

Рамиля Габдулхаевна Миахова

Уфимский университет науки и технологий, кандидат филологических наук,  
доцент кафедры лингводидактики и перевода, Уфа, Россия  
e-mail: miftahovar@yandex.ru

## Решения в сфере переводческих технологий

**Аннотация.** В статье проанализированы принципы работы, процесс эксплуатации систем автоматизированного перевода; представлены текущее состояние и перспективы развития с целью ориентировать крупные компании при выборе переводческих решений, профессиональных и начинающих переводчиков при выборе инструментов перевода. Статья во многом опирается на большой опыт в этой сфере его автора.

**Ключевые слова:** CAT-инструменты, системы автоматизированного перевода, память переводов, база терминов, управление переводами, машинный перевод, автоматизированный перевод.

Ramilya G. Miftakhova

Ufa University of Science and Technology, Candidate of Philological Sciences,  
Associate Professor of the Department of Linguodidactics and Translation, Ufa, Russia  
e-mail: miftahovar@yandex.ru

## Solutions in the Field of Translation Technologies

**Abstract.** The article analyses the principles of operation, the process of operating automated translation systems, presents the current state and development prospects in order to guide large companies in choosing translation solutions, as well as professional and novice translators in choosing translation tools. The article is largely based on the author's extensive experience in this field.

**Keywords:** CAT-tools, automated translation systems, translation memory, database of terms, translation management, machine translation, automated translation.

---

### Введение (Introduction)

Системы автоматизированного перевода (Computer-Assisted Translation), так называемые CAT-инструменты, начали активно развиваться в начале текущего столетия. Параллельно развивался и машинный перевод, но он не мог конкурировать с автоматизированным из-за недостатков статистической модели и той, что основана на правилах. С появлением нейронных технологий в переводе ситуация изменилась кардинально. Качество перевода стало практически соответствовать требованиям, и теперь уже над автоматизированными системами нависла угроза невостребованности. Разработчикам CAT-инструментов ничего не оставалось, как последовать фразе, приписываемой Юлию Цезарю: *Если не можешь победить своего врага, сделай его своим другом*. Так появились системы автоматизированного перевода с интегрированными в них системами машинного перевода. Вместе с тем изменились требования к навыкам, которыми должны обладать переводчики.

Современные высокотехнологичные переводческие услуги базируются на платформе управления гибкими переводческими решениями, которая создана специально, чтобы

поддерживать бесперебойную сквозную локализацию. Благодаря своей структуре, она может легко управлять большими объемами текста и способна адаптироваться к любому, даже малейшему изменению контента, а также к часто меняющимся требованиям к итеративному переводу, профессиональному компаниям, ориентированным на современные цифровые технологии.

Гибкие базы памяти переводов должны разрабатываться так, чтобы перевод легко встраивался в бизнес-процесс — уровень совмещения, делающий бизнес более гибким и эффективным. Как минимум, должен быть стандарт интерфейса для поддержки пользовательской настраиваемой интеграции (см. рис.). В идеале необходима платформа с большим и широким спектром связующих звеньев контента для упрощения взаимодействия с бизнес-системами.

Решением такой задачи могут служить гибридные средства автоматизированного перевода. Они зарекомендовали себя в течение нескольких десятилетий. Но только совсем недавно были оснащены инструментами для бесшовного переключения между онлайн- и офлайн-режимами и позволили переводчикам использовать различные

The screenshot shows the Memsource CAT tool interface. At the top, it displays 'CAT translate-en-ru-T.mxliff - Веб-редактор' and 'Anastasia / CAT translate-en-ru-T.mxliff'. The main area shows a table of translation memory entries. The first few rows are:

	Source (en)	Target (ru)
6	880771c06dd19d29-DME	880771c06dd19d29-DME
7	IP:	IP:
8	62.76.247.2	62.76.247.2
9	UTC time:	UTC время:
10	2024-05-08T06:37:59+00:00	2024-05-08T06:37:59+00:00
11	Browser:	Браузер:
12	Mozilla/5.0 (Windows NT 6.3; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/86.0.4240.193 Safari/537.36	Mozilla/5.0 (Windows NT 6.3; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/86.0.4240.193 Safari/537.36
13	<a href="#">View recommended browsers</a>	Увидеть рекомендуемые браузеры.

On the right side, there is a sidebar titled 'CAT' with a list of items:

- 1 [View recommended browsers](#) 101 Увидеть рекомендуемые браузеры.
- 2 view TB увидеть
- 3 browser TB браузер
- 4 [View recommended browsers](#) 100 Увидеть рекомендуемые браузеры.

Below the table, there are tabs for 'Context note' and 'Preview'. The 'Preview' tab is selected. It shows the details of the current segment, including the ID, IP, UTC time, and browser information. There is also a link to 'View recommended browsers'.

At the bottom of the interface, it says 'Confirmed segments 13 / 13 · words 41 / 41 · chars 311 / 311 | Active chars 26 / 31 · created by nastya\_polyakova\_97 · modified by nastya\_polyakova\_97'.

## Интерфейс системы автоматизированного перевода

устройства для доступа к своей работе в любое время и с любого места. Для достижения максимальной гибкости и эффективности обеспечивается глубокий уровень интеграции между гибридными системами автоматизированного перевода, гибкими базами памяти переводов и лингвистическими интеллектуальными системами. Всё это должно быть встроено в одно переводческое решение для достижения максимальной производительности.

### Методы (Methods)

Методология исследования САТ-инструментов в данной статье заключается в изучении их заметно расширявшегося функционала. Сравнительный, сопоставительный и описательный анализ позволяют погрузить читателя в новейшие технологии и решения в сфере перевода. Представители крупных международных компаний делятся преимуществами, которые они получили в результате использования современных систем автоматизированного перевода. Анализ отечественных инструментов перевода показывает, что развитие этой сферы заметно отстает и отечественным компаниям приходится пользоваться зарубежными продуктами в ущерб безопасности и конфиденциальности данных.

### Результаты и обсуждение (Results and Discussion)

Машинный перевод — одна из первых систем, которая начала использовать технологии искусственного интеллекта (далее — ИИ). Однако бизнес, который стремится усилить свой контент и не испортить свою репутацию, не может полагаться на качество, предлагаемое бесплатными общедоступными инструментами машинного перевода. Для предприятий разрабатываются отдельные системы машинного перевода. Они характеризуются безопасностью, легко адаптируются и масштабируются, таким образом обес-

печивая правительственный и коммерческим организациям защиту данных, доступ к новым рынкам и дальнейший рост. По сведениям компании RWS, более 80 % из 100 мировых брендов, среди них все топовые фармацевтические компании, инвестиционные банки и юридические компании, используют адаптивный машинный перевод Language Weaver [1]. Такая технология легко развертывается в облачных сервисах предприятия, а пользовательский интерфейс для перевода текстов и документов базируется на браузере. API (интерфейс прикладного программирования) интегрирует перевод в режиме реального времени в существующие системы и рабочие процессы, кроме того, может подключаться к популярным приложениям, таким как Microsoft Office, Chrome, Edge, Firefox и др. Таким образом, машинный перевод коренным образом меняет то, как люди разных стран с разными культурами могут совместно развиваться и постигать новые рынки.

Технологии в области перевода растут стремительно. Только вышеупомянутая компания зарегистрировала более 50 патентов, связанных с нейросетевыми инновациями в лингвистике, что позволяет ей поддерживать 3000 языков и более 150 языковых пар. При применении такой технологии для одной языковой пары производительность перевода превышает 2000 слов в минуту, в масштабе 150 языковых пар достигает 500 000 слов в минуту. Система настраивается специально для конкретной организации, учитывает контекст домена и терминологию. С увеличением объема перевода система максимально адаптирует нейросетевые возможности для всей организации. Несмотря на мощный «движок» искусственного интеллекта, интерфейс системы достаточно прост для переводчика, что позволяет легко управлять масштабными переводческими проектами. Сегодня на рынке систем автоматизированного перевода

## ЯЗЫКОЗНАНИЕ

около 25 компаний, есть среди них и автономные. Однако можно прогнозировать, что высокие технологии, необходимость в больших базах данных и конкуренция оставят на рынке не более пяти из них в ближайшие годы, остальные, скорее всего, будут поглощены более продвинутыми компаниями или вовсе уйдут с рынка. Среди высокотехнологичных и быстроразвивающихся онлайн-систем на момент написания статьи следует отметить Phrase, Smartcat и XTM.

Более того, машинный перевод не является единственной формой искусственного интеллекта в лингвистике. Искусственный интеллект может быть использован не только для непосредственного перевода, но и для анализа контента, оценки качества и улучшения процесса постредактирования, всё это обеспечивает гибкость и эффективность. «С помощью искусственного интеллекта переводчики могут создавать креативные выражения, сохраняя при этом смысл исходного текста» [2, р. 29] (здесь и далее перевод наш. — Р. М.). ИИ расширил сферу применения CAT-приложений в мультимедийном пространстве, обеспечив перевод текста с рисунков и потока речи [3, р. 249].

Согласно Ч. С. Петеру, компания Trados летом 2024 г. запустила инновационную надстройку AI Essentials, которая упрощает внедрение искусственного интеллекта в переводческий процесс. Надстройка объединила два прорывных функционала, объявленных чуть ранее в этом году, — «Порождающий перевод» и «Умный редактор». Они были разработаны для улучшения качества перевода, ускорения процесса перевода и сокращения его стоимости без необходимости подписываться на большие языковые модели LLM [4].

Память переводов в системе автоматизированного перевода должна быть глубоко интегрирована в корпоративную систему машинного перевода и другие лингвистические функции ИИ. Кроме того, автоматизированный перевод должен внедрять ИИ там, где в этом есть смысл. ИИ положительно влияет на процесс перевода в CAT-инструментах, но на многих уровнях роль человека трудно переоценить. Приложения ИИ хорошо справляются с задачами, которые не под силу человеку. Это поиск необходимых сегментов текста в огромной базе данных или автоматизация задач, которые человек склонен забывать или делать в них ошибки. Но ИИ тоже имеет ограничения и уязвимости. Ему тяжело даются метафоры и культурные нюансы, к тому же большие языковые модели могут галлюцинировать. Оптимальной является «коллаборация» человека с искусственным интеллектом. Как же должна выглядеть такая коллaborация на практике в контексте перевода?

Введение нейронного машинного перевода значительно преобразовало индустрию перевода и, как следствие, роль переводчика. Переводчику нет необходимости задумываться о структуре предложения, он может сфокусироваться на его редактировании. Редактор, в свою очередь, может взять на себя функции контроля качества, обеспечивая адекватный, соответствующий контексту и учитывающий культурные особенности перевод. Нейронные системы выполняют перевод, базируясь на уже переведенных текстах. «Эта технология хорошо зарекомендовала себя в условиях крупномасштабных параллельных корпусов, но плохо справляется

с малоресурсными языками» [5, р. 734]. Согласно Гуо М. (Guo M), они также некорректно обрабатывают профессиональную терминологию и большие тексты [6, р. 109]. При переводе художественной литературы, маркетинговых или юридических документов, там, где особенно важны нюансы и стиль, необходимо участие человека. Выполняя такое постредактирование в современной онлайн-системе автоматизированного перевода, переводчик невольно улучшает качество перевода нейронных систем. Это, в свою очередь, будет сокращать необходимое количество переводчиков-редакторов в будущем. Сегодня переводчику важно обладать не только лингвистическими знаниями, но и определенными навыками, чтобы использовать нейросетевые технологии и работать с переводческими инструментами. Некорректное использование может привести к этическим проблемам, связанным с персональными данными и конфиденциальной информацией. Таким образом, внедрение новых технологий освободит переводчиков от рутинной работы, а это приведет к тому, что заметно возрастет востребованность узких высококвалифицированных специалистов.

Базовые элементы всех CAT-инструментов состоят из памяти переводов и базы терминов. Память переводов сохраняет ранее переведенные сегменты; если в процессе перевода новых сегментов обнаруживаются похожие фразы или предложения, они автоматически всплывают, и переводчик вправе использовать их или игнорировать. Такая функция повышает эффективность работы и качество перевода [7, р. 132]. Процедура использования памяти перевода сводится к тому, что необходимо создать пустую базу данных памяти перевода или импортировать уже имеющуюся, она обычно имеет расширение .tmx и универсально совместима с любыми CAT-инструментами. Затем подлежащий переводу текст автоматически делится на сегменты, которые при необходимости лингвист может корректировать. Когда переводчик переходит на новый сегмент, память переводов начинает искать идентичный текущему сегменту переведенный блок. Если таковой найден, она предлагает его в качестве перевода в отдельном окне, его можно отредактировать или применить. Затем готовый сегмент подлежит сохранению в память перевода для последующего использования. Например, предложение *Контракт был подписан между двумя компаниями-поставщиками* было переведено как *A contract was signed between the two companies-suppliers* и сохранено в память переводов. Когда в процессе перевода лингвист столкнется, например, с предложением *Соглашение было подписано между двумя компаниями-поставщиками*, система предложит *A contract was signed between the two companies-suppliers*, покажет процент совпадения и выделит красным цветом слово *контракт* как подлежащее редактированию. Это значительно облегчает процесс перевода, его скорость и качество.

Следующим важным компонентом CAT-инструментов является база терминов. Кто-то возразит, ведь есть много специализированных словарей, зачем создавать новый. Но дело не только в правильном переводе терминов, а в том, что они должны быть последовательны по всему документу. Для перевода технической литературы это критически важно. Предположим, в начале документа предложение *Open*

*the cover* было переведено как *Откройте крышку*, а через несколько страниц документации логическое *Close the cover* стало *Закройте люк*. И люк, и крышка являются подходящими переводами для исходного *cover*, однако конечного пользователя могут ввести в заблуждение разные лексемы: возникнет вопрос, тот ли это люк или это совсем другая крышка. Процедура использования базы терминов заключается в том, что первоначально создается пустая база терминов или импортируется готовая. Базу терминов необходимо пополнить актуальными для задачи перевода словами и их переводами. При переходе к переводимому сегменту слова, содержащиеся в базе терминов, подчеркиваются, чтобы лингвист их не пропустил и воспользовался при необходимости их переводом в новом окне. Технология проверки качества фиксирует примененные и отвергнутые термины, количество сегментов с частичным или полным совпадением, которые были вставлены из базы переводов. Некоторые авторы присваивают ей функцию исправления орфографических ошибок [8, р. 86]. Но это некорректное утверждение. Для проверки орфографических ошибок системы автоматизированного перевода интегрируют или подключаются к специальным приложениям «спеллчекерам» (Spell Checker), а технология проверки качества (Quality Assessment) в CAT-инструменте проверяет только технические ошибки, а не орфографические: лишние пробелы, длину исходного текста и длину его перевода, пунктуацию (только в конце предложения), цифры, ссылки, адреса электронной почты, теги (отвечают за форматирование) и др. Функция может также генерировать финальный отчет. Такая статистика позволяет редакторам и менеджерам проекта быстро оценить качество и стоимость работ.

Процедура использования любого CAT-инструмента сводится к следующему:

- Создается проект. Проект характеризуется тем, что для него выбирается языковая пара и могут подключаться/отключаться дополнительные функции, например машинный перевод. Языковая пара в проекте может работать только в одном направлении. Кроме того, необходимо обратить внимание на расширения языков. При использовании разных расширений возможны проблемы при импортировании памяти переводов из одной системы в другую.

- В рамках проекта создается/импортируется память переводов и активируется.

- В рамках проекта создается/импортируется база терминов.

- В рамках проекта импортируется подлежащий переводу файл.

Открытие файла приводит к интерфейсу с сегментированным текстом, а также с множеством инструментов, необходимых в процессе перевода, например переход на новый сегмент, перенос исходной строки в целевую, сохранение переведенного сегмента в память перевода и др.

CAT-приложения прекрасно подходят для перевода технической документации, руководств к использованию оборудования или программного обеспечения, где обязательно должны быть единый стиль, последовательная терминология и достоверный перевод. Они не предназначены для перевода художественной литературы, потому что послед-

нее — это творческая работа, куда автор может привнести свое отношение к исходному произведению. Более того, такая литература может содержать сложные по структуре предложения, культурный контекст или риторические фразы, с которыми системы автоматизированного перевода трудно справляются [9, р. 46]. Они также малоэффективно обрабатывают многозначные слова и метафоры [10].

Какими навыками должен обладать переводчик для работы в системе автоматизированного перевода, интегрированной в ИИ? Сегодня роль переводчика претерпела изменения, профессиональные переводчики и редакторы, а скорее лингвисты, как их принято сегодня называть, это специалисты в области высоких технологий, креативные лингвисты и консультанты-культурологи. Современные технологии позволяют им не только преобразовать большие объемы текста с одного языка на другой, но и творчески подходить к построению предложения, оптимизировать контент, преобразовывать его для разных форматов и устройств, всё это находит отклик у целевой аудитории. Они прекрасно ориентируются в хитросплетениях и жаргонах бизнеса, для которого переводят, и работают с разными видами исходного материала — помимо традиционных документов, с мультимедийными данными или приложениями, — предоставляя большой спектр услуг от перевода до транскреации и оптимизации поисковых систем.

Современные системы автоматизированного перевода решают не только задачи непосредственного перевода, но и задачи управления контентом, управления переводом [11]. Например, компании Emirates необходимо поддерживать 168 веб-сайтов на 28 языках с последовательным, точным, быстро обновляемым контентом, который должен соответствовать бренду. До появления автоматизированных систем большим вызовом был не только огромный объем поддерживаемого и переводимого контента, но и стиль, культурные особенности полилингвальных сайтов. Необходимо было создать централизованное управление контентом и процессом опубликования, в то же время дать возможность региональным отделениям эффективно обновлять свой контент на 28 языках, включая арабский и японский. Решением стала комплексная система автоматизированного перевода, позволившая упростить все сложности, связанные с созданием, управлением и доставкой многоязычного контента в разные ресурсы. В результате работы 1,8 млн слов сохранены в память перевода и только 260 000 слов были совершенно новыми. О масштабах перевода только для одной этой компании говорят следующие цифры: в день 40 000 переводческих заданий для 370 000 страниц веб-ресурсов. Каждый из 168 веб-сайтов с 2200 страницами система позволяет легко обновлять централизованно в несколько кликов.

Даника Шай, руководитель отдела переводов компании Xiaomi, отмечает: «С комплексными решениями систем автоматизированного перевода мы развиваем бизнес без необходимости нанимать новых сотрудников, мы поставляем больше последовательных переводов с высочайшим качеством, с меньшими усилиями и себестоимостью» [12]. Команда переводчиков компании занимается локализацией

веб-ресурсов, маркетинговых материалов, технической документации, пользовательских интерфейсов и других материалов на 70 языков, всего около 10 млн слов в год. Технологии автоматизированного перевода с интегрированным блоком управления контентом избавили компанию от необходимости повторно создавать один и тот же или похожий контент, который приходилось обновлять и отдавать на перевод по несколько раз разным переводческим командам. Всё это приводило к разным уровням качества и непоследовательной терминологии. Благодаря использованию новых переводческих решений, Xiaomi высвободила 2000 рабочих дней, процесс локализации ускорился на 70 %, на 75 % сократилось время опубликования материалов на веб-сайтах разных стран, заметно улучшились качество и последовательность переводов.

### Заключение (Conclusion)

Среди быстроразвивающихся CAT-инструментов заметно выделяется чешский Phrase. Западноевропейским компаниям принадлежат еще с десяток крупных полнофункциональных инструментов для автоматизированного перевода, как с веб-интерфейсом, так и стационарные. Несмотря на внешние отличия, всех их объединяет единый функционал: это память переводов, база терминов, технология сегментации, а также средства проверки орфографических и технических ошибок и, конечно, интеграция с системами нейронного машинного перевода. Значительными показателями на рынке CAT-переводов отметились китайские компании, такие как Xue Ren CAT, Transmate и Ya Xin CAT. Российский рынок технологических решений в сфере перевода заметно отстает. Относительно недавно созданный «Катмант» от компании «Амберит» пока мало известен не только за рубежом, но и в России. Отзывы и комментарии по нему автору статьи тоже не удалось найти.

Для успешного решения задач международного экономического взаимодействия необходимо создавать отечественные системы автоматизированного перевода с возможностью интеграции с другими решениями, связанными с процессом локализации и перевода. Среди основных таких решений — машинный перевод на основе ИИ, который был может использовать глоссарий CAT-инструмента для последовательного перевода терминов, а также адаптироваться к переводимому домену без дополнительных настроек. «Искусственный интеллект не только помогает переводчикам улучшать эффективность и качество перевода, но и способствует автоматизации переведческого процесса» [13, р. 72]. Следующим важным элементом, который должен быть встроен в CAT-инструмент, является блок управления и автоматизации рабочего процесса. Он должен отслеживать проекты в режиме реального времени, автоматизировать такие задачи, как запуск проекта, выбор лингвистов, предварительный перевод, расчет стоимости, генерирование и отправка электронных сообщений, распределение заданий.

Среди других платформ, подлежащих интеграции с системами автоматизированного перевода, — масштабируемая платформа локализации программного обеспечения, веб-сайтов, игр. Сегодня российские компании пользуются зарубежным программным обеспечением для реализации переводческих проектов, что может негативно отразиться на безопасности и защите данных, так как весь переведенный контент сохраняется на серверах владельцев приложений и баз данных. Между тем большая интеграция, автоматизация, гибкость и искусственный интеллект, встроенные в передовые переводческие сервисы, означают, что любой участник этого процесса волен добавлять креатив. Он становится быстрым и расторопным, способен переводить больший объем за меньшее время, может уделять больше внимания там, где действительно важно творчество.

1. Secure Your AI Advantage for Today and Beyond // RWS : [site]. URL: <https://clck.ru/3QXLDQ> (дата обращения: 08.08.2025).
2. Xue J. The impact of Chat GPT on AI-assisted Translation // Engl Square. 2024. No. 6. P. 27–30. URL: <https://clck.ru/3QXLHP> (дата обращения: 08.08.2025).
3. Semantic-Guided Natural Language and Visual Fusion for Cross-Modal Interaction Based on Tiny Object Detection // Xian-Hong Huang, Hui-Kai Su, Chi-Chia Sun, Jun-Wei Hsieh. 2025. DOI: 10.48550/arXiv.2511.05474
4. Trados Announces AI Essentials Alongside Latest Round of Feature Releases // RWS : [site]. URL: <https://clck.ru/3QXLJS> (дата обращения: 08.08.2025).
5. Unsupervised Neural Machine Translation Model Based on Pre-Training / Q. Xue, J. Li, Zh. Gong, D. Xu // Computer Engineering & Science. 2022. No. 44 (4). P. 730–736. The electronic version. URL: <http://joces.nudt.edu.cn/CN/abstract/abstract17365.shtml> (дата обращения: 08.11.2025).
6. Research on the Application of Artificial Intelligence in Machine Translation / M. Y. Guo, X. L. Zhang, H. L. Tang [et al.] // Journal of Henan University of Science and Technology (Natural Science). 2021. No. 42 (3). P. 97–104. The electronic version. URL: <https://clck.ru/3QXLNS> (дата обращения: 08.11.2025).
7. Wang H., Wang Y. A Research Review of Computer-Aided Translation in China (1980–2021). URL: <https://clck.ru/3QXLQf> (дата обращения: 08.08.2025).
8. Wei Ch. Computer-Assisted Translation Software: Current Problems and Development Trends in the Context of Artificial Intelligence // International Journal of Advanced Academic Studies. 2024. No. 6 (11). P. 84–90. DOI: 10.33545/27068919.2024.v6.i11b.1299
9. Yu Q. Research on the Innovation and Development of Computer-Assisted Translation Models Under the Background of Big Data // Journal of Huaihai Institute of Technology. 2019. No. 17 (12). P. 45–47.
10. Xu Y. The Integration of Computer-Assisted Translation Technology in Translation Education: Opportunities, Challenges, and Strategies for Innovation // International Journal of Education and Humanities. 2024. Vol. 17, no. 2. P. 113–116.

11. Alliance Data Centralizes Content Management with Tridion Sites // RWS : [site]. URL: <https://clck.ru/3QXLTx> (дата обращения: 20.08.2025).

12. Xiaomi Saves 2,000 Work Days with Tech-Enabled Content Management // RWS : [site]. URL: <https://clck.ru/3QXLUj> (дата обращения: 08.08.2025).

13. Li J. Design of Computer-Assisted Translation Software Dased on Artificial Intelligence Technology // Information Technology and Informatization. 2022. No. 3. P. 70–73. URL: <https://clck.ru/3QXLVe> (дата обращения: 20.11.2025).