

ARTIFICIAL INTELLIGENCE AND ORIENTAL CLASSICS

Methodological Issues of of AI using ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И КЛАССИКА ВОСТОКА

Методологические вопросы использования ИИ

Научная статья

Филологические науки

УДК 81'25 + 004.8 + 2-1

<https://doi.org/10.31696/2618-7043-2025-8-3-466-480>

Перевод сакральных текстов в эпоху генеративного ИИ

Александр Вячеславович Костыркин

Институт востоковедения РАН, Москва, Россия,

kost@ivran.ru, <https://orcid.org/0009-0002-8167-1529>

Аннотация. Машинный перевод сегодня во многих аспектах превосходит человеческий, однако перевод сакральных текстов — задача, в которой машина не может заменить человека. В статье кратко излагаются основные этапы развития машинного перевода и параллельно — систем поддержки переводческой деятельности. Технологический баланс в последних, по мнению автора, пока не в пользу человека, и ключевую роль в улучшении положения переводчика автор видит в применении генеративных нейросетевых моделей, которые демонстрируют способности к обработке текста на качественно новом уровне. Автор предлагает концепцию эпистемической экосистемы для исследования и перевода канонических текстов, которая включает в себя коллаборативный процесс с участием ИИ, переводческого коллектива и сообщества носителей традиции. Предложенная концепция должна создать механизм генерации знаний и управления открытиями. Обсуждаются функции и возможности наполняющих экосистему ИИ-ассистентов, которые могут предоставлять переводчикам альтернативные пути развертывания перевода, исправлять переводческие искажения, давать интерпретационный прогноз по восприятию перевода. В заключении автор прогнозирует дальнейшее вытеснение человека-переводчика из отрасли и позиционирует представленный комплекс интеллектуальных агентов как средство смягчения последствий этого процесса и гармонизации системы человек-машина.

Ключевые слова: машинный перевод, генеративные нейросетевые модели, сакральные тексты



Контент доступен под лицензией Creative Commons «Attribution-ShareAlike» («Атрибуция-СохранениеУсловий») 4.0 Всемирная.



Благодарности: Исследование выполнено по проекту Института востоковедения РАН «Лаборатория цифровых исследований современного Востока (FMNN-2024-0001)».

Для цитирования: Костыркин А. В. Перевод сакральных текстов в эпоху генеративного ИИ. *Ориенталистика*. 2025;8(3):466–480. <https://doi.org/10.31696/2618-7043-2025-8-3-466-480>.

Original article

Philology studies

<https://doi.org/10.31696/2618-7043-2025-8-3-466-480>

Translating Sacred Texts in the Generative AI Era

Alexandr V. Kostyrkin

Institute of Oriental Studies, Russian Academy of Sciences, Moscow, Russia

kost@ivran.ru, <https://orcid.org/0009-0002-8167-1529>

Abstract. Machine translation today surpasses human translation in many aspects; however, the translation of sacred texts is a task where machines cannot replace humans. The article briefly outlines the main stages in the development of machine translation and, in parallel, computer-assisted translation (CAT) systems. The author argues that the current technological balance in the latter is still not in favor of the human translator and sees the key role for improving the translator's position in the application of generative neural network models, which demonstrate text processing capabilities at a qualitatively new level. The author proposes the concept of an epistemic ecosystem for the study and translation of canonical texts. It involves a collaborative process including AI, a translation team, and a community of tradition bearers. The proposed concept is intended to create a mechanism for knowledge generation and discovery management. The functions and capabilities of the AI assistants populating the ecosystem are discussed. These assistants are supposed to provide translators with alternative translation pathways, correct translation distortions, and offer interpretive forecasts regarding how the translation might be received. In conclusion, the author predicts the further displacement of human translators from the industry and positions the presented complex of intelligent agents as a means to mitigate the consequences of this process and to harmonize the human-machine system.

Keywords: machine translation, generative neural network models, sacred texts

Acknowledgments: The study was carried out under the project of the Institute of Oriental Studies of the Russian Academy of Sciences “Laboratory for Digital Studies of the Modern East (FMNN-2024-0001)”.

For citation: Kostyrkin A. V. Translating Sacred Texts in the Generative AI Era. *Orientalistica*. 2025;8(3):466–480. (In Russ.). <https://doi.org/10.31696/2618-7043-2025-8-3-466-480>.



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).



Введение

Затянувшийся спор о способности машины заменить живого переводчика исчерпал свой познавательный ресурс. Не нужно повторно доказывать преимущества машинного перевода перед ручным: машина опередила человека, и переводчикам пора с достоинством и гордостью, не лишенной горечи, признать и принять это величайшее достижение человечества.

Есть, однако, одно серьезное исключение. Тексты, хранящие свидетельство столкновения с запредельным, составляют предельную семиотическую сложность. Ожидание сакральным текстом своего переводчика напрягает пространство интерпретаций словно гравитационное поле. Текст прокладывает сложный путь через эпохи, языки и души. Встретив на этом пути переводчика, текст коллапсирует в нем, рождая новую вселенную. Перевод — не интерпретация, а ветвление реальности. Переводчик — не пассивный передатчик линейной последовательности кодов, а со-творец новой реальности. Обычный текст мучителен для переводчика технически, сакральный мучителен экзистенциально. Каждое решение о переводе — это лавирование между кошунством упрощения и бессмысленностью буквализма. Задача переводчика сакрального — раскрыть непрозрачность как проявление трансцендентности. Машина способна лишь искусно имитировать эмпатию, интуицию, понимание, смыслообразование, этическое суждение, но не может отвечать за смысловую, культурную адекватность и сакральную точность перевода.

Текст — это исполняемый языковой объект, среда его исполнения — сознание читателя. Текст управляет действиями, интеллектуальным и эмоциональным состоянием. Подобно компьютерной программе он запускает в читателе внутреннюю трансформацию вплоть до изменения мировоззрения. Пока заложенная в тексте программа личностной трансформации не отработает, перевод не состоится. Ни один переводчик не может быть готов к переводу сакрального текста, он должен до него дорасти (*wachsen*), но даже завершив перевод, переводчик может остаться к нему неготовым. Этот рост не поддается планированию и оценке производительности переводческого труда: озарение не измеряется в нормочасах.

Никогда в истории человечества перевод не достигал такой степени демократизации, как сейчас. Везде, где есть интернет, перевод доступен всем и бесплатно. Возможность машины неограниченно продуцировать переводы изменила правила игры и ценность переводческого продукта: порождение единственного текста сменилось отбраковкой сгенерированных вариантов перевода, варьирующих по настраиваемым функционально-стилистическим параметрам — прагматическому, информационному, политическому, эстетическому. Во многих контекстах сегодня под переводом понимается уже только машинный или машинный с постредактированием.

Последние достижения в машинном обучении открывают беспрецедентные возможности для делегирования машине рутинных операций и получения от нее оперативной информационной помощи. Переводчики и переводческие агентства давно используют множество цифровых инструментов — от текстовых процессоров, электронных словарей и корпусов до сетевых многопользовательских систем поддержки переводческой деятельности (CAT), направлен-



ных на организацию, оптимизацию и стандартизацию переводческих процессов. Какое место в этом трансформирующем процессе занимают новые эпистемические машины, известные под торговой маркой «искусственного интеллекта»? Что из этого арсенала можно согласовать с задачей перевода сакральных текстов? Что еще не предусмотрено и ждет своей реализации?

Свет с другой стороны

В этом контексте история разработок, предвалявших появление генеративного ИИ, заслуживает специального обсуждения.

Несмотря на междисциплинарный характер искусственного интеллекта как научной дисциплины, уже на первых этапах в нем произошло предметное и методологическое разделение, чему способствовало перенесение проблематики из исследовательских лабораторий в более строгие рамки первых учебных программ. В 1968 г. американская «Ассоциация по вычислительной технике» (Association for Computing Machinery) в своих рекомендациях по формированию учебных программ по компьютерным наукам предложила программу учебного курса «А9. Искусственный интеллект и эвристическое программирование» (Artificial Intelligence and Heuristic Programming), которая несомненно оказала влияние на преподавание ИИ в мире и, как следствие, на распределение задач внутри исследований. В этой программе, в частности, предусмотрены подразделы, в числе которых: «Распознавание образов и обработка изображений», «Диалоговые программы», «Нейронные сети», «Последние достижения в области машинного перевода языков и обработки естественного языка» [АСМ, 1986].

Машинный перевод стал одной из первых задач ИИ. Колоссальные усилия были вложены в решение нескольких основных классов трудных лингвистических задач. Это в общем хорошо известные любому переводчику проблемы: снятие лексической омонимии (в широком смысле), построения гипотез о значении неизвестного системе слова, снятие синтаксической омонимии, установление анафорических связей, выбор правильной трансформации и подходящих переводных эквивалентов в целевом языке. Долгое время эти задачи решались по лингвистическим правилам, сформулированным лингвистами и заложенным в алгоритм перевода.

В 1983 г. московский лингвист М. А. Балабан [Балабан, 1983], а год спустя, в 1984 г., японский лингвист Макото Нагао [Nagao, 1984] открывают для машинного перевода новый лингвистический ресурс, предлагая использовать не искусственные правила, а уже имеющиеся образцы переводных соответствий, которые можно извлекать из параллельных текстов. Появляется новое направление — перевод, основанный на примерах (Example-based MT), или «прецедентный» перевод (термин предложен М. А. Балабаном). Начинается накопление параллельных корпусов. В 1988 г. выходит статья специалистов из Джорджтаунского ун-та, предложивших новый метод перевода, статистический, — еще один шаг к автоматическому извлечению лингвистических знаний из текста.

Общие идеи, стоящие за современным генеративным машинным переводом, появились довольно давно. Искусственные сети нейронов предложены



Мак-Каллоком и Питтсом в 1943 г., а машинное обучение с вознаграждением предложено в 1960-х гг. советским математиком М. Л. Цетлиным. Тем не менее долгое время перенесение этих методов в лингвистику не представлялось возможным или целесообразным. Только в середине 2010-х гг. нейронный перевод вышел на качественно новый уровень, оставив позади перевод статистический. Наконец, главным событием в мировой компьютерной лингвистике стало появление в 2022 г. продукта фирмы OpenAI под уже ставшим нарицательным именем ChatGPT (generative pre-trained transformer, генеративный предобученный трансформер). Генеративный машинный перевод превзошел перевод нейронный, он стал одной из эмергентных, исходно незапрограммированных способностей больших языковых моделей (LLM).

В IT-отрасли всем очевидно, что за последние три года генеративные модели совершили революцию. С ее плодами придется считаться всем переводчикам и переводческой отрасли в целом. ChatGPT и его многочисленные производные объявили эпохальные перемены в мировой индустрии и культуре перевода. Наверняка, в переводческой отрасли каждый успел почувствовать новые веяния, но пока не все поняли важность случившегося и оказались готовы принять перемены, тем более — предвидеть все последствия.

Генеративный перевод демонстрирует беспрецедентную для машин гибкость в обращении с текстом во всех его аспектах. Поражает не только связность текста перевода, владение идиоматикой и выдерживание заданного стилистического режима, но и способность удерживать в памяти и корректно выражать внутритекстовые связи, например, не меняя по ходу пол действующих лиц или референцию местоимений. В зависимости от содержания запроса перевод может быть сгенерирован в том или ином регистре, снабжен комментариями, указанием на источники приводимых сведений. Возможен перевод прозы в поэтическую форму внутри одного языка и с переводом на другой. Наряду с полноформатным переводом может быть сгенерировано резюме текста, выделен список действующих лиц, их характеристик и совершённых по сюжету действий. Разумеется, в полной мере это относится прежде всего к переводам с английским языком, другие языки и пары пока отстают, но их реванш — лишь вопрос времени. Проблема качества машинного перевода с малоресурсных языков — отдельная задача, для которой в новой парадигме уже появляются решения.

В основе ChatGPT и подобных ему систем — статистические языковые модели, обученные на очень больших массивах текстов, приближающихся ко всему текстовому запасу человечества. При должном запросе и примерах модели способны решать разнообразные лингвистические задачи. Сегодня сенсация не только в том, что одна и та же модель способна переводить, аннотировать, перефразировать, комментировать, рецензировать тексты, т. е. решать задачи, которые прежде ставились в разных разделах ИИ и компьютерной лингвистики и решались независимыми коллективами разными методами. Главная сенсация в том, что модели нового типа обучаются только на текстовых образцах и совершенно обходятся без подготовленных лингвистами описаний языка, и, что обиднее всего, без теории языка. Моделирование языковой деятельности многие десятилетия считалось невозможным без лингвистиче-



ской теории, это не оспаривалось даже в периоды самых пессимистических дискуссий о задачах прикладной лингвистики.

В середине 1990-х гг. в кругу московских лингвистов прошла дискуссия о несбывшихся ожиданиях от компьютерной лингвистики, но тем не менее оппоненты не ставили под сомнение примат лингвистической теории над практической реализацией, что было четко сформулировано А. Е. Кибриком: «Чтобы эффективно моделировать общение человека с машиной, надо, как минимум, иметь теорию, которая объясняет, как пользуется естественным языком человек» [Кибрик, 1996, с. 231]. Как выяснилось, современным языковым моделям уже не требуются ни формальные грамматики, ни машинные словари, которые кропотливо собирались десятилетиями. В случившемся есть доля трагизма, ведь сквозь 70 лет работ по компьютерной лингвистике и теории машинного перевода исследователи пронесли с собой совсем другие представления о том будущем, современниками которого мы стали. «Горьким уроком» назвал случившийся разрыв ожиданий с практикой в области ИИ специалист по машинному обучению Ричард Саттон [Sutton, 2019]. Его оценка пятилетней давности вполне применима к сегодняшнему положению дел в компьютерной лингвистике: стремление заставить компьютеры имитировать языковую деятельность в соответствии с теорией оказались контрпродуктивными. Сейчас мы наблюдаем ровно то, что Р. М. Фрумкина в середине 1990-х гг. описала как кризис в этой лингвистической дисциплине: «теория не востребована, в то время как приложения методологически не обеспечены» [Фрумкина, 1996, с. 80]. Схожие оценки недавно дал в своей статье «Горький урок АБВУД: как лингвисты проиграли последнюю битву за NLP» Даниил Скоринкин в связи с событиями в крупной (некогда российской) лингвистической компании АБИ [Скоринкин, 2014].

Все это имеет непосредственное отношение к организации труда переводчика, которая — параллельно и вместе с машинным переводом — десятилетиями обсуждалась в теоретических работах, но мало сказывалась на переводческой практике.

Нерабочее место переводчика

Прототипы гибридных человеко-машинных переводческих систем появились еще в конце 1970-х гг., получив название «интерактивных переводческих систем» (Interactive Translation System) и «личного секретаря переводчика» (Translator's Amanuensis Мартина Кэя [Сыртыповая и др., 2025]). В это время создание «полностью автоматизированного высококачественного машинного перевода» (FANQT) все еще представлялось недостижимым, в области машинного перевода популярным стал отказ от универсальных переводческих систем в пользу специализированных по отдельным отраслям, а многие переводчики не имели доступа или не умели пользоваться компьютером. В ограничениях того времени родилась и затем надолго прижилась идея, как сейчас видно, *ограниченного симбиоза*, выраженная в 1981 г. лингвистом и переводчиком Аланом К. Мелби [Melby, 1981]. Отправной точкой стал функционал текстового процессора, от которого рабочая станция переводчика (translator's work station) должна была отличаться двумя надстройками: словарной справкой и интерактивным переводом, который сводился к разрешению перевод-



чиком обнаруженных системой неоднозначностей. Мелби вообще сомневался в способности машин к творческому переводу, беря за критерий сложность перевода художественных текстов.

Переводческое сообщество с запозданием, медленно и нерешительно, но все же перенимает лучшие технологии и организационные практики софтверной индустрии. Между этими мирами нет непреодолимых границ, в конце концов инструменты для переводчиков пишут программисты. Однако между ними налицо давний и глубокий технологический, концептуальный и социологический разрыв: мало кто из переводчиков знает, как организован труд программиста, программисты же имеют лишь поверхностное представление о подлинных трудностях перевода.

Программист буквально окружен заботливой средой интегрированной разработки (IDE), которая снабжает его всеми возможными услугами и подсказками, упреждая каждый каприз. Для программиста совершенно прозрачен и структурирован весь корпус файлов, который может достигать десятков тысяч разнородных единиц. Ввод текста с первых символов тут же сопровождается автоматическими подсказками с документацией всех программных объектов. Программист обзрывает хронологию правки всех текстов, видит их авторство и статус. Средства совместной работы, синхронизации и сверки версий тонко подстраиваются под социальную конфигурацию группы разработчиков, позволяя схватывать сетевую структуру рабочего коллектива.

К этим удобствам с начала 2024 г. добавились так называемые ассистенты программиста (AI-assistant, co-pilot) — специально обученные на программном коде языковые модели, способные генерировать код для поставленной задачи и предсказывать продолжение текста на основе глобального и локального контекста. Внедрение ИИ-ассистентов и в другие отрасли уже показывает рост продуктивности и экономическую отдачу. Повышение продуктивности интеллектуального труда достигается делегированием ИИ-ассистенту рутинных аналитико-синтетических операций, таких как семантическое индексирование документов и поиск по ним, составление обобщающих аннотаций, извлечения конкретных фактов, примеров, контрпримеров и др. Циклы взаимодействия ассистента с человеком легко встраиваются в различные рабочие контексты за счет использования естественного языка и высокой адаптивности.

В переводческой индустрии пока еще немного предложений такого класса, при этом они не выходят за рамки помощи в стилистической настройке текста перевода (что тоже полезно). Появление по-настоящему интеллектуальных и глубоко интегрированных в процесс перевода ассистентов следует ожидать в ближайшем будущем. Пока же в области переводческого инструментария с появления в начале 1990-х гг. первых популярных систем в жанре CAT (Computer Assisted Translation) наблюдается многолетний застой, который годами приучает переводчика к аскезе и смирению перед неизбежностью рутины. Никто не заботится об оптимизации даже самых элементарных операций редактирования текста¹.

¹ Любопытный по-русски знает цену смены одного рода на другой в контексте длинных согласованных по роду цепочек: в предложении «Это **новый** переработанный плагин, который потребует новой версии браузера» замена слова «плагин» на «расширение» вызовет три правки подчеркнутых окончаний.



Переводчик незаслуженно обделен и нуждается в по-настоящему комфортной и человеко-ориентированной среде перевода. Это тем более верно, если говорить о задачах переводчика-исследователя. Перевод фундаментальных текстов культуры по всем параметрам должен быть организован как научно-исследовательский проект, т. е. обеспечен сопоставимыми по сложности и объему экономическими, политическими и социальными условиями.

Переводческая майевтика

Прорыв в области машинного обучения и генеративных языковых моделей стал как нельзя более своевременным. Это адекватный ответ на назревшие потребности цивилизации, которая лавинообразно продуцирует тексты. Впрочем, генеративные модели готовы принять и старые филологические вызовы, но сводить их применение к одному переводу было бы не практично — поле филологических задач для ИИ видится значительно шире. Далее предлагается концепция *эпистемической экосистемы* для исследования и перевода канонических текстов в форме коллаборативного процесса с участием ИИ, переводческого коллектива и сообщества носителей традиции. Сказанное нужно понимать как прогноз и одновременно техническое задание для разработчиков.

Ранее уже говорилось, что природа текста сродни программному коду, а чтение сопоставимо с исполнением инструкций в процессоре сознания. В этом сравнении мы лишь вернули компьютерную метафору к ее истоку. На этот процедурный язык легко перекладываются многочисленные сбои, возможные в ходе чтения: программа может не запуститься, а запустившись, может не отработать, а может отработать нештатно, с ошибками. Собственно контролю над выполнением «текстовых инструкций» и учат годами в школе: правильно входить в процесс, отрабатывать его и завершать, не уходить в ложные ветки, не зацикливаться, выходить из тупиков.

Подобно тому как продукт генеративной сети детерминируется ее весами и запросом, перевод человека направляется не только текстом, но и предпосылочным знанием, активная роль которого не всегда осознается. С учетом программирующе-трансформирующей функции текста и скрытого влияния предпонимания в гибридном человеко-машинном переводе задачу перевода можно ставить по-новому — как эффективнее и надежнее выполнять в сознании переводчика программу, заданную текстом? И как при этом учитывать и корректировать его неявные установки? ИИ здесь может стать больше, чем просто еще одним переводчиком, его новая роль — смысловое родовспоможение. ИИ не сообщает переводчику смысл, а помогает запускать в его сознании реакцию смыслопорождения. ИИ — человеко-машинная операционная система, супервизор исполнения переводческих процессов и тредов. Его задача как навигатора по смысловому пространству — привести переводчика к пониманию кратчайшим путем, просчитав наперед все мыслимые в данной культуре интерпретации, к которым может прийти человек, вплоть до абсурдных. Их не обязательно исчислять в полном объеме заранее, большая часть может быть построена, исходя из частичного перевода, как его возможное развитие



или преобразование, т. е. в рамках более сильных ограничений, чем можно предусмотреть на чистом листе.

Экосистема названа эпистемической, потому что населяется эпистемическими машинами (будем называть их по сложившейся традиции ассистентами), превращающими знание из объекта накопления и хранения в интерфейс к областям, лежащим за когнитивным пределом. Их совокупная сила не в понимании сакрального, а в генерации вариантов исследовательских парадигм и способности идти на несколько шагов впереди человека, обозревая все рабочее поле шире, чем доступно человеческому сознанию и памяти.

Экосистема встречи

Проект можно описать в инженерных терминах «экспертных систем» или «рабочих станций» (ускорение рутинных операций, выявление регулярных паттернов и латентных связей, навигация по массивам текстов, поддержка принятия решений, обогащение контекста справочными материалами и т. п.), но это не выразит его сути. Мы проектируем не «автоматизированное рабочее место переводчика», а *экосистему встречи с другим*, которая создаст механизм генерации знаний, управления открытиями, получения инсайтов, когда правильно направленный поисковый импульс делает случайность открытия необходимостью. Глубокая контекстуализация переводимого текста должна показать переводчику то, что он никогда не стремился или не имел возможности увидеть, причем в нескольких измерениях — историческом, культурном, лингвистическом, богословском, ритуальном.

Прежде чем запускать переводческий проект, через партнерства с библиотеками, университетами, религиозными институтами должна быть аккумулирована в цифровом виде вся наличная литература, начиная со всех версий самого исходного текста, до его переводов, комментариев к основному тексту, комментариев к переводам, комментариев к комментариям, по возможности вся исследовательская литература, критика, популярные переложения. Это необходимое капиталовложение, которое окупится в работе, и при современных технических возможностях не должно представляться невыполнимым.

Следующий шаг — картографирование ландшафта переводческой и толковательной традиции, связанной с исходным текстом, и полная систематизация знаний о нем. Это значит, что во всем электронном массиве документов должны быть установлены уже известные межтекстовые связи (какой документ является списком, компиляцией, фрагментом, толкованием, переводом какого/каких документов), обозначена степень гипотетичности этих связей и указание на подтверждающие исследования. В дополнение к этой сети отношений, отражающей исследовательскую традицию и накопленные в ней представления о структуре документов, может быть построена и вторая — уже машинная.

Таким образом, переводчик должен быть обеспечен всеми доступными информационными ресурсами, но наша задача — продвинуться дальше и сообщить корпусу новое качество, сделав его эвристическим корпусом, т. е. инструментом управления эпистемическими эффектами. В этом качественном росте ключевую роль должны сыграть генеративные модели. Их мощности позво-



ляют удерживать в «окне внимания» весь рабочий корпус текстов, а отображение текстов в общее многоязычное семантическое пространство (semantic space) позволит обнаруживать скрытые смысловые пласты, делать прозрачными интертекстуальные связи на пространстве всего рабочего корпуса текстов, а при возможности и шире — всей традиции. Это позволит находить и показывать сюжетные переключки, заимствования, точные, искаженные, неполные, скрытые цитаты, помогать в атрибуции текстов, установлении авторства, раскрытии преемственности текстов, периодизации.

Трансляция разноязычных текстов в единое семантическое пространство дополнит, а где-то полностью заменит словарный и тезаурусный подходы к сравнению и, соответственно, поиску текстов, заменив его поиском концептов. Топология семантического пространства не описывается сетевой метафорой гипертекста, межтекстовые переходы и операции определения близости выполняются иначе. Любой фрагмент документа — связный или нет — может работать как потенциальная отсылка к другим документам, и механизм этой адресации может быть сложнее буквального текстового повтора, т. е. опосредован переходами внутри семантического пространства. Новый тип моделей позволит находить альтернативные интерпретации текста и вообще строить многослойное представление текста, в котором для оригинала дан буквальный автоматический перевод, авторитетные канонические переводы (при их наличии), контекстуальные пояснения (от ИИ или от экспертов), комментарии и интерпретации, лингвистический разбор, историческая и культурологическая справка, связи с другими текстами. Все это может быть дополнительно отображено на «Карту уверенности», которая визуализирует степень согласия источников и экспертов в интерпретации каждого фрагмента.

Заметим, что корпус способен не только снабжать переводчика информацией, но и создавать в определенных условиях невыносимое для одиночного сознания семиотическое давление. Постановка корпуса под управление генеративной модели позволит демпфировать и дозировать это давление соразмерно возможностям человека.

Цифровая келья

Рабочий объект переводчика — порождаемый и редактируемый им текст. Пока этот динамический объект не завершен, он содержит разнородные лингвистические дефекты — несогласованные определения, структурные разрывы, неоконченные фрагменты, перечисление альтернативных переводов, просто опечатки. Работа традиционного лингвистического процессора с таким неполноценным текстом затруднена или просто невозможна, в то время как нейросетевые модели достаточно толерантны к подобным дефектам. Перевод любой степени готовности можно оценивать и автоматически сопоставлять его с оригиналом, прокладывая для переводчика сеть возможных маршрутов развертывания перевода.

Таким же оперативным должно быть выявление в тексте описок, переводческих искажений, пропусков, лишних вставок и советы по их устранению, либо просто автоматическое устранение, если исправленный вариант оценивается как единственно возможный.



Предварительный автоматический перевод текста множеством способов может подсказывать альтернативные пути развертывания перевода в любой точке и при любой степени его готовности. Предъявление альтернатив происходит локально, по запросу или срабатыванию определенных триггеров. В каждой точке текста ассистент должен давать варианты редакторских трансформаций текущего фрагмента на оценку переводчика, без чьего творческого сопротивления языковая модель была бы обречена на бессмысленный brute-force перебор.

Значительное внимание переводчика тратится на выискивание в тексте оригинала или перевода нужных фрагментов для их сверки и актуализации в оперативной памяти. Структурные языковые отличия приводят к смещениям переводных фрагментов относительно оригинала, нарушая параллелизм текстов. Поэтому все типовые лингвистические объекты должны быть автоматически предсказуемо выделены, хорошо различимы и соотнесены с переводом. Оригинал и перевод предъявляется переводчику не как монолитный текст, а как динамически настраиваемая структура, которая подсказывает узловые точки разных аспектов структуры текста и разных уровней иерархии этих структур, это может быть — на выбор — построение деревьев составляющих, зависимостей, валентностных и анафорических связей для основных и второстепенных членов предложения, для составляющих сложных предложений и т. п. Все регистрируемые в переводе типовые трансформации тоже могут получить автоматическую маркировку, которая будет сообщать переводчику о типе и качестве его собственных решений, а ИИ-ассистенту давать материал для обобщений. Можно предположить, что такая оснастка приведет к корректировке разработанной в переводоведении теории переводческих трансформаций.

Ассистент-критик станет первым читателем перевода, взяв на себя роль редактора и виртуального коллективного читателя. Он будет давать интерпретационный прогноз по восприятию перевода в общем контексте произведения (и даже шире) и показывать, какими способами может быть прочитан и понят каждый фрагмент будущим читателем, как изменится его интерпретация с изменением контекста, как изменить контекст, чтобы стимулировать выбор нужной интерпретации и блокировать нежелательные. Автокритик должен уметь решать задачи качественного стилистического анализа, а именно — сообщать, как оригинал или перевод соотносятся с нормами языка, которые могут быть заданы широко как нормы эпохи или узко как нормы определенного корпуса текстов. Отклонения от этих норм должны сопровождаться предложениями по их стандартизации.

Перечислю дополнительные функции, которые следует возложить на группу вспомогательных ассистентов.

В большом проекте обязательна документальная фиксация всех этапов работы. Необходимо протоколирование всех версий перевода и сопроводительных текстов и внутривидеопроектных коммуникаций: комментариев, введений и послесловий, примечаний переводчика и редактора, контекстной привязки использованных словарей и справочной литературы, и т. п. Например, должны фиксироваться все обращения к словарям в ходе перевода данного фрагмента,



все выбранные из словарей эквиваленты, эта информация анализируется и учитывается на следующих этапах работы.

В контексте организации переводческой работы специальный ассистент должен замыкать контур обратной связи, отслеживая реакцию переводчика на машинные предложения как материал для дообучения. Как показывает опыт внедрения ИИ-ассистентов в рабочие среды программистов (IDE), критичное значение в правильной организации работы с ассистентом играют самые мелкие детали, вплоть до скорости появления предложений от нейронной сети. Включенность пользователя в этот режим и степень кооперативности с ассистентом могут зависеть от тона и модальности формулировок, предъявляемых переводчику. Имеет значение длина переводческой подсказки, продолжительность ее демонстрации, способы ее принятия или отказа. В одних контекстах уместна всплывающая подсказка, в других — внутритекстовая (inline). Правильное конструирование диалога ассистента с пользователем — критически важная часть комфортной и продуктивной работы, при этом помощь никогда не должна переходить ту грань, за которой она начинает отвлекать и тормозить работу.

Ни один из существующих цифровых инструментов переводчика не обладает подлинной адаптивностью. Адаптация сейчас фактически всегда односторонняя — человек приспосабливается к машине и вырабатывает операционные привычки, которые бы позволили сохранять продуктивность вопреки машине, а не с ее помощью. Движение в обратном направлении, если и происходит, то с запозданием: изменение функционала невозможно без разработчика софта. Генеративные модели должны компенсировать недостаток кооперативности и гибкости, научиться принимать во внимание жизненный опыт человека, его телесность, эмоции.

Перевод на практике всегда подчиняется накопленным с опытом индивидуальным приемам и привычкам переводчика, которые, хотя и называются индивидуальным стилем, по сути приводят к шаблонности перевода. Причиной нередко становится банальная усталость. Задача будущего — научиться подстраиваться не только к переводному заданию, но и к квалификации и текущему физическому состоянию человека. Речь не о настройках пользовательского интерфейса, а о способности среды «чувствовать» состояние переводчика и различать фазы его сосредоточенности и потери концентрации внимания, определять степень сложности для него переводной задачи, диагностировать нештатные лингвистические ситуации, обнаруживать в переводе лексические, грамматические, стилевые стереотипы.

Заключение

Начавшееся вытеснение человека-переводчика из отрасли, по всей вероятности, продолжится, и с неизбежностью вторгнется в зону сакрального. Представленный здесь комплекс интеллектуальных агентов — это в том числе и способ смягчить этот процесс, скорректировав роли в системе человек-машина. Прогнозируемый выигрыш от гибридизации перевода с реализацией всего комплекса информационных и адаптивных функций будет разноплановым. В аспекте эргономики ожидается сокращение или устранение отвлека-



ющих, раздражающих и утомляющих непродуктивных операций, которые в совокупности снижают скорость и качество работы переводчика.

В аспекте эпистемологии машина поможет переводчику установить и сделать явными скрытые составляющие его собственного предпонимания. Ценности, интуитивные догадки, телесные навыки, эмоциональные реакции, культурные стереотипы получают канал экспликации для своевременной корректировки. Взамен от системы мышления переводчика потребуются активная рефлексия в отношении к новым интеллектуальным инструментам, так как без нее любые возникающие в этом симбиозе эффекты будут в восприятии переводчика случайностями.

Переводческий опыт станет интеллектуально и эмоционально более насыщенным и потребует средств обеспечения эмоциональной устойчивости. Особо болезненным для профессионального переводческого самовосприятия станет неизбежная трансформация общественной оценки, когда сегодняшнее недоверие к автоматическому переводу и безусловное доверие к переводчику-человеку поменяются местами.

Не следует опасаться утраты переводчиком своей индивидуальности, она сможет проявляться другим способом — через его вопрошание к генеративной модели, которое отразится в неповторимой истории активации весов модели.

Переводческий инструмент на основе генеративных моделей способен как расширять возможности переводческого поиска, так и нарушать его. Это гуманитарный инструмент, в который вложена колоссальная субъективная и коллективная составляющая, а следовательно — неконтролируемые и не полностью предсказуемые потенциальные свойства, которые в совокупности с большими объемами данных не позволяют верифицировать корректность их работы. Кроме того, пластичность цифровой среды устраняет четкие границы между инструментом и моделью, исходными данными и их интерпретацией, методом и предметом, затрудняя тем самым анализ и оценку инструментария. Все это — пока не получившие должного ответа составляющие «цифрового вызова». В целом нужно констатировать, что гуманитаристика оказалась в серьезной методико-гносеологической зависимости от компьютерной инженерии, и переводческая отрасль здесь не исключение, но именно в ней видны перспективы для успешной гармонизации человека и машины.

Список литературы / References

1. Балабан А. М. Прецедентная база данных для машинного перевода. *Международный семинар по машинному переводу (Москва, 5–10 декабря 1983 г.). Тезисы докладов*. М.: ВЦП научно-технической литературы и документации, 1983. С. 32–34 [Balaban A. M. Precedent Database for Machine Translation. *International Seminar on Machine Translation (Moscow, December 5–10, 1983). Proceedings*. Moscow: VCP of Scientific and Technical Literature and Documentation, 1983, pp. 32–34 (in Russian)].
2. Кибрик А. Е. О «невыполненных обещаниях» лингвистики 50–60-х годов. В: Перцов Н. В. (сост.) *Московский лингвистический альманах*. Вып. 1: *Спорное в лингвистике*. М.: Языки русской культуры, 1996. С. 230–233 [Kibrik A. E. On the “Unfulfilled Promises” of 1950s–60s Linguistics. In:



- Pertsov N. V. (ed.) *Moscow Linguistic Almanac*. Issue 1: *Controversies in Linguistics*. Moscow: Languages of Russian Culture, 1996, pp. 230–233 [in Russian]].
3. Скоринкин Д. Горький урок ABBYY: как лингвисты проиграли последнюю битву за NLP [Skorinkin D. ABBYY's Bitter Lesson: How Linguists Lost the Final Battle for NLP (in Russian)]. — Электронный ресурс: URL: <https://sysblok.ru/blog/gorkij-urok-abbyy-kak-lingvisty-proigrali-poslednjuju-bitvu-za-nlp/> (дата обращения: 06.07.2025).
 4. Сыртыпова С.-Х. Д., Костыркин А. В., Парфёнов М. В. Автоматизация тибетско-русского перевода — к проекту перевода Тибетского буддийского канона. *Проблемы сохранения объектов культурного наследия и новейшие цифровые технологии. Сборник материалов круглого стола. Санкт-Петербург, 22 мая 2025 г.* М.: ИВ РАН. С. 153–169 [Syrtyпова S.-Kh. D., Kostyrkin A. V., Parfyonov M. V. Tibetan–Russian Translation Automation — towards the Tibetan Buddhist Canon Translation. *Preservation Challenges for Cultural Heritage Sites and Cutting-Edge Digital Technologies: Proceedings of the Roundtable Discussion*. Saint Petersburg: Institute of Oriental Studies, 2025, pp. 153–169 (in Russian)].
 5. Фрумкина Р. М. «Куда ж нам плыть?..». В: Перцов Н. В. (сост.) *Московский лингвистический альманах*. Вып. 1: *Спорное в лингвистике*. М.: Языки русской культуры, 1996. С. 67–81 [Frumkina R. M. “Where should we sail?..”. In: Pertsov N. V. (ed.). *Moscow Linguistic Almanac*. Issue 1: *Controversies in Linguistics*. Moscow: Languages of Russian Culture, 1996, pp. 67–81 (in Russian)].
 6. *ACM Curriculum Committee*. 1968, March. Vol. 11. No. 3, p. 196.
 7. Melby A. K. Translators and Machines — Can they Cooperate? *Meta*. 1981;26(1):23–34.
 8. Nagao M. A Framework of a Mechanical Translation between Japanese and English by Analogy Principle. In: Elithorn A., Banerji R. (eds) *Artificial and Human Intelligence*. Elsevier Science Publishers B. V., 1984.
 9. Sutton R. *The Bitter Lesson*. 2019. — Электронный ресурс: URL: <http://www.incompleteideas.net/IncIdeas/BitterLesson.html> (дата обращения: 06.07.2025).

Список сокращений

АБИ — см. ABBYY.

ИИ — искусственный интеллект.

ABBYY — некогда российская (основана в Москве в 1989 г.), сейчас американская технологическая компания, специализирующаяся на автоматизации обработки документов с использованием искусственного интеллекта.

ACM — Association for Computing Machinery.

AI — Artificial Intelligence (искусственный интеллект).

FAHQT — Fully Automated High Quality Translation (полностью автоматизированный высококачественный машинный перевод).

IDE — Integrated Development Environment (интегрированная среда разработки).



IT — Information Technology (информационные технологии).

CAT — Computer-Assisted Translation (компьютерная поддержка переводческой деятельности).

GPT — Generative Pre-trained Transformer (генеративный предобученный трансформер).

LLM — Large Language Models (большие языковые модели).

NLP — Natural Language Processing (обработка естественного языка).

Информация об авторе

Костыркин Александр Вячеславович — кандидат филологических наук, старший научный сотрудник Лаборатории цифровых исследований современного Востока, Институт востоковедения РАН, Москва, Россия; kost@ivran.ru, <https://orcid.org/0009-0002-8167-1529>.

Раскрытие информации о конфликте интересов

Автор заявляет об отсутствии конфликта интересов.

Информация о статье

Статья поступила в редакцию 04.02.2025; одобрена рецензентами 15.05.2025; принята к публикации 15.05.2025; опубликована 10.10.2025.

Автор прочитал и одобрил окончательный вариант рукописи.

Information about the author

Alexander V. Kostyrkin — PhD (Philology), Senior Researcher, Institute of Oriental Studies of the Russian Academy of Sciences, Moscow, Russia; kost@ivran.ru, <https://orcid.org/0009-0002-8167-1529>.

Conflicts of Interest Disclosure

The author declares that there is no conflict of interest.

Article info

The article was submitted 04.02.2025; approved after reviewing 15.05.2025; accepted for publication 15.05.2025; published 10.10.2025.

The author has read and approved the final manuscript.