Круглый стол.
ПЕРЕВОДГОТОВКА ПО СПЕЦИАЛЬНОСТИ ПЕРЕВОДЧЕСКАЯ ДЕЯТЕЛЬНОСТЬ ТЕОРЕТИЧЕСКИЕ И ПРИКЛАДНЫЕ АСПЕКТЫ

Н. В. Егоров

КОНФИДЕНЦИАЛЬНОСТЬ ОНЛАЙН ИИ-ИНСТРУМЕНТОВ ДЛЯ ПЕРЕВОДА

Информационная безопасность представляет собой обширную область, которая включает в себя применение облачных технологий для решения разнообразных задач. Особое внимание уделяется сервисам для работы с документами, таким как сопоставление версий, конвертация в различные форматы и собственно онлайн-перевод.

В современном мире инструменты для машинного перевода становятся все более популярными как среди обычных пользователей, так и в государственных структурах. Они позволяют быстро обмениваться сообщениями на разных языках, получать доступ к содержанию документов на иностранном языке, создавать презентации и отчеты, а также анализировать информацию о конкурентах на любом языке мира. Эти технологии делают работу современного специалиста более эффективной и продуктивной.

Онлайн-инструменты для перевода могут показаться привлекательными своей доступностью, создавая иллюзию безопасности. Однако следует помнить, что вся загружаемая информация сохраняется в логах онлайн-сервисов. Эта информация может стать общедоступной через поисковые системы или быть проанализирована с помощью статистических и лингвистических методов, что позволяет выявить частоту и особенности употребления определенных слов и выражений.

Одна из популярных онлайн-платформ для переводчиков, DeepL, в своей политике конфиденциальности прямо указывает, что при бесплатном использовании следует вводить только те тексты, которые пользователь готов передать на серверы разработчика на ограниченный период времени для обучения и совершенствования нейронных сетей и алгоритмов перевода. Не все пользователи в полной мере осознают, что, прибегая к услугам машинного перевода онлайн, они автоматически принимают условия пользовательского соглашения и тем самым дают согласие на передачу информации сервису.

Защита персональных данных в эпоху искусственного интеллекта (ИИ) приобретает все большее значение в связи с растущим использованием технологий, которые в процессе перевода обрабатывают большие объемы информации.

В законодательстве Республики Беларусь под ИИ понимается комплекс технологических решений, способных имитировать когнитивные функции человека и получать результаты, сопоставимые с результатами интеллектуальной деятельности человека при выполнении конкретных задач. Этот комплекс включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение, процессы и сервисы, предназначенные для обработки данных и поиска решений.

Основные проблемы, с которыми люди сталкиваются при использовании ИИ, весьма разнообразны и требуют внимания как разработчиков, так и законодателей.

Первый важный аспект – обработка больших массивов данных, необходимых для сервисов ИИ, включая машинный перевод и сопутствующие услуги перефразирования, редактирования, суммаризации и так далее. Это создает трудности с получением согласия пользователей и требует обеспечения возможности обезличивания данных для снижения рисков их утечки.

В рамках цифровизации экономики вероятность утечки персональных данных возрастает. В этой связи бюро переводов, использующим при оказании своих услуг ИИ-инструменты, в частности на безвозмездной основе, следует принимать меры по защите данных своих клиентов и предотвращению их утечки, а также обеспечивать соблюдение законодательства, регулирующего защиту персональных данных.

Еще одной трудностью, выходящей за рамки национальных юрисдикций, является прозрачность алгоритмов ИИ. Для рядового потребителя ИИ подобен «черному ящику», что затрудняет для первого понимание процессов принятия решений последним. При создании собственных переводческих платформ в целях последующей коммерциализации собственного продукта разработчики должны объяснять, какие данные собираются и как в дальнейшем используются.

Одним из способов защиты персональных данных является федеративное обучение. Этот метод позволяет обучать модели машинного обучения ИИ непосредственно на устройствах пользователей, таких как смартфоны или устройства интернета вещей, без необходимости передачи личных данных на центральные серверы разработчика. Такой подход обеспечивает конфиденциальность выполняемых переводов и дает пользователям контроль над персональными данными.

Еще один метод, известный как дифференцированная приватность, заключается в добавлении контролируемого количества шума к данным. Это делает невозможным идентификацию конкретных пользователей при анализе больших массивов информации.

Одним из ключевых моментов в обеспечении безопасности личной информации является разработка и внедрение законодательных инициатив в области ИИ. В Европейском союзе основным нормативно-правовым актом,

регулирующим защиту персональных данных, является «Общий регламент по защите персональных данных» 2018 года (General Data Protection Regulation). Данный регламент устанавливает нормы, которые были учтены и в Законе Республики Беларусь «О защите персональных данных». В нем определены как общие требования к защите персональных данных, так и требования к разработчикам ИИ, которые непосредственно участвуют в обработке персональных данных. Однако, учитывая стремительное развитие технологий ИИ, законодательное регулирование защиты персональных данных не является исчерпывающим. В упомянутом выше законе нашей страны отсутствуют специальные требования при обработке ИИ персональных данных.

Обязательным аспектом для правомерности обработки личной информации является точное соблюдение установленных правил и процедур обработки данных. В соответствии с упомянутым выше Регламентом ЕС обработка персональных данных должна осуществляться в строгом соответствии с принципами законности, справедливости, прозрачности, ограничения целью, минимизации, точности, ограничения хранения, целостности и конфиденциальности данных.

Из-за прогресса ИИ в области перевода возникают новые вызовы: требуется непрерывное совершенствование правовых норм, а также разработка систем защиты персональных данных пользователей. В этому следует относить как глубокий анализ влияния технологических новшеств на сохранность данных, так и активную работу по снижению вероятных угроз при использовании ИИ-инструментов. Особое внимание уделяется соблюдению принципов открытости процессов обработки информации и ограничению объемов собираемых персональных сведений – как со стороны разработчиков, так и законодательных органов.

В условиях господства ИИ в переводческой индустрии обеспечение безопасности персональных данных клиентов становится задачей комплексного характера. Это предполагает не только внедрение передовых технических решений для защиты информации, но также разработку и принятие новых правовых инициатив, направленных на защиту конфиденциальности конечного пользователя переводческих сервисов.