

Доклад по итогам Переводческого саммита TAUS
Амстердам, 22-24 марта 2017 г.



СЕТЬ ЯЗЫКОВЫХ ДАННЫХ

Перевод доклада ***The Translation Industry in 2022. A report from the TAUS Industry Summit. Amsterdam, March 22-24, 2017*** был подготовлен группой студентов переводческого факультета (www.pfnn.ru) Нижегородского государственного лингвистического университета имени Н. А. Добролюбова (www.lunn.ru) в рамках производственной практики в компании «АКМ-Вест» (руководитель практики – **Берендьяев Максим Викторович**, www.akmw.ru) осенью 2017 года. Работа была организована поэтапно с применением Memsource: за подготовку (распознавание) и редактирование текста отвечал **Богданов Владислав**, переводчиками выступили **Батова Антонина**, **Волкова Алиса**, **Килина Ольга**.

Translation of ***The Translation Industry in 2022. A report from the TAUS Industry Summit. Amsterdam, March 22-24, 2017*** from English into Russian was provided by a team of students of School of Translation and Interpreting (www.pfnn.ru) at Nizhniy Novgorod State Linguistic University named after N.A. Dobrolyubov (www.lunn.ru) as part of their Practical Training at AKM-West Translation Company (www.akmw.ru) in September 2017. The team's work was supervised by **Maxim V. Berendyaev**, Tech & Quality Manager of AKM-West, and carried out in stages, involving application of Memsource computer-assisted translation tool. The text was translated by **Antonina Batova**, **Alisa Volkova** and **Olga Kilina**, prepared and proofread by **Vyacheslav Bogdanov**.

Автор: Изабелла Массардо, Яп ван дер Меер

Рецензент: Анне-Май ван дер Меер

COPYRIGHT © TAUS 2017 Все права защищены.

Никакая часть данной публикации не может быть воспроизведена, сохранена в информационно-поисковой системе любого типа, передана или опубликована в любой форме и любыми способами – в том числе посредством электронных, механических, фотокопировальных, записывающих и других устройств – без предварительного письменного разрешения ассоциации TAUS. Любые нарушения авторского права будут преследоваться по закону.

Несмотря на тщательную проработку и редактуру, в данной публикации могут быть обнаружены ошибки или недочеты. Авторы, редакторы и ассоциация TAUS не несут ответственности за возможные последствия.

Оформление: Анне-Май ван дер Меер

Издательство TAUS BV, Де Рейп, Нидерланды

Вы можете задать вопрос, написав на электронный адрес info@taus.net

Содержание	4
1. Основные положения	5
2. Структура доклада	6
3. История создания данного доклада	7
3.1 Консультации	7
3.2 Выбор направлений	7
4. Краткий обзор эволюции переводческой индустрии	9
4.1 Современное состояние	9
4.2 Обзор будущего переводческой индустрии	12
5. Шесть факторов изменений	16
5.1 Введение	16
5.2 Машинное обучение	16
5.3 Машинный перевод	19
5.4 Менеджмент качества	21
5.5 Данные	22
5.6 Функциональная совместимость переводческих инструментов	24
5.7 Академия	26
6. Десять инновационных аспектов	30
6.1. Алгоритмическое управление	30
6.2. Датафикация перевода	32
6.3 Вопросы качества	25
6.4 Конвергенция	32
6.5 Речь	33
6.6 Короткие и длинные переводы	34
6.7 Стремительный рост	34
6.8 Особые заказы	35
6.9 Плата за пользование	36
6.10 Ориентирование на пользователя	36
6.11 Эпилог	37
7. Больше информации по ссылкам:	38
8. Об авторах	39

1. Основные положения

По прогнозам ассоциации TAUS, в ближайшие пять лет процесс автоматизации в индустрии переводов будет идти еще интенсивнее.

Мы ожидаем стремительное внедрение автоматизации с помощью машинного обучения как в сам процесс перевода (машинный перевод и «умные» системы ПП), так и в сферу управления переводами (включающую в себя, например, управление проектами, контроль качества и организацию работы поставщиков).

Процесс автоматизации будет сопровождаться появлением новых перспектив и новых трудностей. Главное преимущество автоматизации заключается, разумеется, в повышении производительности. Трудности же будут связаны с изменением профессиональных обязанностей, с необходимостью обмениваться данными, получать из данных информацию и работать в облачных системах. К новым перспективам также относятся инновации, внедрение которых стало возможным благодаря автоматизации переводов.

Сильные стороны	Слабые стороны
Повышение производительности	Необходимость обмениваться данными
Новые перспективы	Угрозы
Получение информации из данных Работа в облачных системах Конвергенция технологий	Изменение профессиональных обязанностей

Все это становится возможным благодаря конвергенции: конвергенции технологий (таких как машинный перевод и технология обработки речи), а также конвергенции рынков и бизнес-моделей: условно-бесплатного и платного ПО, потребителей и компаний, любителей и профессионалов.

Потребность в переводе скоро станет рядовой потребностью. Покупателям и поставщикам профессиональных переводческих услуг придется пересмотреть свою стратегию с учетом изменяющейся ситуации на рынке переводов.

2. Структура доклада

По структуре данный доклад представляет собой хронологически выстроенный отчет о консультациях с членами TAUS, об анализе полученных результатов, об их обобщении в Консультативном совете и о мозговом штурме в формате World Café, организованном на площадках Переводческого саммита TAUS с 22 по 24 марта. Приведенная далее диаграмма свидетельствует о том, что наша работа началась с просьбы ко всем членам ассоциации TAUS предложить идеи по внедрению инноваций в работу переводчиков.

Мы собрали шестьдесят различных идей и путем голосования определили их важность и актуальность. После тщательного анализа и обобщения результатов список идей был сокращен до шести основных направлений (шести факторов изменений), которые легли в основу программы Переводческого саммита TAUS, состоявшегося 22-24 марта 2017 г. в Амстердаме.

На саммите TAUS началось обсуждение радикальных изменений, которые охарактеризуют индустрию переводов в ближайшие пять лет. Обсуждение затрагивало десять инновационных аспектов, каждый из которых имеет свои особенности, но в то же время тесно связан с остальными.

Данный доклад начинается с главы об истории его создания: в главе 3 описываются консультации и объясняется, каким образом мы выделили шесть факторов изменений.

В главе 4 представлен краткий обзор эволюции переводческой индустрии с технологической точки зрения как в прошлом, так и в будущем.

В главе 5 подробнее рассматриваются шесть факторов изменений: машинное обучение, машинный перевод, контроль качества, данные и датафикация (преобразование в цифровую форму) – в том числе датафикация речи, функциональная совместимость инструментов и образование (подготовка кадров). Вышеперечисленные главы опираются на анализ реальных данных.

В главе 6 на основе фактов и существующих тенденций высказываются предположения о состоянии индустрии переводов в 2022 году, которые подразделяются на десять подглав. Умозрительно? Возможно. Правдоподобно? Определенно. Необходимо? Без сомнений. Потому что это единственный способ подготовиться к стремительным изменениям в индустрии переводов.

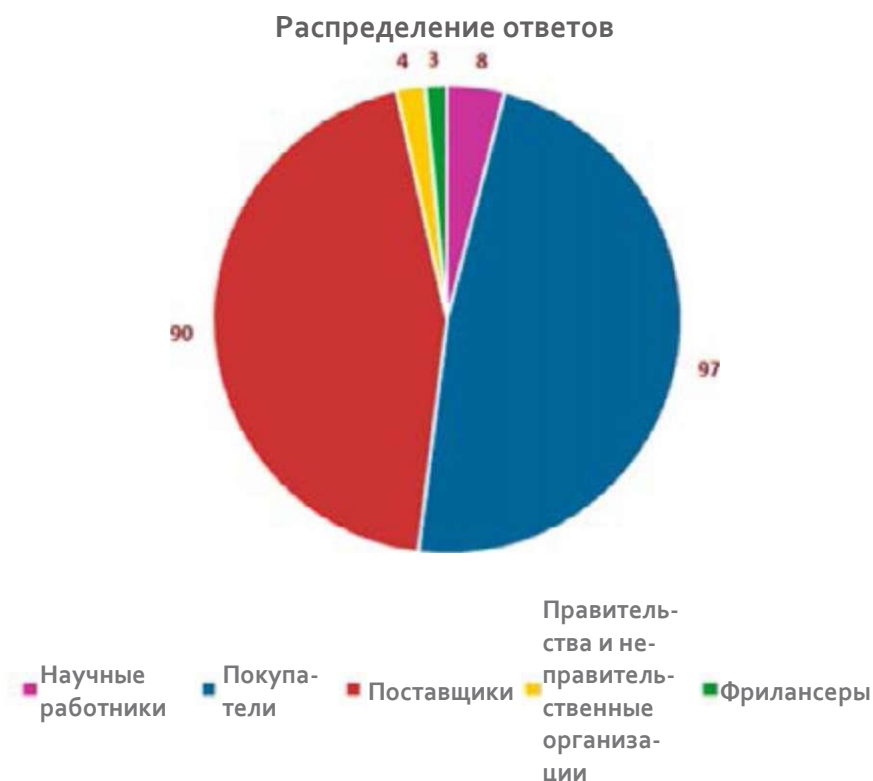


3. История создания данного доклада

3.1 Консультации

С ноября 2016 г. по январь 2017 г. включительно ассоциация TAUS проводила опрос своих членов.

Их просили предложить идеи по внедрению инноваций. В результате было собрано шестьдесят различных идей, которые были сгруппированы в тринадцать категорий. Затем участников попросили составить рейтинг этих тринадцати категорий в зависимости от их важности. Было получено 202 ответа, причем все – от руководителей, т.е. людей на уровне директоров и владельцев компаний. На приведенной ниже диаграмме отражено распределение ответов.



Приведенная ниже диаграмма отображает в порядке убывания степень важности каждой из тринадцати категорий по мнению членов ассоциации. Числа в скобках указывают на количество идей, включенных в соответствующую категорию.

Явным лидером стало машинное обучение, за ним следуют панель управления качеством (Quality Dashboard с инструментом оценки качества TAUS DQF) и машинный перевод (МП) на втором и третьем месте соответственно.

3.2 Выбор направлений

После консультаций и составления общего рейтинга категорий было выбрано шесть направлений, которые легли в основу программы Переводческого саммита TAUS, состоявшегося 22-24 марта 2017 г. в Амстердаме. Эти шесть направлений охватывают десять из тринадцати категорий, упомянутых в рейтинге.

Три категории были исключены из-за низкой позиции в общем рейтинге или из-за недостаточной релевантности. Категории «Нормативы» и «Инструменты с открытым исходным кодом» выходят за рамки исследования, по крайней мере, для членов ассоциации TAUS. Категория «Продвижение ПО», напротив, косвенно присутствует в каждом аспекте.

Рейтинг категорий по степени их важности для членов ассоциации

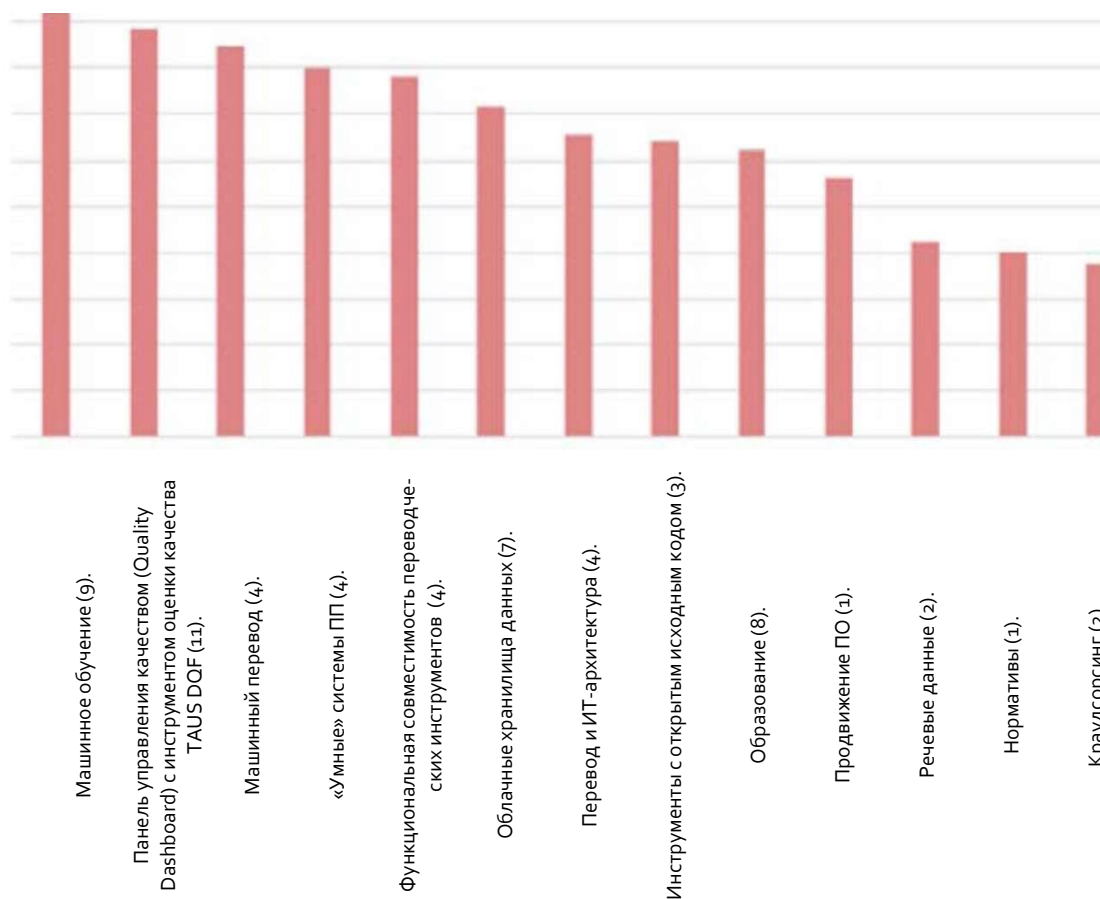


Диаграмма: тринадцать категорий, на которые подразделяются идеи. В скобках указано число идей, включенных в соответствующую категорию.

Шесть выбранных направлений являются ключевыми для планирования стратегии каждого участника переводческого рынка – вне зависимости от того, является ли он потребителем или поставщиком услуг. Все вместе эти направления составляют шесть факторов изменений переводческой индустрии.

Именно эти факторы обусловят изменения индустрии переводов в течение ближайших пяти лет. Перед управляющими встанет непростая задача – работать с инновациями, внедрение которых уже идет полным ходом, так что придется либо задавать новые тенденции, либо следовать уже существующим.

Шесть факторов изменений:

1. Машинное обучение
2. Машинный перевод
3. Контроль качества (в том числе инструмент TAUS DQF и панель управления качеством)
4. Данные (в том числе речевые данные)
5. Функциональная совместимость инструментов
6. Образование (подготовка кадров)

4. Краткий обзор эволюции переводческой индустрии

Благодаря глобализации мы можем получить любую доступную информацию на все большем количестве языков. Мы надеемся, что переводческие технологии помогут нам снизить языковые барьеры.

За последние несколько лет переводческие технологии, охватывающие различные аспекты, стали динамично развивающимся сектором, который, с точки зрения инноваций, привлекает как участников индустрии, так и разработчиков за ее пределами: постоянно создаются новые инструменты, платформы и программные решения, призванные повысить наш многоязычный потенциал.

Изначально основной задачей переводческих технологий было повышение продуктивности. В 2007 году, когда компания Google представила миру сервис Google Translate, она, сама того не желая, подорвала стабильность переводческого рынка. Google Translate не только смог пережить десять лет постоянных упреков в якобы плохом качестве переводов, но и обзавелся постоянными пользователями, чье количество исчисляется сотнями миллионов, а также стал вдохновением для множества стартапов.

До Google Translate перевод всегда был профессиональной услугой, которую обычно могли позволить себе исключительно корпоративные клиенты. Приход Google Translate открыл непосредственный доступ к переводу конечным пользователям. Google Translate повысил важность и значимость перевода, а также вызвал беспрецедентную волну инноваций во всей переводческой индустрии.

Достижения прошлого десятилетия стали прелюдией к эре конвергенции, в рамках которой отдельные технологии начинают обмениваться ресурсами и синергетически взаимодействовать друг с другом на основе общего интерфейса на разных устройствах. Когда конвергенция достигнет своего расцвета, перевод будет доступен на каждом экране, в каждом приложении и для каждой вывески.

Можно сказать, что движение в направлении конвергенции началось в январе 2017 года, когда компания Google объявила, что сервис Google Translate стал доступным на 103 языках и ежемесячно обслуживает 500 млн. пользователей, переводя по 140 млрд. слов в день. Иными словами, за день Google Translate в одиночку переводит больше слов, чем все люди-переводчики в мире за год.

4.1 Современное состояние

История переводческих технологий насчитывает всего несколько десятилетий. Она начинается в конце 1966 года, когда Консультативный комитет по автоматической обработке языка (ALPAC) в своем докладе призвал остановить спонсирование разработок в области машинного перевода (МП).

В докладе ALPAC МП характеризуется как «слишком дорогой, слишком затратный по времени и слишком неточный» – у общественности и остального научного сообщества должно сложиться представление о безнадёжности вышеупомянутой технологии. С тех пор МП всегда встречал стойкое неприятие со стороны переводчиков, которые боялись, что новая технология отнимет у них работу.

Выводы, содержащиеся в докладе, в основном объясняются недостаточным количеством людей-переводчиков в то время, относительно малым количеством текстов для перевода и дешевизной переводческого труда по сравнению с затратами на исследования ученых – все это сделало МП экономически невыгодным. В докладе ALPAC, впрочем, рекомендуется

разрабатывать инструменты, которые помогут улучшить производительность переводчиков.

В конце 80-ых годов, когда по всему миру распространилось доступное компьютерное оборудование и резко увеличилось производство контента, рекомендации ALPAC превратились в жизнеспособную, практичную и доступную технологию. По сути, увеличение производства контента подстегнуло развитие глобального ПО и глобального контента, которые нужно было переводить на десятки языков, в результате чего появилась индустрия локализации. С тех пор переводческие технологии стали развиваться параллельно с индустрией переводов.

Благодаря развитию технологий в 90-е годы уже существующие инструменты получили новые возможности, и одновременно с этим появились новые инструменты с разными программными пакетами, функциями, операционными системами и ценами, что предоставило покупателям более широкий выбор.

2000-е стали десятилетием стремительного роста, а последние несколько лет в особенности ознаменовались глобальным развитием переводческих технологий. Наиболее заметной технологией в это время стал именно МП. Несмотря на выводы, содержащиеся в докладе ALPAC, исследования в области обработки естественного языка продолжались, и развитие МП было в их числе. Честно говоря, развитие МП происходило так, что сегодня большинство пользователей даже не вспомнят о тех мучительных усилиях, что предшествовали появлению Google Translate. Но даже в наши дни, несмотря на весь прогресс и успех МП, многие представители индустрии переводов считают его ошибки неизбежными.

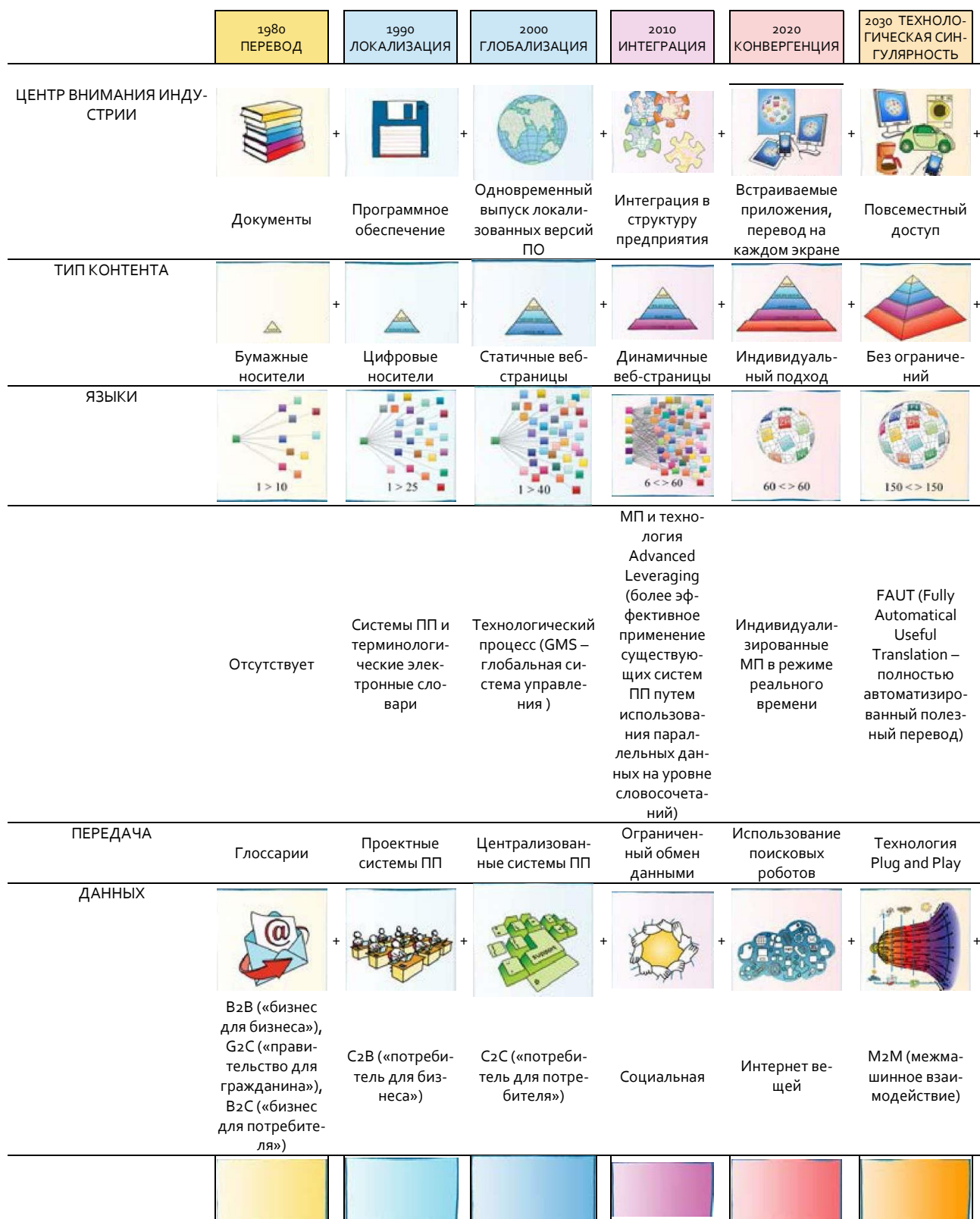


Рисунок: эволюция индустрии переводов

Важнейшие инновационные идеи обычно предлагаются разработчиками, не относящимися к переводческой индустрии, а большинство современных технологий основываются на проектах, которые ранее финансировались из государственного бюджета. Как верно отмечает Мариана Маццукато в своей книге «Государство-предприниматель», Интернет, системы GPS и виртуальные помощники с голосовым управлением разрабатывались вооруженными силами США, а сенсорный экран и язык программирования HTML появились благодаря исследованиям научных работников из университетов и лабораторий, финансируемых государством – и это лишь несколько примеров.

Индустрия переводов и индустрия переводческих технологий не так сильно отличаются друг от друга: сразу же напрашивается пример компании Systran, которая смогла пережить депрессию на рынке МП благодаря финансовой поддержке Евросоюза. Переводческое сообщество уже несколько десятилетий преуспевает благодаря технологиям, льстя себе утверждением, что переводчики работают в высокотехнологичной отрасли, несмотря на то, что большинство используемых бизнес-моделей все еще остаются устаревшими.

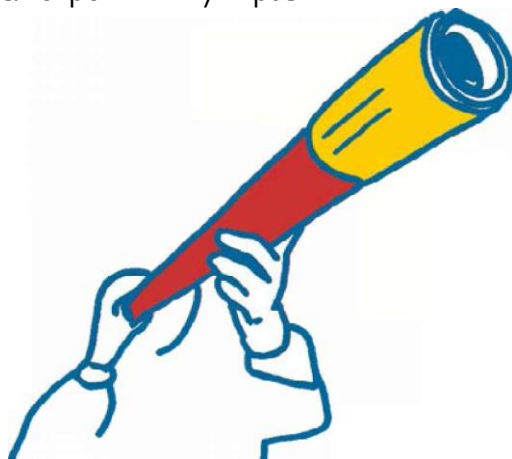
В рамках разработки переводческих технологий были созданы системы управления переводами (translation management system – TMS), которые отвечают потребностям стандартного переводческого бизнеса: сокращают операционные расходы и одновременно имитируют традиционный порядок работы. Результатом стало любопытнейшее усеченное применение платформ управления проектами и планирования ресурсов предприятия (enterprise resource planning – ERP), совершенно непохожих на аналогичные стандартные системы, которые можно встретить на предприятиях, относящихся к любому другому сектору экономики. Вспомните компании Basecamp, Compuware, Infor, JD Edwards (сейчас: Oracle), Jeeves, PeopleSoft, Plex, SAP, Tally: все они являются разработчиками ПО, выпускающими продукты, которые можно было бы адаптировать для переводческого бизнеса. Поскольку вышеупомянутые системы слишком сложны и затратны для переводческих микропредприятий, то стали разрабатываться стандартные системы ПО для управленческой деятельности, ориентированные на переводческий бизнес.

4.2 Обзор будущего переводческой индустрии

Можем ли мы наверняка знать, что произойдет в будущем? История столь богата на примеры ложных прогнозов, основанных на случайных фактах, что в Викицитатнике есть целая страница, посвященная неверным предсказаниям.

Кажется, очень немногие исследователи и представители переводческой индустрии согласны с прогнозом Рэя Курцвейла о том, что в 2029 году МП будет выдавать качество, аналогичное качеству переводов, выполняемых человеком. Всего лишь десять лет назад многие члены переводческого сообщества были убеждены, что МП – это просто мошенничество, но сегодня они вынуждены признать, что он совершенствуется с каждым днем.

По сути, в наши дни переводится лишь малая часть контента, который требует перевода, а объемы этого контента стабильно растут, в результате чего спрос на пригодные для публикации переводы физически не может быть удовлетворен одними лишь людьми-переводчиками. Многие люди, так или иначе связанные с работой профессиональных переводчиков, утверждают, что переводчики уже не отвечают задачам современности, к тому же, в их переводах обнаруживается все больше и больше ошибок, чем раньше. С момента своего появления МП прошел огромный путь развития.



На самом деле, когда сбываются прогнозы, когда-то казавшиеся нереалистичными, мы зачастую забываем, что эти прогнозы были сделаны после тщательного анализа ситуации на момент их создания и небольшого промежутка времени, предшествовавшего этому моменту. Лишь несколько современных тенденций с большой вероятностью продолжают нарастать в ближайшие несколько лет и смогут проложить дорогу в будущее: облачные технологии, всеобъемлющая датафикация и машинное обучение.

Машинное обучение будет все больше применяться к переводу, особенно в областях МП и автоматического перевода устной речи (S2S, или speech-to-speech translation): оно откроет новые перспективы для переводческого бизнеса и поднимет оценку качества на новый уровень.

Сегодня в индустрии переводов наблюдается мода на нейронный машинный перевод (НМП). НМП представляет собой новую технологию МП, которая исследуется и разрабатывается по меньшей мере с 2012 года, а с 2016 года появилась в открытом доступе. Организации, которым довелось поработать с этой технологией, убеждены, что она значительно выигрывает у основанного на переводе словосочетаний СМП в плане целостности и точности переводного текста.

В отличие от статистического машинного перевода (СМП), который для «запоминания» словосочетаний использует таблицы соответствий, НМП учится переводить предложения путем использования технологии под названием «нейронные сети» – это несколько слоев связанных «нейронов», чья работа в симбиозе имитирует когнитивные процессы, происходящие в мозгу человека при узнавании и запоминании информационных моделей. Как и СМП, НМП обучается на параллельных текстах, однако нейросети обеспечивают так называемое «глубокое обучение», благодаря которому движки НМП располагают куда большим потенциалом построения предложений, чем движки СМП.

Сейчас эту технологию оценивают специалисты по разработке МП. Некоторые разработчики систем МП уже предоставляют своим клиентам программные решения, включающие НМП, а поставщики лингвистических услуг проводят вебинары и другие мероприятия для ознакомления с результатами своих разработок. Начальные затраты относительно невысоки, поэтому все представители индустрии МП находятся в равных условиях для внедрения новой технологии. В будущем же основное внимание разработчиков обратится к созданию вокруг существующей базовой технологии индивидуализированных оболочек, которые отвечали бы требованиям в конкретных отраслях, и к оценке работы систем НМП во всех случаях использования их перевода.

Несмотря на всеобщий энтузиазм, НМП не является панацеей, благодаря которой будет создано универсальное программное решение для всех видов перевода. Однако многие из тех, кто использует эту новую технологию, говорят о коренном сдвиге в производительности систем МП.

Исследования в области нейросетей, машинного обучения, глубокого обучения и т.п. будут продолжаться главным образом из-за невероятной притягательности подобных технологий. Велика вероятность, что эти исследования будут финансироваться как из государственного бюджета, так и из частных источников. С конца Второй мировой войны и вплоть до наших дней разработка новых технологий процветает благодаря солидным государственным инвестициям, которые позволили достичь невероятных результатов с не менее невероятными масштабными последствиями.

Результатом исследований, связанных с обороной, может стать конвергенция, например, (нейронного) МП и речевых технологий. И как всегда, фантастические сюжеты (вроде

фильма «Прибытие») породят совершенно новые ожидания. Эволюция переводческих технологий, однако, пойдет разными путями, которые почти не будут пересекаться.

Пока что инновации в сфере переводческих технологий не привели к радикальному изменению самого процесса перевода. В этом плане результатом улучшения качества МП и появления большого количества повсеместно доступных платформ стала их интеграция во все профессиональные переводческие инструменты, что проложило путь к практике столь же устаревшей, как и сам МП – к постредактированию.

Результаты разных исследований свидетельствуют о разных долях постредактирования в конечном переводческом продукте, просто потому что в так называемом интерактивном режиме постредактирование является стандартным рабочим процессом для большинства профессиональных переводчиков, и это останется так, по крайней мере, в краткосрочной перспективе.

Традиционно для распространения статистического постредактирования требуется некоторое время, в течение которого покупатели учатся эффективнее использовать свои данные, применяя расширенные терминологические базы для сопоставления субсегментов и «обучая» движки СМП, что делает возможным получение результата из частично совпадающих фрагментов. Пока НМП не заменит СМП полностью, постредактирование останется главным способом обнаружения проблем в производительности систем СМП, ключом к пониманию этих проблем и к их решению в целях совершенствования движков СМП.

Современные системы управления переводческой памятью используют редакционное расстояние (edit distance) для измерения семантической разницы между двумя сегментами – алгоритм, который совершенно не похож на суждение человека о семантической схожести сегментов. Сегодня в процессе разработки находятся более «умные» аналогичные метрики, основанные на результатах исследований в области обработки естественного языка.

В ближайшие несколько лет, скорее всего, будет стимулироваться внедрение еще большего количества инноваций в область управления переводческой памятью, чтобы сделать последнюю более «умной» и охватить разные формы отображения информации, особенно фонетические языковые данные – путем расширения потенциала метаданных. Когда представителям индустрии переводов задают вопрос о возможных грядущих инновациях, они обычно говорят о развитии машинного обучения, что позволит повысить качество переводов, считающееся главной проблемой, до уровня, на котором МП можно будет доверять.

Машинное обучение может спровоцировать отмирание систем управления переводами (TMS), или, по крайней мере, той их формы, которая нам знакома. В случае с системами управления переводами дело уже не в так называемом «запланированном устаревании» ПО – TMS устарели концептуально. Они не смогли помочь в избавлении от посредников, потому что их разработчики старались скорее удовлетворить требования этих самых посредников, а не потребности конечных потребителей. Однако стремление к избавлению от посредников будет только нарастать: по нашим ожиданиям, процесс автоматизации будет идти все быстрее, даже в таких областях управления переводами, как управление проектами, контроль качества и организация работы поставщиков.

И при этом большую роль снова сыграет машинное обучение. Алгоритм машинного обучения можно развивать, используя профилирование контента и пользователей, производительность, распознавание и накопление терминологии, прогнозирование и оценку качества, информацию о финансах и о проектах для автоматического распределения ресурсов в целях их наиболее эффективного использования и оценивания переводчика, а также для выбора из доступных ресурсов наиболее пригодных для выполнения конкретного задания.

Адаптация движков МП к индивидуальным потребностям также будет происходить автоматически через алгоритмы машинного обучения, что позволит системе определять наиболее подходящие данные.

В этом случае значительную роль сыграет связующее ПО, которое через программные интерфейсы приложений (API) будет контролировать передачу контента от одной системы к другой. Поэтому мы ожидаем, что будут разработаны новые форматы сообщений, не зависящие от языка программирования.

5. Шесть факторов изменений

5.1 Введение

В своей книге «Сигналы говорят» американский футуролог Эми Уэбб описывает метод, позволяющий обнаружить новые тенденции, пока они еще развиваются, и отслеживать эти тенденции после того, как они станут мейнстримом. Исследовательница называет свой метод словом «CIPHER» — акронимом, составленным из английских слов contradictions (противоречия), inflections (изгибы), practices (практики), hacks (подсказки), extremes (крайности), rarities (редкости). Уэбб обращается к прошлому для подкрепления своих выводов и призывает людей сохранять скептический взгляд на вещи и проверять свои умозаключения. В конце концов, умение прогнозировать будущее заключается лишь в способности отличать тенденции от просто популярных в данный момент вещей.

Из шести факторов изменений, выделенных нами в ходе опроса, в ближайшие несколько лет скорее всего продолжит нарастать только одна тенденция – датафикации. Данные стимулируют развитие всех тех технологий, что кардинально меняют нашу жизнь и способы ведения бизнеса. И именно данные позволят нам перейти на новый этап развития искусственного интеллекта с помощью машинного обучения.

Для датафикации вспомогательной технологией является облачная. Поскольку количество ресурсов огромно, облачная технология необходима здесь как инфраструктура для распределения данных и обмена ими, для нейросетей и глубокого обучения.

Ни один прогноз не обходится без определенной вероятности ошибки: если общие показатели технологического прогресса всегда неточны, то справедливо утверждать, что сейчас развитие идет рекордными темпами, а потому кривая на графике будет становиться только круче. Рассмотрим в перспективе: если бы производители автомобилей внедряли инновации так же, как и представители ИТ-сектора, то современные автомобили проезжали бы 42 000 км на литре топлива и стоили бы 50 центов.

Как было указано ранее, путем опроса членов ассоциации TAUS мы выделили шесть факторов изменений переводческой индустрии, которые должны учитываться при планировании бизнес-стратегии любого участника переводческого рынка, вне зависимости от того, является ли он потребителем или поставщиком услуг. Итак, к факторам изменений относятся:

1. Машинное обучение
2. Машинный перевод
3. Контроль качества
4. Данные
5. Функциональная совместимость инструментов
6. Образование (подготовка кадров)

5.2. Машинное обучение

Машинное обучение (Machine Learning) появилось в 50-ых годах прошлого века как разновидность искусственного интеллекта, позволяющая программам «учиться» на данных, выделять в них определенные структуры и основе этого делать выводы.

Избыток данных в открытом доступе – благодаря более дешевой и более мощной компьютерной технике – дарит новую жизнь старым алгоритмам, что приводит к появлению рентабельных сфер применения новых технологий. Однако большинство из нас не может назвать даже три области применения машинного обучения: на ум приходят только беспилотные автомобили или голосовые помощники, встроенные в наши смартфоны. Глубокое обучение

– это уровень технологий машинного обучения, который основывается на нескольких слоях нейронных сетей, работающих в симбиозе и развивающихся за счет изобилия вычислительных ресурсов в облаке. Большинство алгоритмов машинного обучения никак не связаны с нейросетями, а искусственного интеллекта не хватает для того, чтобы компьютеры вели себя, как в фантастических фильмах. Во всяком случае, пока.

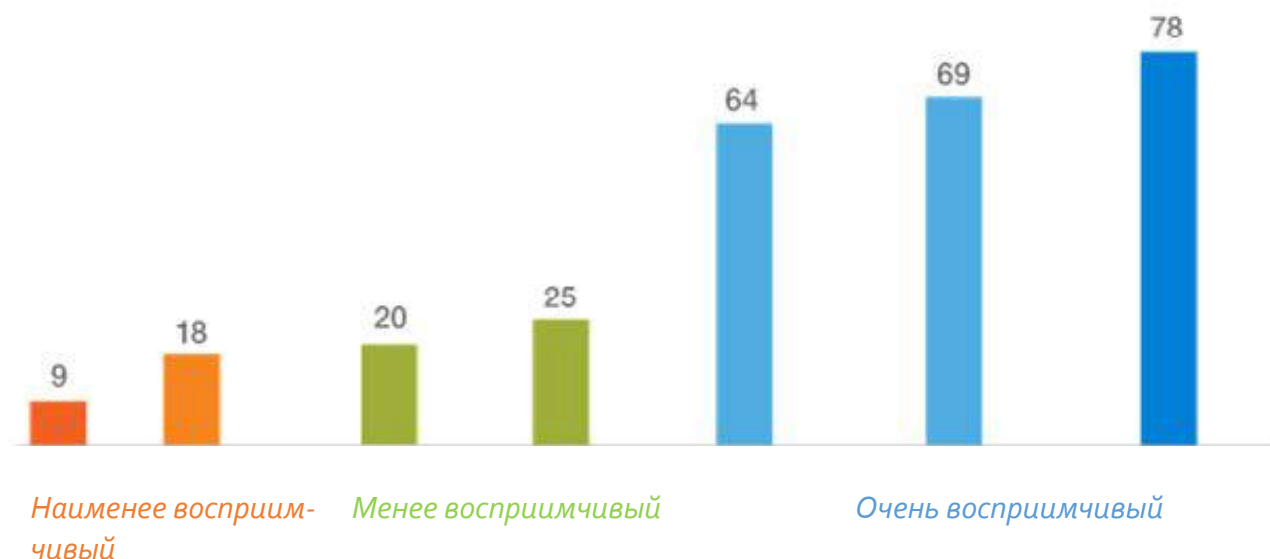
В 2016 году машинное обучение получило невероятное распространение, в результате чего появилось множество сфер его применения, о которых несколько лет назад нельзя было даже мечтать: распознавание изображений и беспилотное вождение – всего лишь часть подобных примеров.

[По мнению компании McKinsey, на графике зрелости технологий от аналитической компании Gartner машинное обучение уже располагается на той части кривой, которая идет вниз](#), на отрезке перехода от пика завышенных ожиданий к нижней точке разочарования. Это значит, что машинное обучение будет оказывать серьезное влияние на стратегическое планирование всех компаний, так как эта технология продолжает оставаться многообещающей в плане выдачи конкурентоспособного качества в ближайшие 5-10 лет.

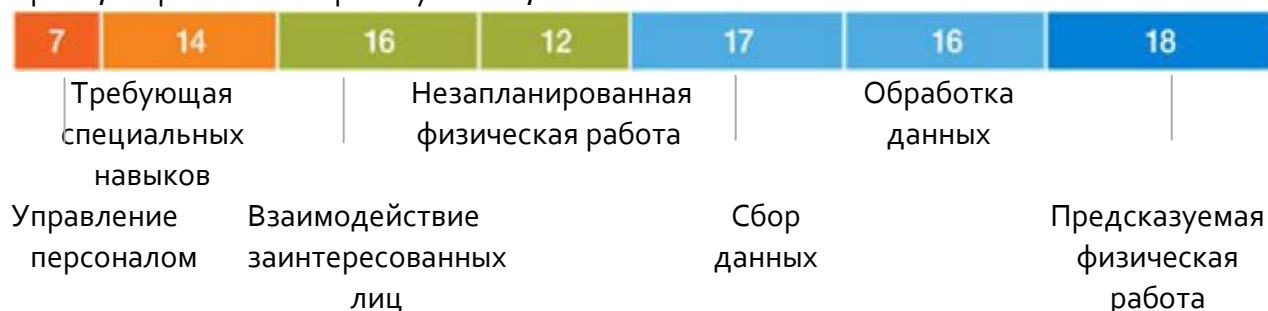
За счет мощности компьютеров и беспрецедентного прогресса в разработке глубоких нейросетей технологии работы с данными станут действительно революционными. Это может случиться и с МП. По сути, СМП – это сфера применения машинного обучения, которая десятилетиями находилась на передовой языковых технологий.

В результате автоматизации искусственный интеллект будет играть все большую роль в повседневной жизни: его влияние будет совершенно разным в зависимости от секторов экономики и видов деятельности. Как правило, рост искусственного интеллекта расценивается как угроза для многих профессий, так как люди будут заменены «умными» системами. Несмотря на то, что в следующем десятилетии искусственный интеллект приведет к исчезновению крайне малого числа профессий, он все равно окажет влияние на все профессии – в большей или меньшей степени, в зависимости от выполняемых задач.

В 2013 году [экономисты из Оксфорда Карл Бенедикт Фрей и Майкл А. Осборн провели исследование](#), в котором оценили вероятность замены 702 конкретных должностей машинами: шансы лишиться работы для корректоров и составителей технической документации составили 84% и 89% соответственно. В 2014 году экономист У. Брайан Артур предположил, что машины могут заменить 100 миллионов рабочих мест в США к 2025 году. В свою очередь, [эксперты компании МакКинзи считают, что автоматизация повлияет на 45% оплачиваемой деятельности](#). Будут автоматизированы 30% операций в 60% всех отраслей.



Время, затраченное на работу в США, %



Источник: McKinsey & Company

Большинство управляющих позиций не будут затронуты автоматизацией, в то время, как многие вспомогательные позиции исчезнут. Люди, занимающиеся креативной работой, находятся в относительной безопасности, в то время как работники, занимающиеся продажами, подвержены риску. В сфере ИТ прогноз выглядит точно так же: управляющие позиции по-прежнему будут востребованы, однако программисты и работники службы поддержки могут быть заменены.

Утверждение о том, что деятельность, требующая ловкости рук, творчества и социальных навыков, сложнее всего компьютеризировать, уже устарело, так как новые разработки в области глубокого обучения делают машины более способными, чем ожидалось, особенно в областях, связанных с творчеством и социальным взаимодействием. Несмотря на то, что возможность технической осуществимости является необходимым предварительным условием для автоматизации, это вовсе не означает, что определенная деятельность все же будет автоматизирована. Стоимость рабочей силы и ИТ-инфраструктуры являются двумя основными факторами, которые влияют на возможность автоматизации. Поэтому автоматизация возможна только в том случае, если относительно недорогая рабочая сила будет в избытке.

Например, в настоящее время почти пятая часть времени, затрачиваемого на рабочих местах в США, включает в себя выполнение определенных действий в хорошо известной последовательности, а изменения в такой деятельности относительно легко предвидеть. Такая работа называется предсказуемой физической работой. Сложнее всего автоматизировать ту деятельность, которая связана с управлением и развитием нематериальных ресур-

сов (9% потенциала автоматизации) или же ту, где необходим опыт для принятия решений, планирования или творческих процессов (18%).

На текстильных фабриках в конце 19-го и начале 20-го вв. замена паровых двигателей на электрические привела лишь к небольшому увеличению производительности. Один из самых больших и ожидаемых технологических прорывов в нашем веке произойдет, если (или когда) машины смогут понимать естественные языки так же, как и человек. Компьютеры, способные распознавать концепции повседневного общения, смогут увеличить технический потенциал автоматизации с 53% всего рабочего времени до 60% в розничной торговле, с 43% до 66% в сфере финансов, страхования и здравоохранения.

В 2013 году Управление по контролю за продуктами и лекарствами одобрило машину для введения седативных средств пациентам без анестезиолога. В такой сфере, как радиология появляется новая область – компьютерная диагностика. Недавнее исследование, опубликованное Королевским обществом, показало, что компьютеры выявляли рентгенопрозрачность (появление темных изображений) более последовательно (почти в десять раз), чем радиологи.

У. Брайан Артур назвал «второй экономикой» часть экономики, в которой компьютеры осуществляют деловые операции только с другими компьютерами. Эта вторая экономика находится над нами, экономика, наполненная оптимистическими предпринимателями, уже порождающая новое поколение миллиардеров, вероятно, во многом способствующая росту в ближайшие десятилетия. Артур также предполагает, что в 2025 году эта вторая экономика может быть такой же большой, как первая экономика в 1995 году. Если она достигнет такого темпа роста, то заменит работу примерно 100 миллионов людей в США (нынешняя общая занятость гражданской рабочей силы в США составляет 146 миллионов работников).

В индустрии перевода, как и в других отраслях, множество функций будет расширено или заменено с помощью машинного обучения.

5.3 Машинный перевод

Как люди добавляют ценность? Множество неверных предсказаний о компьютерах означает только то, что их рассматривали не с той точки зрения. Традиционный подход к вопросу заключается в определении того, какую работу компьютер никогда не сможет выполнять, – так люди пытаются выяснить, как же им оставаться ценными.

SMT был впервые реализован в конце 1980-х и начале 1990-х годов. В последнее время это способствовало значительному возрождению интереса к машинному переводу и значительным изменениям в восприятии. Перед выпуском онлайн-модуля SMT Google в 2007 году, скептики заявили, что отличного качества машинного перевода не удастся достигнуть на протяжении веков; пять лет спустя они уже говорили 200 лет; в начале этого года – 20 лет.

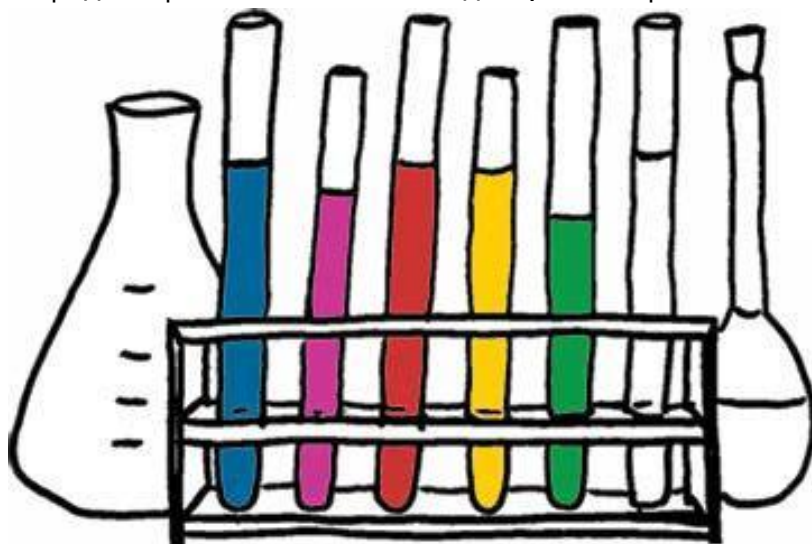
Нынешним успехам предшествовал целый ряд неудач, с которыми в дальнейшем снова придется столкнуться и которые заставят сдаться одних и что-то предпринять других. Проблема заключается в том, что слишком многие пользователи по-прежнему ждут немедленных результатов, особенно если их ожидания завышены, и они готовы инвестировать только тогда, когда поставщики доказывают свою надежность и надежность своих продуктов. К сожалению, большинство пользователей машинного перевода не располагают правильными данными. Это связано с все еще преобладающим нежеланием многих производителей данных защищать и хранить свои данные, которые они рассматривают как закрытый актив, а не продукт для продажи.

В последние годы нейронный машинный перевод все больше совершенствуется, и все чаще его представляют в качестве замены статистическому машинному переводу. Достижения в области искусственных нейронных сетей дают большие надежды на то, что точность нейронного машинного перевода довольно быстро может перегнать точность статистического машинного перевода. Независимые эксперты считают, что переводы с использованием нейронного машинного перевода более гладкие и более точные с точки зрения порядка слов по сравнению с переводами, которые осуществляются с использованием фразового машинного перевода. Они также более точны при создании словоформ. Тем не менее, нейронный машинный перевод менее эффективен при переводе очень длинных предложений. Для исправления используются так называемые «сети внимания».

Машинный перевод лучшего качества может использоваться в различных сферах и для перевода широкого спектра документов. Например, нейронный машинный перевод особенно успешно применяется для перевода более естественного контента, такого, как созданный пользователем текст, который менее структурирован, чем тексты, написанные по установленным правилам. Это говорит о том, что нейронный машинный перевод может быть очень полезен для перевода устной речи, поскольку разговорный ввод, как правило, больше похож на контент, созданный пользователями, и менее похож на профессионально написанные тексты. Это также указывает на то, что нейронный машинный перевод может помочь улучшить качество перевода лекций и других видов подробного разговорного дискурса.

В свою очередь, нейронный машинный перевод будет способствовать дальнейшему расширению технологий преобразования речи в речь (S2S), которые сейчас в основном доступны как одноязычные системы на английском языке. Для того чтобы превратить эти одноязычные системы в многоязычные, потребуется много глубоких и дорогостоящих изменений. Большинство технологий S2S все еще находятся на стадии разработки и их использование ограничено университетскими лабораториями. Нейронный машинный перевод поможет сделать устройства с поддержкой речи широкодоступными.

Поскольку поставщики переводческих услуг уже давно не желают инвестировать в технологии и инновации, они, скорее всего, ждут того момента, когда нейронный машинный перевод станет коммерчески жизнеспособным для того, чтобы ускорить процесс перевода и снизить затраты. Статистический машинный перевод будет все больше и больше развиваться вместе с постредактированием. На самом деле, несмотря на изменения в восприя-



тии, большинство игроков в индустрии перевода все еще не ожидают, что качество машинного перевода может в любой момент достичь уровня перевода, сделанного человеком.

Модели ведения бизнеса в индустрии перевода столько же лет, как и самому переводу, и при этом рентабельность постоянно уменьшается, несмотря на рост спроса. Для многих предприятий объем инвестиций для развития многообещающих технологий слишком велик. Таким образом, машинный перевод приведет к окончательному краху индустрии перевода, так как только премиальный сегмент – художественный и, возможно, медико-биологический перевод останется востребованным.

5.4 Менеджмент качества

Последние 50 лет благодаря космической гонке принесли значительное количество инноваций и изменений, которые позволили достичь развития технологий в геометрической прогрессии. Технологии развиваются именно так, потому что результаты предыдущей стадии развития используются для создания следующей. Приближается тот момент, когда развитие технологий достигнет такого уровня, что они будут влиять на развитие нашей цивилизации. Этот момент называется технологической сингулярностью.

А вот сингулярность в переводе может даже наступить раньше, чем в 2029, и тогда предсказание Рэя Карцвайла не сбудется. После этого момента машины уже не нужно будет обучать, так как они смогут самосовершенствоваться. Пусть и не идеальное, но качество машинного перевода во многом приемлемо уже сегодня. Кроме того, машинный перевод настолько удобен, что пользователи готовы мириться с таким качеством.

Оценка уровня производительности в цепочке поставок и уровня навыков в организации жизненно важно для каждой стратегии инновации и автоматизации. И все же качество перевода остается чем-то расплывчатым и неопределенным, а при его оценке основным фактором является восприятие. Информационная асимметрия еще больше затрудняет всеобъемлющую оценку качества перевода, а учитывая расширение количества языковых пар и увеличение объема перевода, качество стало главной заботой для клиентов.

Из-за неопределенностей, связанных с оценкой качества перевода, а также из-за фиксированности соответствующих концепций в индустрии перевода, пользователи теперь, похоже, готовы пользоваться машинным переводом достаточно хорошего качества, особенно в случае большего объема перевода. Кроме того, машинный перевод доступен в режиме реального времени. Перевод хорошего качества, который не требует постредактирования, компания TAUS назвала полностью автоматизированным полезным переводом уже в 2007 году.

Инвестирование в разработку инструментов поддержки качества перевода стало неотъемлемым для получения уникальных результатов в процессе перевода и для использования машинного перевода с пользой. С увеличением доступности и уменьшением стоимости машинного перевода потенциал данных также распространяется на оценку качества перевода. Аналитические методы и обычные методы оценки перевода используются для оценки качества перевода. Серьезную проблему представляет количество и качество данных перевода, то есть данных проекта. Данные перевода не имеют ничего общего с языковыми данными. Они действительны только в течение срока проекта и в определенный момент времени они устаревают и становятся бесполезными.

Рассмотрим это на конкретном примере. Американский гигант розничной торговли Уолмарт проводит около миллиона транзакций в час и заносит их в базу, рассчитанную на хранение более чем 2,5 петабайт данных. В свою очередь, десять основных игроков переводческой индустрии обрабатывают примерно 3,8 миллиарда слов в год, что в байтах составляет меньше половины объема, обрабатываемого Уолмартом. А это значит, что в год они обрабатывают 0.005% всех слов, переведенных с помощью Google Translate или 1.6% слов, переведенных с использованием сервисов SDL.

Использование машинного обучения для оценки качества перевода с использованием данных приведет к сдвигу в восприятии оценки качества перевода. Оценка качества перевода будет осуществляться не прескриптивно, а дескриптивно, не нормативно, а эмпирически. Для разных типов контента будут использоваться разные характеристики.

Приложения для оценки качества перевода, основанные на обработке данных, смогут сортировать документы и начислять очки за соблюдение стиля. Оценка качества может быть основана на сравнении или же быть автоматической. Эти приложения смогут заниматься выборкой контента или автоматически находить и идентифицировать ошибки. Подход к оценке качества перевода, основанные на обработке данных, потребует качественного сдвига в отношении. Как правило, поставщики переводческих услуг не очень хорошо понимают, зачем нужны корпоративные данные. Они не хотят делиться этими данными, даже тогда, когда это могло бы принести пользу. К тому же, обычно они принимают на работу людей, чьи резюме показывают, что они будут прекрасно работать в стандартных должностях. Для перехода к системе оценки качества перевода, основанной на анализе данных, необходимо развивать эту систему вне самой переводческой индустрии.

5.5. Данные

За последние годы потенциал датафикации стал заметен, и сейчас она идет полным ходом.

Пятьдесят лет спустя все, что описано в рассказе Филипа Киндред Дика «Особое мнение», стало реальностью. Место мутантов заняли машины, которые нашептывают что-то в компьютеры, которые преобразовывают их болтовню в предсказания. Теперь мы больше не пользуемся перфокартами, а обрабатываем большие объемы данных и используем математические, аналитические и прогнозирующие алгоритмы для составления прогнозов погоды и для обеспечения работы полиции. В то же время благодаря массовому внедрению машинного обучения в разных сферах мы получили много новинок, от беспилотных автомобилей до идентификаторов речи и лица.

Данные – это топливо автоматизации. После того, как мы на полном ходу влетели в эру автоматизации, нам пришлось столкнуться с множеством проблем.

Если взглянуть на все это с точки зрения бизнесмена, то каждое задание, которое можно систематизировать и автоматизировать, должно быть систематизировано и автоматизировано. Автоматизация монотонной работы на производстве находится на стадии реализации на протяжении нескольких десятилетий, и ни одной из сфер производства не удастся ее избежать, так как платформы машинного обучения становятся все более доступными. Нам придется столкнуться с парадоксом увеличивающегося в геометрической прогрессии количества продуктов и услуг, которые будут производиться и оказываться со все большей и большей эффективностью. Но в то же самое время безработица и частичная безработица будет только расти, что приведет к падению реальных зарплат и заморозит уровень жизни.

Самым главным вызовом в этой ситуации станет переосмысление экономических моделей для того, чтобы приспособить их для будущего. А так как это будущее приближается все быстрее, нам необходимо действовать как можно скорее.

Переводческие данные, как правило, относятся к метаданным. Это данные о процессе перевода, которые можно собирать вплоть до выдачи готового перевода заказчику. К таким данным относят тип контента, языковую пару, сферу и тему перевода, количество символов/слов/строк, цену, расчетное время завершения проекта, использованные при переводе технологии, переводческую статистику (места в исходном тексте, которые можно переве-

сти с помощью памяти переводов, автоматическая нумерация страниц, машинный перевод – редактируемый или не редактируемый), дата и время последнего сохранения и т.д.

Анализ переводческих данных может помочь понять, как протекает процесс перевода, и найти наилучший ресурс для работы, решить, что переводить и какие технологии использовать для какого типа контента.

Все чаще и чаще переводческие данные генерируются при помощи различных алгоритмов. Все больше данных будет давать оценка персонала и КПЭ (ключевой показатель эффективности). Эти виды данных будут поступать от платформ управления переводом, которые после внедрения машинного обучения не будут требовать участия человека.

Данные, связанные с источниками текстов на разных языках, ошибочно называют переводческими данными. На самом деле, языковые данные состоят из корпусов переводческой памяти, а также лексикографических и терминологических баз. Конечно, все эти источники также содержат метаданные, которые можно было бы использовать.

Оба этих типа данных являются движущей силой процесса автоматизации перевода и их ценность становится стратегически важной. Однако, поиск хороших данных – это нелегкая задача. Прежде всего из-за того, что источники данных неоднородны, кроме того сбор таких данных до сих пор происходит вручную и, наконец, как правило, эти данные плохо формируются и их целостность плохо поддерживается.

Основная проблема, связанная с переводческими данными даже больше чем с языковыми, – это их объем и величина. Они имеют короткую «жизнеспособность» и в определенный момент они становятся устаревшими, неточными и не имеющими особого значения.

Также распространение показателей производительности с индустрии перевода представляется проблематичным, так как поставщики услуг перевода не понимают важности переводческих данных. Участники этого рынка должны стать более открытыми и начать делиться своими переводческими метаданными для того, чтобы сделать их большими данными всей индустрии перевода. Это и есть еще одно новое испытание. Обратная связь становится обязательной для производителей общедоступных данных, и они довольно часто используют повторно.

Наиболее острыми вопросами в этом отношении остаются образование и тренировка, которые становятся все более и более важными для управления автоматически-генерированными данными и для внутреннего и внешнего сопоставительного анализа, который используется для принятия управленческих решений.

Существует большая необходимость в аналитиках, специалистах и ученых, которые занимались бы обработкой данных, но эти должности все еще отсутствуют в индустрии перевода. Поставщики переводческих услуг должны бы искать таких специалистов, которые могут собирать и использовать данные для автоматизации производства. Скорее всего, автоматизация приведет к дальнейшему сокращению количества переводческих компаний, которые остаются конкурентоспособными и процветающими на все более и более конкурентном рынке.

Данные стали новым товаром, однако компании все еще не готовы платить за огромные базы «холодных» данных, для обработки и анализа которых необходимы специфические знания и инструменты, за которые также нужно платить.

С появлением нейронного машинного перевода переводческая индустрия больше не сможет поставлять большие (языковые) данные. С другой стороны, данные переводческих проектов будут необходимы, даже если это малые данные.

Малые данные – это то, что раньше называлось просто данными, и появление больших данных не делает их бесполезными. В повседневной жизни мы в основном сталкиваемся с малыми данными. Определение «малые» означает лишь то, что их объем и формат достаточен для понимания человеком, то есть они информативны, доступны и их можно реализовать. Другими словами, большие данные для компьютеров – это то же самое, что и малые данные для людей. Мы получаем малые данные благодаря результатам анализа больших данных. Человек может воспринимать малые данные и принимать решения, руководствуясь ими. Руфус Полок из организации Open Knowledge утверждает, что вся эта шумиха вокруг больших данных бесполезна. Что действительно представляет интерес, так это данные, связанные с малыми данными. Будущее за распределенными, а не централизованными моделями, нас ждет сотрудничество, а не контроль, и важным аспектом всего этого будут малые, а не большие данные. Если цель анализа больших данных – это обнаружение взаимосвязей, то анализ малых данных выявляет причины этих связей.

В следующие несколько лет основной проблемой может стать стандартизация переводческих данных для того, чтобы сделать их удобными для пользователей и извлечь из них максимальную пользу.

5.6 Функциональная совместимость переводческих инструментов

Совместимость – это возможность двух разных систем работать вместе и взаимодействовать посредством общего языка или интерфейса. Она подразумевает открытые стандарты, а ее разработка не является первостепенной задачей.

Широкомасштабная системная совместимость во всей индустрии необходима для того, чтобы процесс перевода стал практичнее. Совместимость инструментов для перевода и обмен файлами через разные программы, системы управления переводческими проектами и контентом без потери данных остаются одними из основных проблем в переводческой индустрии.

Пока другие индустрии процветают благодаря стандартизации, которая привела к совместимости, автоматизации процессов и инновациям, индустрия перевода всегда страдала от недостатка совместимости. В течение нескольких лет этот недостаток обходился довольно дорого как клиентам (из-за стоимости перевода), так и поставщикам переводческих услуг (из-за размера налогов).

Отсутствие совместимости – это та цена, которую приходится платить за сотрудничество с разными поставщиками и использование разных программ. На самом деле, проблема никогда не заключалась в отсутствии или недостаточной разработанности стандартов. Основным препятствием было и остается отсутствие вышестоящей организации, которая могла бы направить объединить усилия и следить за соблюдением требований.

Многое изменилось с тех пор, как в 2011 году [компания TAUS опубликовала доклад о расходах, возникающих из-за несовместимости в индустрии перевода](#). Многие считают, что отсутствие совместимости связано со стандартами формата обмена, которые препятствуют совместимости. Поэтому никто больше не верит, что совместимость в индустрии перевода вообще может быть достигнута благодаря ознакомительным программам, образованию и сертификации. Стандарты совместимости должны устанавливаться вышестоящей органи-

зацией или объединением, а не лидером на рынке перевода. К сожалению, многие исследования этого вопроса нельзя назвать всесторонними.

Создатели различных переводческих инструментов всегда придерживались стратегии закрытости, а не пытались добиться совместимости. Правда, иногда они обеспечивают переводческие инструменты программными интерфейсами для решения проблем с интеграцией или же поддерживают пользователей, которые пытаются решить эту проблему. И, конечно же, за дополнительную плату.

Распространение машинного перевода привело к прорыву в вопросе совместимости и вызвало множество попыток обновления переводческих инструментов. И все же, большинство этих попыток ограничивается программными интерфейсами. Так как XML годами оставался основным языком разметки, то и форматы TMX и XLIFF считались вполне достаточными.

Пока что большинство программных интерфейсов предназначены для того, чтобы упростить организацию процесса перевода, сократить различные затраты, среди которых расходы на менеджеров проектов. Лишь некоторые программные интерфейсы позволяют отказаться от посредников в работе переводчика и упростить доступ к услугам. Ранее предпринималось несколько попыток решить эту проблему. На самом деле, автоматизация процесса перевода – это один из основных способов улучшения производительности, но почему-то самыми яркими сторонниками повышения эффективности выступают клиенты, а не поставщики переводческих услуг. Опыт показывает, что ни одна из инициатив в переводческой индустрии не будет успешной, пока основные игроки заинтересованы лишь в сохранении и защите существующего положения.

Новый ожидаемый прорыв в области машинного перевода (нейронный машинный перевод) может помочь нерасторопным игрокам переводческой индустрии предугадать процесс освобождения от посредников и договориться о том, какой интерфейс считать стандартным.

В таком случае, мы можем ожидать, что наиболее важные клиенты переводческой индустрии выдвинут свои требования, а разработчики программного обеспечения при наличии технических и денежных ресурсов смогут удовлетворить эти требования или даже предложат новые решения. Такое уже происходило раньше. С другой стороны, клиенты уже столько инвестировали в разработку новых технологий, что сами приложили немало усилий для решения этой проблемы. Кроме того, они немало инвестировали и в поиск решений для совместимости с устаревшими программами.

Во время переноса данных совместимость программ обеспечивается с помощью использования программного интерфейса, который определяет функции, независимые от соответствующей реализации. Как только основные проблемы, связанные с описанием данных, будут решены, количество программных интерфейсов сократится.

На самом деле, стоит быть готовыми к потере данных во время обмена языковыми данными с помощью программных интерфейсов, так стандарты будут отставать от конкретных программных решений, создаваемых разработчиками переводческих инструментов, которые будут стараться избежать каких-либо изменений. Прорыва можно добиться, если уменьшить спектр элементов метаданных, а также стандартизировать их характеристики.

В 2016 году компания TAUS передала все свои разработки по вопросу создания программного интерфейса для открытых интернет-сервисов перевода компании GALA. Компания GALA приняла эстафету и [перезапустила проект под названием TAPICC](#).

5.7 Академия

В книге «Вы не гаджет. Манифест» Джарон Ланье пишет: «Любой навык, как бы тяжело ни было его развить, может стать ненужным с развитием машин.»

Переводческое образование является профессиональным, так как люди готовятся работать как переводчики в соответствующей отрасли. Ни один из навыков, приобретаемых студентом-переводчиком, не является чем-то сложным. Не стоит забывать, что в 1964 печально известный Консультативный комитет по автоматизированной обработке в языке заявил, что средняя зарплата переводчика была значительно ниже, чем зарплата ученых, работающих на правительство, что количество, предлагающих свои услуги переводчиков, значительно превышало спрос, и что существовала необходимость в фундаментальных исследованиях компьютерной лингвистики.

Сельскохозяйственные навыки большинства людей так и оставались на том же уровне, с того момента, когда 12000 лет назад сельское хозяйство только зарождалось, до расцвета промышленной революции в середине 18 века. На западе переход к промышленной экономике и соответствующий сдвиг в иерархии навыков произошел в течение последние 100 лет. Последующий переход к экономике, использующей средства и методы искусственного интеллекта, происходил на протяжении большей части 20 века.

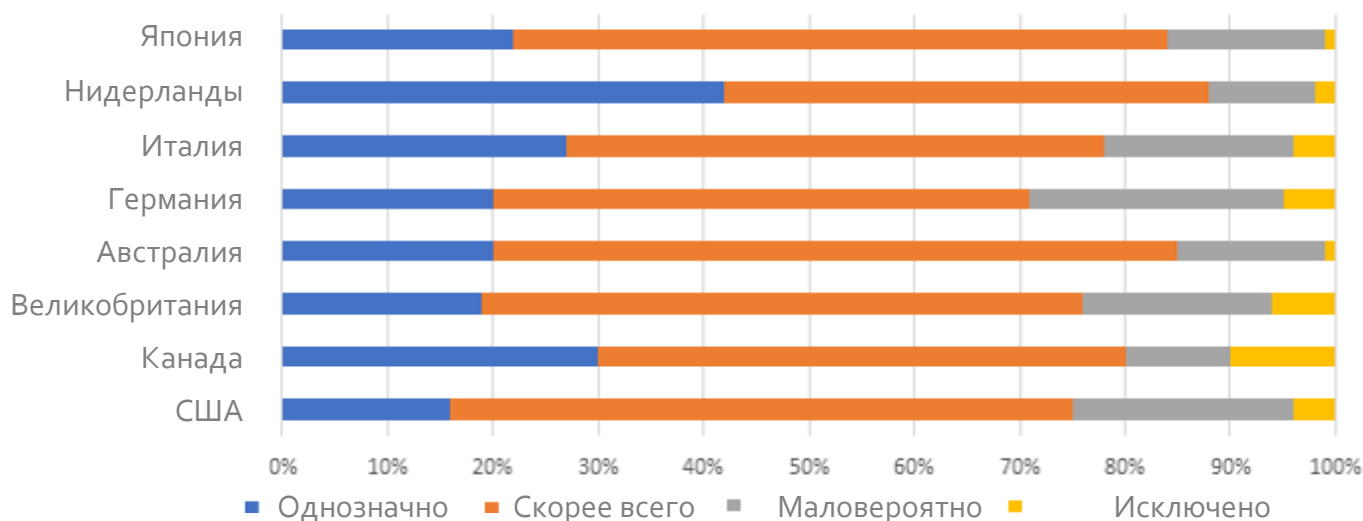
Сегодня многие игроки переводческой индустрии жалуются на недостаток хороших переводчиков, но они игнорируют тот факт, что переводческое образование, больше, чем любое другое академическое образование, строится на устаревших моделях, которые все еще предназначены для 20 века.

Что еще хуже, пропасть между академическим миром и реальной ситуацией в переводческой индустрии настолько велика, что выпускники, начинающие искать работу, быстро понимают, что на самом деле они не так уж и много знают о реальной работе переводчика. Они также узнают о том, что их основные навыки никому не нужны. Только представьте, [что 100 миллионов слов переводятся Google Translate в день](#) и еще миллион слов переводится ежедневно с помощью других переводческих инструментов. Даже все переводчики мира, объединившись, не смогли бы осилить такое количество слов в день.

Система образования никогда не поспевает за новейшими трендами, потому что все притворяются, что инновации жизненно необходимы, несмотря на то, что они дорого обходятся и редко приносят быстрые результаты. Именно поэтому переводческим образованием часто занимаются только государственные учреждения.

С другой стороны, перевод становится предметом обсуждения, только если он плохой. Учитывая вездесущность необходимости перевода из-за глобализации, его воспринимают как обобщенно удовлетворительный. Исследование, [проведенное в мае 2015 переводческим агентством One Hour Translation](#) показало, что большинство пользователей не способно определить, какой перевод является лучшим.

Вы думаете, что сможете отличить перевод,
сделанный человеком от перевода, сделанного
машиной?



Переводчики могут стать ненужными в будущем, по крайней мере, в том качестве, в котором они выступают сейчас, так как аудитория готова принимать перевод все менее и менее хорошего качества.

Высокоавтоматизированная локализованная среда зависит от навыков оценки качества, сопоставления контента, советов относительно культуры, анализа данных, компьютерной лингвистики, и все меньше и меньше от постредактирования. Простой перевод текстов будет все менее востребован.

С другой стороны, согласно недавнему исследованию [отдела карьерного консультирования Университета ДеВри](#) в рамках национальной образовательной конференции SXSWedu в Техасе, в сфере технологий, как это ни печально, существуют пробелы в образовании. Лишь 11% организаций уверены в том, что университеты выпускают специалистов со всеми необходимыми для работы по профессии навыками.

По словам исполнительного директора компании Zuckerberg Media Рэнди Цукерберг, технические навыки нужны не только для программирования, кодирования, разработки и управления; они необходимы практически для любой работы, будь то малый бизнес или маркетинг. Как в сфере предпринимательства, так и в общественном секторе растет спрос на работников, владеющих навыками анализа бизнеса или процедурных проблем и умением решить их с помощью технологий.

Неудивительно, что по словам Криса Вендта, программного менеджера Microsoft по машинного переводу, он бы не рекомендовал своим детям выбирать перевод в качестве профессии, а обратил бы их внимание на работу в качестве консультанта по языку и культуре. По его мнению, транскulturация и адаптация между культурами останутся необходимыми для еще одного поколения, пока различия между культурами нашей планеты не будут сведены к незначительным дельтам, которые машины могут преодолеть.

Никто не может позволить себе роскошь быть слишком довольным собой: технический ландшафт меняется каждые 2-5 лет, и навыки, которыми Вы обладаете сейчас, не обязательно будут полезны через пять лет.

Успех любой инновации зависит от людей, которые будут ее создавать, развивать и применять. Поэтому во времена, когда инновации развиваются семимильными шагами, очень важно, чтобы образование ускорило адаптацию в этих условиях и подготовило следующее поколение работников. Работодатели должны играть важную роль в заполнении пробелов в знаниях, постоянно осуществляя профессиональную подготовку кадров.

Никогда не рано готовиться к будущему. Обширные изменения в трудовых ресурсах организационной структуры необходимы, чтобы кардинально поменять устаревшие бизнес-модели и связанные с ними процессы.

Любое экономическое развитие связано с явлением, названным австро-американским экономистом Шумпетером «созидательным разрушением» («процесс индустриальной мутации, который непрерывно реконструирует экономическую структуру изнутри, разрушая старую структуру и создавая новую»). Глобальные экономические силы невозможно остановить, как и развитие самих технологических инноваций. Разве стоило отказываться от банкоматов, чтобы защитить банковских служащих?

Улучшение подготовки – таков ответ упрощенного подхода. Но при современных темпах изменений усовершенствования образования постоянно проводятся с большим опозданием. Хотя в академических кругах продолжают следовать творческому подходу в обучении переводу, лингвистика требует более научного подхода. Смелый футуристический образ, в котором перевод представлен в фильме «Прибытие», обладает всеми чертами вымысла и имеет мало общего с реальностью.

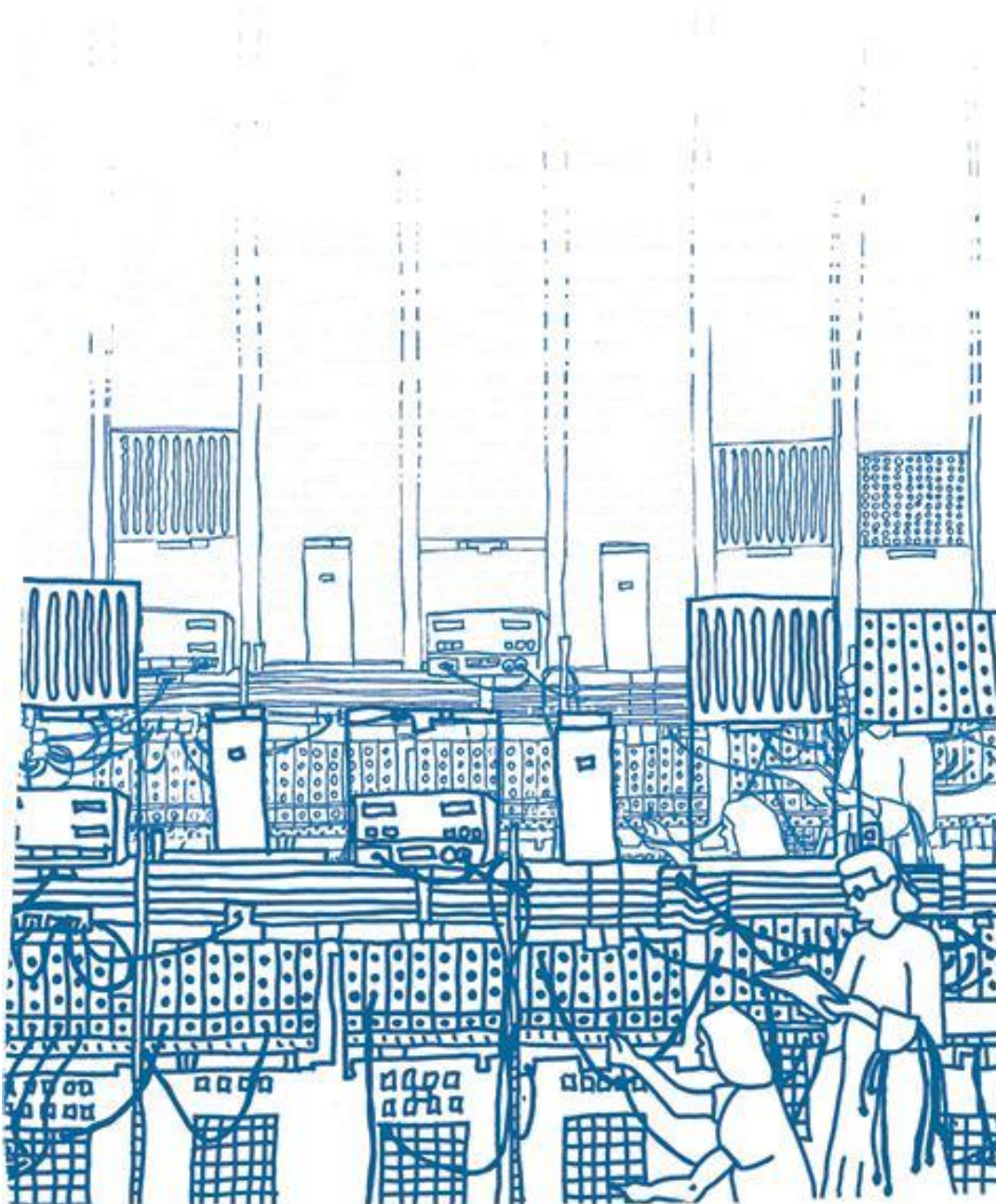
Языковая индустрия должна развиваться не за счет переводов, а скорее за счет появления более широких возможностей, а это осуществимо только за счет появления реальных инноваций. По мере совершенствования технологий зарплаты переводчиков будут становиться все меньше и меньше, а из-за атавистической раздробленности переводческой индустрии, ее развитие все так же будет зависеть от клиентов. Лингвистические исследования могут сделать более значительный вклад в развитие инноваций, чем исследования в области переводов. К тому же, лингвисты, вероятно, получают гораздо больше возможностей, чем переводчики.

В ближайшем будущем основные игроки переводческой индустрии, возможно, станут содействовать развитию новых видов аккредитаций, чтобы повысить авторитет отрасли и увеличить «рабочую трансверсальность», способствующую мобильности работников в разных отраслях: на сегодняшний день менеджера переводческих проектов вряд ли можно поставить на какую-нибудь другую должность.

В своей книге [«Людей недооценивают»](#) Джефф Колвин сообщает о том, что, при проведении исследования в 2013 году, оксфордские экономисты Карл Бенедикт Фрей и Майкл А. Осборн задали работодателям вопрос: какие навыки им больше всего потребуются в ближайшие пять-десять лет. В ответах не были упомянуты деловая проницательность, аналитическое мышление и управление бюджетом, зато были названы умение строить отношения, умение работать и творить в команде, мозговой штурм, культурная чувствительность и умение управлять разнородным коллективом. Это означает, что попытки победить машины в игре по их правилам – бессмысленны. Что делает людей особенными, так это их врожденная склонность к социальному взаимодействию. Именно она больше всего отличает их от логических процессоров. Эмоциональный интеллект всегда будет одним из самых ценных человеческих навыков.

Сегодня языковые услуги стоят на пороге разрушительной трансформации, которая подвергнет переоценке работу профессиональных лингвистов, заставляя их ежедневно совершенствовать свои навыки.

В учебных заведениях при обучении по-прежнему основное внимание уделяется решению



маловажных задач, которые требуют огромного количества времени и энергии. Студентов по-прежнему учат находить решения, им задают вопросы и оценивают по качеству найденных ответов. По мере их продвижения в профессиональной деятельности, все так же будут анализироваться предложенные ими решения, а не заданные вопросы.

6. Десять инновационных сфер

Шесть факторов перемен, которые мы рассмотрели в предыдущей главе, также определяют десять инновационных сфер, которые мы выбрали на ближайшие несколько лет. Благодаря новым достижениям в области машинного обучения, машинного перевода, управления качеством, работы с данными, функциональной совместимости и развития вузов мы можем сделать огромные шаги в нашей отрасли. Новые проекты появляются каждый день и повсюду, знакомя мир с новыми технологиями и новыми способами их использования.

Можно сказать, что между человеком и машиной идет настоящая битва, верх в которой одерживает машина. И если копнуть глубже, на данный момент в переводческой отрасли наблюдается увлекательная интрига: широко обсуждаются революционные технологии и тенденции, а также возможные результаты их развития, которые озадачат даже тех, кто считает, что на сегодняшний день они стоят у руля отрасли.

В 2022 году сегодняшние переводческие компании уже не будут такими, как прежде. Мы станем свидетелями раскола между технологиями перевода и творческими сетями, фабриками данных и сторителлингом, платформами и специализированными компаниями, которые, вероятно, еще будут изредка работать в общих рамках, но уже с четко разграниченными функциями. Знакомая история? Пожалуй, на ум приходит идея о смене парадигмы в индустрии рекламы и маркетинга. Когда-то такое решение считалось очень изобретательным, и имело свое уникальное место в сфере автоматизации и делопроизводства. Но теперь, спустя несколько десятилетий информационных бурь, деятельность престижных рекламных агентств существенно поменялась.

Теперь маркетинг автоматизирован, основан на базах данных и управляется щелчком мышки. Невероятно быстрый рост интернет-рекламы, четкой маркетинговой политики и сферы платных трансляций в таких компаниях, как Google и Facebook, серьезно изменил текущую обстановку. Такие легендарные имена, как Саатчи и Саатчи, Макканн Эриксон, Дж. Уолтер Томпсон, вызывают ностальгические воспоминания о временах сериала Безумцы (Mad Men – сериал о рекламном агентстве), но теперь креативные директора обо всем отчитываются перед гигантскими холдинговыми компаниями, действующими под такими банальными именами, как Omnicom, WPP, Interpublic и Publicis.

В ближайшие пять лет подобные слияния и поглощения, вероятно, произойдут и в относительно небольшой переводческой отрасли, и сближение с другим творческим сектором, павшим жертвой информационных бурь – индустрией рекламы и маркетинга, будет иметь большой смысл.

6.1. Алгоритмическое управление

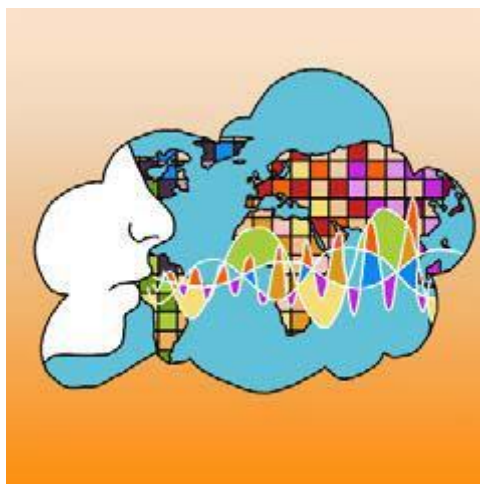
Машины превзойдут людей практически во всех видах деятельности. Ожидается, что, как и во множестве других отраслей, роботы появятся в переводческой индустрии, чтобы повысить качество нашей работы, расширить ее и в конечном итоге заменить нас. «Восстание машин» началось очень просто: с подсчета слов исходного документа, который нужно перевести. Теперь оно разрослось до поиска подходящих для переводчика заказов, определения новой терминологии, оптимизации использования памяти переводов, профилирования контента, построения выборки для проверки, определения типа оценки качества и типов ошибок, которые необходимо проверить, отслеживания состояния и отчетности, выставления счетов и доставки.



Скоро роботы будут проверять качество и производительность труда переводчика, и даже просчитывать качество выполнения предстоящих заданий. Они будут отслеживать окупаемость инвестиций (ROI) при переводе каждого отдельного сообщения или сегмента в зависимости от того, сколько пользователей просмотрели перевод. И это еще даже не говоря об осуществлении самих переводов, переводе устной речи, проведении оплаты по факту выполнения заказа и остальных инновационных технологиях, использующих алгоритмы. Эти технологии описаны в других мини-главах этой статьи. Уже известно, что многие стартапы в переводческой отрасли позиционируют себя как суперпереводчики. Так что вот, пожалуй: автоматический перевод будет нормой в 2022 году.

6.2. Датафикация перевода

Датафикация перевода началась со статьи [«Необоснованная эффективность данных»](#), написанной учеными Google Фернандо Перейро, Питером Норвигом и Алоном Халеви в 2009 году или, возможно, даже раньше, когда в 2008 году было запущено облачное хранилище данных TAUS. Перевод основывается на информации из баз данных. На начальном этапе базы данных фактически функционировали согласно принципу «чем больше, тем лучше». Система машинного англо-французского перевода Google использовала в работе библиотеку из 100 миллиардов слов. Теперь, с появлением нового поколения нейронного



[машинного перевода, огромные базы данных стали пережитком прошлого. Погоня за высококачественными переводческими базами данных бросит вызов системам их охраны и открывает возможности для интернет-пиратства.](#)

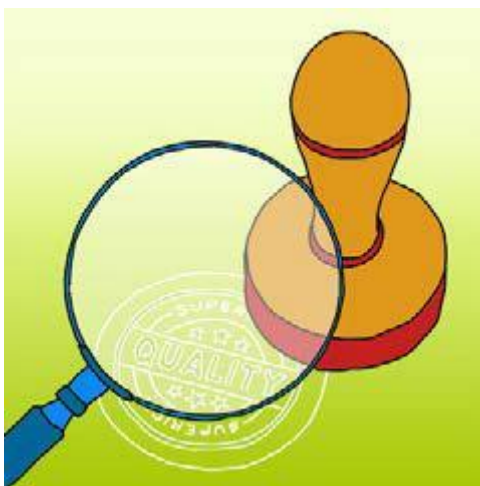
Все просто помешались на данных, по крайней мере, в сфере перевода. И это касается не только данных памяти переводов. Речевые данные тоже очень важны. Еще нужны правки и аннотации как к машинным, так и человеческим переводам, а также атрибуты типов контента, отраслевых секторов, местоположений переводчиков, применяемых методов и тех-

нологий. И почему бы не сопоставить все это с прогнозами погоды, социальными графами переводчиков и с технологией отслеживания их взгляда? Работая с новыми данными, всегда можно чему-то научиться.

У интернет-гигантов было конкурентное преимущество в сфере переводческих данных, но они потеряли его, загрязнив собственные рыболовные угодья машинным переводом. Теперь открыта охота на новые рынки данных. Европейская комиссия уже инвестирует средства в Фонд по соединению Европы. Однако также следует обратить особое внимание на очень молодые китайские венчурные компании, работающие с переводческими данными. А можно взглянуть еще ближе к США: рынок данных TAUS.

6.3 Вопросы качества

Не стоит позволять данным и технологиям дурачить нас. Повсеместная доступность немногого перевода подогреет лишь спрос на контент и идеи, которые стимулируют воображение пользователя, помогают создать клиенту собственный имидж и привлекают глобальные и разнородные сообщества. В данном случае синергия между переводческой и рекламной и маркетинговой отраслями имеет наибольший смысл. В условиях глобализации открытая экономика нуждается в обновлении контента. В сети достаточно некачественного контента, который генерируется и переводится автоматически.



Данные служат прекрасным подспорьем для понимания рынков и клиентов: что им нравится и не нравится, куда щелкают мышью, а куда – нет. Это понимание позволяет принимать обоснованные решения, куда инвестировать, а куда нет, когда дело доходит до создания контента и творческого перевода. Будущее переводчиков представляется таким: они станут писателями, журналистами и рассказчиками, консультантами по культуре, промоутерами мировых брендов.

6.4. Конвергенция

Конвергенция – это слияние технологий, бизнес-моделей или рынков. И происходит она, когда на их пересечении появляются крупные инновации. До сих пор лучшим примером этого является успех Google Translate, который послужил серьезным тревожным звонком для переводческой отрасли. Google не собирался нарушать работу переводческой индустрии и даже выразил обеспокоенность по этому поводу. Однако огромная популярность кнопки автоматического перевода в сочетании с Поиском Google открыла перспективы, о которых некоторые первопроходцы машинного перевода мечтали еще в 1970-х годах ([см. статью о том, как Жан Гашо предложил использовать Systran Translate на базе системы Minitel на конференции в Париже](#)).

В течение следующих пяти лет конвергенция станет двигателем самых интересных и оригинальных инноваций. Японские мегафон-переводчик Megaphonyaku и носимый гаджет-переводчик Wearable Translator служат хорошими примерами того, как конвергенция машинного перевода и речевых технологий может помочь туристам и путешественникам справиться с иностранной языковой средой. Олимпиада 2020, которая пройдет в Токио, скорее всего, принесет еще более удивительные инновации в переводе. И совсем несложно представить подобные инновации в деловой и промышленной сфере. А как насчет автоматического перевода, синтеза речи (из текста) и видеоконференции, одновременно транслируемых в очках смешанной реальности Microsoft Hololens, чтобы обеспечить вьетнамского эксплуатационного инженера техники фирмы John Deere прямой связью с заводом в Германии если он обнаружит дефект в тракторе? И аналогичным образом, не будет ли китайской лаборантке удобнее работать с аппаратом для анализа крови фирмы Roche Diagnostics, когда устройство зачитывает ей инструкцию по эксплуатации и отвечает на ее вопросы на китайском языке?

Такая конвергенция будет особенно эффективна в тех ситуациях, когда делаются переводы, которые никто не читает, потому что они никогда не оказываются в нужное время в нужном месте.

6.5 Речь

Согласно прогнозам, сделанным в недавнем докладе компании [TAUS по технологиям синхронного перевода](#), ожидается, что качественные рабочие решения по технологии синхронного перевода появятся в течение ближайших пяти лет. Переводчик Skype и японские носимые устройства, упомянутые выше – это только начало. Люди ленивы и в целом предпочитают слушать и говорить, а не читать и писать. Технология есть, и она работает.



Сейчас основная задача – поиск голосов на требуемых языках, а также обработка данных и автоматизация систем синхронного перевода. Это еще одно новое поле деятельности поставщиков языковых услуг для расширения и развития новых услуг, поиска и найма талант-

ливых работников. Устный перевод появится во многих приложениях, в наручных браслетах, очках. Его будут встраивать в продукцию производителей программного обеспечения, автомобилей, медицинских устройств и онлайн-услуг.

Чтобы поднять уровень эффективности синхронного перевода до уровня текстового перевода, компаниям, разрабатывающим технологии и предоставляющим услуги в переводческой отрасли, придется разработать новые методы обработки данных, рабочие процессы, наборы данных и инструменты. Таким образом, открываются перспективы для новых специализаций.

6.6 Короткие и длинные переводы

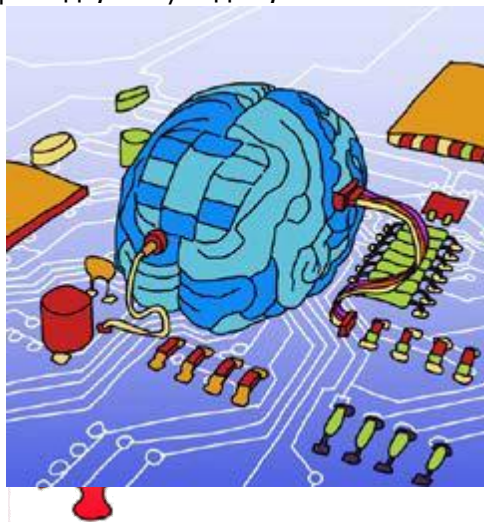
В течение ближайших пяти лет станет отчетливее разграничительная линия, между короткими и длинными переводами, между локализацией продукта и творческой адаптацией перевода. Локализация продукта имеет свои особенности. Прошли те времена, когда продукты выпускались крупной партией один раз в год. Сейчас продукты выпускаются постоянно. Объем задач и заказов становится все меньше и меньше, вплоть до единичного количества. Концепция проекта становится размытой. Методы обработки данных становятся более гибкими и уже применяются на этапе разработки продукта.

В этом помогает переход на качественно новый уровень в сфере машинного перевода. Но вне зависимости от того, используется машинный перевод или нет, перевод осуществляется постоянно, практически в режиме реального времени. По мнению поставщиков услуг и технологий, интеграция платформ неизбежна. За последние несколько лет на рынке появился ряд инновационных облачных платформ с функцией перетаскивания перевода с разным уровнем качества. Такие платформы изменяют существующее положение вещей, и многим поставщикам и покупателям придется ими пользоваться.

Но что, если от человека потребуется нечто большее, чем просто клик мыши, например, когда контент нужно обновить или адаптировать для другого языка и другого сообщества? Переводчики становятся писателями и культурными консультантами, промоутерами торговых марок. Они становятся решающим фактором успеха продукта на новом рынке. Многоступенчатые цепи производства и обработки продуктов уже испытывают трудности, и отказ от посредничества снова станет актуальной темой. Новые предприятия, такие как «Translate and Create», бросят вызов признанным переводческим компаниям. Естественным выходом из положения могло бы стать слияние переводческих услуг с издательскими, рекламными и маркетинговыми. Будет любопытно взглянуть, как станет развиваться рынок в условиях увеличивающегося разрыва между автоматическим и творческим переводами.

6.7. Стремительный рост

В 2022 году, оглядываясь на ту марафонскую дистанцию, которую преодолела технология Нейронного Машинного Перевода, мы увидим, насколько высоким стало качество машин-



ного перевода благодаря ей. Если представить этот скачок в процентном соотношении, то, по мнению экспертов, прогресс, который будет достигнут за пять лет использования Нейронного машинного перевода, будет равен прогрессу, достигнутому за предыдущие двадцать лет применения Статистического машинного перевода. Это звучит пугающе, потому что зачастую даже исследователи не понимают, что вызвало такой быстрый прогресс. Машины превратились в самообучающиеся системы и стали находить тысячи решений, чтобы добиться лучших результатов.

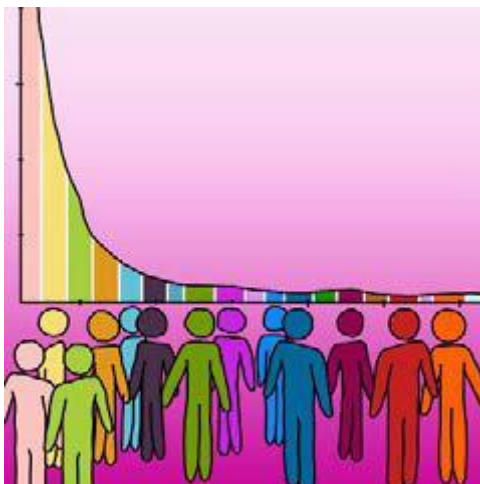
Более того, технология Deep Learning применяет любые виды данных – одноязычные, двуязычные, звуковые и даже видеоролики, для создания универсальных движков, которые могут читать (и переводить) по губам и даже переводить между языками, для которых нет напрямую доступных двуязычных данных. Сейчас машинные переводы осуществляются быстрее и выглядят более естественно, чем те, которые мы привыкли видеть до недавнего времени, так как машины скрывают потенциальные неточности перевода, на которые мы не обращаем внимания. Технология нейронного машинного перевода положительно повлияла на развитие новых приложений для синхронного перевода благодаря увеличению плавности речи и уменьшения загрузки памяти.

Все это приведет к тому, что весь публикуемый контент будет по умолчанию переведен как минимум на пятьдесят языков или, по крайней мере, практически переведен, то есть он будет доступен по запросу, в режиме реального времени и, возможно, бесплатно. Это гигантский шаг вперед от современной реальности, хотя на сегодняшний день (2017) общий объем машинных переводов уже в 500 раз превысил объем работы, выполняемой всеми вместе взятыми людьми-переводчиками.

По слухам, успех нейронного машинного перевода сопровождается новым прорывом в технологиях машинного перевода, который некоторые уже называют квантовым машинным переводом. Поколение квантового машинного перевода может повысить точность машинного перевода до уровня, необходимого для того, чтобы устранить недостаток точности в современных системах нейронного машинного перевода.

6.8 Особые заказы

В 2022 году индустрия перевода будет гораздо менее ограничена в отношении распространения языков. Непрерывная глобализация продолжит расширять рынки. Популистские тенденции в политике вряд ли остановят стремление предприятий по всему миру к большему количеству клиентов, как из стран Запада, так и из Азии. Фактически мы увидим, с какой легкостью электронная коммерция привлекает клиентов, в том числе к малым и средним компаниям, стимулируя дальнейший рост мировой торговли.



Например, в рамках программы «One-Belt One-Road» в Китае основное внимание уделяется странам, расположенным вдоль исторического Шелкового пути, добавив к нему множество языков, активно задействованных в переводческой отрасли. На сегодняшний день глобальное предприятие охватывает в среднем 25 языков, можно предположить, что это число удвоится в ближайшие пять лет. Повсеместная доступность машинного перевода может в значительной степени восполнить этот пробел.

Новое поколение систем Neural MT не только предлагает более качественный перевод, но и ускоряет добавление новых языков, поскольку в меньшей степени зависит от количества языковых пар. Так называемый подход с нулевым ударом для построения механизмов машинного перевода позволяет разработчикам создавать двигатели из непарных данных различных, но родственных языков. Как ни странно, мы считаем, что переводчики не потеряют свою работу, т.к. повсеместная доступность машинного перевода на новых языках будет стимулировать спрос на творческие объемные переводы. Благодаря технологиям будет возрастать интерес к другим культурам и языкам.

6.9 Плата за пользование

Повсеместная доступность переводческих услуг, часто бесплатных, но не всегда адекватных или достаточно хороших, запускает новые модели бизнеса и ценообразования. Почему бы не взимать плату лишь за то, что пользователь нажимает на кнопку мыши и получает перевод? Почему бы не дифференцировать цены в зависимости от вида контента, количества кликов мыши? Или, возможно, стоит ввести метрику, которая поможет определить, когда необходим качественный человеческий перевод.



Компании, предлагающие языковые услуги, привыкли к развитию своего бизнеса и обновлению своих профильных услуг, что и будут продолжать делать в следующие пять лет, только быстрее и радикальнее. Фиксированная цена за перевод слова исчезнет. Переводческие фирмы будут взимать плату за использование своих платформ. Креативные бюро переводов введут почасовую оплату в зависимости от требуемых знаний.

6.10 Ориентирование на пользователя

В частности, в индустрии перевода мы слишком часто забываем о том, зачем мы переводим, для кого мы переводим, и как наш перевод будет использоваться. Слишком часто переводы выполняются в качестве обязательного элемента в старомодных моделях с активным источником данных или для публикаций, и то, как они будут использоваться, во внимание не принимается. Это должно измениться, если мы будем следовать тенденциям к более демократическим и ориентированным на пользователя бизнес-моделям. Поставщики услуг в секторе перевода, вероятно, будут играть более важную роль в консультировании по вопросам использования перевода и культурного разнообразия. В самом деле, привлечение пользователей к определению функций, оценке перевода и набора услуг и продуктов на

новых рынках может стать огромным преимуществом. У пользователей появится новая роль.

6.11 Эпилог

В 2022 году, оглядываясь на историю переводческой индустрии за последние пять лет, мы не увидим такого плавного путешествия по десяти главам, выделенным в этой статье. Конечно, нет. Будут небольшие загвоздки и крупные сбои, испытания и ошибки, а также жесткая конкуренция. Поставщики в индустрии переводов продолжают тратить чрезмерную долю дохода на старомодные продажи. Промышленность будет страдать от дилеммы Бодо*: изобилие инструментов, технологий, данных и инновационных решений в сочетании с острой нехваткой талантов.

Данные становятся все более важными, в то время, как пределы развития технологии уменьшаются в результате более открытых исходных решений и обмена последними достижениями через обучающие программы. С одной стороны, основные тенденции в области передачи данных, совместного использования данных, машинного обучения и облачных технологий приведут к инновациям, росту и развитию, но с другой стороны, вызовут озабоченность по поводу конфиденциальности и безопасности. Мы можем рассматривать блокировку перевода как явление, подобное добавлению блокировки для защиты компаний и сетей поставок от нежелательных и непрошенных бесплатных переводов.

Некоторые провайдеры будут преуспевать, предлагая безопасные и закрытые услуги перевода и технологии. Но, в конечном счете, в следующие пять лет, возможно, облако окажет-



ся непреодолимым даже для самых параноидальных, особенно когда новая технология квантовых компьютеров обеспечит максимальную безопасность. Кто знает?

*Дилемма Бодо названа в честь Bodo Vahldieck, менеджера по качеству в VMware, который на отраслевом саммите TAUS выразил обеспокоенность тем, что не может найти молодые таланты, которые хотят и могут работать с инновационными технологиями в сфере перевода в его компании.

7. Больше информации по ссылкам:

Доклады и статьи TAUS

<https://www.taus.net/think-tank/reports/translate-reports/taus-speech-to-speech-translation-technology-report>

<https://www.taus.net/think-tank/reports/translate-reports/taus-machine-translation-market-report-2017>

<https://www.taus.net/think-tank/reports/translate-reports/taus-translation-technology-landscape-report-2016>

<https://www.taus.net/think-tank/reports/translate-reports/lack-of-interoperability-costs-the-translation-industry-a-fortune>

<https://www.taus.net/blog/is-there-a-market-for-mt>

Внешние источники:

<http://www.nytimes.com/2010/01/15/books/15book.html?pagewanted=all&mcubz=o>

<http://ieeexplore.ieee.org/abstract/document/4804817/>

<https://www.amazon.com/Humans-Are-Underrated-Achievers-Brilliant-ebook/dp/BooOZoTLBK>

<http://www.mckinsey.com/~media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx>

http://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

8. Об авторах

Изабелла Массардо

Изабелла Массардо работала в сфере перевода с 1988 года в качестве переводчика, преподавателя и писателя. Она регулярно пишет статьи на тему бизнеса перевода, уделяя особое внимание технологиям и инновациям. Ее статьи включают маркетинговые материалы для целого ряда глобальных клиентов, от поставщиков языковых услуг до технологических компаний. Изабелла является сертифицированным ECQA менеджером по терминологии. Получила степень магистра по русскому языку и литературе в Университете Пармы, Италия.

Яп ван дер Меер

Яп ван дер Меер был основателем и генеральным директором ряда крупнейших глобальных компаний по переводу и локализации услуг в 1980-х и 1990-х годах. В 2005 году он основал Общество сторонников автоматизации перевода (TAUS). TAUS – это инновационный аналитический центр и платформа для отраслевых услуг для глобального сектора перевода и локализации. Многие крупнейшие IT-компании, государственные отделы по переводу и их поставщики переводческих услуг, технологий и локализации являются членами TAUS. TAUS предлагает, среди прочего, платформу для оценки качества перевода и проведение сопоставительных оценочных испытаний, а также платформу для объединения и обмена данными памяти переводов. За многие годы работы в переводческой индустрии Яп ван дер Меер написал о ней множество статей.

Рецензент

Анне-Май ван дер Меер

Обозреватель

О TAUS

TAUS, сеть языковых данных, является независимой организацией. Мы развиваем сообщества посредством реализации программ мероприятий и организации онлайн-групп пользователей, делясь знаниями, метриками и базами данных, которые помогут всем участникам переводческой отрасли повышать качество своих услуг. Мы предоставляем услуги передачи данных покупателям и поставщикам языковых и переводческих услуг.

Обмен знаниями и данными помогают членам TAUS разрабатывать эффективные стратегии локализации продуктов. Метрики помогают осуществлять более эффективную обработку данных и стандартизацию оценки качества. Базы данных повышают уровень автоматизации перевода.

TAUS разрабатывает программные интерфейсы приложений, которые предоставляют участникам сообществ доступ к таким сервисам, как DQF, Quality Dashboard и TAUS Data Market через их собственные переводческие платформы и инструменты. Метрики и базы данных TAUS уже используются в большинстве крупных переводческих технологий.