



**МЕМОРАНДУМ
АССОЦИАЦИИ ПЕРЕВОДЧЕСКИХ КОМПАНИЙ**

**ПО ВОПРОСАМ ИНСТИТУЦИОНАЛИЗАЦИИ, СТАНДАРТИЗАЦИИ И
СОЗДАНИЯ НОРМАТИВНО-ПРАВОВОЙ БАЗЫ ДЛЯ ПРИМЕНЕНИЯ
МАШИННОГО ПЕРЕВОДА И ПРОЧИХ ТЕХНОЛОГИЙ
АВТОМАТИЧЕСКОЙ ГЕНЕРАЦИИ ТЕКСТА
В КАЧЕСТВЕ ПРОФЕССИОНАЛЬНЫХ ИНСТРУМЕНТОВ В ОБЛАСТИ
ПЕРЕВОДА И ЛОКАЛИЗАЦИИ**

Утвержден и рекомендован компаниями АПК





Разработчики:

Берендяев Максим Викторович (АКМ-Вест)

Сант Сергей Альфредович (Неотэк)

Светова Светлана Юрьевна (Т-Сервис)

Эксперты:

Гладков Сергей Алексеевич (Амберит СНГ)

Куликов Николай Юрьевич (АКМ-Вест)

Мищенко Илья Анатольевич (Литерра)

Ассоциация переводческих компаний:

AG .translate, АКМ-Вест, Амберит СНГ, Априори, Awatera, Литерра, Logrus IT, Марк Бизнес Переводы, Мегатекст, Неотэк, Профперевод, Транслинк, Компания ЭГО Транслейтинг, Эффектифф, Janus.





Ассоциация переводческих компаний (АПК) и нижеподписавшиеся представители переводческой отрасли признают, что научно-технический прогресс, стремительное развитие компьютерной техники и компьютерных сетей, широкое распространение Интернета, использование больших данных и нейронных сетей, сетевизация и цифровизация практических всех сфер деятельности человечества от бытовых до научных и производственных, совершенствование технологий машинного обучения (так называемого «искусственного интеллекта») и, в частности, обработки естественного языка коренным образом изменили профессию переводчика и всю отрасль переводов и локализации в целом.

Это привело к тому, что технологии и системы автоматической генерации текста (включая машинный перевод и генеративные диалоговые модели) следует рассматривать как один из инструментов при оказании лингвистических и переводческих услуг, в особенности на фоне гигантских объемов информации, которые в настоящее время, под влиянием глобализации и научно-технического прогресса, требуется переводить на другие языки или которые создаются или должны создаваться на естественных языках.

Распространение и доступность машинного перевода (МП) в быту привели к изменению представлений в обществе, бизнесе и политике о ценности, целях и принципах перевода как продукта и вида профессиональной деятельности, отчасти исказив отношение к важности аспекта качества перевода в целом, а также к некоторой деградации и девальвации норм родного (русского) языка, то есть снижению внутренних требований к грамотности текста на родном языке и, следовательно, к повышенной опасности его утраты, в особенности в части сокращения выразительных средств. При этом напрямую связанный с увеличением объема контента высокий спрос на перевод вызвал повышенный интерес к нему со стороны как IT-сообщества, так и многих сторонников замещения человеческого труда полностью автоматизированными решениями, тем самым усилив инвестиционную привлекательность технологических решений для межъязыкового посредничества. Все это дает АПК основание полагать, что мы станем свидетелями дальнейших интенсивных изменений в области автоматизации перевода.

В связи с этим АПК считает насущной задачей обеспечить институализацию, стандартизацию и разработку нормативно-правовой базы для профессионального применения средств автоматической генерации текста (включая машинный перевод и генеративные языковые технологии) с учетом интересов представителей профессии и потребителей (заказчиков) перевода, в том числе для оценки качества результата и рисков применения этих технологий, а также привлечение к участию в этой работе представителей разработчиков автоматизированных систем. Кроме того, АПК отмечает острую необходимость использования всеми участниками рынка переводов и заинтересованными сторонами единой профессиональной терминологии и выработки общего подхода к пониманию основ и принципов применения средств автоматической генерации текста в области профессионального перевода как среди потребителей переводческих услуг, так и среди переводческих компаний, переводческих подразделений компаний и учреждений, отдельных переводчиков, а также, что крайне важно, у представителей научного и академического сообществ и разработчиков..





АПК предлагает всем желающим принять деятельное участие в организации системного обмена опытом и накопления знаний; проведении исследований; изучении процессов и результатов оценки качества и оценки рисков; формировании стандартов и иной нормативно-правовой базы в области профессионального применения технологий и средств автоматической генерации текста. С этой целью в рамках АПК создается постоянно действующий рабочий комитет (рабочая группа) с условным названием «Профессиональное применение МП и прочих средств автоматической генерации текста» (Комитет АПК по МП), задачей которого будут являться систематизация и обработка специфических отраслевых знаний, опыта, рассмотрение этических кейсов, рисков, оценка качества, разработка и принятие стандартов, норм и рекомендаций для разработчиков технологий и заказчиков переводческих услуг, проведение отраслевых совещаний и конференций по МП и связанным с ним технологиям, решениям, услугам и процессам, взаимодействие с научным и академическим сообществами.

Первый состав участников Комитета АПК по МП (компаний и назначенных представителей) указан в Приложении 1 к Меморандуму.

Одной из наиболее важных задач в настоящее время является согласование единого отраслевого подхода к используемым терминам и их определениям в области МП, открытый и пополняемый список которых предлагается в Приложении 2 к Меморандуму.

Одним из направлений работы Комитета АПК по МП будет являться поддержка работы научного и академического сообщества и студентов по изучению опыта и научно-практическим исследованиям в области МП.

Комитет также сможет рассматривать и оценивать этические вопросы применения МП в конкретных кейсах по конкретным запросам в АПК и давать свои рекомендации и заключения.

Наряду с первоочередной терминологической работой АПК считает необходимым обратить особое внимание потребителей переводческих услуг и перевода, коллег по отрасли, а также в особенности научного и академического сообщества и студентов на то, что существующие открытые и бесплатные решения, в том числе реализуемые на основе облачных платформ, не должны использоваться в качестве замены профессиональному человеческому переводу, комплексным переводческим и лингвистическим услугам.

Одним из недостатков таких открытых и бесплатных (и условно бесплатных) систем является их обучение на общих языковых данных в наиболее распространенных языковых парах без специальной фокусной подготовки данных для обучения, без учета локали, целевой аудитории, предметной области, управления терминологией, специальной оценки качества результата и т. п. Кроме того, непрофессиональные потребители МП и(или) результатов работы иных систем автоматической генерации текста, как правило, не в состоянии оценить эквивалентность и адекватность полученного перевода как произведения, основанного на исходном тексте, и нередко, в силу отсутствия профессиональных переводческих и лингвистических компетенций, оценивают его качество по грамматической и орфографической верности на языке перевода, удобочитаемости и соответствию ожиданиям потребителя от содержания и формы перевода, то есть субъективно и вне связи с исходным текстом, его смыслом, интенциями и задачами.





На основании полученного опыта внедрения систем машинного перевода АПК также заявляет о том, что профессиональное применение МП предполагает, в соответствии с имеющимся опытом, по меньшей мере обучение, подбор или настройку МП, формирование, проверку, очистку данных для их обучения, определение применимости и критериев оценки качества МП в зависимости от конкретного задания заказчика, измерение качества МП, оценку и уменьшение рисков (управление рисками), корректирующие мероприятия, эксперименты и замеры, постоянный контроль качества результатов.

Профессиональное применение МП включает и работу с исходным текстом, его оцифровку и лингвистическую переработку, называемую также предредактированием текста для МП, последующее редактирование МП (постредактирование) с привлечением квалифицированных операторов (переводчиков-постредакторов МП) и многое другое.

Технологическая работа профессиональных поставщиков лингвистических и переводческих услуг, в частности, обязательно включает в себя определение порядка работы на проектах с МП, подготовку терминологических глоссариев, разработку перечней типов проверок и технологий обеспечения качества, интеграцию работы с МП в существующие технологии ручного перевода и перевода с использованием памяти перевода, то есть формирование гибридного сценария работы специалистов-переводчиков с комплексным использованием ручного перевода, результата работы систем МП и полных или неполных совпадений из памяти перевода, а также глоссариев, терминологических баз и автоматических и полуавтоматических проверок качества и т. п.

Указанные выше отличия профессионального применения МП де-факто уже привели к формированию целого спектра новых профессий в отрасли – постредакторы, специалисты по обучению МП и т. п. – для которых предстоит в кратчайшие сроки определить рамку компетенций и заложить основы обучения и проверки квалификации и аттестации в системе высшего и дополнительного образования.

В рамках комплексного использования технологий МП, памяти перевода и ручного перевода все большее значение приобретает изучение рисков и определение правовых основ такой работы, с одной стороны, с точки зрения авторских прав на ручной перевод, память переводов, МП и получаемый результат их комплексной обработки и, с другой стороны, с точки зрения ответственности за такой перевод, его качество и возможные искажения и смысловые неточности, маскировку ошибок и генерирование внешне связного, но бессмысленного или не соответствующего действительности текста.

АПК обращает внимание на то, что 100-процентное исключение ошибок в результатах нейронного МП невозможно в силу принципиальных особенностей этой технологии, и поэтому ожидания появления безупречного полностью автоматического перевода преждевременны, не имеют под собой никаких оснований и несут в себе огромные риски и вред как для потребителей перевода, так и для самой отрасли переводов и локализации.

Кроме того, АПК обращает внимание на тот факт, что любая информация, передаваемая открытым системам для получения результата или для обучения, становится доступной операторам и разработчикам таких систем, то есть по сути общедоступной, что может повлечь за собой причинение ущерба безопасности





государства и(или) бизнеса, нарушение режимов секретности, конфиденциальности и законодательства в сфере защиты персональных данных и регулирования их трансграничной передачи.

При этом профессиональное использование МП, как правило, предполагает использование данных заказчика, его предыдущих проектов и переводов для обучения и дообучения систем МП, что также возможно только при наличии согласованной нормативно-правовой базы, стандартов и договорных отношений между потребителями и заказчиками услуг перевода, переводческими предприятиями и разработчиками, операторами систем и платформ МП.

В настоящее время гражданское законодательство РФ предоставляет исполнителям перевода свободу выбора инструментов для оказания своих услуг, то есть формально никак не ограничивает их право использовать для оказания услуг бесплатные, открытые системы автоматизации, в том числе и системы МП. Также не существует норм и стандартов, требующих от исполнителей специально отмечать («маркировать») переводы, созданные с использованием каких-либо средств автоматизации. Это касается и баз памяти перевода, которые могут быть «загрязнены» результатами МП, в том числе непрофессионального, не проверенного и не отредактированного.

Указанные правовые вопросы, риски и правила также требуют изучения и отражения в стандартах и нормах.

АПК предлагает всем своим членам и компаниям-партнерам предложить своих представителей для регулярной работы в составе Комитета АПК по МП и(или) принять участие в мероприятиях и проектах Комитета, планирует провести в 2023 году «Научно-практическую конференцию Ассоциации переводческих компаний по машинному переводу и технологиям автоматической генерации текста» с участием переводчиков-практиков, представителей переводческого бизнеса, преподавателей и студентов, а также описать Рамку квалификаций специалистов по настройке МП, предредактированию и постредактированию МП (Приложение 3 к Меморандуму), чтобы сформировать базу знаний и создать на базе АПК отраслевой центр компетенций, сформулировать предложения по форме и составу российских стандартов в области автоматизации перевода и инициировать их принятие на государственном уровне.





Ассоциация
Переводческих
Компаний

Приложение 1. Предложение по составу Комитета АПК по МП





Приложение 2. Термины и определения (образец, в алфавитном порядке)

автоматизированный перевод
автоматический перевод
базовый МП
галлюцинации обученной модели
глубокое обучение
контролируемый язык
корпоративный МП
легкое постредактирование
максимально допустимая плотность ошибок
машинное обучение
машинный перевод (МП)
метрика оценки качества перевода
модуль/сервис машинного перевода
обработка естественного языка
онлайн-машинный перевод
полное постредактирование
пользовательская настройка/обучение
поставщик машинного перевода
постредактирование машинного перевода
постредактирование МП
предвзятость обученной модели
предметная область
расстояние редактирования
результат машинного перевода
риск потребителя
риск производителя
сегмент с машинным переводом
сегмент, переведенный с помощью МП
система машинного перевода (МП)
совпадение из МП
специализированный модуль МП
средняя плотность ошибок





Приложение 3. Рамка квалификации специалистов по МП

РОЛЬ И КВАЛИФИКАЦИЯ СПЕЦИАЛИСТА ПО МТ И ПОСТРЕДАКТОРА МП

1. Основные требования
2. Образование и опыт
3. Инструментарий постредактора

ОБУЧЕНИЕ СПЕЦИАЛИСТОВ ПО МП И ПОСТРЕДАКТОРОВ МП

1. Основные требования
2. Пример типовой программы
3. Уровни обучения

ПОДГОТОВКА ТЕКСТА ДЛЯ МП И ПОСТРЕДАКТИРОВАНИЕ МП КАК ПРОФЕССИОНАЛЬНАЯ УСЛУГА

1. Основные требования
2. Уровни постредактирования
3. Стратегии постредактирования
4. Инструкции по постредактированию и типологии ошибок МП
5. Постредактирование в САТ-системах
6. Редактирование после постредактирования
7. Расчет себестоимости и стоимости услуги

НАСТРОЙКА МП КАК УСЛУГА

1. Основные требования
2. Уровни настройки и средства проверки и измерения качества
3. Источники данных для настройки МП
4. Обучение, переобучение и дообучение
5. Расчет себестоимости и стоимости услуги по настройке применения МП

СТАНДАРТЫ ПРИМЕНЕНИЯ МП В ПРОФЕССИОНАЛЬНОМ ПЕРЕВОДЧЕСКОМ ПРОЦЕССЕ

1. Основные преимущества
2. Основные ограничения
3. Применимость МП, измерение уровня качества МП, и результатов постредактирования МП
4. Выбор МП и настройка МП и метрики измерения качества перевода как базовая услуга
5. Предредактирование текста на ИЯ для МП как вспомогательная услуга (общие вопросы)
6. Постредактирование МП как базовая услуга
7. Автоматическое постредактирование





Ассоциация
Переводческих
Компаний

Подписанты

Настоящий Меморандум подписан представителями АПК и компаниями-партнерами 23 июня 2023 года.

Представители АПК

Компании-партнеры

