения.

Таким образом, анализ выявленных вызовов подчеркивает необходимость системного подхода к внедрению ИИ в образовательный процесс. Каждый из этих аспектов требует внимательного рассмотрения и проработки, чтобы обеспечить успешную интеграцию технологий ИИ в образование.

Кроме перечисленных задач, существуют также фундаментальные междисциплинарные вызовы:

– В области компьютерных наук необходимо разрабатывать алгоритмы и модели ИИ, адаптированные для образовательных задач, а также обеспечивать их безопасность и надежность [1–2].

– В области педагогики (в том числе в области когнитивистики и изучения влияния цифровых технологий на физическое и ментальное здоровье обучающихся) важно проводить исследования влияния ИИ-технологий на когнитивные процессы, мотивацию и успеваемость, учитывая индивидуальные особенности студентов, их адаптацию к новым условиям и возможные негативные последствия использования цифровых технологий [3].

– В области социальной философии в настоящее время рассматриваются условия возникновения общего (сильного) искусственного интеллекта, вопросы сознания человеческих и нечеловеческих акторов, особенности техно-умвельта, вопросы о статусе коммуникации с искусственными агентами [4–5].

– В области прикладной этики обсуждаются вопросы соответствия целей и ценностей ИИ человеческим ценностям (AI alignment), конфиденциальности и безопасности данных, риски позитивной и негативной дискриминации алгоритмами и наборами данных [6]. Предлагаются рамочные решения в виде этических кодексов для образовательных организаций [7].

– Отдельно стоит отметить важность нерешенных юридических вопросов касательно авторского права и субъектности ИИ [8].

Менеджмент вузов оказался в ситуации необходимости принятия решений о легализации и внедрении технологий искусственного интеллекта в ситуации крайней неопределенности, когда не решены фундаментальные педагогические, философские, этические и правовые вопросы.

Постановка проблемы. С прикладной точки зрения менеджменту университетов требуется ответить на вопросы: какие функции ИИ может выполнять более эффективно, чем человек (преподаватель), и какие дилеммы и управленческие вызовы возникают на пути разработки и внедрения данных технологий.

Цель исследования: оценить возможные эффекты, преимущества и риски внедрения технологий искусственного интеллекта, разработанных на основе больших языковых моделей, в образовательный процесс (на уровне учебной дисциплины). Чтобы достичь цели исследования, мы сосредоточимся на трех задачах:

– разработаем системно-коммуникативную модель образовательного процесса, фундированную теорией ассамбляжей Мануэля Деланда;

– опишем профессиональные задачи преподавателя вуза и представим системно-коммуникативную модель синхронного учебного курса;

– проведем сравнительный анализ функциональности больших языковых моделей с задачами преподавателя и предложим критерии и методы оценки эффективности делегирования задач преподавателя системам на основе ИИ.

Системно-коммуникативная модель образовательного процесса

В основе образовательного процесса лежит процесс коммуникации [9]. Следует признать, что трансляционная модель коммуникации, когда преподаватель является носителем непреложного знания, источником экспертизы, а студент – объектом воздействия, уходит в прошлое. Субъект-объектная педагогическая модель уже не работает. Само представление о коммуникации претерпевает изменения. От теории сетей М. Кастельса [10, *496*], который пишет о том, что «сетевое общество – это такое общество, в котором ключевые социальные структуры и деятельность его членов организованы *вокруг* сетей электронных коммуникаций», мы переходим к акторно-сетевой теории Б. Латура [11] и М. Каллона [12], где цифровые технологии трактуются как актанты, т.е. *источник* действия,

[8] Artificial Intelligence in Education. 2023 Survey Insights. URL: https://www.holoniq.com/notes/artificial-intelligence-in-education-2023-survey-insights (дата обращения: 09.02.2025).

а также к эпистемическим культурам К. Кнорр-Сетины, которая ставит под вопрос принципиальную возможность единых ценностей и норм [13]. При этом теория коммуникативного действия Ю. Хабермаса указывает нам пути поиска основ, которые способны обеспечить взаимопонимание и интеграцию социальных субъектов [14].

От понятия образовательного субъекта и объекта мы переходим к коммуникационной модели взаимодействия социальных акторов, опосредованное цифровыми системами. Цифровое образовательное пространство определяет взаимодействие социальных акторов и создает контекст партнерских отношений. Отказ от понятия субъекта и объекта образовательного процесса и смещение внимания с иерархичной модели педагогических отношений на вопрос об эффективной коммуникации в цифровом образовательном пространстве дают нам возможность найти пути выхода из того сложного положения, в котором оказались преподаватели вузов в современном постиндустриальном обществе.

Концепция сетевого общества, осмысленная в работах П. Бурдье и Ж. Делеза, с появлением ИИ получает новый смысл: коммуникация становится гетерархичной и не только человеческой. Сейчас нам требуется новая онтология социального, позволяющая помещать в модель коммуникации всех акторов независимо от их материального носителя. Поэтому теоретической рамкой нашего исследования выбрана теория ассамбляжей, разработанная Мануэлем Деланда [15]. М. Деланда постулирует ассамбляж как оформление всех реальных объектов независимо от их материального (вещи-предметы) и антропологического (люди-существа) статуса. В модели социальной онтологии Деланда для ассамбляжей характерны отношения экстериорности, т.е. элементы системы не обладают органическим единством. Эти отношения предполагают, что «…составная часть ассамбляжа может быть отделена и помещена в другой ассамбляж, с иными формами взаимодействия» [15, *20*]. Проиллюстрируем разницу сущностей в виде ассамбляжей и в виде тотальностей, используя классическую метафору Ж. Делеза[9] об осе и орхидее. Если мы рассматриваем технологии ИИ и человека через отношения интериорности, метафорически это можно представить в виде кентавра: отделенная от целого часть перестает быть собой и не может существовать автономно. Ассамбляжи, в отличие от тотальностей, могут быть представлены как отношения осы и орхидеи, когда они взаимодействуют и удовлетворяют свои потребности, при этом сохраняя свои идентичности. Эта метафора помогает понять, как различные модели взаимодействия технологий и человека могут влиять на их развитие и функционирование.

Для осмысления новой модели коммуникации требуется поэтапно рассмотреть, какие именно изменения произошли в образовательной сфере.

Во-первых, на этапе информатизации произошел переход к открытой информационно-образовательной среде, когда вузы начали активно использовать информационно-коммуникационные технологии для проведения занятий, а студенты получили доступ к электронным учебникам, справочным материалам и онлайн-библиотекам (рис. 1). На этом этапе авторами контента всегда являлись человеческие акторы, и ответственность за поиск и анализ информации также лежала на них. Информационные источники в Интернете были доступны, но не обладали агентностью [16].

Во-вторых, появление ИИ-агентов привело к возникновению сетевой коммуникационной

[9] Оса и орхидея взаимодействуют, создавая уникальную экосистему. Ассамбляжи представляют собой временные и нестабильные сборки различных элементов, которые могут быть как человеческими, так и нечеловеческими. Они взаимодействуют друг с другом, создавая новую сущность, которая не сводится к сумме своих частей. См.: Делез Ж., Гваттари Ф. Тысяча плато: Капитализм и шизофрения. М., 2010. 895 с.

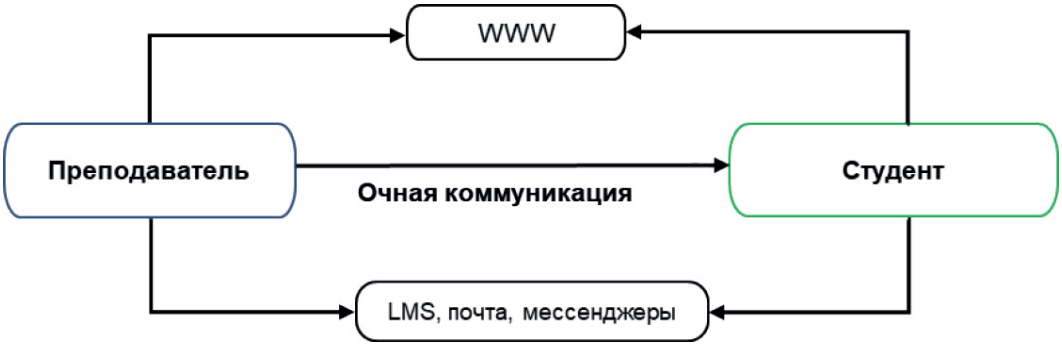

Рис. 1. Модель коммуникации на этапе информатизации
Fig. 1. Communication model at the stage of informatization

модели, в которой возникают синергетические эффекты, такие как ускорение обработки информации, персонализация обучения, расширение доступа к знаниям. Согласно теории ассамбляжей Деланда [15], свойствами связей в сети являются: взаимность, сила, плотность, стабильность. Тот узел в сети, который имеет больше связей с другими узлами, обретает большую важность: например, сила связи между преподавателем и ИИ-ассистентом может определяться частотой их взаимодействия и объемом совместно решаемых задач. В контексте нашего прикладного исследования теоретические положения приобретают особое значение. В сетевой модели, где взаимодействуют агенты искусственного интеллекта, возникают новые формы сотрудничества и конкуренции, то есть люди и ИИ-агенты могут как сотрудничать между собой, так и конкурировать. Появление агентов искусственного интеллекта в этой среде может иметь как положительные, так и отрицательные последствия для образовательного процесса. Для повышения качества образования следует обеспечить адаптацию образовательных систем к новым формам сотрудничества и конкуренции агентов искусственного интеллекта и людей в системе коммуникаций. Это включает в себя разработку надёжных и безопасных алгоритмов, а также обучение преподавателей и студентов работе с агентами искусственного интеллекта и стимулирование преподавателей работать «в команде» со многими ИИ-агентами, увеличивая количество связей в сети, тем самым повышая свою роль в системе.

На рис. 2 представлена модель коммуникации в цифровой образовательной среде с участием ТИИ. На рисунке изображены как внешние ИИ-агенты, так и ИИ-агент, валидированный вузом (с ним взаимодействуют и преподаватель, и студент) и за счет сотрудничества имеющий большее значение в сети.

Роль университетов в основном сохраняется, однако ИИ может существенно трансформировать формы и методы их работы в части образовательной, научной и инновационной деятельности. Для примера рассмотрим, как именно ИИ-агенты встраиваются в образовательную коммуникацию, и оценим возможность внедрения инструментов ИИ на уровне образовательной дисциплины (синхронного учебного курса). Научная и административная деятельность будет рассмотрена в других публикациях.

Следует отметить, что в настоящее время в сфере очного синхронного обучения жесткое разделение на «офлайн» и «онлайн» теряет свою актуальность. Это обусловлено не столько цифровизацией вузов, сколько увеличением роли технологий в повседневной жизни. Мы проанализируем профессиональные задачи преподавателя, который действует в гибридной (phygital) реальности, то есть взаимодействие со студентами происходит как в очном режиме, так и посредством интернет-коммуникаций.

Если преподаватель разрабатывает и реализует учебный курс совместно с ИИ-агентом, особое внимание следует уделить выбору *способа проектирования* учебного курса. Согласно теории ассамбляжей, мы предполагаем, что преподаватель будет взаимодействовать с несколькими ИИ-агентами, работать с ними в гибридной «команде». Это значит, что все материалы преподавателя-человека должны быть организованы понятным образом для машинного интеллекта.

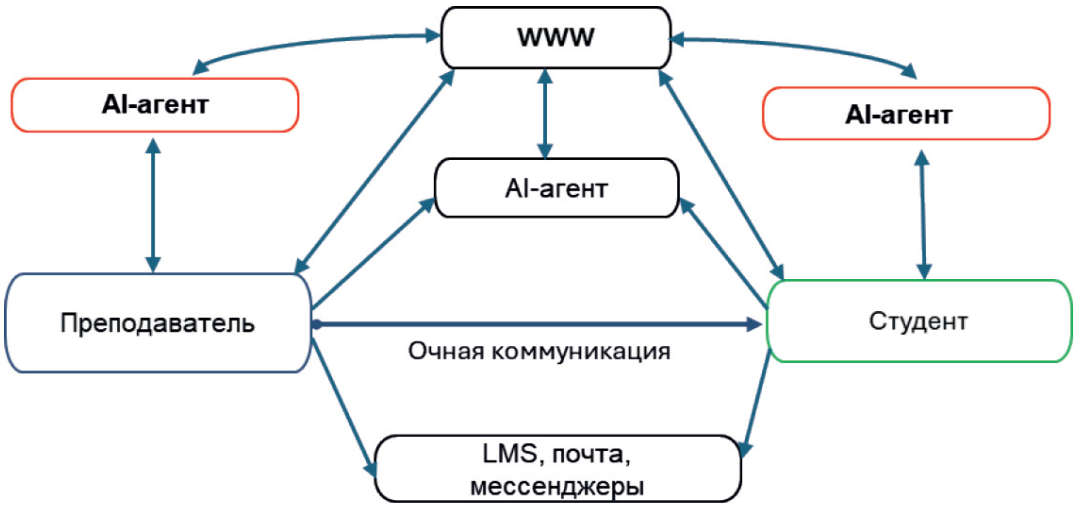

Рис. 2. Модель коммуникации в цифровой образовательной среде с ИИ-агентами
Fig. 2. Communication model in a digital educational environment with artificial intelligence agents

С точки зрения этапов разработки и реализации учебного курса можно выделить следующие шаги: проектирование курса, создание контента курса, реализация учебной дисциплины, рефлексия и корректировка по итогам реализации курса. На каждом из этапов использование ТИИ имеет свои особенности.

Задачи преподавателя при реализации учебного курса с участием ТИИ

Технологии искусственного интеллекта, получившие широкое распространение для обработки различных данных, о которых идет речь в настоящей статье, относятся к алгоритмам машинного обучения на основе больших языковых моделей (Large Language Model, LLM). Истоки современных достижений в области обработки естественного языка можно найти в статье А. Васвани и его коллег, опубликованной в 2017 году под названием «Attention is all you need» [17]. В этой статье авторы описали использование архитектуры Transformer для задач *перевода.* Результатом этой работы стало улучшение качества не только нейросетевого перевода, но и решения задач классификации текста, выделения именованных сущностей и других задач, связанных с пониманием естественного языка.

В период 2017–2024 гг. появилась языковая модель, которая быстро обрела мировую известность, – ChatGPT. Помимо разработки команды OpenAI были созданы языковые модели семейств Llama, Gemini, GigaChat, Yandex GPT [18; 19].

Рассмотрим на конкретном примере, какие действия преподавателя приведут к повышению качества образовательного процесса. В настоящий момент ТИИ на LLM широко представлены и доступны как в виде веб-сервисов, так и в виде локальных решений. Для того, чтобы применение ТИИ было эффективным, требуется предварительное проектирование образа конечного результата.

После проведенных экспериментов авторы могут обозначить два способа проектирования учебного курса: предметные онтологии (создание графовых баз данных, т.н. GraphRAG[10]) и дидактические матрицы (n-мерные матрицы на основе дидактических решений). Разберем на конкретном примере, как выглядит предметная онтология курса.

Этапы создания модели курса, адаптированного для реализации совместно с ТИИ:

0 этап. Разработка ОХОП на основе ФГОС и СУОС вуза, проектирование компетентностной модели выпускника программы и формирование рабочей программы дисциплины (РПД). Данный этап в нашей статье мы рассматривать не будем, так как применение LLM на данном этапе авторами еще не проводилось.

1 этап. Формирование структуры учебной дисциплины на основе полученной РПД. На входе – РПД, на выходе – структура и наполнение курса (содержание лекционных и практических занятий), организационно-методические материалы дисциплины.

2 этап. Формирование структуры лекционных занятий в зависимости от объема дисциплины, целевой аудитории и т. д. Определение формата лекций (живые, ВКС, видеолекции). Подготовка презентаций, интерактивных элементов лекций (синхронные опросы, конспекты или схемы).

3 этап. Формирование структуры практических занятий. Определение формата практических занятий, которые могут обеспечить достижение заявленных умений и навыков. Подготовка заданий для письменной и устной текущей аттестации. Фиксация критериев оценки для текущей аттестации, подготовка инструкций и методических материалов для самостоятельной работы студента.

4 этап. Подготовка заданий для промежуточной аттестации (зачет, экзамен) в письменном или устном формате. Фиксация критериев оценки для промежуточной аттестации, подготовка инструкций и методических материалов для самостоятельной работы студента.

В итоге педагогический дизайн и модель учебного курса выглядят как уровень компетенций, структуры и формата учебного материала (рис. 3). Проектирование учебного курса в виде цельной модели позволяет предоставлять модели LLM необходимый контекст, связывая все элементы курса в единую модель.

Учебный курс на рисунке представлен в виде трехслойной модели. Поясним содержание каждого уровня.

Первый уровень, **компетентностная модель**, описывает когнитивную структуру: знания, умения, навыки и мотивационный компонент компетенций (ценности, установки, нормы и шаблоны поведения). Данный уровень отражен в рабочей программе дисциплины.

Второй уровень, **структура учебного материала**, описывает структуру учебного контента: теоретические и практические занятия, учебные задания (самоконтроль и формирующее оценивание), контрольные задания и организационно-методические

[10] Структура онтологии в базе данных: Универсальная компетенция → Результат обучения → Задание → Знания → Лекция → Модуль → Курс. Инструменты: Llama 3.1-405B + Neo4j.

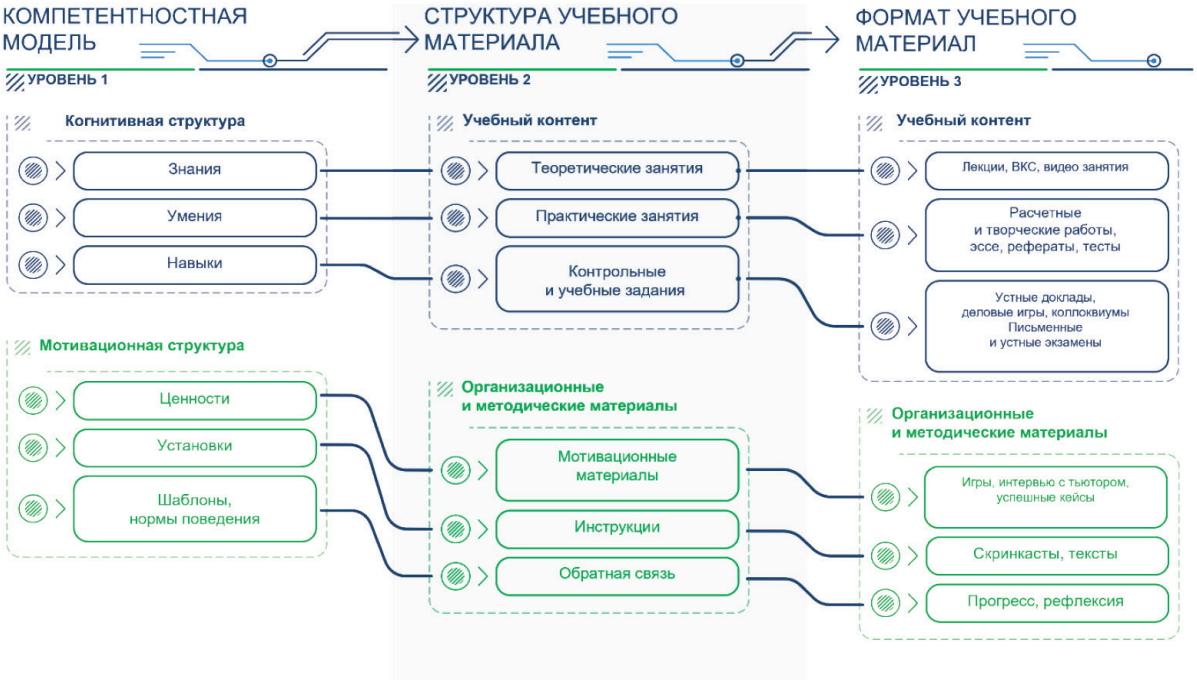

Рис. 3. Трехслойная модель учебного курса: уровень компетенций, структуры и формата учебного материала

Fig. 3. Three-layer training course model: competence level, structure and format of the training material

материалы: мотивационные материалы, инструкции, обратная связь.

Третий уровень, **формат учебного материала**, содержит разные варианты формата учебного контента и организационно-методических материалов (живые лекции и ВКС, тексты, презентации, письменные расчетные и творческие работы, эссе, рефераты, тесты, скринкасты и тексты с инструкциями и т.д.).

Каждый элемент учебного курса должен быть согласован с другими по вертикали и по горизонтали. Каждый элемент первого уровня связан с элементами на втором и третьем, например, знания – теоретические занятия – лекции – видеолекции в записи. Умения и навыки – практические занятия – эссе и доклад с презентацией – письменный экзамен. По горизонтали знания определяют умения и навыки, на практике реализуемые на основе ценностей, установок и норм поведения.

Итак, представленная модель учебного курса описывает этапы проектирования и создания контента таким образом, чтобы все этапы можно было делегировать технологиям ИИ при условии создания непротиворечивых требований к контенту учебных курсов и правил оркестрации. Безусловно, при делегировании каждого этапа ТИИ разработчики технологии должны иметь инструменты оценки качества и эффективности работы LLM.

5 этап. Реализация учебного курса. На этапе реализации преподаватель взаимодействует со студентами во время контактной работы, а также осуществляет контроль самостоятельной работы студента и проводит мероприятия текущей и промежуточной аттестации. На этом этапе мы выделяем следующие рутинные функции преподавателя:

– Проверка письменных домашних заданий студента;

– Устные коллоквиумы, зачеты и экзамены.

6 этап. Сбор обратной связи от студентов. Анализ обратной связи, рефлексия преподавателя и подготовка плана корректировки курса. Каждая их этих задач может быть делегирована ТИИ и осуществляться под контролем преподавателя, который может воспользоваться существующими вебсервисами. Для улучшения качества работы LMM могут применяться различные технологии.

Дообучение, расширение контекста и оценка качества работы LLM

Сущность работы больших языковых моделей состоит в том, чтобы после получения входного промпта продолжать последовательность слов на естественном языке. Магия начинается, когда LLM «дообучается» под конкретную задачу. Для эффективного решения преподавательских задач

следует использовать различные подходы расширения контекста, включая дообучение LLM. К технологиям расширения контекста, которые открывают возможности для использования больших языковых моделей, обученных на больших корпусах данных, для решения частных задач, не обучая модели с нуля, относятся: промпт-инжиниринг, RAG, адаптивные методы (например, LoRA, PEFT, адаптация на примерах). Кратко рассмотрим данные технологии и их значимость.

Промпт-инжиниринг, или составление инструкций для LLM, значительно повышает качество работы языковых моделей для отдельных задач и позволяет использовать большую языковую модель для решения новых задач без процедуры дополнительного обучения (fine-tune). Техники промпт-инжиниринга основаны на зависимости ответа модели от исходного промпта [20]. Например, если в промпт с условиями математических задач добавить не только пошаговое описание, но и примеры решения, языковая модель сможет лучше понять, как решать подобные задачи. Появились специальные техники работы с LLM, которые «заставляют» модели не просто генерировать необходимый ответ, а рассуждать, делать предварительные выводы и только потом выдавать окончательный ответ [21].

Использование таких инструкций позволяет дообучать модели решению конкретных задач. Например, классическое применение LLM – это общение с пользователем в формате диалога с промптом “you are helpful assistant”. Современные языковые модели значительно улучшили возможности ведения диалогов с пользователями, хотя и имеют ограничения.

Стоит отметить, что большие языковые модели могут галлюцинировать – выдавать внешне правдоподобный ответ, несущий некорректную информацию. Например, может генерироваться ссылка на несуществующую статью, несуществующая цитата философа и т.д. [22]. Одним из способом борьбы с галлюцинированием больших языковых моделей является техника *RAG (retrieval-augmented generation*) [23]. Основная идея состоит в создании векторной базы знаний из документов, которые соотносятся с конкретной задачей (например, нормативная документация компании, готовый перечень вопросов и ответов, источник информации по типу сайта Wikipedia, в нашем случае – модель учебного курса со всеми учебными материалами и списком литературы). Затем, когда в языковую модель поступит промпт, будет проведён поиск по векторной базе данных. В результате будет найден фрагмент, наиболее точно связанный с полученным запросом, который дополнит исходный запрос. Это поможет уменьшить количество ошибок, возникающих в языковой модели. Кроме того, такой подход позволяет адаптировать общую языковую модель к конкретной задаче.

Другим эффективным способом дообучения большой языковой модели под локальную задачу является подход *LoRA* [24]. В рамках этого подхода веса исходной модели не обновляются, вместо этого происходит обучение относительно небольших матриц-адаптеров (насадок), которые заточены на решение конкретной подзадачи.

Для решения комплексных задач также может применяться *мультиагентный подход*. В этом случае одна или несколько языковых моделей запускаются с определенной ролью (агент) для выполнения определенных задач (проверка или обобщение текста, ответы на вопросы, написание кода) [25]. Агенты выполняют заранее прописанный план действий, который может включать коммуникацию с другими агентами или инструментами. Отмечается, что при использовании такого подхода несколько небольших языковых моделей (на 7 миллиардов параметров) могут выдавать результаты, сопоставимые с большими моделями (на сотни миллиардов) [26].

Как правило, эффективность LLM для решения конкретных задач оценивается с использованием бенчмарков (benchmark). Бенчмарк подразумевает создание стандартизированного набора данных и задач к нему, а также метрик оценивания качества. Примерами бенчмарков для оценки общего качества понимания текста LLM являются GLUE [27] и RussianGLUE [28].

В области педагогики и образования важным шагом в применении языковых моделей является создание собственного бенчмарка. Можно назвать следующие укрупненные этапы этой задачи:

1. Сбор данных для дообучения (тренировочный набор данных) и валидации модели (валидационный набор данных), что позволяет адаптировать модель к конкретной задаче и выявить её сильные и слабые стороны.

2. Разработка метрик для оценки модели, адаптированных под конкретную задачу. Это поможет лучше отразить специфику задач и обеспечить более точную оценку производительности модели.

Для оценки модели на валидационном наборе данных важно четко понимать, какую задачу мы решаем. В зависимости от этого следует выбрать подходящий тип валидации: общую или частную. Общая валидация применима, когда необходимо получить общее представление о производительности модели на разнообразных задачах. Частная

валидация необходима, когда важно сосредоточиться на узких задачах или специфических группах данных. Это особенно актуально, если модель применяется в контексте, где важны специфические требования или условия, например, в вузах.

Для оценки текста важно правильно подбирать метрики, соответствующие конкретным задачам. В задачах классификации полезны такие метрики, как precision, recall, accuracy, f1-score, которые помогают оценить качество предсказаний модели. В контексте машинного перевода также часто используются BLEU и METEOR, позволяющие оценивать соответствие перевода оригиналу.

Для оценки развернутых ответов, например, ответов на вопросы или эссе, в данный момент используется ручная оценка преподавателем. Она является необходимой, поскольку существуют задачи, которые требуют глубокого понимания темы от проверяющего. Рассмотрим примеры таких задач:

– Творческое письмо: в заданиях, связанных с поэзией или литературным анализом, оценка оригинальности, стиля и эмоциональной глубины текста требует человеческого взгляда.

– Эссе: когда речь идет о сложных текстах, необходимо учитывать структуру аргументации, логические связи и качество обоснований. Преподаватель может оценить, насколько убедительно автор представляет свои идеи и насколько хорошо они связаны друг с другом.

С развитием языковых моделей становится возможной автоматизация процесса оценки текстов. Так, итоговая метрика для оценки работы модели будет рассчитываться как процент совпадений между оценкой, выставленной языковой моделью, и оценкой человека. Такой подход позволяет измерять согласованность между автоматизированной и человеческой оценками. Это открывает новые горизонты в сфере образования, однако возникают важные особенности, которые необходимо учитывать:

5. *Необходимость создания критериев оценки*: для того, чтобы языковая модель могла эффективно оценивать текст, необходимо четко определить критерии, которым он должен соответствовать. Это могут быть грамматическая правильность, логичность, глубина анализа и соответствие заданной теме. Четкие критерии помогут обеспечить надежность оценки.

6. *Вопрос доверия к ИИ*: при использовании языковых моделей для оценки текста возникает вопрос доверия к таким системам. Языковые модели могут допускать ошибки в интерпретации текста или не учитывать сложный контекст, что требует постоянной проверки их выводов.

Важно не только сформулировать критерии оценки, но и разработать систему доверия к информации, сгенерированной языковой моделью. Данная система должна учитывать изменяемые критерии в зависимости от типа контрольного задания и обратную связь от преподавателей. Под обратной связью подразумевается необходимость проведения анализа оценок, выставляемых преподавателями по каждому заданию, что позволит модели выявить тенденции в их оценках. Создание ориентированной на преподавателя системы оценки позволит повысить доверие к LLM в образовательном процессе. Важно, чтобы система доверия была гибкой и адаптивной, учитывающей различные требования к заданиям.

На практике уже применяются подобные подходы. Например, один из методов [29] обеспечивает корректировку оценок, выставленных языковой моделью, учитывая распределение оценок от преподавателей. Этот подход показал, что модель, оценивающая сгенерированные тексты, начинает подражать человеку. В другой работе [30] рассчитывается коэффициент доверия. Основная идея заключается в том, что, если значение коэффициента превышает критический уровень (заданный человеком), человек может с уверенностью полагаться на оценки, выставленные LLM.

В таблице 1 указаны критерии оценки качества, преимущества и риски для отдельных технологий ИИ. Оценивать и управлять качеством сгенерированной информации (текстовых лекций, видеороликов, ответов на работы студента и т.д.) можно автоматически с помощью другой LLM или обратной связи от эксперта.

Таким образом, задачи преподавателя при реализации учебного курса с участием ТИИ сводятся к следующим крупным блокам: создать модель курса (предметную онтологию), сформулировать непротиворечивые критерии, которым должен соответствовать каждый элемент учебного курса, создать базу материалов, которые таким критериям соответствуют (лучшие примеры, соответствующие разработанным критериям), выбрать конкретные LLM и дообучить их в соответствии с собственными задачами, проконтролировать полученный результат.

На уровне управления кафедры, департамента, института в связи с внедрением ИИ появляются следующие управленческие задачи:

– проводить всестороннее обсуждение возможностей применения ТИИ для разработки методических подходов и адаптации преподавателей к учебному процессу с использованием ТИИ;

Таблица 1

Возможные способы применения ТИИ и рекомендации по контролю рисков

Table 1

Possible ways to use artificial intelligence technologies and recommendations for risk control

Задача	Преимущества	Вызовы и риски	Рекомендации для менеджмента вуза
Формирование структуры учебного курса	Создание индивидуальных образовательных траекторий и персонализированного контента на основе данных об успеваемости и интересах. Технологии расширения контекста позволяют генерировать ответ с учетом локальных баз знаний.	Предвзятость алгоритмов, недостаток данных для обучения, необходимость постоянного обновления образовательного контента, галлюцинации модели.	1. Использовать разнообразные и сбалансированные наборы данных для обучения LLM, собирать достаточные объемы качественных обучающих данных для дообучения. 2. Обеспечить систему регулярного обновления и проверки образовательного контента, используемого LLM. 3. Внедрить этапы проверки сгенерированного контента, включая экспертную оценку людьми (например, на соответствие учебной программе). 4. Human oversight: сохранять экспертный контроль преподавателей, выявлять и править ошибки, чтобы избежать распространения недостоверной информации.
Создание текстовых лекций на разных языках	Современные LLM мультиязычны.	Качество перевода (особенно специализированных терминов и идиом), культурные несоответствия, галлюцинации модели (неправильная интерпретация фактов).	1. Использовать специализированные модели, натренированные на конкретных областях знаний. 2. Привлекать экспертов-лингвистов для вычитки переводов, особенно для сложных терминов и культурных нюансов.
Создание учебных визуальных материалов	Автоматизация создания визуальных материалов.	Снижение качества дизайна (несоответствие корпоративному стилю, устаревший дизайн), необходимость соответствия визуального содержания целевой аудитории, ресурсозатратность (значительная нагрузка на GPU).	1. Тщательно разрабатывать корпоративные промпты, уделяя внимание стилю, композиции и целевой аудитории. 2. Использовать инструменты для постобработки изображений для улучшения качества дизайна и соответствия бренду. 3. Использовать изображения с минимально необходимым разрешением для снижения нагрузки на ресурсы.
Создание обучающих видео	Автоматическое создание видео делает контент более доступным и интересным.	Снижение качества дизайна (несоответствие стандартам видеопроизводства), необходимость соответствия визуального содержания целевой аудитории, ресурсозатратность (время и вычислительные мощности для рендеринга видео).	1. Разработать шаблоны и гайдлайны для создания видео, чтобы обеспечить единообразие и качество дизайна. 2. Предусмотреть возможность ручной доработки видео для улучшения синхронизации и добавления элементов, повышающих вовлеченность. 3. Экспериментировать с разными моделями и настройками для снижения затрат на генерацию видео.
Преобразование текстов лекций в аудио	Доступность контента для студентов с ограниченными возможностями и предпочитающих аудиообучение.	Качество озвучивания (неестественное звучание, ошибки произношения), необходимость соответствия темпа и интонации контексту.	1. Выбирать модели синтеза речи с наиболее естественным и выразительным звучанием. 2. Тонко настраивать параметры синтеза речи (темп, интонация, паузы) для соответствия контексту лекции. 3. Предусмотреть возможность редактирования аудио для исправления ошибок и улучшения общего качества (например, добавление музыкального фона).

Окончание табл. 1
Table 1 finishes

Задача	Преимущества	Вызовы и риски	Рекомендации для менеджмента вуза
Создание тестов и учебных заданий	Создание персонализированных тестов и заданий на основе индивидуальных потребностей студентов. Технологии расширения контекста позволяют генерировать ответ с учетом локальных баз знаний.	Предвзятость в оценке знаний (оценка на основе устаревших данных или шаблонов), необходимость постоянного обновления заданий, галлюцинации модели (создание некорректных или неразрешимых заданий).	1. Привлекать преподавателей для проверки сгенерированных тестов и заданий на предмет корректности, релевантности и соответствия учебным целям. 2. Постоянно обновлять базу знаний и алгоритмы LLM, чтобы избежать устаревания и предвзятости.
Создание организационно-методических материалов курса	Автоматизация упрощает процесс создания материалов. Технологии расширения контекста позволяют генерировать ответ с учетом локальных баз знаний.	Качество и актуальность материалов (устаревшие шаблоны, общая информация, отсутствие специфики), галлюцинации модели (несоответствие стандартам учебного заведения).	1. Разработать шаблоны для организационно-методических материалов, чтобы обеспечить единообразие и качество. 2. Привлекать опытных преподавателей и методистов для оценки сгенерированных материалов. 3. Постоянно обновлять информацию и шаблоны, используемые LLM.
Проверка домашних и контрольных заданий	Быстрота, объективность, масштабируемость. Обработка и анализ большого количества текстов, выявление типичных ошибок, предоставление комментариев. Проверка текстов на разных языках.	Использование персональных данных и защита приватности студентов, галлюцинации модели (неправильная интерпретация ответов, неспособность оценить креативное мышление), риск предвзятости (оценка на основе стереотипов), необходимость создания системы доверия к ИИ.	1. Использовать методы анонимизации для защиты персональных данных студентов, особенно при дообучении моделей. 2. Обеспечить прозрачность алгоритмов проверки, чтобы студенты могли понимать, как оцениваются их работы. 3. Проводить регулярную проверку алгоритмов на предвзятость и принимать меры для ее устранения. 4. Использовать LLM в сочетании с ручной проверкой преподавателями для обеспечения максимальной точности и справедливости.
Проверка видео с докладами	Быстрая автоматическая транскрипция позволяет преподавателю сосредоточиться на содержательной оценке доклада. Современные модели могут обрабатывать доклады на разных языках. Создание текстовых версий улучшает доступ к материалам для людей с нарушением слуха.	Ошибки в распознавании речи (особенно в присутствии шумов, акцентов или технической лексики), неспособность различить омофоны (слова, звучащие одинаково, но имеющие разное значение).	1. Использовать хорошее оборудование для повышения качества входного сигнала. 2. Рассмотреть возможность адаптации модели распознавания речи или постобработки текста под конкретные акценты и терминологию, используемую в докладах.
Сбор обратной связи от студентов	Автоматическая генерация рекомендаций по улучшению курса, адаптированных под конкретную группу студентов. Отслеживание изменений в уровне удовлетворенности и вовлеченности студентов со временем.	ИИ может неправильно интерпретировать контекст, сарказм или сложные эмоции в отзывах. Обработка персональных данных требует соблюдения стандартов безопасности и защиты конфиденциальности. Предвзятость анализа (оценка на основе тональности, а не содержания).	1. Формулировать четкие и конкретные вопросы для сбора обратной связи. 2. Разрабатывать алгоритмы, способные учитывать контекст и эмоциональную окраску отзывов. 3. Обеспечить соблюдение всех требований по защите персональных данных студентов. 4. Комбинировать автоматический анализ отзывов с ручным анализом преподавателями.

– инициировать пилотные внедрения по применению ТИИ и анализировать применение ТИИ студентами и преподавателями, корректируя процесс с учетом нового опыта;

– внедрить систему регулярного обучения работе с data-driven подходами, разработать комплексную программу повышения квалификации в области ИИ, создать систему мотивации для освоения новых компетенций;

– разрабатывать этические стандарты использования ИИ, в том числе обеспечить прозрачность работы ИИ-систем при их использовании для проверки компетенций студентов.

Успех цифровой трансформации и внедрения ТИИ на уровне вуза будет зависеть от развития технической инфраструктуры, так как необходимо создать современную техническую базу, инвестировать в высокопроизводительные вычислительные кластеры, разработать систему сбора и хранения структурированных данных, обеспечить безопасность данных при работе с ИИ-моделями.

Выводы

В условиях цифровой трансформации общества изменяется и образовательный процесс. Цифровые технологии изменяют коммуникацию, превращая её в гетерархичную и сетевую, где важна эффективность, а не иерархические отношения. Для анализа новой реальности требуется пересмотреть онтологию социального, ориентируясь на такие теории сетевого общества, как теория ассамбляжей Мануэля Деланда и плоские онтологии Б. Латура, которые предлагают новый подход, где все акторы, независимо от их материального или антропологического статуса, могут быть включены в модель коммуникации.

Внедрение ИИ в образование ставит перед менеджментом вузов сложные задачи, связанные с необходимостью принятия решений в условиях, когда многие этические, правовые и педагогические аспекты еще не до конца проработаны.

Выявленные вызовы, такие как трансформация компетенций преподавателей, необходимость разработки комплексных систем и вопросы этики, требуют системного подхода к внедрению ИИ. Делегирование профессиональных задач преподавателей вузов технологиям ИИ (с сохранением контроля преподавателя) может повысить качество обучения и сделать процесс более творческим. Важно не только определить, какие задачи преподавателя могут быть выполнены более эффективно с применением ТИИ, но и разработать надежные критерии для оценки использования различных моделей ИИ. В соответствии с теорией ассамбляжей предполагается, что преподаватель будет взаимодействовать с несколькими агентами искусственного интеллекта, работая с ними в гибридной «команде». Это подразумевает, что все материалы, подготовленные преподавателем-человеком, должны быть структурированы таким образом, чтобы быть доступными для понимания машинным интеллектом.

Технологии ИИ могут эффективно применяться для повышения качества и эффективности образовательного процесса уже сегодня. Автоматизация оценки текстов домашних и контрольных заданий студентов открывает новые возможности в образовании, но требует тщательной проработки критериев и сбора обратной связи от преподавателей, обеспечивая доверие к результатам, получаемым от LLM. Системы оценки, основанные на такой обратной связи, могут значительно повысить надежность и точность оценок, предоставляемых языковыми моделями. Интеграция этих технологий в образовательные процессы способствует более эффективному взаимодействию преподавателей с ИИ, что, в свою очередь, поддерживает высокие стандарты обучения и способствует развитию образовательной среды.

В настоящий момент еще не накоплен достаточный опыт применения ИИ-технологий в высшей школе. С учётом стремительного развития и внедрения технологий ИИ в образовательный процесс важной проблемой становится скорость их интеграции и тестирования. Авторы планируют продолжить исследование влияния искусственного интеллекта на качество образования, результаты обучения и трансформацию профессиональных навыков преподавателей.

Список литературы

1. *Brown N. B.* Enhancing Trust in LLMs: Algorithms for Comparing and Interpreting LLMs //arXiv preprint. 2024. 15 p.

2. *Pappachan P. et al.* Transparency and Accountability //Challenges in Large Language Model Development and AI Ethics. IGI Global, 2024. P. 178–211.

3. *Гаркуша Н. С., Городова Ю. С.* Педагогические возможности ChatGPT для развития когнитивной активности студентов // Профессиональное образование и рынок труда. 2023. Т. 11, № 1. С. 6–23. DOI: 10.52944/PORT.2023.52.1.001

4. *Резаев А. В., Трегубова Н. Д.* Философия общения и искусственный интеллект: опыт сравнительного анализа дискуссий в отечественной и зарубежной литературе // Эпистемология и философия науки. 2024. Т. 61, № 2. С. 134–156.

5. *Лешкевич Т. Г.* Искусственный интеллект в контексте философского осмысления // Вопросы философии. 2023. № 5. С. 50–60.

6. *Eden C. A. et al.* Integrating AI in education: Opportunities, challenges, and ethical considerations // Magna Scientia Advanced Research and Reviews. 2024. Vol. 10, nr 2. P. 6-13.

7. *Chen H.* The ethical challenges of educational artificial intelligence and coping measures: A discussion in the context of the 2024 World Digital Education Conference // Science Insights Education Frontiers. 2024. Vol. 20, nr 2. P. 3263–3281.

8. *Черных Н. А*. Обзор, концепции, точки зрения к возможности правового регулирования технологии искусственного интеллекта // Бюллетень науки и практики. 2024. Т. 10, № 5. С. 519–525. DOI: 10.33619/2414-2948/102/67

9. *Луман Н*. Социальные системы. Очерк общей теории. СПб.: Наука. 2007. 648 с.

10. *Кастельс М*. Информационная эпоха: экономика, общество и культура. М.: ГУ ВШЭ, 2000. 608 с.

11. *Латур Б*. Пересборка социального: введение в акторно-сетевую теорию. М.: Изд. дом Высшей школы экономики, 2014. 384 с.

12. *Каллон М*. Акторно-сетевая теория. М.: Высшая школа экономики, 2014. 10 с.

13. *Knorr Cetina K*. The Scientist as an Analogical Reasoner: A Critique of the Metaphor Theory of Innovation // Social Process of Scientific Investigation. Dordrecht: Reidel, 1981. P. 49–67.

14. *Хабермас Ю.* Моральное сознание и коммуникативное действие. СПб.: Наука, 2001. 380 с.

15. *Деланда М*. Новая философия общества: Теория ассамбляжей и социальная сложность. Пермь : Гиле Пресс, 2018. 170 с.

16. *Глуховский А. С., Дурнев А. Д., Чирва Д. В*. Распределенная моральная ответственность в сфере искусственного интеллекта // Этическая мысль. 2024. Т. 24, №. 1. С. 129–143.

17. *Vaswani A*. Attention is all you need // Advances in Neural Information Processing Systems. 2017. Vol. 30. 11 p.

18. *Dubey A. et al.* The llama 3 herd of models // arXiv preprint. 2024. 92 p.

19. *Team G. et al*. Gemini: a family of highly capable multimodal models // arXiv preprint. 2023. 90 p.

20. *Brown T. B*. Language models are few-shot learners // arXiv preprint. 2020. 75 p.

21. *Wei J. et al*. Chain-of-thought prompting elicits reasoning in large language models // Advances in neural information processing systems. 2022. Vol. 35. P. 24824–24837.

22. *Zhang Yu. et al*. Siren's song in the AI ocean: a survey on hallucination in large language models // arXiv preprint. 2023. 32 p.

23. *Lewis P. et al*. Retrieval-augmented generation for knowledge-intensive nlp tasks // Advances in Neural Information Processing Systems. 2020. Vol. 33. P. 9459–9474

24. *Hu E. J. et al*. Lora: Low-rank adaptation of large language models // arXiv preprint. 2021. 24 p.

25. *Wu Q. et al*. Autogen: Enabling next-gen LLM applications via multi-agent conversation framework // arXiv preprint. 2023. 28 p.

26. *Wang K. et al*. Adapting LLM agents with universal feedback in communication // ICML 2024 Workshop on Foundation Models in the Wild. 2024. 18 p.

27. *Wang A*. Glue: A multi-task benchmark and analysis platform for natural language understanding // arXiv preprint. 2018. 20 p.

28. *Shavrina T. et al*. RussianSuperGLUE: A Russian language understanding evaluation benchmark // arXiv preprint. 2020. 10 p.

29. *Hashemi H. et al*. LLM-Rubric: A Multidimensional, Calibrated Approach to Automated Evaluation of Natural Language Texts // arXiv preprint. 2024. 29 p.

30. *Jung J., Brahman F., Choi Ye*. Trust or Escalate: LLM Judges with Provable Guarantees for Human Agreement // arXiv preprint. 2024. 25 p.

References

1. Brown N. B. Enhancing Trust in LLMs: Algorithms for Comparing and Interpreting LLMs. *arXiv preprint*, 2024, 15 p. (In Eng.).

2. Pappachan P. et al. Transparency and Accountability. In: Challenges in Large Language Model Development and AI Ethics, IGI Global, 2024, pp. 178–211. (In Eng.).

3. Garkusha N. S., Gorodova Yu. S. Pedagogicheskie vozmozhnosti ChatGPT dlya razvitiya kognitivnoi aktivnosti studentov [Pedagogical capabilities of ChatGPT for the development of cognitive activity of students]. *Professional'noe obrazovanie i rynok truda*, 2023. vol. 11, nr 1, pp. 6–23. doi 10.52944/ PORT.2023.52.1.001 (In Russ.).

4. Rezaev A. V., Tregubova N. D. Filosofiya obshcheniya i iskusstvennyi intellekt: opyt sravnitel'nogo analiza diskussii v otechestvennoi i zarubezhnoi literature [Philosophy of communication and artificial intelligence: experience of comparative analysis of discussions in domestic and foreign literature]. *Epistemologiya i filosofiya nauki*, 2024. vol. 61, nr 2, pp. 134–156. (In Russ.).

5. Leshkevich T. G. Iskusstvennyi intellekt v kontekste filosofskogo osmysleniya [Artificial intelligence in the context of socio-humanitarian philosophical understanding]. *Voprosy filosofii*, 2024, nr 5, pp. 50–60. (In Russ.).

6. Eden C. A. et al. Integrating AI in education: Opportunities, challenges, and ethical considerations. *Magna Scientia Advanced Research and Reviews*, 2024, vol. 10, nr 2, pp. 6–13. (In Eng.).

7. Chen H. The ethical challenges of educational artificial intelligence and coping measures: A discussion in the context of the 2024 World Digital Education Conference. *Science Insights Education Frontiers*, 2024, vol. 20, nr 2, pp. 3263–3281. (In Eng.).

8. Chernykh N. A. Obzor, kontseptsii, tochki zreniya k vozmozhnosti pravovogo regulirovaniya tekhnologii iskusstvennogo intellekta [Review, concepts, points of view on the possibility of legal regulation of artificial intelligence technology]. *Byulleten' nauki i praktiki*, 2024, vol. 10, nr 5, pp. 519–525. doi 10.33619/2414-2948/102/67 (In Russ.).

9. Luman N. Sotsial'nye sistemy. Ocherk obshchei teorii [Social systems. An essay on general theory]. St. Petersburg, Nauka, 2007, 648 p. (In Russ.).

10. Kastel's M. Informatsionnaya epokha: ekonomika, obshchestvo i kul'tura [Information Age: Economy, Society and Culture]. Moscow, GU VShE, 2000, 608 p. (In Russ.).

11. Latur B. Peresborka sotsial'nogo: vvedenie v aktorno-setevuyu teoriyu [Reassembling the social: An introduction to actor-network-theory]. Moscow, Izd. dom Vysshei shkoly ekonomiki, 2014, 384 p. (In Russ.).

12. Kallon M. Aktorno-setevaya teoriya [Actor network theory]. Moscow, Izd. dom Vysshei shkoly ekonomiki, 2014, 10 p. (In Russ.).

13. Knorr Cetina, K. The Scientist as an Analogical Reasoner: A Critique of the Metaphor Theory of Innovation. In: Social Process of Scientific Investigation, Dordrecht, Reidel, 1981, pp. 49–67. (In Eng.).

14. Khabermas Yu. Moral'noe soznanie i kommunikativnoe deistvie [Moral consciousness and communicative action]. St Petersburg, Nauka, 2001, 380 p. (In Russ.).

15. Delanda M. Novaya filosofiya obshchestva: Teoriya assamblyazhei i sotsial'naya slozhnost' [New Philosophy of Society: Assemblage Theory and Social Complexity]. Perm', Gile Press, 2018, 170 p. (In Russ.).

16. Glukhovskii A. S., Durnev A. D., Chirva D. V. Raspredelennaya moral'naya otvetstvennost' v sfere iskusstvennogo intellekta [Shared moral responsibility in the sphere of artificial intelligence]. *Eticheskaya mysl'*, 2024, vol. 24, nr 1, pp. 129–143. (In Russ.).

17. Vaswani A. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017, vol. 30, 11 p. (In Eng.).

18. Dubey A. et al. The llama 3 herd of models. *arXiv preprint*, 2024, 92 p. (In Eng.).

19. Team G. et al. Gemini: a family of highly capable multimodal models. *arXiv preprint*, 2023, 90 p. (In Eng.).

20. Brown T. B. Language models are few-shot learners. *arXiv preprint*, 2020, 75 p. (In Eng.).

21. Wei J. et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems,* 2022, vol. 35, pp. 24824–24837. (In Eng.).

22. Zhang Yu. et al. Siren's song in the AI ocean: a survey on hallucination in large language models. *arXiv preprint*, 2023, 32 p. (In Eng.).

23. Lewis P. et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems,* 2020, vol. 33, pp. 9459–9474. (In Eng.).

24. Hu E. J. et al. Lora: Low-rank adaptation of large language models. *arXiv preprint*, 2021, 24 p. (In Eng.).

25. Wu Q. et al. Autogen: Enabling next-gen LLM applications via multi-agent conversation framework. *arXiv preprint*, 2023, 28 p. (In Eng.).

26. Wang K. et al. Adapting LLM agents with universal feedback in communication.*ICML 2024 Workshop on Foundation Models in the Wild*, 2024, 18 p. (In Eng.).

27. Wang A. Glue: A multi-task benchmark and analysis platform for natural language understanding. *arXiv preprint*, 2018, 20 p. (In Eng.).

28. Shavrina T. et al. RussianSuperGLUE: A Russian language understanding evaluation benchmark. *arXiv preprint*, 2020, 10 p. (In Eng.).

29. Hashemi H. et al. LLM-Rubric: A Multidimensional, Calibrated Approach to Automated Evaluation of Natural Language Texts. *arXiv preprint*, 2024, 29 p. (In Eng.).

30. Jung J., Brahman F., Choi Ye. Trust or Escalate: LLM Judges with Provable Guarantees for Human Agreement. *arXiv preprint*, 2024, 25 p. (In Eng.).

Информация об авторах / Information about the authors

Орешкина Татьяна Анатольевна – кандидат социологических наук, доцент кафедры социологии и технологий государственного и муниципального управления Школы государственного управления и предпринимательства Института экономики и управления Уральского федерального университета; ORCID 0000-0003-0004-7709; t.a.oreshkina@urfu.ru

Долганов Антон Юрьевич – доцент кафедры радиоэлектроники и телекоммуникаций Институт радиоэлектроники и информационных технологий Уральского федерального университета; ORCID 0000-0003-2318-9144; anton.dolganov@urfu.ru

Маяцкая Екатерина Александровна – инженер кафедры радиоэлектроники и телекоммуникаций Институт радиоэлектроники и информационных технологий Уральского федерального университета; ORCID 0009-0007-8732-8103; ekaterina.maiatskaia@urfu.ru

Артюгин Олег Юрьевич – исполнительный директор-начальник центра, Центр развития технологий AI во благо общества; ORCID 0009-0006-1888-6630; oyartyugin@sberbank.ru

Tatiana A. Oreshkina – PhD (Sociology), Associate Professor, Department of sociology and technologies of public administration, Institute of public administration and entrepreneurship, Ural Federal University; ORCID 0000-0003-0004-7709; t.a.oreshkina@urfu.ru

Anton Yu. Dolganov – Associate Professor of the Department of Radio Electronics and Telecommunications, Institute of Radio Electronics and Information Technology, Ural Federal University; ORCID 0000-0003-2318-9144; anton.dolganov@urfu.ru

Ekaterina A. Mayatskaya – Engineer of the Department of Radio Electronics and Telecommunications of the Institute of Radio Electronics and Information Technology, Ural Federal University; ORCID 0009-0007-8732-8103; ekaterina.maiatskaia@urfu.ru

Oleg Yu. Artyugin – Executive Director, SberAI; ORCID 0009-0006-1888-6630; oyartyugin@sberbank.ru