Лесаев Августин Николаевич
бакалавр, студент
Игнатович Илья Денисович
бакалавр, студент
Андриенко Владислав Владимирович
бакалавр, студент
Научный руководитель
Миролюбова Наталия Алексеевна
старший преподаватель
ФГБОУ ВО «МИРЭА – Российский технологический университет»
г. Москва

ИСПОЛЬЗОВАНИЕ АССИСТЕНТОВ НА БАЗЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА (CHATGPT, CLAUDE) ПРИ СЕМАНТИЧЕСКОМ АНАЛИЗЕ И ПЕРЕВОДЕ ТЕХНИЧЕСКОЙ ДОКУМЕНТАЦИИ НА АНГЛИЙСКОМ ЯЗЫКЕ СТУДЕНТАМИ ТЕХНИЧЕСКИХ ВУЗОВ

Аннотация: *в статье рассматривается проблема формирования у студентов технических вузов умения работать с англоязычной технической документацией в условиях широкого распространения генеративных языковых моделей. Цель исследования – теоретически обосновать и разработать методику, в которой ассистенты на базе искусственного интеллекта (ChatGPT, Claude) выступают инструментом семантического анализа и критической верификации перевода, а не источником готового текста. Методологической основой послужили деятельностный и контекстный подходы, положения теории перевода и научные данные об ограничениях больших языковых моделей. Применены методы теоретического анализа, педагогического проектирования и систематизации типичных ошибок автоматического перевода. Предложены четырёхэтапный алгоритм работы с документацией, структура учебного запроса к ассистенту, классификация ошибок моделей и система критериев оценки резуль-*

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

тата. Показано, что обязательный этап верификации и сопоставление ответов нескольких ассистентов превращают модель из заменителя мышления в средство его развития. Материалы могут быть использованы при проектировании курсов профессионально ориентированного иностранного языка в технических вузах.

Ключевые слова: *искусственный интеллект, большие языковые модели, ChatGPT, Claude, перевод технической документации, семантический анализ, терминологическая компетенция, верификация перевода, технический вуз.*

Англоязычная техническая документация (спецификации, руководства пользователя, стандарты, технические отчёты) остаётся основным источником актуальных знаний для будущего инженера. Работа с ней требует не только владения лексикой, но и умения устанавливать смысловые связи в насыщенном терминологией тексте. Появление ассистентов на базе больших языковых моделей изменило практику студентов: перевод фрагмента занимает секунды, однако получаемый текст часто принимается без проверки. Возникает противоречие между доступностью автоматического перевода и риском фактических и терминологических искажений, которые в технической области могут иметь практические последствия [8; 9].

Проблема исследования состоит в недостаточной разработанности методики, которая позволяла бы использовать ChatGPT и Claude для развития переводческой и аналитической компетенции, а не для её подмены. Цель работы – теоретически обосновать и спроектировать такую методику. Задачи: проанализировать теоретические основы смыслового анализа текста и особенности работы языковых моделей; систематизировать типичные ошибки моделей при переводе документации; разработать алгоритм и учебный запрос для работы студентов; определить критерии оценки эффективности методики.

В отечественном переводоведении перевод трактуется как процесс, начинающийся со смыслового анализа оригинала, который включает выявление комму-

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

никативной цели, терминологического поля и логической структуры высказывания [3]. Л. С. Бархударов подчёркивает, что выбор эквивалента определяется контекстом, а не словарным соответствием [1], а Я. И. Рецкер выделяет контекстуальные замены как необходимый приём при передаче значений многозначных единиц [5]. Для технической документации это означает обязательное обращение к предметной области, отражённой в стандартах и параллельных текстах.

Психолого-педагогической основой исследования служит теория деятельности: усвоение происходит через собственные действия обучающегося, а не через потребление готового результата [4]. Контекстный подход А. А. Вербицкого предполагает моделирование в обучении предметного и социального содержания будущей профессиональной деятельности [2]; работа с реальной документацией отвечает этому требованию.

Современные языковые модели построены на архитектуре трансформера, использующей механизм внимания для учёта контекста [10], и способны выполнять переводческие задачи по нескольким примерам в запросе без дообучения [6]. История автоматического перевода показывает, что качество систем неизменно оценивалось и как научная, и как практическая проблема [7]. Вместе с тем модели порождают гладкие, но не подтверждённые источниками утверждения, что в литературе обозначается как «галлюцинации» [8]. Обзор образовательных рисков отмечает опасность чрезмерного доверия к ответам моделей и ослабления самостоятельного мышления, одновременно указывая на возможность персонализации обучения [9]. Побуждение модели к пошаговому рассуждению повышает качество решения сложных задач и делает ход рассуждений доступным для проверки [11]; этот принцип был использован при построении запросов.

Исследование носит теоретико-методический характер. Использованы анализ научной литературы по переводоведению, педагогике и компьютерной лингвистике, педагогическое проектирование и систематизация типичных ошибок автоматического перевода, описанных в литературе и выявляемых при сопоставлении ответов ассистентов. Разработанная методика предназначена для занятий по профессионально ориентированному иностранному языку в технических вузах;

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

учебным материалом служат фрагменты технических описаний электронных компонентов, руководств по программному обеспечению и текстов стандартов.

Алгоритм работы студента включает четыре этапа: 1) самостоятельный смысловой анализ фрагмента с выделением терминов и логических связей; 2) обращение к ассистенту со структурированным запросом; 3) верификация по глоссариям, стандартам и параллельным текстам с сопоставлением ответов ChatGPT и Claude; 4) ведение рефлексивного журнала ошибок. Структурированный запрос содержит пять элементов: роль ассистента (технический переводчик в заданной предметной области), исходный фрагмент, требование сохранить различие между обязательными и рекомендательными положениями, просьбу привести для каждого термина альтернативные варианты с указанием контекста их употребления и требование отметить места, в которых модель не уверена в переводе. Такой формат ограничивает возможность подмены анализа и делает ответ пригодным для последующей проверки.

Для оценки эффективности методики при её апробации предусмотрена система критериев, представленная в таблице 1.

Таблица 1

Система критериев оценки эффективности методики

Критерий	Показатель	Инструмент оценки	Этап
Терминологическая точность	Доля верно переведённых терминов, %	Сверка с глоссарием и стандартом	3
Семантическая адекватность	Оценка передачи смысла, баллы (0–5)	Экспертная оценка двумя преподавателями	3
Критичность к ответу модели	Доля непроверенных заимствований, %	Анализ запросов и рефлексивного журнала	2–4
Качество запроса	Полнота пяти элементов запроса, баллы	Контрольный перечень	2

Для сравнения результатов предполагается формирование контрольной группы, работающей с документацией без ассистентов, и экспериментальной группы, работающей по предложенному алгоритму; обработка данных может выполняться непараметрическими критериями для независимых выборок.

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

Систематизация показала четыре группы ошибок, характерных для ассистентов при переводе технической документации. Первая – неверный выбор значения полисемичного термина при недостаточном контексте (например, bearing как «направление» вместо «подшипник»). Вторая – нивелирование модальности: глаголы shall и should, различающие обязательное требование и рекомендацию в стандартах, передаются одним и тем же словом «должен». Третья – разрушение связей в длинных предложениях с цепочками атрибутивных групп. Четвёртая – уверенные пояснения, не опирающиеся на текст оригинала, что соответствует описанным в литературе галлюцинациям [8]. Знание этих групп позволяет студенту целенаправленно проверять именно те места текста, где риск ошибки наиболее высок.

Принципиальное значение имеет сопоставление ответов ChatGPT и Claude. Расхождение между переводами становится сигналом для обращения к первоисточникам, а не поводом выбрать более убедительный вариант. Тем самым студент осваивает действие, которое составляет основу профессиональной переводческой деятельности: обоснование выбора эквивалента с опорой на контекст [1; 5]. Предложенный запрос, требующий пояснений и указания неуверенности модели, реализует принцип пошагового рассуждения [11] и делает проверяемым ход выбора термина.

Теоретическим обоснованием методики служит деятельностный подход: ценность работы определяется тем, какие действия выполняет студент [4]. Если модель выдаёт готовый перевод, анализ замещается копированием; если от неё требуют обоснований, а ответ подлежит обязательной проверке, студент вынужден самостоятельно устанавливать смысл. Практической значимостью отличается и опора на реальную документацию, что соответствует контекстному подходу [2]. Ограничениями методики являются зависимость ответов моделей от версии и формулировки запроса, необходимость предварительной подготовки глоссариев и недостаточная изученность долговременного влияния такой практики на языковую компетенцию.

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

Проведённый анализ позволяет сформулировать педагогические условия продуктивного использования ассистентов на базе искусственного интеллекта при работе с технической документацией: самостоятельный смысловой анализ до обращения к модели; структурированные запросы с требованием обосновать выбор терминов; обязательная верификация по авторитетным источникам; сопоставление ответов нескольких ассистентов; рефлексия типичных ошибок. Реализация этих условий позволяет рассматривать языковую модель как объект критического анализа и инструмент развития терминологической и аналитической компетенции. Перспективой исследования является экспериментальная проверка методики по предложенным критериям и уточнение классификации ошибок на материале различных предметных областей.

Список литературы

1. Бархударов, Л. С. Язык и перевод: вопросы общей и частной теории перевода / Л. С. Бархударов. – Москва : Международные отношения, 1975. – 240 с.

2. Вербицкий, А. А. Активное обучение в высшей школе: контекстный подход / А. А. Вербицкий. – Москва : Высшая школа, 1991. – 207 с.

3. Комиссаров, В. Н. Современное переводоведение / В. Н. Комиссаров. – Москва : ЭТС, 2002. – 424 с.

4. Леонтьев, А. Н. Деятельность. Сознание. Личность / А. Н. Леонтьев. – Москва : Политиздат, 1975. – 304 с.

5. Рецкер, Я. И. Теория перевода и переводческая практика / Я. И. Рецкер. – Москва : Международные отношения, 1974. – 216 с.

6. Language models are few-shot learners / T. Brown, B. Mann, N. Ryder [et al.] // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 1877–1901.

7. Hutchins, W. J. An introduction to machine translation / W. J. Hutchins, H. L. Somers. – London : Academic Press, 1992. – 362 p.

8. Survey of Hallucination in Natural Language Generation / Z. Ji, N. Lee, R. Frieske [et al.]. – 2023. – ACM Comput. Surv. 55, 12, Article 248 (December 2023), 38 p. – DOI: 10.1145/3571730.

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

9. ChatGPT for good? On opportunities and challenges of large language models for education / E. Kasneci, K. Sessler, S. Küchemann [et al.] // Learning and Individual Differences. – 2023. – Vol. 103.

10. Attention is all you need / A. Vaswani, N. Shazeer, N. Parmar [et al.] // Advances in Neural Information Processing Systems. – 2017. – Vol. 30. – P. 5998–6008.

11. Chain-of-thought prompting elicits reasoning in large language models / J. Wei, X. Wang, D. Schuurmans [et al.] // Advances in Neural Information Processing Systems. – 2022. – Vol. 35. – P. 24824–24837.

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)