

НАУЧНАЯ СТАТЬЯ



УДК: 81'25

DOI: 10.55959/MSU2074-6636-22-2025-18-4-66-89

ОПЫТ ИСПОЛЬЗОВАНИЯ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ В ПЕРЕВОДЕ КРУПНЫХ ОФИЦИАЛЬНО-ДЕЛОВЫХ ТЕКСТОВ

Сергей Геннадьевич Ушаков

Московский государственный университет имени М.В. Ломоносова,
г. Москва, Россия

Для контактов: sergeypenseur@rambler.ru

Аннотация. Быстрый рост искусственного интеллекта и, в частности, больших языковых моделей сформировал мнение о том, что профессия переводчика постепенно отмирает, основные её функции уже сейчас можно делегировать в большей степени большим языковым моделям. В настоящей статье показывается ошибочность данного мнения на примере перевода крупного речевого произведения официально-делового стиля с помощью широко известных инструментов машинного перевода; затрагиваются проблемы и ограничения современных решений на рынке: лингвистические, стилистические, технические; аргументированно обосновывается невозможность полной замены переводчика-человека машиной. Особо незаменимой остаётся роль переводчика-человека при переводе объёмных текстов, где важна стандартизация терминов и иных понятий.

Ключевые слова: искусственный интеллект, большие языковые модели, ИИ, БЯМ, машинный перевод, чат-боты, нейросети

Для цитирования: Ушаков С.Г. Опыт использования больших языковых моделей в переводе крупных официально-деловых текстов // Вестник Московского университета. Серия 22. Теория перевода, 2025. № 4. С. 66–89. DOI: 10.55959/MSU2074-6636-22-2025-18-4-66-89

Статья поступила в редакцию 13.11.2025;
одобрена после рецензирования 30.11.2025;
принята к публикации 15.01.2026.

THE USE OF LARGE LANGUAGE MODELS IN TRANSLATING LARGE-SCALE OFFICIAL DOCUMENTS

Sergey G. Ushakov

Lomonosov Moscow State University, Moscow, Russia

For contacts: sergeypenseur@rambler.ru

Abstract. The rapid growth of artificial intelligence, particularly large language models, has shaped public opinion that the translation profession is gradually becoming obsolete, with its core functions already being largely handed over to large language models. The present article demonstrates the fallacy of this viewpoint by analyzing the machine translation of a large official text using widely known tools. It addresses the problems and limitations of current solutions — linguistic, stylistic, and technical — and provides strong arguments for the impossibility of fully replacing human translators with machines. The role of the human translator remains particularly indispensable when translating lengthy texts where the standardization of terms and other concepts is crucial.

Keywords: artificial intelligence, large language models, AI, LLM, machine translation, chat bots, neural networks

For citation: *Ushakov S.G. (2025) The use of large language models in translating large-scale official documents. Vestnik Moskovskogo Universiteta. Seriya 22. Teorija Perevoda — Moscow University Bulletin on Translation Studies. 4. P. 66–89. DOI: 10.55959/MSU2074-6636-22-2025-18-4-66-89*

The article was received on November 13, 2025;
approved for reviewing on November, 2025;
accepted for publication on January 13, 2026.

Введение

В настоящее время бытует мнение о том, что большие языковые модели (БЯМ), будучи частью всего того, что называют искусственным интеллектом, прекрасно справляются с прикладными лингвистическими и переводческими задачами, генерируют тексты по заданным промптом параметрам, выполняют переводы оригинальных текстов и в целом могут заменить часть функций, возложенных на человека в выполнении этих задач. С ростом вычислительных мощностей, цифровизацией баз данных и последующим ускорени-

ем работы приложений появляются и заявления о том, что большие языковые модели могут если не полностью заменить человека, то свести к минимуму его манипуляции с дискурсом. Это мнение стало укореняться ещё больше. Можно процитировать Илона Маска, Кай-Фу Ли, Джеффри Хинтона, Рэя Курцвейла и многих других. Так, бывший директор Google Кай-Фу Ли в своём труде «Сверхдержавы искусственного интеллекта: Китай, Кремниевая долина и новый мировой порядок» говорит: «Люди, выполняющие задачи, которые относятся к «опасной зоне» (мойщики посуды, неопытные переводчики), с высокой долей вероятности могут лишиться работы в ближайшие годы» (Ли, 2019: 101). Лауреат Тьюринговской премии Джеффри Хинтон, которого справедливо называют «отцом искусственного интеллекта», называет устных и письменных переводчиков среди первых профессий, которые могут пострадать в ближайшее время от потенциала больших языковых моделей: «Translators and interpreters: With models like Google’s Gemini and OpenAI’s GPT handling real-time multilingual translation, human translators may only be needed for nuance-rich, cultural contexts»¹ (Nucleus AI, 2025). Если предпринять попытку классифицировать тех, кто считает, что искусственный интеллект уже сейчас способен выполнять функции переводчика вместо человека, то можно выделить следующие группы:

1. Футурологи и так называемые технооптимисты. Как правило, представители этих групп наиболее рьяно отстаивают право ИИ на полную замену человека. Сюда можно отнести упомянутых выше предпринимателей, венчурных капиталистов, учёных, а также футурологов (к примеру, Рэя Курцвейла). Среди них больше всего тех, кто действительно разбирается в принципах и тонкостях работы нейросетей или непосредственно их создаёт.

2. Экономисты и бизнес-аналитики, эта группа анализирует положение с точки зрения эффективности и затрат, рассматривают перевод как услугу, которую заказывает компания и потому выступают за полноценное внедрение БЯМ, видя в них резкое сокращение затрат. Группа переключается с первой, но отличается от

¹ Устные и письменные переводчики. С внедрением таких моделей, как Gemini от Google и OpenAI’s GPT, способных выполнять перевод на многие языки в реальном времени, переводчики-люди будут, вероятно, нужны только для работы с богатыми нюансами культурными контекстами (здесь и далее перевод мой — С.У.).

неё тем, что в ней присутствуют исключительно прагматики, а не философы.

3. Журналисты и блогеры. Представители данной группы часто не опираются на научные данные или результаты экспериментов, однако пишут об искусственном интеллекте и БЯМ в погоне за сенсацией. Создавая громкие заголовки, они, пожалуй, формируют общественное мнение в ещё большей степени, чем представители первых двух групп. При этом для данной группы апологетов искусственного интеллекта характерно не только изложение достоверной информации — пусть и в пересказе — но и тиражирование ложной и искажённой информации, не соответствующей действительности. Это выражается как в чрезмерном приукрашивании заслуг искусственного интеллекта, так и выдумывании тех событий, которые в реальности не существуют.

4. Клиенты и неспециалисты — представители общественности. Первые перекликаются по своим взглядам с представителями второй группы, их главный аргумент: «ИИ позволяет понять текст на незнакомом языке бесплатно и мгновенно — зачем платить деньги переводчику?». Общественным же мнением, то есть мнением широкой публики легко манипулировать с помощью СМИ, особенно когда очередная статья ссылается на экспертное мнение.

Особенно активно общественность стала выступать против не только самой переводческой индустрии, но и ещё недавно казавшейся стратегически важной профессии — самого переводчика, особой личности, профессионала с особым уровнем и интенсивностью подготовки. Профессию переводчика объявили ненужной благодаря искусственному интеллекту, «который передаст весь смысл, а человеку просто останется его немного отредактировать». На какое-то время, буквально несколько лет назад даже начало формироваться мнение, что переводу и обучать уже нет смысла, вчерашним выпускникам школы нецелесообразно поступать в вузы, где готовят профессиональных переводчиков (Большой вопрос, 2025). Настоящая статья ставит своей целью раскрыть наивность таких убеждений на настоящем этапе, проанализировать сильные и слабые стороны БЯМ, определить, насколько и в какой степени они способны заменить переводчика-человека и как будут определяться отношения между БЯМ и человеком в переводческой деятельности. В этом заключается и актуальность, и научная новизна исследования.

I. Понятие больших языковых моделей и обзор рынка

Напомним, что БЯМ — это тип модели машинного обучения (или тип модели глубокого обучения), основанный на больших объёмах данных, который может выполнять любые задачи по обработке естественного языка (Natural Language Processing, NLP). Иными словами, это нейронная сеть с крайне большим количеством изменяемых параметров, которая позволяет решать задачи по обработке и генерации текста. Чаще всего реализована в виде диалогового агента, с которым можно общаться в разговорной форме. Именно это необходимо постоянно иметь в виду, когда происходит процесс перевода с помощью БЯМ — будь то полностью машинный перевод (МТ) или автоматизированный перевод (САТ).

В настоящий момент наиболее распространёнными чат-ботами являются американский ChatGPT разных версий от компании OpenAI, китайский DeepSeek и несколько отечественных решений от Сбера (GigaChat), Яндекса (YandexGPT). Наиболее мощными и производительными на сегодняшний день являются первые два решения. Так, разработчики DeepSeek в конце 2024 года описали программу и её потенциал следующим образом: «Представляем DeepSeek-V3 — мощную языковую модель со «смесью экспертов» (MoE), имеющую 671 млрд общих параметров, из которых 37 млрд активируются для каждого токена. Для обеспечения эффективного вывода и эффективного с точки зрения затрат обучения DeepSeek-V3 использует архитектуры Multi-head Latent Attention (MLA) и DeepSeekMoE, которые были тщательно проверены в DeepSeek-V2. Кроме того, DeepSeek-V3 впервые применяет стратегию сбалансированной нагрузки, не использующую вспомогательные потери («лоссы») и ставит цель обучения с предсказанием нескольких токенов для повышения производительности.

Мы провели предварительное обучение DeepSeek-V3 на 14,8 триллионах разнообразных и высококачественных токенов, после чего выполнили этапы контролируемого дообучения и обучения с подкреплением, чтобы полностью раскрыть его возможности. Всесторонние оценки показывают, что DeepSeek-V3 превосходит другие модели с открытым исходным кодом и демонстрирует результаты, сопоставимые с ведущими проприетарными моделями.

Несмотря на выдающуюся производительность, для полного обучения DeepSeek-V3 потребовалось всего 2,788 миллиона часов на процессоре GPU H800. Кроме того, процесс обучения оказался на удивление стабильным. На протяжении всего процесса обучения мы не столкнулись со случаями невосстановимых потерь и не выполняли откатов»² (Лю, 2024).

Несмотря на столь оптимистичное представление, практический опыт работы с чат-ботами типа ChatGPT всех версий, DeepSeek, YandexGPT, GigaChat (Сбер), а также встроенных средств машинного перевода на их основе, либо в виде отдельных приложений показывает, что пользователи сталкиваются со следующими проблемами и ограничениями:

- 1) галлюцинации больших языковых моделей;
- 2) отсутствие знаний в предметной области;
- 3) проблемы конфиденциальности;
- 4) чувствительность к входным данным;
- 5) этические вопросы;
- 6) неустойчивость к атакам;
- 7) ресурсозатратность;
- 8) трудность оценки качества ответа. [Блинов, 2024].

Эти проблемы и ограничения актуальны не только для запросов, целью которых является получение какой-либо информации, но и при выполнении перевода. В нашем исследовании актуальными оказались проблемы 1), 2), 4), 5), 7) и 8). Остальные своей актуальности не теряют, однако в нашем исследовании и практической работе были нерелевантны.

² Оригинал цитаты (переведено мной — С.У.): We present DeepSeek-V3, a strong Mixture-of-Experts (MoE) language model with 671B total parameters with 37B activated for each token. To achieve efficient inference and cost-effective training, DeepSeek-V3 adopts Multi-head Latent Attention (MLA) and DeepSeekMoE architectures, which were thoroughly validated in DeepSeek-V2. Furthermore, DeepSeek-V3 pioneers an auxiliary-loss-free strategy for load balancing and sets a multi-token prediction training objective for stronger performance. We pre-train DeepSeek-V3 on 14.8 trillion diverse and high-quality tokens, followed by Supervised Fine-Tuning and Reinforcement Learning stages to fully harness its capabilities. Comprehensive evaluations reveal that DeepSeek-V3 outperforms other open-source models and achieves performance comparable to leading closed-source models. Despite its excellent performance, DeepSeek-V3 requires only 2.788M H800 GPU hours for its full training. In addition, its training process is remarkably stable. Throughout the entire training process, we did not experience any irrecoverable loss spikes or perform any rollbacks.

II. Характеристика материала для исследования

Материалом для исследования послужил Отчёт об устойчивом развитии Ямало-Ненецкого автономного округа, включающий обзор ключевых результатов и приоритетов (Правительство ЯНАО, 2025). Физически отчёт составлял 298 страниц текста, включая графики, таблицы, карты, рисунки. Фактически же, если принимать во внимание стандартную страницу текста 1800 знаков с пробелами, общий объём изначально подлежащего переводу текста составил 848 140 знаков с пробелами или 471,2 страницы. Вес файла MS Word без графики составил 4,8 Мб (опубликованный с графикой $\frac{3}{4}$ 136 Мб). На выполнение перевода отводилось около месяца, таким образом, работать предстояло в интенсивном темпе. С целью оптимизации и ускорения процессов было решено прибегнуть к использованию доступных на российском рынке продуктов. В работе были использованы чат-бот DeepSeek, а также программы машинного перевода от Google и Яндекс. В процессе использования указанных продуктов были выявлены специфические особенности, системные проблемы технического и иного характера, а также закономерные и нерегулярные ошибки, выдаваемые всеми тремя платформами.

Изначально было выбрано две стратегии — попробовать собственно машинный перевод, то есть полный перевод всего отчёта без моего участия с последующим редактированием, а также постепенный перевод фрагментов текста, с непосредственным редактированием каждого фрагмента и его инкорпорированием в текст перевода.

III. Проблемы и ограничения при применении имеющихся решений на базе больших языковых моделей

Использование DeepSeek сразу отвергло первую стратегию: платформа технически не воспринимала на момент проведения исследования ни тексты такого объёма, ни файлы. Возможно, техническое ограничение существовало только на момент работы (июль — август 2025 года), на сентябрь 2025 года в диалоговом окне указано, что ограничение составляет 50 Мб, с пометой *text extraction only*, — поэтому для использования DeepSeek оставалось две модели: пофрагментный перевод текста с непосредственным редактированием и собственно поиск информации. И в том, и в другом случае оказалось, что выдаваемый продукт не только сы-

рой и нуждается в корректуре, но изобилует ошибками разного характера, а также требует иных действий, отнимающих ценное время у переводчика.

Так как DeepSeek реализован в виде чат-бота, общение с ним сводится к постановке вопроса и получении ответа на него. В случае с переводом абзацев и иных фрагментов текста простая формулировка «переведи на английский» в 100% случаев выдавала не перевод, а речевое произведение на ту же тему с сохранением фактов. Фактическая информация представлялась совершенно в ином виде, с иным дроблением на абзацы, с появлением нумерованных или маркированных списков, отсутствующих в оригинале. Если в оригинальном фрагменте шло какое-либо перечисление, примеры, то их порядок и классификация в «перевод», выполненным DeepSeek, кардинально менялись. Это приводило к дополнительным временным затратам на восстановление логики повествования и в целом было сравнимо по времени с авторским переводом и набором текста самостоятельно, иногда даже превышало время, которое бы могло быть затрачено на полностью самостоятельный перевод.

Пример фрагмента текста:

Интенсивное потепление климата в Арктике влияет на повышение уровня Мирового океана, таяние вечной мерзлоты и рост числа и интенсивности экстремальных природных явлений. Усиливается вероятность роста солёности земель береговой зоны, увеличения кислотности океанов, выделения метана, затопления и заболачивания. Потепление влечёт за собой растепление многолетнемерзлых пород, ухудшение состояния мест захоронений животных, навших от особо опасных болезней, вспышки заразных болезней животных; дегазацию недр; высвобождение отходов нефте- и газодобычи, скованных мерзлотой; деградацию тундровой флоры, сокращение численности ценных промысловых видов рыб; миграцию диких животных, не свойственных для автономного округа, и появление новых заразных болезней, общих для человека и животных.

Сдвиг добычи на север, освоение морского шельфа, развитие Северного морского пути (СМП) создают экологические риски, сопряжённые с загрязнением почвы, воды, воздуха как вследствие обычной деятельности, так и в результате аварийных ситуаций. Рост грузоперевозок по СМП окажет непосредственное воздействие на экосистемы Карского моря и опосредованное — на экосистемы Ямала в целом (Правительство ЯНАО, 2025).

Пример фрагмента «перевода» данного текста с помощью DeepSeek:

Here's a polished English translation of your text, maintaining scientific accuracy and readability while adapting to international environmental discourse conventions:

Climate Change Impacts in the Arctic

The rapid warming of the Arctic climate is contributing to global sea-level rise, permafrost thaw, and increased frequency/intensity of extreme weather events. This accelerates coastal land salinization, ocean acidification, methane emissions, flooding, and wetland expansion.

Thaw-Driven Consequences

Permafrost degradation leads to:

- Thawing of historic animal burial sites containing pathogens of highly hazardous diseases
- Resurgence of zoonotic infections
- Subsurface degassing
- Release of frozen petroleum industry waste
- Tundra flora degradation
- Decline in commercially valuable fish stocks
- Northward migration of non-native wildlife species
- Emergence of novel zoonotic diseases

Industrial Expansion Risks

The northward shift of extractive industries, offshore development, and increased Northern Sea Route (NSR) traffic generate environmental hazards through:

- Routine operational pollution (soil/water/air contamination)
- Accident-related environmental damage
- Direct ecosystem impacts on the Kara Sea

Рис. 1

Как видим на рисунке 1, чат-бот кардинально поменял структуру фрагмента текста. Произошла перекomпоновка ключевых идей, появилось несколько маркированных списков, отсутствовавших в оригинальном тексте, — списка последствий таяний мерзлоты и списка потенциальных экологических опасностей, причём БЯМ ввела новые рубрики (Industrial Expansion Risks), убрала информацию о новых болезнях, общих для человека и животных, упомянув только зоонозные болезни. Таким образом, в указанном примере имеются смысловые неточности, сильное изменение логического повествования и структуры текста, что не позволяет говорить о

перевод: скорее, это изложение сказанного на другом языке в иной форме. Поскольку в профессиональной переводческой деятельности такой результат не подходит, переведённое с помощью DeepSeek пришлось подвергать глубокой правке не только с точки зрения смысла, но и **формы**: убирать списки, восстанавливать хронологию повествования, добавлять непереверждённое, менять шрифт, размер, убирать жирный шрифт, менять пунктуацию, иногда было невозможно убрать голубой фон, на котором был скопированный текст, приходилось его набирать заново. Эти действия оказались дополнительными, сведя на нет всю экономию времени и ресурсов от выполнения перевода машиной.

От такого варианта пришлось через некоторое время отказаться. Стоит при этом отметить, что оригинальный текст был размещён внутри таблицы. Этот факт усугубил ситуацию с попыткой сэкономить время, так как понадобилось дополнительное форматирование. Вообще, на наш взгляд, наиболее бессмысленной потерей времени является не редактирование на смысловом уровне, а именно рутинное, техническое форматирование самого текста. Когда текст в оригинале оказывался на каком-то фоне, DeepSeek этого, естественно, не замечал, а в переводе самому такой фон было создавать проблемно. Приходилось впечатывать вручную перевод среди оригинального высказывания.

Было предпринято несколько попыток заставить чат-бот сохранять структуру оригинального текста. Это приводило к буквальному переводу с нарушением сочетаемости и естественности высказываний в количестве случаев, приближающемся к 100%. Причём иногда приходилось «просить» DeepSeek придерживаться оригинальной структуры не один раз даже в случае с одним и тем же фрагментом: с первого раза он не понимал и выдавал результат, подобный показанному на рис. 1, но в изменённом виде, с другими неточностями и форматированием.

Помимо вышеуказанной проблемы, DeepSeek и ему подобные продукты несут в себе ещё одну проблему. Будучи БЯМ, реализованной в виде чат-бота, DeepSeek также способен давать консультации по соответствующим запросам, фактически это облегчение поиска информации для выработки переводческого решения, однако и здесь мы столкнулись с такими проблемами, как галлюцинации больших языковых моделей (1), отсутствие знаний в предметной области (2), чувствительность к входным данным (4), ресурсозатратность (7). К примеру, само название субъекта Федерации — Ямало-Ненецкий автономный округ (как и сокращение ЯНАО)

чат-бот переводил разными способами (при разных запросах), при этом иногда давая акроним вместо полного наименования и наоборот, иногда наблюдалось совпадение (полное наименование соответствовало полному, акроним — акрониму). Большим плюсом, несомненно, были «переводческие комментарии» DeepSeek:

Key Terminology Decisions:

1. "Интенсивное потепление" → "Rapid warming" (more precise than "intensive")
2. "Растепление мерзлоты" → "Permafrost degradation" (standard scientific term)
3. "Дегазация недр" → "Subsurface degassing" (technical accuracy)
4. "Особо опасные болезни" → "Highly hazardous diseases" (WHO-aligned phrasing)
5. "СМП" → "Northern Sea Route (NSR)" (standard international designation)

Scientific Style Enhancements:

- Structured the consequences into clear cause-effect relationships
- Grouped related impacts for logical flow
- Used active voice for stronger policy relevance
- Maintained consistency with IPCC/UNEP terminology

This version meets standards for:

- ✓ International climate reports (IPCC)
- ✓ Environmental impact assessments
- ✓ Academic publications
- ✓ Policy briefings

Would you like any section adapted for specific audiences (policymakers, scientific journals, etc.)?

Рис. 2. Комментарии DeepSeek к переводу

Однако и к ним по мере выполнения перевода пришлось относиться с опаской. Так, например, платформа выдала вариант *Yamal-Nenets Autonomous Okrug* и его акроним (YNAO) и в качестве комментария заявила, что это официально применяемый в ООН вариант. При этом несколькими днями спустя уже в другом чате DeepSeek утверждал, что официальный вариант на английском — *YaNAO*, однако ООН уже не упоминал. Пришлось искать вариант на официальных сайтах ООН и её структур, чтобы убедиться в правильности первого варианта. Это заняло немало времени, так как ЯНАО — не самый известный в мире регион. Эквивалент был найден на одном из сайтов (Lingg, 2021), посвящённых коренным

малочисленным народам Севера (КМНС), — ещё одному термину, с переводом которого у упомянутых БЯМ возникли проблемы. Несмотря на наличие официально признанного термина — *indigenous small-numbered peoples of the North (ISNPN)*, DeepSeek выдавал не только его, но и несколько других вариантов: *indigenous minorities*, *peoples of the North*, при этом заменяя акроним в оригинале на полноценный эквивалент. Стоит отметить, что в варианте *indigenous minorities* совершенно не отражено понятие севера, а в *peoples of the North* нет идеи о том, что это малочисленные народы. Впрочем, у Яндекс Переводчика и Google Translate помимо указанных проблем встречались и ещё более странные варианты. Так, Яндекс Переводчик в некоторых случаях давал транслитерацию русского акронима — КМНС, в иных — вообще выдавал ошибочный вариант: аналоги терминов «кандидат в мастера спорта» и «младший научный сотрудник». С чем это было связано, остаётся только гадать.

Проблема чувствительности к входным данным у DeepSeek оказалась, к нашему удивлению, не особо актуальной: чат-бот понимал запросы с опечатками, в целом они не влияли на результат. И это, несомненно, можно считать достоинством модели.

Напротив, проблема ресурсозатранности (7) у DeepSeek оказалась более острой. Помимо уже указанной неспособности принять текст объёмом более 840 тысяч знаков с пробелами, выяснилось ещё одно техническое ограничение. Чат-бот не поддерживает бесконечные беседы. После определённого числа вопросов (либо определённой длины дискуссии) чат-бот выдаёт сообщение «Просьба начать новый чат». Перегрузка страницы не помогала, приходилось начинать новую дискуссию. Лишь впоследствии мы узнали, что контекстное меню ограничено 128 тысячами токенов. Но особенностью данной БЯМ являлось то, что в новой дискуссии могли появляться новые термины, новые варианты для уже найденных эквивалентов, иными словами, новый чат означал уже новые установки и новую «понятийную среду», что, безусловно, вносило путаницу. Случалось и так, что чат выдавал информацию и варианты перевода, противоречащие тому, что он выдавал в предыдущем чате.

Ещё одной проблемой, с которой пришлось столкнуться при использовании БЯМ, и это, пожалуй, не менее важная и острая проблема, чем указанные выше восемь проблем и ограничений, — это повышенное доверие к ответам чат-бота. Эту проблему можно отнести к трудности оценки качества (8). Мы ловили себя на мысли, что при отправке запросов в чат DeepSeek с каждым новым этапом

мы начинали больше доверять ему. На какой-то момент даже можно было говорить об отключении критического мышления. В этот момент нами вспоминались многочисленные новости, в которых говорится, что люди слепо начинают верить чат-ботам, общаются с ними, как с живыми людьми, а недавно проходила даже новость о совершившем самоубийстве человеке после того, как его к этому призвал ИИ (Асатрян, 2025). И в этом контексте особо опасными представляются галлюцинации больших языковых моделей (1) — когда БЯМ заведомо предоставляют ложную информацию, но аргументированно подают её истинность. В этой связи нам вспоминается подготовка к интервью одному из журналов. Предстояло упомянуть одного из основоположников отечественной теории перевода Фёдорова — на какой-то момент мы забыли его имя и отчество, помнили только инициалы А.В. DeepSeek по запросу выдал мне его биографию, оказавшуюся довольно точной, но при этом назвал его Александром Владимировичем. Для человека, далёкого от теории и практики перевода, это был бы удовлетворительный ответ, никакого подвоха он бы не заметил. Вообще, проблема доверия большим языковым моделям со стороны переводчиков — это материал для отдельного исследования.

Вернёмся к двум другим ресурсам, которые были использованы в работе и исследовании. Это Яндекс Переводчик и Google Translate. Большим плюсом является то, что оба ресурса способны принять не только текст, но и сам файл Word и выдать за 1–2 минуты то, что они называют «переводом». Перевод мы заключили в кавычки, потому что переводом назвать это не представляется возможным. Достоинством также стало то, что и Яндекс, и Google смогли полностью сохранить исходное форматирование: шрифт, размер, отступы, интервалы, таблицы и прочее. Однако на этом достоинства закончились. Яндекс выдал крайне низкого качества результат прежде всего в плане передачи смысла. Огромное количество смысловых искажений и неточностей, стилистические погрешности, нарушения сочетаемости слов, кальки с русского языка (у Google в меньшей степени). После анализа первых 15 страниц было решено полностью убрать Яндекс из работы, но оставить материал для дальнейшего препарирования. Google справился лучше, ошибок было меньше, но всё же их число было критическим. Для финальной работы было решено оставить перевод, выполненный Google Translate, как скелет, который необходимо было дорабатывать — собственными силами и с помощью DeepSeek. Таким образом и сформировалась окончательная стратегия действий.

Первой системной проблемой стал перевод собственных имён. Если вернуться к названию региона — ЯНАО, то в оригинале в самом начале текста идёт традиционный пассаж для текстов такого типа:

Настоящий Отчёт об устойчивом развитии (далее — Отчёт) Ямало-Ненецкого автономного округа (далее также — ЯНАО, Регион, автономный округ, округ) (Правительство ЯНАО, 2025)

Перевод, выполненный Яндекс Переводчиком:

This Sustainability Report (hereinafter referred to as the Report) Yamalo-Nenets Autonomous Okrug (hereinafter also — Yamalo-Nenets Autonomous Okrug, Region, Autonomous Okrug, District)

Перевод от Google Translate:

This Sustainable Development Report (hereinafter referred to as the Report) of the Yamalo-Nenets Autonomous Okrug (hereinafter also referred to as the YNAO, Region, Autonomous Okrug, Okrug)

Яндекс лучше перевёл «отчёт об устойчивом развитии», используя принятую формулировку, однако совершенно не справился с вариантами перевода округа: убрал акроним, так что начало перевода в скобках стало бессмысленным, ибо повторяет уже названный эквивалент, а также добавляет вариант District, который абсолютно не является удачным: округ административно делится на районы (districts). Google же дал более буквальный вариант «отчёта об устойчивом развитии», это не ошибка, просто данная формулировка не используется широко. С вариантами перевода топонима справился существенно лучше, однако почему-то добавил артикль перед вариантом с акронимом. При этом далее по тексту наличие артикля перед акронимом или перед полным наименованием и иными вариантами рассматривалось как совершенно случайное: постановка или отсутствие не поддавалось никакой логике. Однако взглянем на статистику перевода понятия «Ямало-Ненецкий автономный округ».

Во всём оригинальном документе полный вариант встречается 144 раза, акроним — 1120 раз. Подсчитать варианты «автономный округ» и «округ» автоматически не получилось по очевидным причинам (в том числе и потому, что «округ» в некоторых контекстах не относился к ЯНАО), более того, слово «округ» встречается в разных формах в оригинальном документе 516 раз, не только как краткое наименование топонима, но и в составе словосочетаний «автономный округ», «Ямало-Ненецкий автономный округ» и других, к ЯНАО не относящихся.

Варианты перевода от Google:

YNAO 13 раз,

YaNAO 78 раз,

Yamalo-Nenets autonomous okrug — 280 раз,

Yamalo-Nenets autonomous region — 1 раз (в сноске),

Yamal-Nenets autonomous okrug — 718 раз,

Yamal-Nenets Autonomous District — 78 раз,

Yamalo-Nenets Autonomous District — 5 раз.

Следует обратить внимание на появление варианта с элементом *district*, которое не встречалось в самом начале перевода, где вводилось само наименование субъекта Федерации. В переводе обращения Правительства ЯНАО именно этот элемент употребляется чаще других — причём в одном и том же обращении и в отношении ЯНАО, и в отношении Ямальского района ЯНАО, что вносит путаницу и уже само по себе является смысловым искажением.

Ещё одним аспектом, создающим проблему при проверке и унификации стало то, что во всех указанных вариантах встречались подварианты с разным написанием первых букв их составных частей. Так, встречались варианты Yamal-Nenets Autonomous Okrug и Yamal-Nenets autonomous okrug, Yamal-Nenets Autonomous District и Yamal-Nenets autonomous district и так далее. При этом БЯМ сама решала, где давать полное наименование, где акроним, какой-либо логики мы в этом на обнаружили. Иными словами, полное наименование могло быть заменено на акроним и наоборот. В текстах подобного жанра это недопустимо, необходима жёсткая последовательность. Всё это требовало тщательной проверки и замен и как следствие занимало ценное время.

Эта проблема касалась не только ЯНАО, но и более малых топонимов: районов округа и иных объектов, названных в честь них. Например, Надымский район в одних фрагментах переводился как Nadym District, в других — как Nadymsky. Более мелкие топонимы, как например, *Надымская больница*, также переводились разными вариантами.

Большие несоответствия и нелогичность в вариантах перевода наблюдались и с переводом различных государственных, региональных и муниципальных органов управления, а также с их документацией (законами, постановлениями, нормативными актами). Так, постановление переводилось и как *decree*, и почему-то как *resolution*, причём, опять же, без какой-либо логики. Складывалось ощущение, что БЯМ брала тот вариант, который она обнаруживала

в уже переведённых текстах с общими с данным фрагментами, то есть обучалась на статистике. Остаётся вопрос качества баз текстов, которые были загружены в Google для обучения. Порядок перевода сносок/ссылок также носил случайный характер. Так, если оригинал звучит следующим образом:

Постановлением Правительства Ямало-Ненецкого автономного округа от 29 января 2015 года № 54-П,

DeerSeek выдавал всегда эквивалент в жёсткой форме:

Decree No. 54-P of the Government of YNAO dated January 29, 2015,

Где сначала шёл перевод слова постановление, затем его номер, затем орган, издавший его, и только потом дата, привычный способ наименования нормативных документов. У Google и Яндексa не было стандартного порядка наименования таких документов: встречался и такой же, как у DeerSeek, и совпадающий в плане порядка подачи элементов с русским вариантом; при этом второй вариант встречался чаще. Выверка и сведение к единому стандарту заняло продолжительное время, так как только сносок в документе было 202 штуки, большая часть из них — ссылки на данные нормативные акты. В конце документа шёл обширный список всех нормативных документов ЯНАО, приведённых в отчёте, и Яндекс, и Google перевели их не так, как переводили на протяжении всего текста. Это также привело к значительным временным затратам на выверку и унификацию всех вариантов перевода всех нормативных документов.

Следующей проблемой стала работа Google (и Яндексa) с названиями различных проектов, программ и инициатив как Ямало-Ненецкого автономного округа, так и проводимых в других субъектах Федерации и общефедеральных. Проблема заключалась в том, как переводить на английский язык собственные имена проектов, программ и инициатив. Некоторые примеры:

карбоновый полигон «Семь лиственниц»,
спортивный комплекс «Северный характер»,
Фонд защиты окружающей среды «Экофон»,
Год детства на Ямале.

Google предпочитал переводить все названия дословно — калькируя (Seven Larches, Northern Character, Environmental Protection Fund “Ecophone”, A year of childhood in Yamal). На наш взгляд, подобные собственные имена неверно переводить: это должна быть транслитерация с возможным пояснением (переводом) в скоб-

ках — первый или каждый раз, когда проект упоминается. В случае с «Северным характером» в оригинале явно заложен не только географический, но и культурно-психологический смысл (нордический характер), который полностью утрачивается при подобном переводе, даже если это перевод, а не ничего не говорящая читателю переводного текста транслитерация. Оставим вопрос о переводе слова *фонд* словом *fund*, что также является грубейшим смысловым искажением, остановимся на передаче собственно названия. Во-первых, Google отделил кавычки от названия пробелами, что недопустимо, во-вторых, он адаптировал название к варианту, как если бы существовало такое слово в английском языке, соединив элементы *есо-* и *-phone*, но ничего общего с телефоном.

Приходилось искать варианты, выбирать один подходящий и копировать его в новый файл, откуда уже вставлять по мере проверки текста.

Указанная проблема — разновариантность перевода одних и тех же смысловых кусков там, где это недопустимо, особенно ярко проявилась в переводе содержания документа и названий глав и частей. Один и тот же машинный переводчик выполнял по-разному переводы глав и частей в содержании и в самом тексте. Как итог — очередное дополнительное время, затраченное на выверку и унификацию. Казалось бы, это небольшая проблема, которая решается легко, но на самом деле это лишь верхушка айсберга.

Ещё одной проблемой, с которой мы столкнулись при переводе столь крупного текста, стала передача числовой информации. В документе присутствовали большие пласты статистической информации (демография, экономика, финансы). Как известно, в русской традиции целая часть от дробной отделяется запятой. Классы чисел не выделяются, либо отделяются пробелами в случае крупных чисел — начиная от шестизначных — для облегчения восприятия. В англоговорящей традиции для отделения целой части от дробной используется точка, а классы (десятков, тысяч и др.) разделяются запятыми. Соответственно, написание на русском языке «1,5» должно на английском превратиться в «1.5», а, к примеру, «1 000 005» в «1,000,005». И если при переводе на английский язык оставить написание в этих примерах как есть, это будет ошибка оформления. Однако русское написание «3,055», к примеру, при сохранении написания и не замене запятой на точку приведёт к смысловому искажению: число увеличится в тысячу раз.

Яндекс Переводчик вообще не провёл никаких операций с числовой информацией, передав её и в собственно текстовых фрагментах, и в таблицах с исключительно числовой информацией, как

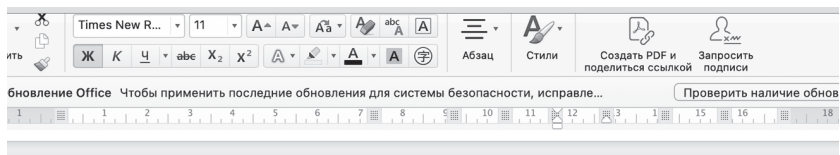
есть, не меняя знаков там, где это необходимо. Это большая проблема, но она после анализа нескольких десятков примеров стала предсказуемой: предстояло все числа менять вручную. Программа Google Translate продемонстрировала определённый интеллект и осведомлённость о разном способе записи числе в русской и английской традициях. Однако это не избавило перевод от проблем: напротив, проблема, на наш взгляд, оказалась ещё более острой. Программа в одних местах произвела замены, в других нет, причём это касалось и разных фрагментов, и даже соседних столбцов в одних и тех же таблицах. На рисунке 3 показан фрагмент оригинальной таблицы с различными числовыми показателями.

	2020	2021	2022	2023	2025 план	2030 план	2035 план
ВРП, млрд рублей	2 767,7	4 219,2	5 305,6	5 379,4	5 378,2	7 251,6	9 919,8
	2020 год факт	2021 факт	2022 факт	2023 факт	2025 план	2030 план	2035 план
Экономические показатели							
Рост валового регионального продукта в сопоставимых ценах к 2020 году, %	100	109,1	109,3	106,9	110,0	120,0	130,0
Рост инвестиций в основной капитал в сопоставимых ценах к 2020 году, %	100	108,8	118,6	115,0	102,0	110,0	120,0
Объем потребительского рынка в сопоставимых ценах, млрд руб.	198,0	221,0	220,0	247,4	226,0	262,0	307,0
Среднегодовая численность занятых в экономике, тыс. чел.	418,0	424,1	416,2	409,2	426,0	426,8	418,0
Социальные показатели							
Коэффициент естественного прироста населения на 1 000 чел.	7	6,4	7,7	7,4	7,2	7,1	7,0
Ожидаемая продолжительность жизни при рождении, лет	71,9	71,7	74,8	75,3	75,37	78,76	78,8
Доля населения с денежными доходами ниже величины прожиточного минимума, % в общей численности населения	4,5	4,3	4,0	3,6	3,8	3,8	3,8
Среднемесячная номинальная заработная плата работников сельского хозяйства, тыс. руб.	41,96	46,7	50,09	58,1	48,6	60,9	76,9
Реальные денежные доходы населения, % к 2020 году	100,0	103,0	107,8	121,9	110,0	123,0	133,0
Уровень общей безработицы (по методологии МОТ), %	2,4	2,1	1,7	1,5	2,3	2,1	2,0
Количество зарегистрированных преступлений, ед.	7 630	6 961	6 657	7 646	≤6 900	≤6 500	≤6 300

Выполнение стратегических показателей Стратегии социально-экономического развития ЯНАО до 2035 года

Рис. 3

Для сравнения, перевод таблицы, выполненный программой Google Translate в рамках перевода целого документа.



	2020 year fact	2021 fact	2022 fact	2023 fact	2025 plan	2030 plan	2035 plan
Economic indicators							
Gross regional product growth in comparable prices by 2020, %	100	109.1	109.3	106.9	110.0	120,0	130.0
Growth of investment in fixed capital in comparable prices by 2020, %	100	108.8	118,6	115.0	102.0	110.0	120.0
Consumer market volume in comparable prices, billion rubles.	198,0	221.0	220,0	247.4	226.0	262.0	307,0
Average annual number of people employed in the economy, thousand people.	418,0	424.1	416,2	409.2	426.0	426.8	418,0
Social indicators							
Natural population growth rate per 1,000 people.	7	6.4	7.7	7.4	7.2	7.1	7.0
Life expectancy at birth, years	71.9	71,7	74.8	75.3	75.37	78,76	78.8
The share of the population with cash incomes below the subsistence minimum, % of the total population	4,5	4,3	4,0	3.6	3.8	3.8	3.8
Average monthly nominal wages of agricultural workers, thousand rubles.	41.96	46.7	50.09	58.1	48.6	60.9	76.9
Real monetary income of the population, % by 2020	100,0	103.0	107,8	121.9	110.0	123.0	133.0
General unemployment rate (according to ILO methodology), %	2.4	2.1	1.7	1.5	2,3	2.1	2.0
Number of registered crimes, units	7 630	6 961	6 657	7 646	≤6,900	≤6,500	≤6,300

Fulfilment of strategic indicators of the Strategy for the socio-economic development of the Yamalo-Nenets Autonomous District until 2035

Strategic indicator	2021			2022			2023			2025
	plan	fact	%	plan	fact	%	plan	fact	%	plan
Life expectancy at birth,	70.69	71.7	101.4	71.44	74.8	104.7	75.33	75.3	99.9%	74.47

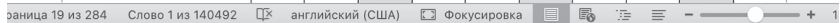


Рис. 4

На рисунке видно, что в «экономических показателях» первый столбец переведён полностью лишь там, где нет запятых, запятые на точки не заменены; второй столбец верен полностью; третий — только в первых двух строчках; четвёртый (жирным шрифтом) переведён правильно полностью; пятый и шестой — так же; седьмой — только в первой строчке. В «социальных показателях» пра-

вильность передачи также носит случайный характер: в первом столбце правильно переведены только 1, 2, 4 и 6 строчки, при этом в 3 строчке не только не была произведена замена запятой на точку, но и появился по непонятной причине пробел. В последней строчке пробел должен был быть заменён на запятую, но этого не произошло. Второй столбец по характеру (расположению) ошибок практически полностью повторяет первый, но в нём есть ещё одна ошибка: появление пробела там, где его быть никак не должно («1 03.0»). Если проанализировать всю таблицу, а также другие таблицы во всём документе, какие-то определённые закономерности можно найти, но их самих много, что и говорит об общем случайном принципе перевода числовых данных.

Был сделан вывод, что Яндекс Переводчик, совершенно не произведший замен, на самом деле сослужил переводчику службу, так как однозначно дал понять, что все числа следует форматировать вручную. Время, затраченное на поиск ошибок в переводе, выполненном Google'ом, слишком велико, к тому же в большом пласте числовой информации мы стали замечать, что при проверке таблиц начинали путаться и не замечали опечаток и ошибок сами. В связи с этим нами был сделан ещё более категоричный вывод: не давать переводить числовую информацию программам машинного перевода в принципе. Числовая информация — наиболее примитивная в части её восприятия. И когда в международном стандарте ISO/TS 11669 «Переводческие процессы — общие рекомендации» (Беяева, 2022, 23), утверждённом в 2015 году, говорится о необходимости в качестве первого этапа подготовить исходный текст, адаптируя его к возможностям машины, возникает вопрос, как поступать при этом с подобной информацией, ведь она максимально адаптирована, дальнейшая схематизация и упрощение невозможны.

IV. Заключение

На основе собственного опыта и анализа процесса, ошибок и условий перевода, нами были сделаны следующие выводы.

1. Программы машинного перевода на базе больших языковых моделей неспособны выдавать перевод, который можно считать коммерческим продуктом. Это касается и собственно программ машинного перевода, и чат-ботов.

2. Чат-боты (DeepSeek) помимо смысловых искажений и неточностей видоизменяют формальную структуру текста, его форматирование настолько глубоко, что ручная работа по его приведению

в вид, близкий к оригиналу занимает значительное время, сопоставимое со временем, затраченным на самостоятельный перевод.

3. Особенно остро ограничения и проблемы перевода с помощью БЯМ проявляются при работе с крупными текстами.

4. Большие языковые модели переводят одни и те же термины или понятия, имеющие только один эквивалент, по-разному даже в рамках одного запроса.

5. Разные контекстные окна в DeepSeek могут давать разные варианты переводческих решений, которые могут противоречить друг другу.

6. Нейросети употребляют акронимы вместо полноценных наименований и наоборот, употребление носит случайный характер и приводит к смысловым ошибкам.

7. Большие языковые модели не умеют «переводить» числовую информацию в комбинации языков русский — английский.

8. В итоге при переводе крупных текстов программы машинного перевода абсолютно неспособны заменить человека. Глубокое редактирование и коррекция полученного текста сводят на нет всю мнимую экономию при переводе текста машиной.

Список литературы

Асатрян А. В США родители покончившего с собой подростка подали в суд на разработчика ChatGPT // Forbes, 2025. [Электронный ресурс]. URL: <https://www.forbes.ru/tekhnologii/544679-v-ssa-roditeli-pokonivsegos-soboj-podrostka-podali-v-sud-na-razrabotcika-chatgpt> (дата обращения 25.09.2025).

Ачкасов А.В. Индустрия локализации и подготовка переводчиков // Индустрия перевода: материалы VIII Международной научной конференции (Пермь, 6–8 июня 2016 г.). Пермь: Пермский национальный исследовательский политехнический университет, 2016. С. 13–19.

Беляева Л. Машинный перевод в современной технологии процесса перевода // Известия Российского государственного педагогического университета имени А.И. Герцена, 2022 № 203. С. 22–30.

Блинов А. GPT на языке бизнеса: особенности применения больших языковых моделей // SpBit.ru. 2024. [Электронный ресурс]. URL: https://spbit.ru/it_class/GPT-na-yazyke-biznesa-osobennosti-primeneniya-bol-shikh-yazykovykh-modelei-297420 (дата обращения 07.10.2025)

Егорова А.Ю., Зацман И.М., Романенко В.О. Машинный перевод с помощью ChatGPT: мониторинг воспроизводимости результатов // Системы и средства информ. 33:3, 2023. С. 117–128.

Журавлёв Д.В., Смолин В.С. Нейросетевая революция искусственного интеллекта и варианты её развития // Проектирование будущего. Пробле-

мы цифровой реальности: труды 6-й Международной конференции (2–3 февраля 2023, Москва). М.: ИПМ имени М.В. Келдыша, 2023. С. 223–244.

Ли Кай-Фу. Сверхдержавы искусственного интеллекта: Китай, Кремниевая долина и новый мировой порядок // М.: Изд-во «Манн, Иванов и Фербер» (серия МИФ), 2019. ISBN: 978-5-00146-163-0.

Перспективно ли сейчас идти учиться на переводчика? Где работать? [Электронный ресурс]. URL: <https://www.bolshoyvopros.ru/questions/4525933-perspektivno-li-sejchas-idti-uchitsja-na-perevodchika-gde-rabotat.html> (дата обращения 05.10.2025).

Устойчивое развитие в Ямало-Ненецком автономном округе: отчёт о результатах и ключевые направления концепции. Правительство ЯНАО. 2025. [Электронный ресурс]. URL: <https://yanao.ru/dokumenty/116823/> (дата обращения 29.08.2025).

Gokul Yenduri, Ramalingam M., Chemmalar Selvi G., Supriya Y., Gautam Srivastava, Praveen Kumar Reddy Maddikunta, Deepti Raj G., Rutvij H. Jhaveri, Prabadevi B., Weizheng Wang, Athanasios V. Vasilakos, and Thippa Reddy Gadekallu. (2025) GPT (Generative Pre-Trained Transformer) — A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions. IEEE Access. PP. 1–1. [Электронный ресурс]. URL: <https://arxiv.org/abs/2305.10435> (дата обращения 25.09.2025).

Kuhn R., De Mori R. (1990) A cache-based natural language model for speech recognition. IEEE transactions on pattern analysis and machine intelligence 12.6, 570–583.

Lingg Christoph. (2019) The world in faces. Сайт ООН [Электронный ресурс]. URL: <https://www.un.org/en/exhibits/exhibit/world-in-faces> (дата обращения 25.07.2025).

Liu Aixin, Feng Bei, Xue Bing, Wang Bingxuan, Wu Bochao, Lu Chengda, Zhao Chenggang, Deng Chengqi. (2024) DeepSeek-AI. *DeepSeek-V3 Technical Report*, arXiv:2412.19437 [Электронный ресурс]. URL: <https://arxiv.org/abs/2412.19437> (Дата обращения: 20.07.2025).

Nucleus AI. (2025) Godfather of AI? Geoffrey Hinton reveals which jobs are at risk from AI and which are safe. [Электронный ресурс]. URL: <https://yourstory.com/ai-story/geoffrey-hinton-ai-jobs-risk-warning> (дата обращения 27.09.2025).

Pyun A., Yu Hao. (2024) How to Augment Language Skills: Generative AI and Machine Translation in Language Learning and Translator Training. [Электронный ресурс]. URL: <https://www.researchgate.net/publication/381988926> (дата обращения 05.10.2025).

References

Asatrjan A. (2025) V SShA roditeli pokonchivshego s soboj podrostka podali v sud na razrabotchika ChatGPT. Forbes. URL: <https://www.forbes.ru/tekhnologii/544679-v-ssa-roditeli-pokonchivshego-s-soboj-podrostka-podali-v-sud-na-razrabotchika-chatgpt>.

Achkasov A.V. (2016) Industrija lokalizacii i podgotovka perevodchikov = The localization industry and the training of translators and interpreters. Industrija perevoda: materialy VIII Mezhdunarodnoj nauchnoj konferencii (Perm', 6–8 June 2016). Perm': Permskij nacional'nyj issledovatel'skij politehnicheskij universitet, pp. 13–19 (In Russian).

Beljaeva L. (2022) Mashinnyj perevod v sovremennoj tehnologii processa perevoda = Machine translation in the modern translation technology. *Izvestija Rossijskogo gosudarstvennogo pedagogicheskogo universiteta im. A.I. Gercena*. No. 203, pp. 22–30 (In Russian).

Blinov A. (2024) GPT na jazyke biznesa: osobennosti primeneniya bol'shikh jazykovykh modelej = GPT in business terms: features of large language models application. URL: [https://ict-online.ru/it class/GPT-na-yazyke-biznesa-osobennosti-primeneniya-bol-shikh-yazykovykh-modelei-297420?hl=ru-RU](https://ict-online.ru/it-class/GPT-na-yazyke-biznesa-osobennosti-primeneniya-bol-shikh-yazykovykh-modelei-297420?hl=ru-RU) (In Russian).

Egorova A.Ju., Zacman I.M., Romanenko V.O. (2023) Mashinnyj perevod s pomoshh'ju ChatGPT: monitoring vosproizvodimosti rezul'tatov = Machine translation using ChatGPT: monitoring the reproducibility of results. *Sistemy i sredstva inform.* No. 33: 3, pp. 117–128 (In Russian).

Gokul Yenduri, Ramalingam M., Chemmalar Selvi G., Supriya Y., Gautam Srivastava, Praveen Kumar Reddy Maddikunta, Deepti Raj G., Rutvij H. Jhaveri, Prabadevi B., Weizheng Wang, Athanasios V. Vasilakos, and Thippa Reddy Gadekallu. (2025) GPT (Generative Pre-Trained Transformer) — A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions. *IEEE Access*. PP. 1–1. URL: <https://arxiv.org/abs/2305.10435>

Kuhn R., De Mori R. (1990) A cache-based natural language model for speech recognition. *IEEE transactions on pattern analysis and machine intelligence* 12.6, pp. 570–583.

Li Kaj-Fu. (2019) Sverhderzhavy iskusstvennogo intellekta: Kitaj, Kremnievaja dolina i novyj mirovoj porjadok = AI superpowers: China, Silicon Valley and the new world order. Moscow: Izd-vo "Mann, Ivanov i Ferber" (Serija MIF). ISBN: 978-5-00146-163-0 (In Russian).

Lingg Christoph. (2019) The world in faces. Сайт ООН [Электронный ресурс]. URL: <https://www.un.org/en/exhibits/exhibit/world-in-faces>.

Liu Aixin, Feng Bei, Xue Bing, Wang Bingxuan, Wu Bochao, Lu Chengda, Zhao Chenggang, Deng Chengqi. (2024) DeepSeek-AI. *DeepSeek-V3 Technical Report*, arXiv:2412.19437 [Электронный ресурс]. URL: <https://arxiv.org/abs/2412.19437>.

Nucleus AI. (2025) Godfather of AI' Geoffrey Hinton reveals which jobs are at risk from AI and which are safe // [Электронный ресурс]. URL: <https://yourstory.com/ai-story/geoffrey-hinton-ai-jobs-risk-warning>.

Perspektivno li sejchas idti uchit'sja na perevodchika? Gde rabotat'? = Are there any prospects in becoming a translator today? Where is there an opportunity to find a job as a translator? (2025) URL: <https://www.bolshoyvopros.ru/>

questions/4525933-perspektivno-li-sejchas-idi-uchitsja-na-perevodchika-gde-rabotat.html (In Russian).

Ustojchivoe razvitie v Jamalo-Neneckom avtonomnom okruge: otchet o rezul'tatah i kljuchevye napravlenija koncepcii = Sustainable development in Yamal-Nenets Autonomous Okrug: performance report and key policy directions. (2025) Pravitel'stvo YNAO. URL: <https://yanao.ru/dokumenty/116823/> (In Russian).

Zhuravljov D. V., Smolin V.S. (2023) Nejrosetevaja revoljucija iskusstvennogo intellekta i varianty ejo razvitija = The AI neural network revolution and its development paths. Proektirovanie budushhego. Problemy cifrovoj real'nosti: trudy 6-j Mezhdunarodnoj konferencii (2–3 fevralja 2023 g., Moskva). Moscow: IPM imeni M.V. Keldysha, pp. 223–244 (In Russian).

ИНФОРМАЦИЯ ОБ АВТОРЕ:

Ушаков Сергей Геннадиевич — кандидат филологических наук, доцент кафедры теории и методологии перевода Высшей школы перевода (факультета) МГУ имени М.В. Ломоносова; 119991, г. Москва, Ленинские горы, МГУ, ГСП-1, дом 1, корп. 51; sergeypenseur@rambler.ru

ABOUT THE AUTHOR:

Sergei G. Ushakov — Cand. Sc. (Linguistics), Associate Professor at the Department of Theory and Methodology of Translation, Higher School of Translation and Interpreting, Lomonosov Moscow State University; Building 51, the 1st Building of Humanities, Leninskie Gory, 1, Moscow, 119991, Russia; sergeypenseur@rambler.ru

Конфликт интересов: положения и точки зрения, представленные в данной статье, принадлежат автору и не обязательно отражают позицию какой-либо организации или российского научного сообщества.

Conflict of interests: the ideas and opinions presented in this article entirely belong to the author and do not necessarily reflect the position of any organization or the Russian scientific community as a whole. The author state that there is no conflict of interests.