

Рабочая группа по правовому развитию Гаагской системы международной регистрации промышленных образцов

**Одиннадцатая сессия
Женева, 12-14 декабря 2022 года**

ИССЛЕДОВАНИЕ НА ТЕМУ «НАЛИЧИЕ ТЕХНОЛОГИЙ ПЕРЕВОДА И ВОЗМОЖНОСТИ ДЛЯ ИХ ИСПОЛЬЗОВАНИЯ»

Документ подготовлен Международным бюро

I. ИСТОРИЯ ВОПРОСА

1. На своей десятой сессии, состоявшейся в 2021 году, Рабочая группа по правовому развитию Гаагской системы международной регистрации промышленных образцов (далее – «Рабочая группа») рассмотрела документы H/LD/WG/10/ 4 и H/LD/WG/10/5, озаглавленные «Обновленное углубленное исследование с описанием финансовых последствий и осуществимости с технической точки зрения добавления китайского и русского языков в Гаагскую систему» и «Обновленный вариант критериев отбора дополнительных языков для включения в Гаагскую систему» соответственно. Рабочая группа поручила Секретариату представить результаты пересмотренного исследования финансовых последствий и технической осуществимости включения новых языков в Гаагскую систему с учетом технических решений, в частности машинного перевода, для рассмотрения Рабочей группой на ее следующей сессии (см. документ H/LD/WG/10/6, пункт 18(ii)).

2. В настоящем документе содержится подробная информация о технологиях и методах перевода, используемых различными переводческими службами ВОИС. Представленная в нем информация призвана облегчить рассмотрение Рабочей группой вопроса о потенциальной способности технологий перевода обеспечить снижение расходов, связанных с включением дополнительных языков в Гаагскую систему.

II. ТЕХНОЛОГИИ ПЕРЕВОДА, ПРИМЕНЯЕМЫЕ В ВОИС

3. Функции перевода в ВОИС выполняются различными глобальными службами по охране интеллектуальной собственности, а также Отделом письменных переводов. Процесс перевода обычно состоит из двух этапов: во-первых, текст, который необходимо перевести, сопоставляется с записями, содержащимися в соответствующей памяти переводов, а во-вторых, переводчик либо редактирует предлагаемые варианты перевода, которые генерируются системой нейронного машинного перевода (НМП) (WIPO Translate), если существует определенная модель, доступная в определенном сочетании предметной области и языка, либо переводит остальной текст на язык перевода. Наконец, определенный процент переводимых материалов подлежит контролю качества, осуществляемым штатным переводчиком-редактором.

WORLDSEVER

4. WorldServer — это система управления процессом перевода, включающая средство автоматизированного перевода, которое управляет рабочим процессом перевода путем автоматизации выполняемых функций. Сначала он заполняет переводы сегментами на языке перевода, извлеченными из памяти переводов, а затем в отношении остальных сегментов система НМП выдает предлагаемые варианты перевода. WorldServer используется Отделом переводов и терминологии Мадридского реестра и Отделом переводов PCT¹.

ПРОГРАММЫ ПАМЯТИ ПЕРЕВОДОВ

5. В основе любой системы перевода лежит память переводов, представляющая собой базу, в которой хранятся ранее переведенные тексты в предметной области (в рамках Гаагской системы, Мадридской системы, PCT или Отделом письменных переводов). Все записи в памяти переводов проверяются профессиональным переводчиком для обеспечения качества. Существуют отдельные памяти переводов для каждой службы и каждой языковой комбинации². Количество записей, содержащихся в соответствующей памяти переводов, зависит от объема ранее переведенного текста. Например, начиная с 2016 года доля заявок, поданных по процедуре Гаагской системы на английском языке, составила более 86%, на французском — около 12%, а на испанском — около 1%³. Следовательно, в памяти переводов в рамках Гаагской системы, где исходным языком является английский, содержится гораздо больше записей, чем в памяти переводов, где исходным языком является французский, в то время как в памяти переводов, где исходным языком является испанский, почти нет записей. Память переводов интегрирована в систему управления процессом перевода (WorldServer).

WIPO TRANSLATE

6. WIPO Translate — это инструмент машинного перевода на основе ИИ, который был разработан ВОИС самостоятельно и обучен для конкретной тематики и в котором используется новейшая технология НМП. НМП — это технология, основанная на моделях нейронных сетей, которые «учатся» на больших объемах ранее переведенного текста⁴. Первоначально WIPO Translate был обучен переводу патентных документов⁵

¹ В ближайшем будущем Отдел письменных переводов ВОИС перейдет на использование WorldServer.

² Например, в настоящее время существуют шесть отдельных блоков памяти переводов для комбинаций трех рабочих языков Гаагской системы (английский-французский, английский-испанский, французский-английский, французский-испанский, испанский-английский и испанский-французский).

³ См. Ежегодный обзор Гаагской системы 2022 года, рисунок 13.

⁴ Например, обучающие наборы данных, используемые WIPO Translate для перевода патентных документов, обычно содержат миллионы документов.

⁵ Система WIPO Translate обучена с использованием текстов заявок, поданных по процедуре PCT, и патентных фондов национальных и региональных ведомств, содержащихся в Patentscope.

на 10 языков публикации PCT и ориентирован на них⁶. WIPO Translate, общедоступный для пользователей Patentscope, предлагает варианты перевода с этих языков на английский и наоборот.

7. Для каждой из секций Отдела письменных переводов ВОИС была разработана индивидуальная модель WIPO Translate⁷. Однако для выполнения высококачественных переводов WIPO Translate необходимо обучать работе с большими объемами текста в соответствующей области для конкретной языковой пары⁸.

8. Кроме того, предлагаемые варианты перевода, генерируемые моделью НМП для комбинаций лингвистически более далеких друг от друга языков, обычно имеют более низкое качество, чем для более близких в лингвистическом отношении языковых комбинаций. Следовательно, для комбинаций более далеких друг от друга языков еще важнее иметь достаточно обучающих данных, чтобы обеспечить надлежащее качество перевода⁹.

ПЕРЕВОД, ПОСТРЕДАКТИРОВАНИЕ И КОНТРОЛЬ КАЧЕСТВА

9. Описанные выше инструменты не могут полностью автоматизировать процесс перевода, поэтому, даже если используются технологии машинного перевода, всегда требуются человеческие ресурсы для постредактирования и контроля качества. Эти переводчики должны знать соответствующие языки обладать навыками перевода на эти языки. В этой связи следует напомнить, что всегда предпочтительно, чтобы постредактирование и контроль качества перевода на язык перевода выполнял переводчик, для которого язык перевода является родным. Контроль качества осуществляет штатный переводчик-редактор.

10. Качество перевода, выполненного с использованием системы НМП, влияет на затраты на постредактирование. Если качество машинного перевода является неудовлетворительным, общие затраты на постредактирование могут быть близки или даже превышать затраты на перевод, выполняемый непосредственно человеком.

III. ПРАКТИКА МАШИННОГО ПЕРЕВОДА В ВОИС

ГААГСКАЯ СИСТЕМА

11. Следует напомнить, что перевод международных заявок в рамках Гаагской системы в настоящее время выполняется Отделом перевода и терминологии Мадридского реестра¹⁰. В этих заявках переводятся указания изделия или продукта, описание отличительных особенностей промышленного образца, формулы изобретения и условные обозначения¹¹. В то время как указание продукта, формула изобретения и

⁶ Арабский, китайский, английский, французский, немецкий, японский, португальский, русский и испанский.

⁷ В настоящее время индивидуализированная модель используется для переводов в рамках Гаагской системы, Мадридской системы, для некоторых языковых пар в PCT и для перевода некоторых документов Отделом письменных переводов.

⁸ Хотя потенциально WIPO Translate можно обучить переводу только в одну сторону, его обычно обучают переводу в обе стороны, поскольку обучение с использованием параллельных данных (содержания на двух языках) значительно повышает качество перевода.

⁹ Например, английский и французский языки лингвистически близки, тогда как английский и арабский, китайский или корейский языки – это далекие друг от друга языки. Английский и русский языки принадлежат к индоевропейской языковой семье; однако у них есть грамматические различия, поэтому для перевода произвольного текста потребуется больше обучающих данных.

¹⁰ В настоящее время перевод (постредактирование) всех документов в рамках Гаагской системы осуществляется внутри организации.

¹¹ См. статью 5(2)(b)(ii) и (iii) Женевского акта (1999 года), правила 7(3)(iv), (4)(b), (5)(a), 11(2) и (3) Общей инструкции к Акту 1999 года и Акту 1960 года Гаагского соглашения (далее именуемой «Общая инструкция») и разделы 403 и 405(с) Административной инструкции по применению Гаагского соглашения.

условные обозначения обычно содержат менее 10 слов, описание представляет собой произвольный текст объемом в среднем 64 слова в 2022 году¹². Отмечается, что доля международных заявок, содержащих описание, увеличилась с 17 процентов общего числа заявок в 2012 году до 65 процентов в 2022 году¹³.

12. Отдел перевода и терминологии, придерживающийся практики прямого перевода, использует шесть отдельных блоков памяти переводов для комбинаций трех рабочих языков Гаагской системы, а также WIPO Translate. В настоящее время блоки памяти переводов ранее переведенных текстовых материалов в заявках, подаваемых в рамках Гаагской системы, содержат в общей сложности около 280 000 записей, причем большинство из них содержится в блоках памяти переводов, в которых английский язык является исходным языком. По оценкам Международного бюро, в переводах с английского на французский или испанский язык совпадения обнаруживаются в памяти переводов примерно в 35 процентах текстовых сегментов, а остальной текст (65 процентов) требует постредактирования выдаваемых WIPO Translate вариантов перевода. Для сравнения, для переводов, где исходным языком является французский или испанский, совпадения в памяти переводов обнаруживаются в гораздо меньшем количестве случаев¹⁴. Поэтому в таких переводах большая часть текста подвергается постредактированию с использованием ограниченного количества предлагаемых вариантов перевода, генерируемых WIPO Translate.

13. Кроме того, отмечается, что, в то время как память переводов постоянно пополняется, процент совпадений растет намного медленнее. По оценкам Международного бюро, в текстовых материалах Гаагской системы степень совпадения в памяти переводов может не превышать 40 процентов, учитывая непрерывное развитие в области промышленных образцов и связанной с ними терминологии, содержащейся в заявках, подаваемых в рамках Гаагской системы (в частности, в текстовых описаниях)¹⁵. Таким образом, Международное бюро исходит из того, что примерно в 60 процентах переводов генерируемые WIPO Translate предлагаемые варианты перевода всегда будут подвергаться постредактированию человеком.

МАДРИДСКАЯ СИСТЕМА

14. Следует напомнить, что в рамках Мадридской системы действует такой же языковой режим, что и в Гаагской системе, с использованием тех же трех языков. Большая часть переводческой работы, выполняемой в рамках Мадридской системы, — это перевод указаний товаров и услуг в международных заявках. Эти указания обычно представляют собой определенный набор стандартизированных терминов и поэтому они отличаются от описаний, содержащихся в заявках, подаваемых в рамках Гаагской системы.

¹² Длина описания существенно различается. Например, в 2021 году, а также в первые три квартала 2022 года 15 процентов описаний содержали более 100 слов, один процент — более 500 слов. В 2021 году 589 заявок, поданных в рамках Гаагской системы, содержали описание длиной от 100 до 500 слов, а в 30 таких заявок количество слов в описании превышало 500.

¹³ Это изменение является результатом присоединения Договаривающихся сторон, ведомства которых являются ведомствами, проводящими экспертизу, и требующими или рекомендующими включение описания. В связи с этим отмечается, что Китай сделал заявление, в котором он потребовал, чтобы международная заявка содержала краткое описание отличительных особенностей образцов, а ведомство Российской Федерации (РОСПАТЕНТ) настоятельно рекомендует представлять краткое описание отличительных особенностей промышленного образца в заявке на его регистрацию.

¹⁴ Совпадение обычно обнаруживается только в том случае, если один и тот же заявитель подает несколько заявок на очень похожие образцы. В этих случаях совпадения могут быть найдены в указаниях продукта или в условных обозначениях, но реже в описаниях.

¹⁵ Причины этого кроются в характере охраны промышленных образцов, которая должна соответствовать критериям новизны и оригинальности в соответствии с законами, действующими в большинстве стран. Регистрация образца является результатом разработки новых продуктов, что приводит к созданию новых образцов как для новых, так и для существующих продуктов.

15. Отдел перевода и терминологии Мадридского реестра использует память переводов текста документов, ранее переведенных в рамках Мадридской системы¹⁶, и WIPO Translate. Блоки памяти переводов в рамках Мадридской системы содержат в общей сложности более шести миллионов записей по шести языковым комбинациям. Степень совпадений в памяти переводов текстов в рамках Мадридской системы составляет около 65 процентов¹⁷, а остальной текст требует постредактирования выдаваемых WIPO Translate вариантов перевода. Высокое качество вариантов перевода, генерируемых WIPO Translate, особенно для перевода документов, исходным языком которых является английский, позволяет Международному бюро сократить расходы на постредактирование на 50 процентов¹⁸.

16. В настоящее время Рабочая группа по правовому развитию Мадридской системы международной регистрации знаков (далее — «Мадридская рабочая группа») изучает возможные варианты некоторых мер в рамках обсуждения вопроса о возможном расширении языкового режима, такие как создание памяти переводов с большим количеством эквивалентных терминов на арабском, китайском и русском языках, обучение системы WIPO Translate таким образом, чтобы она могла выполнять переводы в рамках Мадридской системы на эти языки, и сбор языковых пар применительно к перечням товаров и услуг, которые ведомства Договаривающихся сторон могут иметь в своем распоряжении¹⁹.

17. Тем не менее учитывая разные типы текстовых материалов, содержащихся в заявках, подаваемых в рамках Гаагской и Мадридской систем (описания в виде произвольного текста в сравнении со стандартными терминами), постоянное развитие в области промышленных образцов, различия между Гаагской и Мадридской системами в плане объема записей в памяти переводов (280 000 записей по сравнению с более чем 6 миллионами), и, следовательно, разные степени совпадения в памяти переводов двух систем, а также разные объемы потенциальных обучающих данных для каждой модели WIPO Translate, в контексте Гаагской системы аналогичные меры представляются менее перспективными.

СИСТЕМА PCT

18. Отдел письменных переводов PCT отвечает за перевод заголовков и рефератов заявок по процедуре PCT с 10 языков публикации заявок на английский и французский языки (если они подаются не на этих двух языках), а также за перевод отчетов о международном поиске и заключений международной предварительной экспертизы о патентоспособности на английский язык (если они составлены не на английском языке)²⁰.

19. Название изобретения – это его краткое и точное описание, и рекомендуется, чтобы оно содержало от двух до семи слов²¹. Реферат заявки PCT объемом от 50 до 150 слов²² содержит краткое изложение того, что раскрыто в описании изобретения²³. Отчеты о международном поиске и заключения международной предварительной экспертизы о патентоспособности представляют собой технические документы, составляемые патентными ведомствами в отношении патентоспособности изобретений, и содержат в

¹⁶ Поскольку в Мадридской системе используются те же три рабочих языка, что и в Гаагской системе, Отдел перевода и терминологии Мадридского реестра использует шесть блоков памяти перевода для тех же языковых комбинаций.

¹⁷ См. документ MM/LD/WG/19/7.

¹⁸ Это означает, что расходы на постредактирование составляют 50 процентов расходов, которые потребовались бы, если бы текст полностью переводился без помощи WIPO Translate.

¹⁹ См. документ MM/LD/WG/19/7, пункты 39-50.

²⁰ См. правила 45.1, 48.3(c), 62*bis*.1, 72.1 и 86.2 Инструкции к Договору о патентной кооперации (далее – «Инструкция к PCT»).

²¹ См. правило 4.3 Инструкции к PCT.

²² Если реферат представлен на английском языке или если он переводится на английский. См. правило 8 Инструкции к PCT.

²³ См. правило 8.1(a)(i) Инструкции к PCT.

среднем от 50 до 200 слов (отчеты о международном поиске) и от 550 до 950 слов (заклучения международной предварительной экспертизы о патентоспособности) в зависимости от языка, на котором составляется документ. Тексты этих отчетов существенно отличаются от текстов заявок, подаваемых в рамках Гаагской системы.

20. Отдел письменных переводов РСТ использует память переводов РСТ для 18 языковых комбинаций 10 языков публикации РСТ²⁴, терминологическую базу данных («WIPO Pearl») ²⁵ и WIPO Translate. Поскольку в разные годы подавалось разное количество заявок по процедуре РСТ на соответствующих языках²⁶, в памяти переводов имеется большое количество записей переводов с китайского, немецкого, японского и корейского языков на английский язык, в то время как записей переводов с арабского или русского на английский не так много²⁷. Поэтому документы с арабского и русского на английский язык переводятся профессиональными переводчиками (с использованием весьма ограниченной существующей памяти переводов и Терминологической базы данных).

21. Аналогичным образом, из-за разного объема имеющихся наборов данных, используемых для обучения, WIPO Translate в настоящее время используется только для переводов с китайского на английский и с английского на французский внутри организации. В настоящее время оценивается целесообразность использования WIPO Translate для других языковых комбинаций (в частности, для перевода с французского, немецкого, японского и корейского языков на английский язык).

22. Модель WIPO Translate, адаптированная для перевода документов РСТ внутри организации, не будет использоваться для перевода текстов в рамках Гаагской системы, поскольку, во-первых, в настоящее время она используется только для двух языковых комбинаций, а во-вторых, тексты, на которых она обучалась (патентные документы), относятся к другой области техники.

ОТДЕЛ ПИСЬМЕННЫХ ПЕРЕВОДОВ

23. Отдел письменных переводов ВОИС переводит с шести официальных языков Организации Объединенных Наций (ООН) и на эти языки договоры и другие международные документы, административные функции в отношении которых выполняют ВОИС и УПОВ²⁸, тексты национальных законодательных актов, документов для Ассамблей государств-членов, заседаний других основных органов, комитетов и рабочих групп, включая их рабочие документы и отчеты о заседаниях; публикации, учебные материалы, веб-сайт ВОИС и другие коммуникационные материалы. Таким

²⁴ Однако, поскольку Отдел письменных переводов РСТ переводит документы на английский язык, в памяти почти нет записей перевода с английского на другие языки публикаций РСТ (кроме переводов на английского на французский).

²⁵ Терминологическая база данных РСТ содержит научно-технические термины из патентных документов на 10 языках публикаций РСТ. В настоящее время она содержит около 230 000 терминов. База данных общедоступна по следующему адресу: <https://www.wipo.int/reference/ru/wipopearl/>.

²⁶ См. *Ежегодный обзор РСТ 2022 года*, рисунок С1.

²⁷ Кроме того, отчеты о международном поиске и заключения международной предварительной экспертизы о патентоспособности, составляемые на арабском или русском языках, недоступны в машиночитаемом формате, поэтому переводы этих документов не могут быть включены в память переводов РСТ или использованы для обучения WIPO Translate. Таким образом, имеющиеся данные в памяти переводов РСТ для арабского и русского языков основаны только на переведенных названиях и рефератах, содержащихся в заявках, поданных по процедуре РСТ на этих языках.

²⁸ Услуги письменного и устного перевода также предоставляются УПОВ, который использует четыре официальных языка – английский, французский, немецкий и испанский.

образом, текст, переведенный Отделом письменных переводов, отличается от текста, содержащегося в заявках, подаваемых в рамках Гаагской системы²⁹.

24. Большинство текстов³⁰ переводятся Отделом письменных переводов ВОИС с английского на другие официальные языки ООН, в то время как небольшое количество переводов выполняются с французского и испанского языков на английский, и совсем немного текстов переводятся с других официальных языков ООН, а также немецкого, португальского и др. на английский и наоборот³¹. Все переводы выполняются профессиональными переводчиками с использованием блоков памяти переводов на соответствующих языках. В этой связи следует отметить, что в памяти переводов, используемой Отделом письменных переводов, есть записи переводов с английского на арабский, китайский и русский языки, однако нет доступных для использования данных по переводу с любого из этих языков на английский язык. В последние пару лет результаты машинного перевода определенных категорий документов, например стенографических отчетов, генерируемые WIPO Translate, подлежат постредактированию.

25. Модель WIPO Translate, адаптированная для внутреннего использования Отделом письменных переводов, не подходит для перевода текстов документов, имеющих отношение к Гаагской системе, поскольку, во-первых, тексты, на которых она обучалась, относятся к другой области техники (документы для заседаний), а во-вторых, большая часть текстов переводится с английского на другие языки, поэтому данные по переводам с других языков на английский отсутствуют.

IV. ДРУГИЕ СООБРАЖЕНИЯ

КАЧЕСТВО ПЕРЕВОДА

26. Важно напомнить, что перевод текстов заявок, подаваемых в рамках Гаагской системы, осуществляется Международным бюро без какого-либо участия или контроля со стороны заявителя³². В частности, указание и описание продукта или их перевод затем учитываются некоторыми указанными ведомствами при проведении экспертизы заявок на регистрацию образцов по существу и остаются актуальными для использования в рамках административных или судебных разбирательств во всех указанных Договаривающихся сторонах в отношении отказов, исков о признании недействительности прав или исков об их нарушении. Таким образом, переводы, сделанные Международным бюро, являются важным элементом любого решения по существу в отношении действительности или объема охраны в юрисдикциях Договаривающихся сторон³³, хотя, в отличие от РСТ³⁴, заявитель не принимает в них участия. Поэтому высокое качество перевода,

²⁹ Хотя Отдел письменных переводов отвечает за перевод связанных текстовых материалов, имеющих отношение к Гаагской системе, таких как документы Рабочей группы, текст этих документов отличается от текста заявок, подаваемых в рамках Гаагской системы.

³⁰ По оценкам Отдела письменных переводов, около 90 процентов.

³¹ Для перевода документов с немецкого и португальского и на эти языки привлекаются внешние подрядчики.

³² См. правило 6(4) Общей инструкции. В этом отношении следует отметить, что, хотя заявитель может приложить к международной заявке предлагаемый перевод любого содержащегося в ней текстового материала, на практике Международное бюро никогда не получает такие переводы.

³³ Договаривающиеся стороны могут использовать описание в том виде, в котором оно приведено в заявке, или его перевод.

³⁴ Реферат заявки по процедуре РСТ, переведенный Международным бюро, просто служит инструментом поиска в определенной области и не принимается во внимание ни для каких других целей, в частности для определения объема испрашиваемой охраны (статья 3(3) РСТ и правило 8.3 Инструкции к РСТ). Объем охраны определяется формулой изобретения, а изобретение раскрывается в его описании. Однако формула изобретения и его описание переводятся заявителем при переходе на национальную фазу в его ведомстве, и на основе этого перевода проводится экспертиза по существу национальным/региональным ведомством.

выполненного Международным бюро, является необходимым условием защиты прав владельцев на территориях указанных Договаривающихся сторон.

ОЦЕНКА ИНСТРУМЕНТОВ НМП

27. Кроме того, Международное бюро провело тестовую оценку возможности использования инструментов НМП, включая WIPO Translate, общедоступный на Patentscope, для перевода текстовых материалов, содержащихся в заявках, подаваемых в рамках Гаагской системы. Тестовая оценка была проведена для перевода текстов, содержащихся в национальных регистрациях промышленных образцов Китая, которые были извлечены из Глобальной базы данных по образцам³⁵. Предварительные выводы штатных переводчиков заключались в том, что качество перевода не является достаточно удовлетворительным для постредактирования, если только системы не обучены работе с текстами в данной предметной области. Однако были отмечены некоторые преимущества WIPO Translate по сравнению с другими инструментами НМП при переводе технических терминов.

28. Возможна дальнейшая оценка целесообразности использования WIPO Translate для перевода текстов Гаагской системы на китайский и другие языки. Однако для этого Международному бюро потребуются образцы регистраций промышленных образцов на других языках, в том числе содержащих различные указания продуктов, условные обозначения и описания³⁶.

УСЛУГИ МАШИННОГО ПЕРЕВОДА, ОКАЗЫВАЕМЫЕ ВНЕШНИМИ ПРОВАЙДЕРАМИ

29. В качестве еще одного варианта услуги машинного перевода могут предоставляться внешними поставщиками. Однако качество перевода может быть не таким высоким, как при использовании WIPO Translate, учитывая, что качество зависит от обучения в соответствующей области с использованием большого объема данных. В этой связи следует напомнить, что модель WIPO Translate для перевода документов в рамках Гаагской системы обучается с использованием совокупности записей памяти переводов таких текстов. Кроме того, использование внешних услуг машинного перевода также может быть более дорогостоящим, чем использование WIPO Translate, разработанного внутри организации. Наконец, учитывая конфиденциальный характер заявок, подаваемых в рамках Гаагской системы, необходимо оценивать возможных поставщиков с точки зрения безопасности и защиты данных.

V. ПРИГОДНОСТЬ СРЕДСТВ МАШИННОГО ПЕРЕВОДА ДЛЯ ИСПОЛЬЗОВАНИЯ В РАБОТЕ С ДОПОЛНИТЕЛЬНЫМИ ЯЗЫКАМИ

30. В настоящее время Международное бюро не располагает памятью переводов для перевода текстовых материалов, содержащихся в заявках, подаваемых в рамках Гаагской системы, на какие-либо языки, кроме трех рабочих языков системы. Кроме того, WIPO Translate не был обучен работе с терминологией Гаагской системы на других языках. По завершении анализа доступности и удобства использования упомянутых выше технологий перевод текстов Гаагской системы с любых языков, кроме трех рабочих, и/или на эти языки в обозримом будущем будет осуществляться профессиональными переводчиками.

31. Кроме того, возможность использования этих инструментов для дополнительных языков будет зависеть от нескольких факторов, в частности от количества заявок, подаваемых на том или ином языке (что определяет объем записей в памяти переводов и обучающих данных для WIPO Translate), имеющегося объема других данных в этой

³⁵ Тестовая оценка проводилась для перевода текстовых материалов с китайского на английский, поскольку WIPO Translate используется для переводов документов РСТ в этой языковой комбинации.

³⁶ В связи с этим следует отметить, что Глобальная база данных по образцам не содержит каких-либо данных на арабском или русском языках.

области по каждой языковой комбинации для обучения WIPO Translate, а также от относительной схожести/несхожести языков в каждой из языковых комбинаций.

СООБРАЖЕНИЯ В ОТНОШЕНИИ СОЗДАНИЯ БАЗ ПЕРЕВОДА

32. Если в Гаагскую систему будут включены дополнительные языки, Международному бюро необходимо будет создавать и пополнять память переводов для каждой новой языковой комбинации.

33. В этом отношении создание корпуса памяти переводов перед введением дополнительного языка может сделать их более удобными для использования и снизить затраты на перевод на этот язык и с него, если прогнозируется достаточное количество заявок на этих языках. Таким образом, в качестве возможного варианта можно заранее перевести имеющиеся записи в памяти переводов на дополнительные языки, и этот вариант в настоящее время рассматривается Мадридской рабочей группой для использования в отношении указаний товаров и услуг. Тем не менее, хотя создание эквивалентных терминов на других языках может быть полезным для более коротких текстовых сегментов (в частности, указаний продуктов или условных обозначений), это может быть не столь полезным для перевода описаний³⁷, то есть для основной части переводческой работы.

34. Кроме того, перевод записей в памяти переводов потребует определенных затрат³⁸. Таким образом, стоимость перевода существующей памяти на дополнительные языки необходимо будет сравнить со стоимостью простого перевода текстов заявок, подаваемых в рамках Гаагской системы, на дополнительные языки и пополнения памяти переводов с течением времени. Если количество заявок меньше определенной пороговой величины, стоимость перевода существующих записей в памяти переводов будет значительно выше, чем перевод текста заявок, подаваемых в рамках Гаагской системы на дополнительных языках.

ОБУЧЕНИЕ WIPO TRANSLATE

35. В настоящее время Международное бюро использует WIPO Translate для машинного перевода текстовых материалов Гаагской системы, по которым нет записей в памяти переводов. Можно использовать WIPO Translate для новых языковых комбинаций.

36. В этой связи, чтобы обеспечить надлежащее качество перевода текстовых материалов Гаагской системы, необходимо будет обучать WIPO Translate работе с большими объемами текста в соответствующей области для каждой новой языковой комбинации. Поэтому может оказаться полезным обмен терминами в комбинациях соответствующих языков, такими как указания продукта, условные обозначения, формулы изобретения и описания. Однако неясно, находятся ли большие объемы таких данных для обучения в самих ведомствах Договаривающихся сторон или где-то еще. Следует также подчеркнуть, что, хотя WIPO Translate можно обучить с использованием записей памяти перевода только в одну сторону (например, с английского на китайский), это не гарантирует такого же качества перевода, если система используется для перевода в обратную сторону (например, с китайского на английский)³⁹. Более того, предварительный перевод соответствующего текстового материала в любом случае будет сопряжен с расходами, если его будет выполнять Международное бюро⁴⁰.

³⁷ Как упоминалось выше, описания представляют собой произвольный текст и сильно различаются. См. также пункт 17.

³⁸ Стоимость перевода записей переводческой памяти в рамках Мадридской системы на арабский, китайский и русский языки оценивается в 2,36–4,31 млн шв. франков. См. документ MM/LD/WG/19/7, пункт 42.

³⁹ См. сноску 8.

⁴⁰ См. пункт 34.

37. Кроме того, следует напомнить, что, несмотря на наличие значительных объемов данных, которые потенциально можно использовать для обучения, в настоящее время в системе PCT WIPO Translate используется только для двух языковых комбинаций⁴¹, тогда как Отдел письменных переводов использует WIPO Translate лишь в очень ограниченной степени⁴². Поэтому на данном этапе потенциальная возможность использования WIPO Translate для перевода текстовых материалов Гаагской системы на другие языки, помимо трех рабочих языков, и с этих языков представляется весьма ограниченной.

38. Наконец, также отмечается, что НМП, в том числе WIPO Translate, продолжает совершенствоваться, и поэтому можно рекомендовать периодически оценивать возможность использования доступных инструментов НМП для перевода текстовых материалов Гаагской системы.

VI. ЗНАЧЕНИЕ СМЕТЫ ЗАТРАТ

39. Текущие затраты на перевод текстовых материалов Гаагской системы, смета которых представлена в документе H/LD/WG/10/4, рассчитаны на основе ставок, по которым оплачивается перевод документов PCT, выполняемый внешними подрядчиками⁴³. Ставки зависят от объема переводимого текста в каждой языковой комбинации, от того, насколько техническим по своему характеру является переводимый текст, и от особенностей каждой языковой комбинации в PCT, и, следовательно, фактическая ставка, которая может применяться для каждого из дополнительных языков, может меняться в контексте общей ситуации в Гаагской системе, если такие языки будут включены в систему. В связи с этим в настоящем исследовании основное внимание уделяется потенциальной возможности снизить текущие расходы на перевод материалов на дополнительные языки за счет использования доступных в настоящее время технологий перевода.

40. Смета расходов в этом документе уже учитывает повышение эффективности за счет постепенной компиляции данных памяти переводов. Даже если Международное бюро переведет весь объем памяти переводов на дополнительные языки, это вряд ли повысит процент совпадений при переводе описаний, а предварительный перевод потребует дополнительных расходов⁴⁴. Фактически, учитывая непрерывное развитие в области промышленных образцов и связанные с этим изменения в терминологии, содержащейся в заявках, подаваемых в рамках Гаагской системы, Международное бюро наблюдает замедление роста степени совпадений в памяти переводов Гаагской системы и, следовательно, предполагает, что около 60 процентов вариантов перевода, генерируемых инструментом НМП, всегда будут подлежать постредактированию профессиональным переводчиком⁴⁵.

41. Кроме того, учитывая приведенные здесь результаты анализа, имеющиеся в настоящее время инструменты НМП в обозримом будущем не смогут обеспечить повышение производительности или эффективности перевода текстовых материалов, содержащихся в заявках, подаваемых в рамках Гаагской системы, на дополнительные языки. Стоимость перевода и контроля качества не может быть снижена без ущерба для качества перевода. В связи с этим подчеркивается, что любой компромисс в отношении качества перевода поставит под угрозу права пользователей⁴⁶. Таким образом,

⁴¹ См. пункт 21.

⁴² Определенные категории документов, такие как стенографические отчеты о заседаниях ВОИС. См. пункт 24.

⁴³ В смете указаны дополнительные расходы на перевод в случае включения китайского и русского языков в соответствии с запросом Рабочей группы (см. документ H/LD/WG/8/8, пункт 29). Хотя можно предположить, что те же текущие расходы на перевод в случае добавления новых языков также могут зависеть от ставок, применимых в рамках PCT, можно провести дополнительную оценку, если это будет одобрено Рабочей группой.

⁴⁴ См. пункты 34 и 36 выше.

⁴⁵ См. пункт 13 выше.

⁴⁶ См. пункт 26 выше.

результаты технического анализа не дают оснований для любого пересмотра сметы расходов на включение дополнительных языков в Гаагскую систему.

42. Наконец, отмечается, что использование технологий перевода может повлиять только на стоимость непосредственно перевода. Все остальные расходы, в том числе на серверы для поддержки любой дополнительной рабочей нагрузки по переводу и контролю качества, будут иметь место независимо от использования технологий перевода, если в Гаагскую систему будут включены дополнительные языки.

43. Рабочей группе предлагается обсудить и прокомментировать содержание настоящего документа.

[Конец документа]