

МАШИННЫЙ ПЕРЕВОД В БЕСПИЛОТНОЙ АВИАЦИИ: ВОЗМОЖНОСТИ И ВЫЗОВЫ

MACHINE TRANSLATION IN UNMANNED AVIATION: OPPORTUNITIES AND CHALLENGES

*E. Isaeva
K. Frolova
A. Popovtseva*

Summary: Modern machine translation technologies have made it possible to achieve a high level of automation in the work of technical translators, whose key tasks now are the creation of a high-quality domain-specific glossary and post-editing of automatically translated text. The Smartcat platform is one of the most popular open-source translation resources. Nevertheless, its capabilities, including the integration of linguistic resources based on artificial intelligence, have not been tested on technical documentation in the new actively developing industry of unmanned aircraft systems. In this paper, the peculiarities of creating the glossary of unmanned aircraft systems are considered, the experiment of machine translation of technical text is conducted, the quality of machine translation performed by online AI-based tools is compared. The results allow us to argue that modern machine translation tools are effective in optimizing the activities of technical translators in the field of unmanned aircraft systems, but do not eliminate them completely from the process. Our work contributes to the development of translation science, expanding the range of its application areas and the technical translator's toolkit.

Keywords: translation of unmanned aircraft systems terms, technical translation, glossary, UAS terminology, unmanned aerial systems, translation studies, Google Neural Machine Translation, DeepL, Yandex, Baidu Translate API, OpenAI GPT.

Исаева Екатерина Владимировна

кандидат филологических наук, доцент, Пермский
государственный национальный исследовательский
университет
ekaterinaisae@gmail.com

Фролова Ксения Владимировна

Пермский государственный национальный
исследовательский университет
kksyufrolova@gmail.com

Поповцева Анна Артемовна

Пермский государственный национальный
исследовательский университет
a.popovtseva@yandex.ru

Аннотация: Современные технологии машинного перевода позволили достичь высокого уровня автоматизации деятельности технического переводчика, ключевыми задачами которого в настоящий момент являются создание качественного отраслевого глоссария и постредактирования автоматически переведенного текста. В этом смысле платформа Smartcat является одной из наиболее популярных открытых переводческих ресурсов. Тем не менее, ее возможности, и в том числе подключение лингвистических ресурсов на основе искусственного интеллекта не были апробированы на технической документации в новой активно развивающейся отрасли беспилотных авиационных систем. В данной статье рассматриваются особенности создания глоссария подобных систем, проводится эксперимент машинного перевода технического текста, сравнивается качество машинного перевода, выполненного с помощью онлайн инструментов на основе искусственного интеллекта. Полученные результаты позволяют утверждать, что современные средства машинного перевода эффективны для оптимизации деятельности технического переводчика в области беспилотных авиационных систем, но не позволяют полностью исключить его из данного процесса. Наша работа вносит вклад в развитие переводоведения, расширяя спектр отраслей его применения и инструментарий технического переводчика.

Ключевые слова: : машинный перевод в области беспилотных авиационных систем, технический перевод, CAT-системы, глоссарий, терминология БАС, беспилотные авиационные системы, переводоведение, Google Neural Machine Translation, DeepL, Yandex, Baidu Translate API, OpenAI GPT.

Системы автоматизированного перевода (CAT – computer-aided translation) играют важную роль в профессиональной деятельности современного переводчика. Они позволяют получить максимальный результат с минимальными усилиями, за минимальное время в соответствии с техническим заданием [4]. CAT-технологии упрощают процесс перевода в различных предметных областях и могут успешно применяться для перевода технической документации в области бес-

пилотных авиационных систем (БАС). В данной статье рассматривается пример такого применения. Наше исследование позволяет оценить эффективность использования одной из наиболее популярных CAT-систем, Smartcat со встроенными опциями машинного перевода без оформления коммерческих подписок и выявить необходимость и обоснованность добавления дополнительных инструментов, основанных на глубоком машинном обучении и генеративных нейронных сетях.

1. Применение платформы Smartcat для перевода технической документации в области беспилотных авиационных систем

1.1. Обзор возможностей Smartcat для автоматизации технического перевода

САТ-инструменты — это программные приложения, которые помогают повысить эффективность и согласованность переводческой работы благодаря таким функциям, как память переводов, управление терминологией и машинный перевод. Машинный перевод может служить отправной точкой для переводчика, который затем вносит необходимые корректировки. Система Smartcat создана российской компанией ABBYY Language Services, что обуславливает ее интегрированность с одним из самых надежных словарей и, соответственно, высокое качество машинного перевода. Программа позволяет работать с различными форматами текста и предоставляет сервисы оптического распознавания, обрабатывающие такие файлы, как DOC, JPEG, PDF, JSON, XLIFF и другие [6].

Технология перевода Smartcat, основанная на искусственном интеллекте, особенно эффективна при техническом переводе, который предполагает перевод специализированных текстов, таких как руководства, инструкции и технические документы [6]. Такие тексты часто требуют точного перевода, чтобы обеспечить передачу смысла без потери информации или двусмысленности. Технология искусственного интеллекта Smartcat разработана таким образом, чтобы обучаться на основе правок и переводов, сделанных пользователями, тем самым улучшая качество и согласованность переводов с течением времени. Таким образом, этот сервис предлагает следующий функционал, способствующий оптимизации процесса перевода технического текста: создание глоссариев для поддержания единой терминологии, интеграция с системами машинного перевода для автоматического предварительного перевода, использование памяти перевода для повторного использования фрагментов текста, контроль качества перевода (проверка орфографии, грамматики, терминологии и стиля), форматирование текста для достижения точного визуального соответствия и сохранения мультимедийных элементов в переводном тексте.

1.2. Создание глоссария БАС

Как отмечают авторы статьи, посвященной использованию средств автоматизации для стандартизации терминологии в академическом переводе [8], для обеспечения согласованности терминологии, создание и управление глоссариев должно осуществляться группой экспертов, руководствующихся международными стандартами регулируемыми терминологические метамодели (ISO 16642:2017)[10], категории данных (ISO 12620-1:2022)[15]

и форматы обмена (ISO 30042:2019)[12] и оформления источников терминов (ISO 12615:2004) [9]. Для эффективной работы с платформой Smartcat рекомендуется создать глоссарий соответствующей предметной области. Глоссарий позволяет обеспечить последовательность и точность при переводе терминологии, что имеет большое значение в сфере высоких технологий БАС.

Рассмотрим процесс создания англо-русского глоссария в Smartcat для перевода технической документации, относящейся к отрасли БАС, включивший следующие последовательные шаги:

1. Определение предметной области и целевой аудитории: анализ исходных текстов по БАС для погружения в узус и выявления ключевых терминов относящихся к области беспилотных авиационных систем (национальные и международные стандарты, отраслевые форумы и др. [11, 14, 16, 17]); учет особенностей целевой аудитории статьи (разработчики БАС, летный состав, регулирующие органы, пользователи и др.).
2. Сбор терминологии: выделение всех специализированных терминов из ГОСТа терминов и определений [11]; расширение списка терминов, путем обращения к исходным текстам по БАС, специализированным словарям и справочникам в области авиации и робототехники, а также в ходе консультаций экспертов.
3. Заполнение глоссария: создание нового глоссария в Smartcat, выбрав соответствующую языковую пару (в нашем случае – английский-русский); заполнение терминологической статьи, указав для каждого термина следующие данные: исходный термин, переводной термин, определение или пояснение к термину, контекст использования термина, ссылки на источники, соответствующие нормативные акты или стандарты.
4. Включение глоссария в проект перевода: заключительный шаг по подготовке глоссария – это добавление созданного глоссария к проекту перевода в Smartcat. Глоссарий должен использоваться последовательно на всех этапах перевода документации по БАС. Следует регулярно обновлять глоссарий, чтобы добавить новые термины, синхронизировать его с развитием отрасли и учесть изменения в нормативных документах.

1.3. Анализ примеров перевода технических текстов с использованием Smartcat.

Чтобы показать возможности Smartcat, мы проанализировали несколько примеров перевода фрагментов технической документации, связанной с беспилотными летательными аппаратами и авиамоделями «The Drone and Model Aircraft Code» («Правила эксплуатации дронов и авиамodelей») [13]. Этот документ, выпущенный Управ-

лением гражданской авиации в 2019 г. с целью повысить осведомленность людей в корректном использовании дронов, представляет собой совокупность инструкций и правил, которые регулируют безопасное использование беспилотных летательных аппаратов. Он охватывает такие темы, как юридические требования для полетов, ограничения по высоте полета и расстоянию от людей, правила полетов рядом с аэропортами, меры безопасности до, во время и после полета, защита частной жизни, получение идентификатора пилота и оператора и др. [13]

Рассмотрим два примера перевода отрывков из данного документа с помощью инструмента Smartcat. Пример 1 иллюстрирует особенности перевода терминологии и ссылок на нормативные акты (Таблица 1); пример 2 демонстрирует необходимость постредактирования машинного перевода, выполненного в Smartcat (Таблица 2).

Таблица 1.

Терминология и ссылки на нормативные акты.

Исходный текст	Переводной текст
Drones and model aircraft in the law. Acts and regulations. <...> CAP1789A: Consolidated version of the EU UAS Implementing Regulation. CAP1789B: Consolidated version of the EU UAS Delegated Regulation.	Дроны и авиамодели в законодательстве. Законы и нормативные акты. <...> CAP1789A: Консолидированная версия Регламента ЕС по внедрению БАС. CAP1789B: Консолидированная версия делегированного регламента ЕС по управлению БАС.

Проанализируем результат машинного перевода. Термины *drones and model aircraft* переведены как *дроны и авиамодели*, возможны варианты *беспилотные летательные аппараты* и *авиамодели* более частотные в русскоязычном профессиональном дискурсе БАС. Ссылки на нормативные акты *CAP1789A/B* оставлены без перевода, так как это обозначения конкретных документов. Названия документов переведены корректно.

Таблица 2.

Машинный перевод и редактирование.

Исходный текст	Переводной текст
You must be over 18 to register as an operator. If you're under 18, you'll need to ask a parent or guardian to register for an operator ID.	Вам должно быть больше 18 лет, чтобы зарегистрироваться в качестве оператора. Если вам меньше 18 лет, вам нужно будет попросить родителя или опекуна зарегистрироваться для получения удостоверения оператора.

Данный отрывок демонстрирует стилистические различия исходного и переводного текста. В английском языке официально-деловой стиль в инструкциях и технической документации допускает использование местоимения второго лица – *you*, тогда как в русском языке

местоимения второго лица: *ты* и *вы* характерны для разговорного и публицистического стиля. Таким образом, данный отрывок следует перевести следующим образом – *Регистрация в качестве оператора возможна только для лиц старше 18 лет. На лиц, не достигших совершеннолетия, идентификатор оператора оформляется по заявлению родителя или законного представителя.*

Результаты анализа показывают, что использование Smartcat позволяет оптимизировать процесс перевода технической документации, делая его более эффективным, быстрым и качественным. Интеграция различных функций платформы, таких как глоссарии, машинный перевод, память переводов и инструменты контроля качества, форматирование, помогает улучшить производительность и согласованность переводов. Примеры демонстрируют, как Smartcat помогает справиться с основными сложностями перевода технических текстов, включая работу с терминологией и сохранение стиля. Конечно, машинный перевод все еще требует редактирования, но удобный интерфейс Smartcat позволяет быстро исправить ошибки и добавить новые термины в глоссарий.

2. Выбор дополнительных инструментов для машинного перевода технической документации БАС

В расширенной версии Smartcat есть возможность использования различных модулей машинного перевода, например Google Neural Machine Translation, DeepL, Yandex, Baidu Translate API, OpenAI GPT и др. Сравним потенциал ИИ для технического перевода в области БАС на примере одного предложения из представленного выше текста «The Drone and Model Aircraft Code» (Пример 3).

Пример 3: *Drones are typically powered by electric motors and equipped with navigational systems such as GPS receivers, gyroscopes, accelerometers, magnetometers, altimeters, and other sensors to help them remain stable while in flight.*

Google переводчик:

Дроны обычно приводятся в движение электродвигателями и оснащены навигационными системами, такими как GPS-приемники, гироскопы, акселерометры, магнитометры, высотометры и другие датчики, которые помогают им оставаться стабильными во время полета.

Яндекс переводчик:

Дроны обычно приводятся в действие электрическими двигателями и оснащены навигационными системами, такими как GPS-приемники, гироскопы, акселерометры, магнитометры, высотометры и другие датчики, которые помогают им оставаться стабильными во время полета.

DeepL:

Дроны обычно приводятся в движение электродвигателями и оснащены навигационными системами, такими как GPS-приемники, гироскопы, акселерометры, магнитометры, высотомеры и другие датчики, которые помогают им сохранять стабильность в полете.

Baidu Translate API:

БПЛА обычно питаются электродвигателями и оснащены навигационными системами, такими как GPS-приемники, гироскопы, акселерометры, магнитометры, высотомеры и другие датчики, чтобы помочь им оставаться стабильными во время полета.

ChatGPT:

Дроны обычно оснащены электродвигателями и навигационными системами, такими как приемники GPS, гироскопы, акселерометры, магнитометры, высотомеры и другие датчики, которые помогают им оставаться стабильными в полете.

Как представлено в примере 3, машинный перевод, выполненный представленными выше сервисами, сохраняет смысловое содержание грамматические и стилистические особенности исходного текста. Все машинные переводчики, кроме Baidu Translate API, сохранили популярный термин *дрон*. Комплекующие, в т. ч. *electric motors* (*электродвигатели*) и перечень датчиков также приведены в соответствии с их эквивалентами на русском языке. Однако, следует отметить, что термин *powered*, который в данном контексте следует перевести как *приводятся в движение*, корректно представлен только в переводах от Google и DeepL. Яндекс предложил вариант *приводиться в действие*, что более применимо к механизмам, не предполагающим перемещение в пространстве. В переводе Baidu Translate API используется дословный термина *powered* – *питаются*, что имеет эффект олицетворения, искажающего смысл предложения.

Интересное решение предложено генеративной нейросетью, опустившей перевод термина *powered* и объединившей *электродвигатели* и *навигационные системы* в словосочетание с глаголом *оснащены*. В данном случае имеет место упрощение предложения (сема – «быть источником энергии» и «приводить в движение» подразумевается наличием термина *электродвигатель*), что положительно влияет на легкость восприятия текста. Тем не менее, технический переводчик должен быть осторожен при подобных трансформациях, чтобы избежать возможных искажений технических характеристик устройств.

Термин *remain stable* был переведен с применением калькирования почти всеми онлайн переводчиками как *оставаться стабильными*, за исключением DeepL, в котором представлен вариант *сохранять стабильность*, что представляется нам более удачным, т. к. помогает избежать олицетворения. Тем не менее, DeepL единственный допускает ошибку при переводе следующего термина – *in flight* – *в полете*, который правильнее было бы перевести *во время полета*, как представлено в остальных переводах.

Поэтому мы можем сделать вывод, что одной из «слабостей» машинного перевода является сложность соблюдения стилистических и грамматических норм русского языка. Например, во время расширенного анализа мы заметили большое количество тавтологии и плеоназмов «*двигатель должен быть достаточно мощным, чтобы обеспечивать достаточную тягу*», отсутствие согласованности между частями предложений и высокую частотность перевода при помощи калькирования. Таким образом, мы приходим к заключению, что машинные переводчики способны передать основное содержание исходного текста, однако постредактирование компетентным специалистом является неотъемлемым этапом автоматизированного и машинного перевода.

3. Дискуссия об особенностях технического перевода в области БАС с применением Smartcat

Как и ожидалось, Smartcat значительно упорядочивает процесс перевода и, при условии добавления расширенного глоссария терминов БАС, позволяет стандартизировать использование языка для профессиональных целей. Интересное наблюдение, полученное в ходе сравнения инструментов на основе ИИ, показало, что подключение современных средств, таких как DeepL, ChatGPT и др. позволяет достичь «требуемой репрезентативности перевода», сохраняя при этом «неточности перевода в частных случаях на микроуровне» [1, с. 62] и не позволяя полностью автоматизировать процесс перевода технической документации, и требует последующего редактирования автоматически переведенного текста техническим переводчиком. Такое заключение согласуется с выводами других исследователей данного вопроса, полученными на материале других предметных областей. Например, авторы отмечают, что при переводе юридических текстов сохраняется потребность в исследовании и учете сложных нюансов отраслевого языка, улавливании тонкостей профессиональной коммуникации и адаптации к особенностям жаргона [7]. При локализации бизнес-процессов были выявлены «как стилистические погрешности, так и потеря смысловых нюансов, что ведет к мелким и крупным искажениям значения» [2, с. 133]. Продемонстрирована необходимость стилистической правки автоматически переведенных документов, т. к. «стилистические нормы иноязычных

документов различны» [3, с. 111].

Таким образом, «независимо от полноты автоматического словаря системы и его привязки к предметной области, результат машинного перевода требует постредактирования на уровне синтаксической структуры всего предложения, на лексическом уровне для уточнения и/или изменения переводов отдельных слов и словосочетаний, введения переводов для незарегистрированных в словаре единиц, а также для изменения морфологических характеристик рода, числа, падежа, уточнения форм времени и залога, введения корректной пунктуации» [5, с. 46]. При интерпретации результатов следует отметить, что они были получены с помощью открытых версий приложений (без оформления коммерческих подписок) и памяти переводов.

В данной статье мы продемонстрировали потенциал автоматизации деятельности технического переводчика при переводе документации в области БАС. На примере перевода отрывков из документа «The Drone and Model Aircraft Code» была дана оценка машинного перевода, выполненного в web-приложении Smartcat и предложе-

ны возможные авторские правки. Также было проведено сравнение дополнительных инструментов машинного перевода, основанных на ИИ, которые могут быть дополнительно подключены в Smartcat. Все эти средства автоматизации продемонстрировали свою эффективность в переводе текстов отрасли БАС при условии пост-редактуры техническим переводчиком, погрузившимся в данную отрасль и обладающим соответствующей терминологической компетенцией.

Дизайн исследования предполагал минимизацию роли опыта технического переводчика в решении задачи перевода технической документации в области БАС. Полагаем, что оформление коммерческих подписок, последних версий GPT и других моделей генеративных сетей, использование специализированной памяти переводов в области БАС помогли бы довести процесс машинного перевода технического текста практически до совершенства. Наше исследование может быть полезно как начинающим переводчикам, так и опытным, актуализирующим для себя новые возможности машинного перевода или новую активно развивающуюся отрасль БАС.

ЛИТЕРАТУРА

1. Афонасова В.Н. Оценка качества машинного перевода военного текста // Теория и практика военного образования. 2024. № 1(2). С. 56–63.
2. Исаева А.А. Машинный перевод и бизнес: благо или вред? // Язык. Культура. Коммуникация. 2023. № 24. С. 127–138.
3. Каменева Н.А. Технический перевод и особенности англоязычных научно-технических текстов // Вестник Московского Международного Университета. 2024. № 1(1). С. 110–114.
4. Ушакова А.О. Постредактирование машинного перевода технического текста // Евразийский гуманитарный журнал. 2022. № 3. С. 69–76.
5. Belyaeva L.N. Machine translation in the age of digitalization: new practices, procedures and resources // Terra Linguistica. 2023. № 1(51). P. 41–56.
6. Li J. Effectiveness of SmartCat Technology in Artistic, Scientific-Technical and Social-Publicist Translations: How Modern Software Influence the Specificity of Different Types of Translation // Journal of Psycholinguistic Research. 2023. № 5 (52). P. 1841–1854.
7. Moneus A.M., Sahari Y. Artificial intelligence and human translation: A contrastive study based on legal texts // Heliyon. 2024. № 6(10). P. 28106.
8. Serpil H., Durmuşoğlu-Köse G., Erbek M., Öztürk Y. Employing Computer-assisted Translation Tools to Achieve Terminology Standardization in Institutional Translation: Making a Case for Higher Education // Procedia-Social and Behavioral Sciences. 2016. № 231. P. 76–83.
9. ISO 12615:2004 – Bibliographic references and source identifiers for terminology work [Электронный ресурс]. URL: <https://www.iso.org/ru/standard/40359.html> (дата обращения: 15.06.2024).
10. ISO 16642:2017 – Computer applications in terminology – Terminological markup framework [Электронный ресурс]. URL: <https://www.iso.org/ru/standard/56063.html> (дата обращения: 15.06.2024).
11. ГОСТ Р 57258–2016 Термины и определения // 2018. С. 1–8.
12. ISO 30042:2019 – Management of terminology resources – TermBase eXchange (TBX) [Электронный ресурс]. URL: <https://www.iso.org/ru/standard/62510.html> (дата обращения: 15.06.2024).
13. The Drone and Model Aircraft Code | UK Civil Aviation Authority // UK Civil Aviation Authority [Электронный ресурс]. URL: <https://register-drones.caa.co.uk/drone-code> (дата обращения: 16.06.2024).
14. ГОСТ Р 59517–2021 Беспилотные авиационные системы. Классификация и категоризация // 2021. С. 1–7.
15. ISO 12620-1:2022 – Management of terminology resources – Data categories – Part 1: Specifications [Электронный ресурс]. URL: <https://www.iso.org/standard/79078.html> (дата обращения: 15.06.2024).
16. ISO 21384-4:2020 – Unmanned aircraft systems – Part 4: Vocabulary [Электронный ресурс]. URL: <https://www.iso.org/standard/76785.html> (дата обращения: 30.04.2024).
17. Аэронет 2035 [Электронный ресурс]. URL: <https://xn--2035-43d4a7chr0j.xn--p1ai/> (дата обращения: 30.04.2024).

© Исаева Екатерина Владимировна (ekaterinaisae@gmail.com), Фролова Ксения Владимировна (kksyufrolova@gmail.com), Поповцева Анна Артемовна (a.popovtseva@yandex.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»