

Крылов Роман Александрович

магистрант

Научный руководитель

Оношко Вячеслав Николаевич

канд. филол. наук, профессор

ФГБОУ ВО «Вятский государственный университет»

г. Киров, Кировская область

ЛИНГВИСТИЧЕСКИЕ ОСОБЕННОСТИ МАШИННОГО ПЕРЕВОДА НАУЧНОГО ТЕКСТА

***Аннотация:** в статье рассматриваются достоинства и недостатки машинного перевода научного текста. Целью данного исследования является анализ основных проблем, связанных с машинным переводом. В связи с этим нами рассматривалось качество компьютерного перевода, была произведена оценка адекватности машинного перевода, выполненного с помощью программ DeepL и Яндекс Переводчик. По результатам количественного анализа были составлены таблицы. В выводах были подведены итоги сравнения качества работы данных программ и определены дальнейшие перспективы развития программ машинного перевода.*

***Ключевые слова:** машинный перевод, научный перевод, оценка качества машинного перевода, адекватность, классификация, метод общего анализа.*

Введение.

В настоящее время появилось множество технических достижений способных облегчить жизнь и работу человека. Существенного прогресса в данном направлении достигли технологии перевода. Большой популярностью пользуются такие программы для машинного перевода как Яндекс Переводчик и DeepL. Заявлено, что данные программы способны справиться с такой сложной задачей как автоматический перевод сложных научных текстов. В данной статье исследуются современные

программы перевода и технологии затрагивающие, в частности, перевод текста на научную тематику, а также приводится оценка их эффективности.

Объектом исследования является текст оригинала, а также текст переводов, сделанных вручную и с применением технологий машинного перевода.

Целью работы является исследование проблематики машинного перевода научного текста. Для выполнения поставленной цели были выполнены следующие задачи:

- подобрать материал для исследования перевода научного текста, выполненного машиной;
- сравнить между собой и оценить два машинных перевода выполненных с помощью программ DeepL и Яндекс Переводчик;
- выявить перспективы дальнейшего развития средств машинного перевода.

Методы исследования: общий метод оценки качества переводов, метод качественного анализа, сопоставительный метод, метод изучения переводов.

Материалом исследования является научная статья «Misinformation in Social Media: Definition, Manipulation, and Detection».

Прежде чем приступить к анализу качества переводов, необходимо разработать методы оценки качества переводов технических текстов. В данной работе мы будем использовать общий метод оценки качества переводов.

Суть общего метода оценки качества переводов технических текстов заключается в том, что каждый перевод, полученный в рамках системы машинного перевода, оценивается по таким критериям как общая адекватность, общая читабельность, трудозатратность редактирования.

Таблица 1

Критерии для оценки качества перевода
с применением универсального подхода [Батуев 2021]

Критерий	Количество баллов
Общая адекватность	1–5
Общая удобочитаемость	1–5
Трудозатратность редактирования	1–5

Набор данных.

Для достижения целей нашего исследования мы выбрали и перевели вручную с английского языка на русский научный текст. Затем мы произвели перевод данного текста с помощью программ машинного перевода и сравнили результаты машинного перевода с результатами перевода, выполненного человеком.

Далее в статье приведены часть примеров, которые анализировались по данной классификации:

Оригинальная фраза из научной статьи.

«Content extracted from a user's posts has been studied in early research to directly identify misinformation spreaders, and a text classifier can be used to classify malicious users».

DeepL.

«Содержимое, извлеченное из сообщений пользователя, было изучено в ранних исследованиях для прямого определения распространителей дезинформации, а текстовый классификатор может быть использован для классификации вредоносных пользователей».

Яндекс Переводчик.

«Контент, извлеченный из сообщений пользователя, был изучен в ранних исследованиях для непосредственного выявления распространителей дезинформации, а текстовый классификатор может быть использован для классификации злонамеренных пользователей».

Перевод, выполненный человеком

«В более ранних исследованиях чтобы напрямую выявить распространителя дезинформации специалисты изучали содержимое поста пользователя. Текстовый классификатор может использоваться для классификации вредоносных пользователей».

Для представления двух инструментов перевода были использованы две таблицы, в которых оценка качества перевода были выставлена в соответствии с категорией ошибок, к которой они относятся.

Таблица 2

Итоги оценки качества перевода с применением Яндекс Переводчик

Яндекс Переводчик	Общая адекватность	Общая удобочитаемость	Трудозатратность редактирования
Синтаксис	4	4	4
Употребление артиклей	4	5	4
Лексика	3	4	5
Акронимы	4	4	4
Культурологические отсылки	3	3	4
Соотнесенность темы и ремы в предложении	3	4	3
ИТОГО	21	24	28

Таблица 3

Итоги оценки качества перевода с применением DeepL

DeepL	Общая адекватность	Общая удобочитаемость	Трудозатратность редактирования
Синтаксис	4	3	4
Употребление артиклей	4	4	5
Лексика	3	4	4
Акронимы	4	5	5
Культурологические отсылки	4	3	4
Соотнесенность темы и ремы в предложении	4	3	3
ИТОГО	23	22	25

Выводы.

Результаты сравнительного анализа ошибок пролили свет на различия в эффективности перевода двух рассматриваемых инструментов перевода, а именно DeepL и Яндекс Переводчик. DeepL обеспечил в целом лучшую общую эффективность и адекватность перевода, Яндекс Переводчик оказался лучше в плане удобочитаемости, трудозатратности редактирования и транслитерации.

Исходя из результатов, полученных в настоящем исследовании, возможным направлением дальнейшей работы может стать углубленное лингвистическое изучение категорий ошибок, использованных в данной статье в качестве основных критериев оценки качества машинного перевода. Таким образом, можно будет предложить эффективные меры по их исправлению с целью повышения эффективности работы систем машинного перевода.

Список литературы

1. The Institute for Information Transmission Problems named after A. Kharkevich [Электронный ресурс]. – Режим доступа: <http://iitp.ru/en/about> (дата обращения: 06.05.2025).
2. Институт проблем передачи информации им. А.А. Харкевича [Электронный ресурс]. – Режим доступа: <http://www.iitp.ru/> (дата обращения: 06.05.2025).
3. Яндекс. Переводчик [Электронный ресурс]. – Режим доступа: <https://translate.yandex.com/> (дата обращения: 06.05.2025).
4. DeepL Переводчик [Электронный ресурс]. – Режим доступа: <https://www.deepl.com/translator> (дата обращения: 06.05.2025).
5. Wörterbuch Englisch-Deutsch und Suche in einer Milliarde Übersetzungen [Электронный ресурс]. – Режим доступа: <https://www.linguee.de/> (дата обращения: 06.05.2025).
6. Translate launches a hybrid machine translation system [Электронный ресурс]. – Режим доступа: <https://yandex.com/company/blog/one-model-is-better-than-two-yuuyandex-translate-launches-a-hybrid-machine-translation-system/> (дата обращения: 06.05.2025).
7. Mičić S. Languages of medicine – present and future / S. Mičić // JAHR. – 2013. – Т. 4, №7. – С. 217–233.
8. Tiberii P. Dictionary of collocations / P. Tiberii // Zanichelli. – 2018.
9. Wu L. Misinformation in Social Media: Definition, Manipulation, and Detection / L. Wu // ACM SIGKDD Explorations Newsletter, Volume 21, Issue 2. – 2019. – Т. 21, №2. – С. 80–90.
10. Yandex company [Электронный ресурс]. – Режим доступа: <https://yandex.com/company> (дата обращения: 06.05.2025).
11. Батуев А.А. Системы машинного перевода: сравнение качества перевода и возможностей их использования (на примере технической документации в металлургической отрасли) : магистерская диссертация / А.А. Батуев // Уральский федеральный университет имени первого Президента России Б.Н. Ельцина, Уральский гуманитарный институт, Кафедра лингвистики и профессиональной коммуникации на иностранных языках. – Екатеринбург, 2021. – С. 65–69.