

Оценка качества машинного перевода в IT

Р.Г. Азиатцева

Приволжский государственный университет путей сообщения, Самара, Россия

Обоснование. В условиях быстрого развития сферы информационных технологий, особенно в контексте внедрения искусственного интеллекта и нейросетевых технологий, появляется необходимость выявить и оценить существующие сложности, такие как лингвистические, технические, культурные нюансы и качество перевода.

Цель — исследовать и проанализировать основные проблемы машинного перевода с помощью современных AI и нейросетевых технологий, а также оценить отношение и опыт работников IT-отрасли в использовании автоматических переводчиков.

Методы.

1. Анализ литературных источников показывает, что машинный перевод сталкивается с проблемами передачи точной терминологии и стилистики, особенно в узкоспециализированных областях [4, 5]. Традиционным методам с применением нейросетевых технологий пророчат будущее уже в ближайшие годы, однако современные системы нуждаются в доработке для улучшения качества перевода [2]. Важное значение придается роли переводчика как носителя языковой личности и эксперта, способного корректировать машинный перевод [1, 3].
2. Проведение эксперимента с переводом. Практическое сравнение качества перевода специализированного письма от tech lead с использованием разных инструментов: Яндекс переводчик, Google Translate, нейросетевых систем и ручного перевода, учитывающего профессиональный жаргон и сленг носителя языка. Анализ ошибок терминологии и стилистики позволил выявить основные проблемы области. Грамматические ошибки не входили в анализ, так как современные инструменты машинного перевода преодолевают этот барьер с высокой точностью. В ходе эксперимента было выявлено, что основные ошибки машинного перевода связаны с неправильной передачей терминологии (от 6 до 20 % ошибок в зависимости от системы) и стилистическими неточностями (до 10 %). Нейросетевые переводчики демонстрируют более высокую точность (около 84 %), чем традиционные сервисы, однако они все еще не способны в полной мере заменить профессиональный перевод с учетом жаргона и контекста, характерных для IT-текстов. По итогам исследования литературы и описываемого эксперимента стало понятно, что, несмотря на скорость развития AI, мы не можем заменить ручной перевод, учитывающий культурные и профессиональные аспекты человека. Этот вывод подтвердит проведенное социальное исследование.
3. Социологическое исследование (опрос) — сбор и анализ данных от двух групп молодых людей (начинающих специалистов и опытных IT-работников) посредством анкеты, содержащей 13 вопросов. Опрос был направлен на изучение опыта использования AI-переводчиков, предпочтений в методах перевода и частоты необходимости перевода профессиональных материалов. Согласно результатам, большинство работников в сфере IT предпочитают использовать AI при переводе. Они сталкиваются с необходимостью перевода документов довольно часто, особенно в общении с клиентами или работодателями в переписке. Стоит выделить тот факт, что люди с опытом в IT более трех лет сравнивают свой уровень перевода с машинным и не считают его хуже или лучше. Однако при этом показатели все равно смешанные, и участники отмечали как положительные, так и отрицательные стороны работы с онлайн-переводчиками и нейросетевыми системами.
4. Второй эксперимент по оценке понимания текста — участникам предоставлялись восемь документов, включающие профессиональный жаргон и термины на английском языке, после чего они отвечали на вопросы по их содержанию, указывая используемый и предпочитаемый метод перевода.

Это позволило оценить качество восприятия информации в зависимости от применяемого средства и опыта участников.

5. Статистический анализ полученных данных — усреднение результатов экспериментов и обработка данных опроса для выявления закономерностей, подтверждение гипотезы о текущих проблемах машинного перевода и предпочтениях, проблемах IT-специалистов. По этим результатам сделан вывод: люди с большим опытом работы в профессиональной сфере предпочитают самостоятельный перевод; люди без опыта чаще выбирают Яндекс переводчик; а нейросетевые системы используются всеми категориями участников, особенно в сложных моментах, связанных с англоязычным сленгом.

Результаты. Полученные данные подтверждают, что актуальные системы машинного перевода требуют дальнейшего совершенствования для адекватной обработки профессиональной лексики и стилистики в IT. Кроме того, работа выявила необходимость регулярного обновления оценки качества переводов с учетом быстрого развития технологий и специфики профессиональных текстов.

Выводы. В ходе исследования выявлены основные проблемы машинного перевода в IT, связанные с неправильной передачей терминологии и стилистики. Опытные IT-специалисты предпочитают самостоятельный перевод, а нейросети используют для сложных случаев. Результаты подтверждают необходимость дальнейшего улучшения технологий для повышения качества их эффективности в узкоспециализированных задачах.

Ключевые слова: проблемы; машинный перевод; онлайн-переводчики; AI; IT.

Список литературы

1. Бар-Хиллель Y. Some linguistic problems concerned with machine translation // Philosophy of Science. 1953. Vol. 20, N 3. P. 217–225. doi: 10.1086/287266
2. Беляева Л.Н. Машинный перевод в работе переводчика: практический аспект // Вестник ПНИПУ. Проблемы языкознания и педагогики. 2019. № 2. С. 8–20. doi: 10.15593/2224-9389/2019.2.1 EDN: DXNDSD
3. Бабина О.И. Языковая личность переводчика и машинный перевод // Вестник Челябинского Государственного Университета. 2011. № 24(239). С. 191–193. doi: 10.1111/j.1467-8322.2011.00840.x EDN: OXAMDH
4. WTM. The conference on machine translation. В: Machine Translate [Internet]. Режим доступа: <https://machinetranslate.org/wmt> Дата обращения: 27.01.2025.
5. Oxford Dictionaries [Internet]. Режим доступа: <https://en.oxforddictionaries.com/> Дата обращения: 27.01.2025.

Сведения об авторе:

Радмила Георгиевна Азиатцева — студентка, группа ПИб-21, электротехнический факультет; Приволжский государственный университет путей сообщения, Самара, Россия. E-mail: azimilly1016@gmail.com

Сведения о научном руководителе:

Ирина Геннадиевна Баканова — доцент кафедры «Лингвистика»; Приволжский государственный университет путей сообщения, Самара, Россия. E-mail: backanovairina@yandex.ru