

Профессиональный

ПЕРЕВОД

и управление информацией

ПРОФЕССИОНАЛЬНЫЙ ПЕРЕВОД И УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ

Управление

TermWiki: управлять терминологией стало проще

Комментарии специалистов

Повторите, пожалуйста, или DITA для начинающих

Семь стратегий для судебных лингвистов

Вокруг света

Многоязычная Европа: проблемы языковых технологий

№ 6 (53) июнь 2011

**ПРОФЕССИОНАЛЬНЫЙ ПЕРЕВОД
И УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ**

Журнал распространяется по подписке и в розничную торговлю не поступает.

Подписной индекс по каталогу Агентства «РОСПЕЧАТЬ» — 36111
Интернет-подписка — www.ITbook.ru

6

июнь 2011

Содержание

УПРАВЛЕНИЕ

- 3** Язык, технологии и коммуникационный процесс *Джин Шрайвер*
- 8** Локализация программного обеспечения для техники *Франсуа Массьон*
- 16** TermWiki: управлять терминологией стало проще *Уве Мюге*
- 21** Локализация программного обеспечения, веб-сайтов
и международных приложений *Амир Хельцер*

КОММЕНТАРИИ СПЕЦИАЛИСТОВ

- 27** Законы культуры *Кэйт Эдвардс*
- 30** Турботестирование машинного перевода в компании Cisco *Лори Тике*
- 36** Повторите, пожалуйста, или DITA для начинающих *Тони Селф*
- 41** Семь стратегий для судебных лингвистов *Дена Бугель-Шунра*

ВОКРУГ СВЕТА

- 46** Многоязычная Европа: проблемы языковых технологий
Джордж Рехм и Ганс Ускорайт

ИЗ-ПОД ПЕРА МАСТЕРА

- 50** Происхождение (отрывок) *Диана Абу-Джабер*

Главный редактор Сергей Гладков
Зам. гл. ред. по производству Нана Чатынян
Научный редактор Маргарита Бажина,
margaritaba@logrus.kiev.ua
Дизайн и верстка Елена Козлова
Менеджер по подписке Ольга Астахова, itj@rusedit.com

Размещение рекламы reclama@rusedit.com
Журнал «Профессиональный перевод и управление информацией» —
совместный проект издательства «Русская Редакция» и компании «Логрус».

 **РУССКАЯ РЕДАКЦИЯ**

**ООО «Издательство
«Русская редакция»**

Адрес: 123298, Москва,
3-я Хорошевская ул., 11
Телефон: (499) 197-0422
E-mail: info@rusedit.com
Веб-сайт: www.rusedit.com



Компания «Логрус»

Адрес: 115114, Москва,
Дербеневская, д. 20, стр. 16, этаж 3
Телефон: (495) 646-3563
Факс: (495) 646-3562
E-mail: management@logrus.ru
Веб-сайт: www.logrus.ru

Журнал «Профессиональный перевод и управление информацией» содержит материалы из оригинальных изданий MultiLingual Computing & Technology и Tcworld, переведенные на русский язык и опубликованные с разрешения Multilingual Computing, Inc. и Tekom.

Подробнее о журнале см. на веб-сайте www.profitran.ru.



Журнал зарегистрирован в Федеральной службе по надзору за соблюдением законодательства в сфере массовых коммуникаций и охране культурного наследия.

Свидетельство ПИ № ФС77-33439 от 10.10.08

Тираж 2000 экз. Формат 60x90/16. Объем 3,75 усл.-печ. л.

Журнал выходит ежемесячно. Распространяется по подписке.

Подписной индекс Агентства «Пресса России» — 82558.

Интернет-подписка — www.ITbook.ru, телефон (499) 197-04-22, e-mail: ITJ@mail.ru

© Издательство «Русская Редакция», 2011

© Компания «Логрус», 2011

Полное или частичное воспроизведение и распространение материалов, содержащихся в настоящем издании, допускается только с письменного разрешения издателя.

Язык, технологии и коммуникационный процесс

Джин Шрайвер (Gene Schriver)

Джин Шрайвер размышляет об изменении роли технологий и коммуникации. Интегрирование развивающихся технологий и областей их применения в бизнес повысит, по мнению автора, скорость и уровень работы, а также расширит сферу деятельности.

Сегодня можно общаться практически с любым человеком на планете. Можно говорить по телефону, увидеться в скайпе, написать сообщение с помощью электронной почты, телефона, планшетного компьютера или даже видеоигры. В большинстве случаев это бесплатно.

Можно продавать и приобретать таким образом множество товаров и услуг. Так давайте займемся бизнесом! Подождите-ка. Эти люди говорят на других языках. Ладно, вроде пара моих знакомых знает испанский, а может и еще какие-то языки. Интересно, у них найдется время? Можно и в бюро переводов позвонить. Но это, наверное, дорого. И не быстро. А мне нужна помощь. Сейчас.

В нынешний информационный век мы можем обмениваться информацией быстрее, чем когда-либо. Тем не менее, наличие постоянной связи и обмен ин-

формацией — не одно и то же. На самом деле, чем больше возможностей для общения, тем очевиднее потребность в комплексной, качественной языковой поддержке. Технологии могут нас сблизить и делать мир меньше, но для объединения людей нужно общение: между людьми, между людьми и компаниями, между компаниями.

Я много лет трудился над тем, чтобы превратить маленький семейный бизнес в одну из крупнейших в мире компаний, предоставляющих языковые услуги. Новая глобальная экономика только развивалась, и услуги в нашей отрасли по большей части не успевали по скорости, масштабу и доступности за теми изменениями в сфере коммуникаций, которые произошли в связи с развитием телефонии и Интернета. Отчасти целью основания GLOBO было удовлетворение растущих потребностей за счет использования технологий, привлечения талантливых лингвистов и создания мгновенного и надежного доступа на облачной платформе. Но не так важно, предоставляются ли специальные или стандартные языковые услуги, — от применения технологий зависит конкурентное преимущество.

Чем больше работы в облаках, тем больше шансов у бизнеса воспарить над ними

Технологии совместной работы существенно изменили наш подход к деятельности поставщика языковых услуг. Наша отрасль, как и многие другие, построена на запатентованных технологиях и системах, служивших лишь нескольким избранным игрокам, у которых было достаточно финансовых и человеческих ресурсов для инвестирования в такие системы. Сегодня дела обстоят иначе. Технологии совместной работы существенно уравнили шансы крупных и малых игроков. Небольшие компании сейчас лучше готовы к конкуренции, чем раньше. По сути, в настоящий момент мелкие поставщики языковых услуг получили небольшой запас времени, который позволяет обогнать крупные компании за счет построения усовершенствованных рабочих процессов с помощью решений, созданных и обслуживаемых третьими сторонами. Специализированные серверные и облачные системы становятся все более доступными практически для любой организации. Как объясняет Крисси Дэниелсон (Krissi Danielson), благодаря таким системам «детали скрываются от пользователей — им больше не нужно обладать специальными знаниями об инфраструктуре облака, на которой все построено, или управлять ею» (Danielson, 2008).

Например, когда-то решающую роль в устном переводе по телефону играло оборудование: серверные, коммуникаторы, море проводов. Инвестиции в это оборудование требовали значительных

капиталовложений, нужно было также организовать круглосуточное дежурство специалистов. Кроме того, помещения для оборудования — это тоже источник беспокойств: возникают риски прерывания и остановки работы, устаревания аппаратуры, а также необходимость разбираться в оборудовании нового поколения и вкладывать в него средства.

Сегодня в этом отношении мы обладаем преимуществами, мы можем применять такие технологические подходы, каких не было десять, пять лет и даже два года назад. Услуги устного перевода по телефону предоставляются частично или полностью на платформе третьей стороны, у нас есть время и возможность сосредоточиться на самих услугах, а не на серверном оборудовании и мини-АТС. Такая модель обеспечивает доступ к передовым технологиям без дополнительных издержек и труда. Услуги связи предоставляют эксперты, набрать специалистов такого уровня в штат было бы просто невозможно. Мы можем со временем переходить на более совершенное оборудование практически или совсем без проблем, и это не повлияет на текущий рабочий процесс. Безопасность и надежность сегодня тоже лучше, чем когда-либо, поэтому мне все реже звонят среди ночи по экстренным вопросам. Важнее всего, вероятно, то, что мы можем работать быстрее, наши услуги становятся более гибкими и масштабируемыми, наше влияние расширяется.

Тот же принцип использования систем сторонних поставщиков применим и к другим большим областям производства, например платформам управления проектами по переводу. С одной стороны,

многие компании очень гордятся «запатентованными» или «самостоятельно разработанными» системами. Такие системы, возможно, и великолепны, но вообще-то большинство поставщиков языковых услуг — не технологические компании. Даже при наличии в штате высококлассных ИТ-специалистов внедряемые технологии оказываются ограниченными по сравнению с альтернативными вариантами. Я часто слышу, как люди, рекомендуя собственные технологические системы компании, одновременно жалуются, что они застряли на каком-то инструменте, в который вложено много средств, а технологии вокруг продолжают развиваться. Сегодня необходимо работать по моделям, не привязанным к прошлому. В определенном смысле компания-новички находятся в более выгодном положении, поскольку часто начинают внедрять собственные стратегии, применяя чью-то проверенную систему.

Прежнего мира больше не будет — вот и чудесно

Технологии влияют не только на методы работы поставщиков языковых услуг, но и на род деятельности. В лингвистическом сообществе постоянным источником беспокойств является ценность перевода, выполненного человеком, в сравнении с машинным переводом. Спор о человеке и машине не прекращается годами. Многие спрашивают: заменит ли когда-нибудь компьютерный перевод специалистов, работающих в нашей отрасли?

На мой взгляд, никогда. Машинному переводу отводится в нашей отрасли

очень важное место: он создает рабочие места, а не отбирает их. Языковые услуги могут быть недостижимыми для некоторых компаний из-за проблем, связанных со стоимостью, оборудованием и доступностью. Машинный перевод используют компании, которые в противном случае не смогли бы себе позволить переводить материалы, а еще он помогает поддерживать отрасль переводов. Кроме того, он повышает скорость и эффективность работы в отрасли, в которой требуется щепетильное отношение к деталям, этот факт еще больше повышает ценность услуг отрасли. Хотя традиционное представление о работе переводчика, вероятно, просуществует довольно долго, наибольшие возможности роста открываются перед теми лингвистами, которые приспосабливаются к новым технологиям, идущим в ногу с современными методами и темпами создания контента.

В моей компании используется несколько рабочих процессов, в которых объединены знания и умения людей и машин. В каждом из них может быть свое сочетание и своя последовательность действий переводчиков и компьютеров. Например, мы ввели новую методику изучения высказываний о компаниях и брендах в Интернете во всем мире. Для выполнения такой работы необходим компьютер.

Можно считать огромные объемы текста, создаваемого в чатах и других онлайн-средствах общения. Перевод всех комментариев сделал бы задачу невыполнимой из-за временных и финансовых затрат. Но для роста и поддержки международного бизнеса крайне важно узнать мнение клиента. Наши переводчики могут со-

средоточиться на конкретных абзацах или комментариях и на часто встречающихся словах или понятиях. Проведя анализ, мы предоставляем общие и конкретные сведения, касающиеся репутации компании с точки зрения ключевых заинтересованных лиц. Мы можем помочь клиенту провести анализ. Такой подход — не массовое решение, а скорее метод, адаптируемый под каждый конкретный случай. Предприятия всего мира начинают брать его на вооружение.

Поскольку мы стремимся обезопасить свой бизнес, выгодно использовать технологии, мы можем защититься от угроз, создаваемых технологиями и перенасыщением рынка, предлагая услуги, устойчивые к обоим факторам. Например, наша компания устраняет разрыв между маркетингом (рекламой) и услугами по переводу. Часто приходится встречать клиентов, для которых поставщик услуг перевода и рекламное агентство выполняют одну и ту же работу. Обоим исполнителям не хватает компетенции в сфере специализации другого исполнителя, и результаты их труда сильно отличаются. Невозможность рассчитывать на качество подобных услуг создает слишком большой риск для любого бренда. Поэтому в идеале нужно обращаться к поставщику языковых услуг и рекламному агентству, а затем использовать лучшие результаты обоих. Такой гибридный бизнес-процесс представляет собой всем знакомую концепцию, в полной мере используемую находчивыми поставщиками языковых услуг. Она именуется *пересозданием*.

В рамках данной концепции цель переводчиков — преобразовать текст

таким образом, чтобы он не искажал информационное сообщение, но в то же время провоцировал желательную реакцию реципиентов — носителей языка перевода. Пересоздание не является ни точным переводом, ни созданием с нуля. По своей природе это творческий процесс, компьютер с такой работой не справится. Нельзя сказать, что это услуга массового потребления или что с такой работой справится кто угодно.

В конце концов, международный бизнес требует гибкости. Методы ведения бизнеса меняются с каждым днем, они варьируются в зависимости от рынка, страны и культуры. В разных ситуациях требуются нестандартные и эффективные процессы. Необходимо предоставлять масштабируемые услуги, доступные клиенту в нужном месте и в нужное время.

Задача владельцев компаний, занимающихся переводом, заключается в предоставлении портфеля языковых услуг, соответствующих требованиям современного клиента. На наш взгляд, наилучшие перспективы имеют те, кто устраняет разрыв между технологиями и коммуникацией. Наш ассортимент услуг — прямая реакция на изменяющиеся бизнес-требования и технологические разработки на международном рынке, и мы полагаем, что конкурентные преимущества появляются за счет изучения возможностей, которые мы сами не в состоянии создать. Мы хотим использовать те преимущества, которые другие компании могут создать для наших решений, причем не только в языковой отрасли. Аналогично, мы хотим, чтобы наши наработки использовали не только непосредственно клиенты, но и другие

поставщики языковых услуг. Совместный, краудсорсинговый подход часто остается без внимания, но перешагнув через потребность все делать самостоятельно и всем владеть самим, мы поднимем отрасль на новый уровень.

Управлять в наше время компанией — поставщиком языковых услуг — увлекательно. Изменения грядут и в бизнесе, и в коммуникациях, прогресс не стоит на месте. Я стараюсь реагировать на такое развитие гибко и внимательно, продолжая изучать передовые методы отрасли и интегрировать их в нашу модель бизнеса и наш рабочий процесс.

Библиография

Danielson, Krissi (2008-03-26). «Distinguishing Cloud Computing from Utility Computing». Ebizq.net. http://www.ebizq.net/blogs/saasweek/2008/03/distinguishing_cloud_computing. Получено 22 августа 2010 г. от http://en.wikipedia.org/wiki/Cloud_computingcite_note-0

Джин Шрайвер — основатель и генеральный директор GLOBO, международного поставщика комплексных языковых услуг и услуг по межкультурной коммуникации, специализирующегося на международном маркетинге, социальных сетях и услугах языковой поддержки, предоставляемых по требованию. До основания GLOBO г-н Шрайвер руководил развитием небольшого семейного бизнеса, превратив его в одного из крупнейших в мире поставщиков языковых услуг. Регулярно выступает с докладами об инновациях в сфере языковых услуг и технологий по всему миру. Г-н Шрайвер получил степень бакалавра (с отличием) по экономике в Университете Темпл, степень доктора юридических наук в Университете Вилланова. Г-н Шрайвер выражает благодарность Дженнифер Бухгольц за помощь в написании статьи. Дженнифер является директором по коммуникациям компании GLOBO. До перехода в компанию GLOBO г-жа Бухгольц работала в нескольких международных компаниях в сфере рекламы и связей с общественностью. Получила две степени бакалавра по английскому и испанскому языку в Рутгерском университете и степень магистра по профессиональным коммуникациям в Университете ля Салле, где также работала адъюнкт-профессором.

Локализация программного обеспечения для техники

Франсуа Массьон (*François Massion*)

Очевидно, что сейчас нет недостатка в литературе по локализации программного обеспечения. Есть много интересных статей, в которых обстоятельно описывается все, что необходимо для локализации программ, начиная с глобализации программного обеспечения на этапе разработки с учетом требований различных стран и языков и заканчивая созданием интерактивной справки на разных языках. Эти материалы касаются преимущественно офисных приложений для сред Windows, Linux или Mac.

Быстрый прогресс в области автоматизации промышленного производства, широкое использование электроники в повседневной жизни и применение Интернета в качестве средства связи и платформы для сотрудничества накладывают свой отпечаток на традиционный подход к локализации. Развитие ставит новые задачи, как перед переводчиками, так и перед разработчиками программного обеспечения. Около 50 лет назад на рынке появилось первое устройство с числовым программным управлением. Сегодня текст программного обеспечения можно увидеть и в автомобилях, и в операционных залах, и на обычных

бытовых устройствах. И, конечно же, пользователи ожидают, что они смогут прочесть и понять все сообщения и команды на своем родном языке.

До настоящего времени локализация программного обеспечения для техники и встроенных систем никого или практически никого в локализационной отрасли не интересовала даже несмотря на наличие весьма больших проблем (или, может, как раз именно из-за них). Как следствие, многим производителям станков, установок и оборудования с числовым программным управлением приходится просто из кожи вон лезть, тратя слишком много времени и средств, чтобы локализовать свое программное обеспечение. У переводчиков тоже есть свои трудности с текстами для техники. Такая ситуация отчасти возникает из-за того, что специалисты по разработке программного обеспечения мало что знают о процессе интернационализации и локализации, а отчасти из-за отсутствия в этой сфере общепринятых концепций локализации. Тем не менее, многие производители стандартизируют разработку своих программ, как в случае со стандартом IEC 61131-3 или ISO 14649 (STEP-NC).

Экран технического устройства

Компаниям, занимающимся локализацией программного обеспечения для техники, приходится сталкиваться с лингвистическими и организационными проблемами. Чтобы лучше понять, чем локализация программного обеспечения для техники отличается от «обычных» проектов локализации, давайте для начала вкратце рассмотрим классический процесс локализации.

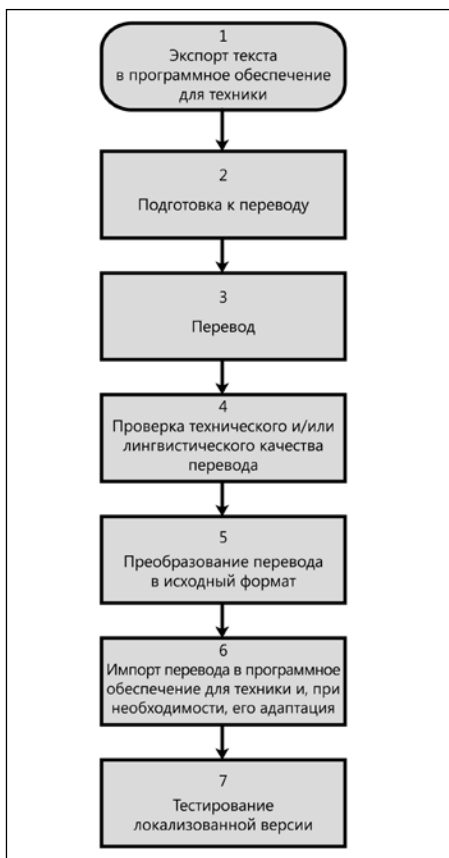
Традиционно или по крайней мере в идеале программное обеспечение изначально должно разрабатываться из расчета на международное использование. Конкретные способы реализации могут в чем-то отличаться в зависимости от языка программирования, однако основной принцип состоит в том, что тексты различных объектов программирования хранятся в отдельных файлах (которые обычно называются файлами ресурсов) вместе с определенной метainформацией, такой как тип объекта (сообщения, пункты меню, кнопки, заголовки диалоговых окон, названия полей и т. п.) и, возможно, идентификатором объекта. В случае приложений для платформы Windows .NET программисты создают так называемые вспомогательные сборки, содержащие переводимые строки. В других языках программирования такие строки хранятся в аналогичных файлах программного обеспечения, например в пакете ресурсов в Java-приложениях. После создания файлов ресурсов их можно обработать с помощью таких средств локализации, как SDL Passolo, Visual Localize или Alchemy CATALYST. В этих программах исполь-

зуется средство синтаксического анализа (фильтр) для двоичных исходных файлов (DLL или EXE) и других форматов (XML-фалов). Они импортируют подлежащий переводу текст в редактор переводов вместе с дополнительной информацией (в случае ее наличия). При работе с такими языками программирования, как Visual C++ или C#, в этих средствах локализации предусмотрена функция предварительного просмотра диалоговых окон в редакторе переводов. Переводчик может откорректировать длину поля или объекта, если для перевода требуется больше места.

При программировании некоторых приложений для техники и устройств также используются такие языки, как C++. Если разработчики придерживаются общепринятых концепций локализации, эти приложения можно локализовать, как и другие программы. На практике, однако, не все разработчики программного обеспечения для техники знакомы с процессом локализации и его методами и не используют все возможности, предлагаемые применяемым языком программирования. Поэтому они иногда выбирают сложные и не всегда надежные подходы к локализации программного обеспечения для техники.

Многие приложения для техники пишутся на специальных языках программирования, например EXHART, COMPACT, Siemens S7, APT, или таких высокоуровневых языках, как С для программирования микроконтроллеров, так что для них используются специальные компиляторы. Поэтому разработчикам и переводчикам сложно обрабатывать

Рис. 1. Локализация программного обеспечения для техники



переведенный текст. Это сопряжено с массой проблем.

Прежде всего, текст нужно сделать доступным для переводчика в редактируемом формате. После перевода текст необходимо импортировать обратно в программное обеспечение для техники.

При этом должна поддерживаться кодировка специальных символов для множества языков, таких как испанский, французский, азиатские языки или

языки с двунаправленным письмом, например арабский и иврит.

Из-за ограничения длины пространство, доступное на экране технического устройства, часто позволяет разместить на нем только определенное количество символов. В зависимости от ситуации некоторые приложения работают с одной или несколькими строками текста в каждом сообщении. Максимальная длина текста может быть задана в символах или пикселях.

Программисты вставляют в подлежащий переводу текст переменные, ярлыки, символы переноса строки, а также ескаре-последовательности, что переводчику приходится принимать во внимание. Все это должно также соответствовать лингвистическому контексту языка перевода.

Многие документы крайне сложно понять вне контекста или без дополнительных объяснений. В большинстве случаев переводчик никогда не видит итоговый результат своего перевода (например, весь текст и все объекты, отображаемые в одном диалоговом окне), и у него нет возможности проверить перевод с учетом контекста.

Каким был бы типичный поток операций в проекте локализации программного обеспечения для техники? Сначала разработчик экспортирует подлежащий переводу текст. Затем этот текст готовится к переводу переводчиком или бюро переводов. По завершении перевода выполняется проверка качества, в рамках которой проверяются как технические, так и лингвистические аспекты проекта. Далее перевод экспортируется обратно в исходный формат и отправляется раз-

работчику, который импортирует его в программное обеспечение для техники и, при необходимости, корректирует переведенные строки по длине и т. п. К сожалению, многие компании не всегда принимают необходимые меры для тестирования локализованных версий программного обеспечения, которое позволило бы убедиться в правильности перевода, как с технической, так и с лингвистической точки зрения (рис. 1).

Если бы условия были идеальными, разработчики бы изначально планировали локализацию текстов. Программисты отделяли бы текст от кода и предоставляли бы дополнительную метаинформацию. Это является необходимым условием рентабельности процесса локализации. Однако в действительности компании часто экспортируют подлежащий переводу текст построчно, вследствие чего переводчику сложно понять, что к чему и что в действительности означают те или иные выражения. Некоторые компании осознают эту проблему и предоставляют переводчику дополнительные полезные метаданные. Для выполнения этой работы требуется много времени, а иногда и сложная цепочка этапов преоб-

разования. Иногда программисты сортируют и группируют строки программного обеспечения по объектам, модулям, разделам и функциям и хранят их в отдельных файлах, в таблице Excel или другой электронной таблице. Из-за этого как разработчику, так и переводчику приходится выполнять дополнительную работу, которой можно было бы избежать, если бы разработчики изначально закладывали в свое программное обеспечение возможность локализации.

Обычно экспортированные данные переводятся с помощью стандартных средств памяти переводов после того, как поставщик языковых услуг отделит текст от кода. Вот тут-то и начинаются первые сложности, которые зависят от использовавшихся принципов программирования и формата экспорта данных. Отделение текста от кода не всегда является тривиальной задачей, поскольку для различения переводимого текста и требующего защиты кода программы необходима единая стандартная структура (рис. 2).

Некоторые экспортированные файлы создают переводчику дополнительные проблемы. Типичный случай — разби-

Рис. 2. Код и подлежащий переводу текст должны быть отделены друг от друга

```
-- Translate the data in the field <SX_Description to a max of 20 characters
-- The symbol ' is a control character and should not be used in the text!
INSERT INTO SX_ActionDataNameDef (S_Lgn ,
                                SX_Action
                                SX_Fieldname
                                SX_Description
VALUES ('E', '678', 'SX_ActionType' , 'ActionType')
GO
```

ние одного предложения на две строки и более, которые хранятся отдельно в программном обеспечении для техники. Пример:

21066, OBJECT MUST BE IN THE
RELEASED OR

21067, IN TRANSPORT STATUS.

Поскольку синтаксис различных языков отличается, такие предложения во многих случаях будут переведены с ошибками при подстановке из системы памяти перевода несоответствующих блоков.

Следующая проблема связана с ограниченностью пространства, доступно для размещения перевода на экране технического устройства. Переводчик получает указания не превышать заданное количество символов или пикселей в отображаемых строках. Но как это реализовать? Проверять длину перевода можно вручную построчно, что весьма затруднительно, или воспользоваться сценарием, макросом либо специализированным приложением, гарантирующим отсутствие превышений максимально допустимой длины. Для этого используется широкий спектр инструментов: от простых формул в Excel, позволяющих выводить в ячейке длину текста, до сложных процедур, вычисляющих ширину текста в пикселях с учетом таких параметров, как тип буквы (узкая, такая как Г, или широкая, такая как Ш) и размер шрифта.

Наиболее серьезные проблемы возникают в случае постоянного изменения правил ограничения длины. Вот пример инструкций для переводчика:

«В конце строк добавлено сокращение и число, например (sl/72). Сокраще-

ние sl означает, что строки не должны превышать длину одной одиночной строки. Сокращение ml означает, что перевод можно разделить на несколько строк. Число соответствует максимальной длине строки в символах (в данном примере это 72). Нужно сохранять пробелы между кавычками».

Подлежащая переводу строка может выглядеть следующим образом:

1873 : ('xmessage', 'search failed', '\n search operation could not be performed because: %(reason)s.\n') (ml/80)

Эти проблемы вполне можно решить, однако на это требуется время, деньги и навыки программирования, которые есть далеко не у каждого переводчика. Сценарии и макросы дают возможность переводчику и редактору узнать, в каких случаях нужно изменить перевод в соответствии с требованиями разработчиков. В случае несоблюдения этих требований существует риск того, что часть текста будет не видна, а это, в свою очередь, может привести к неправильной работе технического прибора или устройства либо станет причиной дорогостоящего процесса внесения исправлений в перевод. Как программисту узнать, что имеющиеся решающее значение слова, которые соответствуют переводу *on* («включено») или *off* («выключено»), в конце предложения на немецком языке не отображаются на экране технического устройства именно потому, что перевод оказался слишком длинным?

Самые же сложные проблемы с пространством возникают тогда, когда клиент требует, чтобы переводчик соблюдал отступы, применяемые в исходном языке. В технических устрой-

ствах и средствах более старых моделей использовались непропорциональные шрифты, то есть все символы были одинаковой ширины, а столбцы в таблицах создавались с помощью обычных пробелов. Из-за различий в длине и количестве слов в разных языках требование такого рода выполнить особенно трудно. Опять-таки, для автоматизации контроля над соблюдением таких инструкций необходим поставщик языковых услуг, умеющий обращаться со сценариями и средствами программирования.

Последняя серьезная техническая проблема связана с правильным представлением специальных символов в иностранных языках. Не во всех программах для техники используется Юникод. Следовательно, не все они могут работать с двухбайтовыми символами, такими как китайские и японские иероглифы. Многие системы в настоящее время поддерживают, по меньшей мере, все европейские языки, включая русский. Однако некоторые из них по-прежнему работают с разными кодовыми страницами, так что программисты и переводчики должны заранее уточнять, какие будут выбраны шрифты и кодировка.

С лингвистической точки зрения локализация программного обеспечения для техники — также отнюдь не простая задача. Многие разработчики программного обеспечения просто отправляют своим переводчикам обычные списки слов. Они считают, что им в кратчайшие сроки предоставят перевод высокого качества, невзирая на отсутствие контекста. Если такое слово, как *device* («устройство») или *support* («под-

держка») рассматривать отдельно, выполнение задачи перевода превращается в русскую рулетку. Первая проблема связана с качеством исходного текста. В отличие от традиционного программного обеспечения или документации к ПО, эти тексты составляются инженерами по разработке программного обеспечения с недостаточной лингвистической подготовкой или без таковой. Нередко в них встречаются грамматические и орфографические ошибки. Но хуже всего то, что программное обеспечение разрабатывается в течение определенного периода времени несколькими программистами, и в нем снова и снова возникают расхождения в названиях одних и тех же элементов, например *Compressed Air Valve On* и *Switch On Air Pressure Valve* для команды «Открыть клапан впуска сжатого воздуха».

В проектах разработки программного обеспечения некоторые понятия пишутся по-разному, в зависимости от объекта программы (заголовок диалогового окна, поле, сообщение), для которого они используются, по-разному сокращаются или даже означают совершенно разные вещи, как в случае со словом *support* («поддержка», «опора», «стойка» и т. д.). В таких случаях традиционные системы памяти переводов могут приводить к ошибкам, если существующие переводы извлекаются из памяти без критической оценки.

В некоторых ситуациях точный смысл выражения невозможно определить без контекста. Что в действительности означает выражение *search term*? Это существительное («условие поиска») или команда («найти термин»)? Было

бы удобно, если бы разработчики экспортировали информацию, которая помогла бы переводчику установить тип объекта, относящегося к строке, и определить, какие части текста связаны между собой. В любом случае важно и необходимо, чтобы у переводчика была возможность задавать вопросы (и чтобы он их действительно задавал), а также чтобы клиент назвал контактное лицо, которое хорошо знает программное обеспечение и продукт и может дать компетентный ответ на эти вопросы.

Переводчики должны адаптировать свой перевод с учетом технических требований клиента. Порядок слов в различных языках отличается, и это иногда влияет на способ перевода текста. Например, разработчик может вставить в центр предложения *escape*-последовательность, такую как `\n`, обозначающую разрыв строки. В этом случае переводчику нужно догадаться, где следует поместить *escape*-последовательность в переводе. В некоторых случаях программисты не знакомы с лингвистическими правилами языка перевода и злоупотребляют переменными. Такая ситуация может возникнуть, когда соседние слова, такие как прилагательные, переводятся по-разному в зависимости, например, от рода существительного.

Из-за ограниченности доступного пространства иногда приходится использовать сокращения или писать слитно несколько слов. Как сократить выражение *Shutoff Valve Not Closed* («запорный клапан не закрыт») всего до десяти символов? В результате сокращения получают конструкции наподобие

ShOffVlvNc, которые сложно понять и на исходном языке даже оператору технического устройства. У некоторых сокращений могут быть разные значения, зависящие от ситуации, в которой их применяют; например, в одном из проектов сокращение *Pos* использовалось для обозначения как слова *positive* («положительный»), так и слова *position* («положение»).

По той же причине смысл написанных слитно слов или команд оказывается совершенно неочевидным. Выражение *Dedusting Solenoid Valves On Duration Timer Setpoint (x0,1s)* остается тайной за семью печатями даже для некоторых клиентов. Тут переводчику необходимо четко понимать принцип, по которому было составлено это выражение. Какова его структура? Какая информация заложена в начале и в конце сообщения? Используются ли для обозначения группы слов такие типографические элементы, как прописные буквы? Чтобы конечный пользователь не путал различные структуры, клиент должен найти в руководстве по стилю лингвистические правила, определяющие принципы создания и составления строк и единообразный подход к написанию сообщений. Это же касается и создания переменных.

Для рядового переводчика соблюдение этих технических и лингвистических требований превращается в кропотливую и крайне трудную задачу, на которую нередко уходит не один час труда. Зачастую эффективным и надежным решением может стать разработка соответствующих сценариев или проверочных процедур. Однако это возможно только при условии активного участия со сторо-

ны разработчиков и только для проектов определенного масштаба. К сожалению, в небольших проектах локализации программного обеспечения для техники по-прежнему значительный объем работы приходится выполнять вручную.

Некоторые аспекты текста можно проверить с помощью функций контроля качества памяти переводов или с помощью независимых программ для контроля качества. К ним относится согласованное использование предопределенной терминологии при условии, конечно, что такая терминология была создана, и проверка правильности использования чисел в программном обеспечении для техники. Однако далеко не все можно проверить с помощью программного обеспечения. Поскольку некоторые переводы выполняются без сведений о контексте, обычно конечный результат должен быть проверен после компиляции локализованной версии, так как изрядное количество строк зависит от контекста.

Мало кто из клиентов задумывается над процессом локализации в целом и, соответственно, редко кто выделяет для этого этапа достаточно времени и средств. Это, к сожалению, приводит к необходимости исправления переведенного текста на более поздних этапах, при поступлении жалоб или когда переведенные строки уже попали к конечному заказчику. Некоторые производители программного обеспечения для техники осознают эту проблему и разрабатывают собственные приложения для виртуализации локализованных диалоговых окон в процессе рецензирования. Благодаря этому рецензент может воочию увидеть весь переведенный текст, который

отображается одновременно в диалоговом окне.

Если клиенты разработали собственные специализированные приложения, позволяющие локализовать их программное обеспечение для техники и перевести строки ПО, производительность и функциональность таких приложений, как правило, далеки от масштабов и возможностей профессиональных средств локализации, которые уже многие годы разрабатываются специально для этой цели.

Учитывая все вышеперечисленное, имеет смысл рассмотреть возможность использования в качестве альтернативы традиционным системам памяти перевода программы, позволяющие локализовать программное обеспечение для техники и встроенных систем. В таких программах для локализации предусмотрены, например, следующие функции:

- считывание доступных метаданных из файлов, сгенерированных программным обеспечением (CSV, TXT, XML), например данных о максимальной длине строки в символах или пикселях;
- проверка максимального количества символов или пикселей в отображаемой строке;
- разработка настраиваемых средств проверки или преобразования файлов с помощью встроенного редактора макросов;
- предоставление переводчику информации о контексте в форме метаданных, изображений или ссылок на внешние сайты;
- связывание единиц перевода с идентификаторами (при их наличии), ко-

торые упрощают перевод зависящих от контекста строк с различными возможными вариантами перевода.

Исправлять всегда дороже, чем сделать правильно с первого раза. Переводчики и разработчики могут избежать множества проблем и сократить объемы работы, сотрудничая в долгосрочной перспективе и вместе планируя все

этапы процесса локализации приложений для техники.

Франсуа Массьон получил степень кандидата наук в области лексикографии в Эрлангенском университете (Германия) в 1986 году. Он — владелец и генеральный директор компании D.O.G. GmbH, а также преподаватель курса по терминологии в Университете Анхальта (Германия).

TermWiki: управлять терминологией стало проще

Уве Мюге (Uwe Muegge)

Все больше и больше специалистов в области технических коммуникаций, перевода и локализации начинают понимать, что многоязычная терминология имеет критическое значение для жизненного цикла продуктов и услуг, в особенности для предлагаемых на международных рынках. Но не все осознают, что обмен глоссариями в виде таблиц по электронной почте создает столько же проблем, сколько и решает. В данной статье рассматриваются основные возможности последнего выпуска TermWiki — бесплатного решения по управлению терминологией для совместной работы, которое теперь поддерживает индивидуальное управление глоссариями.

Управление терминологией с помощью электронных таблиц — как правило, не лучшая идея

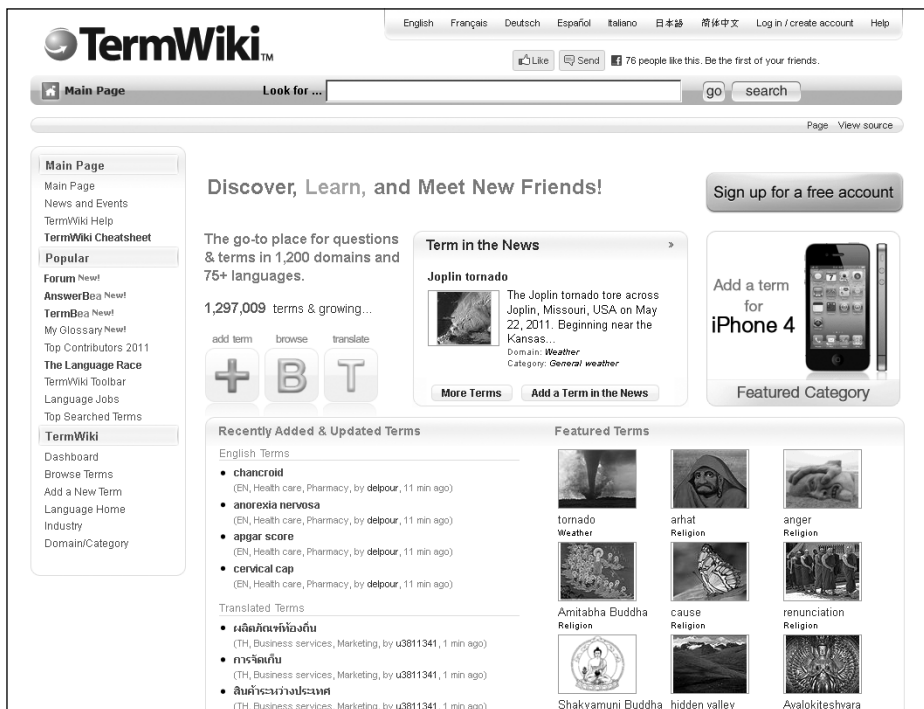
Документы обеспечивают лишь ограниченные возможности совместной работы

Если задуматься, документы любого типа, хранящиеся на жестком диске ком-

пьютера или передаваемые по электронной почте, обладают рядом недостатков. Конечно, сотрудники могут работать и с глоссарием в виде электронной таблицы, но это не облегчает сотрудничество для территориально удаленных друг от друга коллег, особенно если их много.

Кроме того, структурирование и контроль категорий данных, собирае-

Рис. 1. Интерфейс главной страницы TermWiki выглядит привычным для пользователей WikiMedia



мых в глоссарии разными группами в пределах одной организации, могут быть проблематичными. Вспомним старые глоссарии корпорации Майкрософт со значительными расхождениями и несогласованностью отдельных совокупностей данных. Что тогда говорить о процессе редактирования, перевода и утверждения!

Документы могут запросто стать неудобными

Электронная таблица — практичное решение, когда вы начинаете управлять терминологией, данных мало или вовсе

нет. Но как только количество терминов превысит несколько сотен, возникает необходимость классифицировать термины по линейкам продуктов или бизнес-подразделениям. А с увеличением числа языков до десятка и более электронная таблица, которая вначале казалась столь простой и удобной, совершенно неожиданно превращается в неуправляемый ресурс данных, и отслеживать, кто внес какое изменение и по какой причине, управлять отдельными обновлениями и поддерживать синхронизацию разных языков становится практически невозможно.

Рис. 2. Обзор некоторых проблем, связанных с управлением терминологией в таблице Excel



Что меняется при использовании TermWiki?

Ниже приведены три отличия средства TermWiki от традиционных решений по управлению терминологией на основе документов.

Система TermWiki всецело построена на веб-технологиях

В отличие от типичных рабочих станций управления терминологией, в которых применяются программные приложения, устанавливаемые на отдельных компьютерах, TermWiki представляет собой облачную систему — пользователю не приходится устанавливать и поддерживать специальное программное обеспечение. Все, что нужно для поиска, ввода, передачи и загрузки терминологических данных в TermWiki, — это веб-браузер и подключение к Интернету.

При таком подходе любой термин, вводимый в систему, автоматически становится доступным для всех. Централизованное управление терминологией на основе веб-технологий означает, что у всех пользователей есть простой и постоянный доступ не просто к коллек-

ции терминов, но к самой последней ее версии.

Пользоваться TermWiki чрезвычайно просто

Одной из целей проекта TermWiki было сделать новую систему управления терминологией максимально понятной. Пользователям TermWiki знаком интерфейс вики с ориентированной на задачи структурой, в которой все основные элементы находятся на видном месте и доступны для использования.

В Монтерейском институте международных исследований, в котором я преподаю курс управления терминологией, студенты направления перевода и локализации, помимо прочих инструментов, использовали систему TermWiki. От них я постоянно слышу один и тот же комментарий: как приятно было работать с TermWiki после того, как они потратили много времени и сил на создание стандартного средства управления терминологией, в котором необходимо определить все категории данных и создать модели ввода для более сложных записей, чем обычные пары терминов. В системе

TermWiki с ее четко структурированной методикой ввода и доступными списками выбора для многих категорий данных даже новичок сможет создать записи за считанные минуты после первого же входа в систему.

Система TermWiki совершенно бесплатна

Хотя TermWiki не только предоставляет комплексный набор мощных функций для управления терминологией, но и содержит сотни тысяч доступных для поиска, классифицированных терминов с толкованиями более чем на 70 языках, версия TermWiki для сообщества является абсолютно бесплатным ресурсом. Пользователям не досаждают реклама; по сути, на сайте вообще нет рекламы.

Сочетание серверной системы управления терминологией для использования как терминологами, так и неспециали-

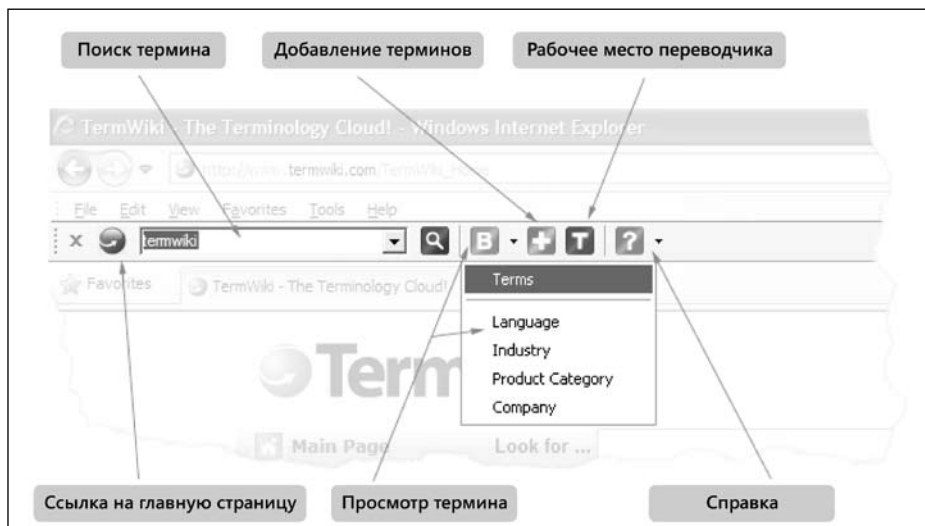
стами, и бизнес-модели, благодаря которой бесплатная версия теперь доступна для международного сообщества людей, работающих в сфере технических коммуникаций, стало беспроигрышным решением. Менее чем за год с момента своего появления в апреле 2010 года система TermWiki объединила большое количество пользователей и активных участников и пополняется на сегодняшний день со скоростью приблизительно 10 000 новых терминов в день.

Новые возможности TermWiki

Система TermWiki поддерживает персональные глоссарии

С помощью функции *My Glossary* («Мой глоссарий») можно создавать и передавать персональные глоссарии в систему TermWiki, управлять ими и выполнять по ним поиск. Пользователи

Рис. 3. Панель инструментов TermWiki упрощает поиск и ввод терминов



даже могут приглашать коллег для рецензирования, редактирования и перевода терминов в своих персональных коллекциях TermWiki. Кроме того, на вкладке обсуждения на странице каждого термина можно обмениваться мнениями и следить за изменениями как всей записи, так и какой-либо ее части. При необходимости можно загружать персональные глоссарии целиком либо частично, например только определенные категории данных, с помощью функции экспорта в файл Excel.

Система TermWiki оснащена собственной панелью инструментов в окне браузера

Совсем недавно была добавлена новая функция — панель инструментов TermWiki. Это устанавливаемая за считанные минуты надстройка для браузера, которая дает возможность выполнять поиск терминов, определений и переводов для конкретных отраслей, языков и продуктов. Панель также упрощает совместное использование терминов и переводов — теперь для этого достаточно лишь нескольких нажатий кнопки мыши. В качестве инструмента, дополняющего саму платформу TermWiki, панель инструментов TermWiki включает также множество экспресс-ссылок на функции, регулярно применяемые пользователями TermWiki.

Наличие панели инструментов TermWiki подтверждает тот факт, что система TermWiki построена на основе концепций и дает возможность выполнять поиск по конкретным отраслям и даже категориям продуктов.

В системе TermWiki предусмотрено рабочее место терминолога

В отличие от многих других систем управления терминологией, в которых термины нужно добавлять по одному, рабочее место терминолога TermWiki дает возможность вводить до десяти терминов в базу данных одновременно с помощью списков выбора для многих категорий данных. Данный метод ввода значительно повышает эффективность и экономит время пользователей, которые управляют терминологией небольшими пакетами.

Чтобы одновременно ввести больше десяти терминов, можно воспользоваться функцией импорта в систему TermWiki из файла Excel для преобразования существующих одно- и многоязычных терминологических баз любого размера.

Подходит ли данное решение для организаций, которые хотят обеспечить конфиденциальность данных?

TermWikiPro для корпораций

В версии TermWiki для сообщества все глоссарии, включая данные «Мои глоссарии», являются общими ресурсами, поэтому решение по управлению терминологией TermWiki предназначено для внештатных переводчиков и небольших организаций, которые хотят предоставить доступ к языковым ресурсам всему сообществу TermWiki.

Для организаций, которым нужно обеспечить полную конфиденциальность данных, спонсор TermWiki, компания CSOFT, предлагает TermWikiPro,

настраиваемую серверную службу, включающую все возможности и функции версии для сообщества, а также ряд дополнительных, которые упрощают управление пользователями, обеспечивая безопасность и масштабируемость.

Библиография и ссылки

1. DEPALMA, D.A. TermWiki Lowers the Barrier to Terminology Management. Global Watchtower [в сети]. 2010, no. 14 January 2010. Статья доступна в Интернете по адресу <http://www.globalwatchtower.com/2010/01/14/csoft-termwiki>.
2. MUEGGE, U. Disciplining Words: What You Always Wanted to Know About

Terminology Management. tcworld, 2007, vol. 2, no. 3, p. 17-19.

3. MUEGGE, Uwe. TermWiki: A New Collaborative Terminology Management Solution. tcworld [в сети]. 2010. Статья доступна в Интернете по адресу <http://www.tcworld.info/index.php?id=188>
4. www.termwiki.com

Уве Мюге — директор MedL10N, подразделения наук о жизни компании CSOFT. В данный момент он является членом TC37 в Международной организации по стандартизации (ISO) и читает курс «Автоматизированный перевод» в Монтерейском институте международных исследований (США).

Локализация программного обеспечения, веб-сайтов и международных приложений

Амир Хельцер (Amir Helzer)

Локализация продукта — отличный способ быстрого расширения бизнеса. В докладе Ассоциации по стандартизации в области локализации (LISA) за 2007 г., например, сообщается, что каждый доллар, инвестированный в локализацию, приносит 25 долларов прибыли.

Свыше 78 % пользователей Интернета находятся за пределами США, на международном рынке, а за прошлое десятилетие частота использования арабского языка в сети увеличилась на 2500 %, китайского — на 12 %, а испанского — на 7 %.

При расширении деятельности в мировых масштабах компании должны ин-

вести средства в полную локализацию своих веб-сайтов и программного обеспечения. Локализация предполагает перевод и другие преобразования и осуществляется с учетом местных особенностей и культуры, обеспечивая создание продукта, подходящего для использования в стране назначения.

Локализация программного обеспечения

Первый этап на пути создания компании по разработке программного обеспечения на нескольких языках — это локализация такого программного обеспечения. Для локализации необходимо определить подлежащее переводу содержимое, разработать стратегию, исходя из предполагаемых продаж, создать график выпуска продуктов для процесса локализации, задействовав для выполнения переводов профессионалов, для которых язык перевода является родным, тщательно управлять процессом локализации, а в конце проверить переводы.

Этап 1 (подготовка). Составление графика проекта, разработка глоссариев, отделение текста и подготовка файлов ресурсов. Ключевым моментом в локализации программного обеспечения является отделение текста приложения от кода программы. Это делает возможным процесс локализации как таковой и упрощает его осуществление.

Этап 2 (перевод). Перевод строк программного обеспечения, интерактивной справки и документации. После создания для программного обеспечения файла ресурсов, содержащего весь подлежащий переводу текст, крайне важно

составить документацию, которая поможет переводчикам понять, что именно они переводят. Объясните переводчикам, для чего предназначено приложение, а также кто и как им пользуется. Это позволит им адаптировать текстовые материалы приложения, обеспечив их естественное звучание на новом языке.

Этап 3 (программирование). Масштабирование диалоговых окон; адаптация подсказок, пунктов меню и кнопок; компиляция файлов справки, а также локализация файлов мультимедиа и встроенных графических элементов, содержащих в себе текст.

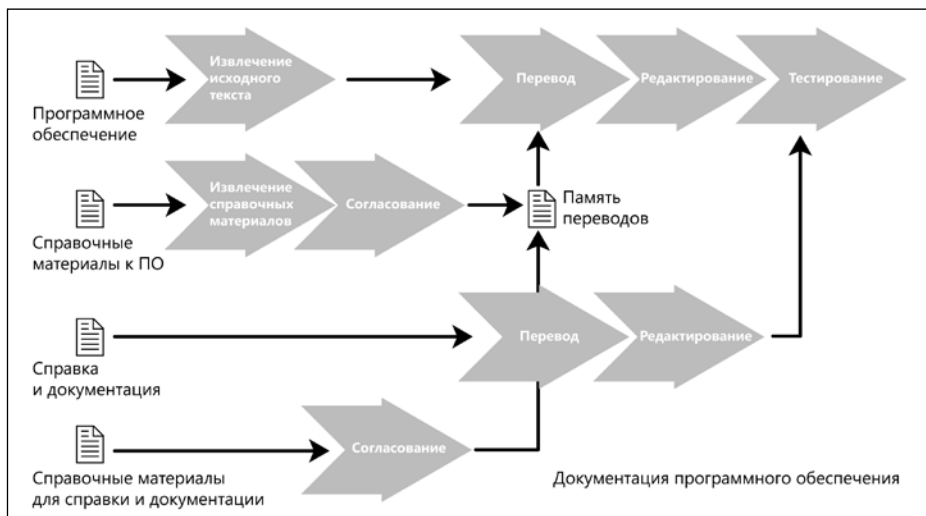
Этап 4 (тестирование). Тестирование локализованного программного обеспечения, включая проверку функций и интерфейса пользователя.

На **рис. 1** показан краткий обзор лингвистических процессов, которые происходят в типичном проекте локализации программного обеспечения.

Локализация веб-сайтов

Как и в случае программного обеспечения, для создания веб-сайтов применяются различные методы: для некоторых из них используется обычный HTML, для других — PHP или ASP, а для третьих — такие системы управления контентом, как WordPress или Drupal. Независимо от способа создания сайта, некоторые принципы всегда остаются неизменными. Для успешной работы веб-сайта на нескольких языках необходима возможность создания переводов и поддержания их в актуальном состоянии при изменении контента сайта; автоматизация построения структуры сайта, которая даст

Рис. 1. Рабочий процесс локализации программного обеспечения



возможность избежать редактирования ссылок и элементов навигации переводчиками; а также предоставление пользователям удобной возможности выбора языка. Нужно разбираться в торговом праве и соответствующих правовых нормах, а также экономике и политике соответствующей страны или региона.

Большинство основных приложений для обработки платежей оснащены функциями расчетов стоимости со встроенной поддержкой нескольких языков. Убедитесь, что страница расчетов отображается на правильном языке. Отсутствие возможности изменить язык — верный способ упустить возможность продать товар. При перенаправлении с вашего сайта в систему оплаты сторонней организации не забудьте добавить язык в ссылку. Если клиенты осуществляют покупку с помощью вашего приложения, программа должна автоматически определить язык пользователя и создать

ссылку на страницу расчета стоимости приобретенных товаров с правильным аргументом для языка. Клиенты также должны получать сообщения электронной почты с подтверждением покупки и указаниями относительно загрузки и активации приобретенной программы. Это сообщение электронной почты должно быть составлено на языке клиента, поэтому следует зафиксировать, на каком языке клиент использует программу.

Для каждого товара необходимо обеспечить предпродажную и послепродажную поддержку. Предоставление поддержки на нескольких языках обеспечит конкурентное преимущество и позволит произвести приятное впечатление на клиентов. Связь можно поддерживать с помощью электронной почты или контактной формы. В случае контактной формы сообщения могут автоматически перенаправляться разным сотрудникам службы поддержки, исходя из используемого языка.

Локализация приложений для мобильных устройств

По данным компании Google, к 2015 г. две трети всех покупок и половина транзакций будут осуществляться с помощью мобильных устройств. К концу 2010 года 1,9 млрд человек во всем мире активно пользовались Интернетом, а пять миллиардов человек являлись абонентами мобильной связи. В январском выпуске Direct Marketing News за 2011 г. приведена цитата гиганта Интернета, прогнозирующего, что к 2020 году пять миллиардов человек будет активно пользоваться Интернетом, а число абонентов мобильной связи достигнет десяти миллиардов.

Свыше половины мирового роста в области мобильной связи приходится на развивающиеся страны, такие как Бразилия, Россия, Индия и Китай. Завершают список десяти самых быстрорастущих рынков мобильных услуг Пакистан, Индонезия, Украина, Нигерия, Вьетнам и Бангладеш. Большинство конечных пользователей на этих развивающихся рынках практически или совсем не говорят на английском и будут покупать только продукты и услуги, доступные на их родном языке. Учитывая такое огромное число пользователей, которые используют мобильные устройства для доступа к Интернету, компаниям нужно будет переводить свои приложения для завоевания этой мировой аудитории. Существует множество средств перевода контента, включая системы машинного перевода.

В 1999 г. Международный научно-исследовательский институт передовых телекоммуникационных технологий в японском Кансайском научном городке

представил решение, которое должно было переводить произнесенные слова на мобильном устройстве и отправлять их в виде голосового сообщения на мобильный телефон другого пользователя. В 2004 году было представлено первое коммерческое программное обеспечение для машинного перевода на портативных устройствах, позволяющее переводить текст, электронную почту и мгновенные сообщения.

В ноябре 2005 г. другая японская компания, NEC Corporation, анонсировала разработку системы переводов, которую можно загрузить в мобильные телефоны. Эта система перевода могла распознавать 50 000 японских и 30 000 английских слов, и ее можно было использовать для простых переводов во время путешествия.

Система машинного перевода выполняет подстановку слов на одном естественном языке вместо слов на другом языке, однако одно только это обычно не позволяет создать хороший перевод текста, поскольку должна быть понятна вся фраза. Существует множество методик, которые позволяют улучшить перевод, в том числе лингвистическая типология, обработка различий в переводе идиоматических и разговорных выражений.

Начиная с 2007 г., количество приложений для перевода, использующих различные мобильные платформы, постоянно растет. Однако, чтобы контент был локализован наилучшим образом, необходимо убедиться в его правильности с лингвистической точки зрения и адекватности с точки зрения культуры. Все знают, что если просто пропустить контент, написанный на одном языке, через систему

переводов, например Переводчик Google, то в результате получатся слова, которые зачастую лишены всякого смысла. Переводчик Google создает статистический машинный перевод, отыскивая структуры в миллионах документов, чтобы выбрать оптимальный перевод. Однако поскольку в Переводчике Google используются средства машинного перевода, не все переводы будут идеальными, а их точность отличается в разных языках.

Как и в случае с любой деятельностью по локализации, лучший способ перевода приложений — это сочетание работы переводчика с компьютерным решением, используемым для управления процессом перевода. Ниже приведена процедура, которой следует придерживаться.

Прежде всего, для подготовки приложения к локализации необходимо провести аудит, чтобы определить потенциальные подводные камни, которые могут встретиться в процессе перевода, такие как встроенный текст и изображения, поддержка кодировок, формат даты и времени, числовые форматы, макет страницы, специфические функции для каждой конкретной страны, поддержка многобайтовых символов, языки с двунаправленным письмом, операционные системы, запрещенный код и многое другое.

Для создания всех приложений для iPhone используется кодировка Юникод, обозначаемая как UTF-16. Программы, моделирующие iPhone, поддерживают этот формат и обеспечивают предварительный просмотр переведенного приложения. Поскольку в конструкцию телефонов iPhone изначально заложена поддержка нескольких языков, если приложение доступно на разных языках,

телефон сам знает, что делать. Он отображает текст, используемый в графическом пользовательском интерфейсе, на правильном языке, в соответствии с настройками телефона. Следующий этап — подготовка ресурсов к переводу.

- Приведите ресурсы к пригодному для локализации виду в средстве Xcode, которое используется для разработки приложений для iPhone. Переместите файл ресурсов в новый каталог для локализации, а затем перетащите его в проект Xcode. Теперь средство Xcode распознает ресурс как локализованный на соответствующем языке. Повторите эту процедуру для остальных файлов ресурсов, которые необходимо локализовать.
- Локализируйте файлы Interface Builder. Прежде всего, извлеките строки из файлов ресурсов Interface Builder, запустив средство `ibtool`, которое устанавливается вместе со стандартным инструментарием разработки приложений для iPhone. `Ibtool` создаст пары значений объект-строка, просмотрев которые можно удалить ненужные строки, не требующие перевода.
- Разберитесь с повторяющимися строками. Файл ресурсов может содержать повторяющиеся строки, например наименования кнопок в различных состояниях. Некоторые из них нужно перевести только один раз, в то время как перевод других будет отличаться. Наименование одной и той же кнопки в различных состояниях, конечно же, должно переводиться одинаково. Однако отдельные строки в одних местах могут передаваться с помощью существительных, а в дру-

гих — с использованием глагольных конструкций. Строки, перевод которых должен быть идентичным, можно удалить из файла ресурсов или оставить их в нем, предоставив переводчику соответствующие инструкции.

- Поместите встроенный текст в файлы ресурсов. Обычно приложения для iPhone содержат тексты двух видов: интерфейс приложения, для управления которым используется Interface Builder, и встроенные строки, которые встречаются в самом коде. Строки интерфейса уже организованы в файл ресурсов, и нужно убедиться, что в самом коде используются только строки, получаемые из этого файла. Просмотрите весь код и определите, какие строки будут встречаться пользователям и подлежат локализации. Не переводите строки, которые пользователь никогда не увидит, например ключи в NSDictionary, если, конечно, эти ключи не выводятся на экран.

Теперь у вас должно быть два файла ресурсов: Interface Builder, созданный с помощью средства ibtool, и файл ресурсов со встроенными строками, созданный с помощью genstrings. Оба файла можно отправлять специалистам, которые переведут их на соответствующие языки.

Существуют специальные средства управления переводами для локализации файлов ресурсов iPhone, которые дают возможность справиться с переводом приложений, не тратя время на технические задачи. Они считывают файлы ресурсов, извлекают подлежащий переводу текст и создают переведенные файлы ресурсов. В них необходимо лишь передать файлы ресурсов с использованием

кодировки UTF-16 или UTF-8. Средства извлекают пары подписей, строк и комментариев. Затем переводчикам нужно будет перевести только строки.

Текст должны переводить профессиональные переводчики с богатым опытом локализации приложений для iPhone. Вам необходимо поддерживать с ними связь, следя за тем, чтобы перевод был не только точным, но и адаптированным к культуре и обычаям других стран. Вам понадобится создать проект перевода для iPhone, передать файлы ресурсов, выбрать язык перевода и, конечно же, оплатить перевод. Когда все это будет сделано, можно загрузить локализованные файлы ресурсов и импортировать их в приложение.

Откройте переведенный файл в Interface Builder и все просмотрите. Зачастую требуется существенно сокращать длину строк. Перевод с английского на немецкий и французский обычно на 50 % длиннее исходного текста. Чтобы все элементы правильно отображались, придется несколько изменить размещение компонентов. Теперь ваше приложение iPhone поддерживает несколько языков! Но есть еще один этап: тестирование переведенного приложения на устройстве iPhone. Для этого на телефоне необходимо изменить языковые настройки. Для смены языка в телефоне iPhone следуйте указаниям в пособии, предоставленном компанией Apple.

Амир Хельцер — генеральный директор и основатель ICanLocalize, компании по управлению переводами, занимающейся переводами программного обеспечения, веб-сайтов, мобильных приложений и типичных текстов для компаний всех размеров.

Законы культуры

Кэйт Эдвардс (Kate Edwards)

Наш мир полон законов, правил, норм и принципов, которые формируют современное общество во всех уголках земного шара. У человеческого общества длинная история создания законов, восходящая к египтянам, шумерам и вавилонянам, считавшим, что крайне важно разработать набор правил, который, по их мнению, должен служить на благо каждого. Пожалуй, одним из самых известных примеров можно считать свод законов вавилонского царя Хаммурапи, не говоря уже о таких религиозных трудах, как Библия и Талмуд. С тех древних времен до наших дней функционирование общества и культуры опирается на широкий круг очевидных законов и неочевидных норм, которые формируют саму суть человечества.

Большинство законов, с которыми мы знакомы, касаются преимущественно вопросов обеспечения общественной безопасности и поддержания мира, а также широкого круга проблем, для которых лучше всего подходит слово «практические» — и я не буду рассматривать конкретные законы, которые никак нельзя таковыми назвать; так что речь пойдет о вещах более общего порядка. Практические законы в основном помогают нашим обществам функционировать в спокойном и размеренном темпе, будь то автомобильное движе-

ние, личные взаимодействия, торговля или международные отношения.

Существует и другая форма права, которая касается управления контентом, особой категории законов, которую было бы интересно рассмотреть в этой статье. Многие из нас знакомы с законами о контенте более общего характера, например с законами о разных типах порнографии, а также с законами об авторском праве, охватывающими проблему подлинности. Большинство из них предназначены либо для защиты молодежи от контента, к которому она не готова, либо для защиты прав интеллектуальной собственности. Но в этом множестве связанных с защитой контента законов есть то, что мне по-настоящему интересно: менее определенное — но от этого не менее реальное — подмножество связанных с контентом законов, касающихся уязвимости, которая является следствием определенного культурного мировоззрения и геополитических представлений.

Думаю, это лучше объяснить на примере, чем с помощью умозаключений, поэтому позвольте мне проиллюстрировать сказанное. Географические карты относятся к изменчивому типу контента, так как некоторые правительства настаивают на весьма специфических требованиях к отображению на них принадлежащей им территории.

Фактическое обладание этой территорией на самом деле значения не имеет: их местное восприятие, политическая позиция или геополитическое представление оказываются куда более приоритетными, чем юрисдикция. Например, в Китае и Индии эти особые требования закреплены в законодательном порядке, поэтому если компания не изображает страну соответствующим образом, она может столкнуться с серьезными последствиями. В случае с Китаем это означает, что Тайвань нужно изображать как часть Китая, наряду со спорными территориями Аксай Чин, Аруначал-Прадеш и всеми островами Южно-Китайского моря. В Индии это означает показывать всю территорию Кашмира как неотъемлемую часть Индии.

К сожалению, такие законы далеко не очевидны для большинства разработчиков контента или даже для многих юристов и специалистов в области юриспруденции, если они не получили специальное образование по международному праву. Проблема с такими законами заключается в том, что многие компании не подозревают об их существовании до того момента, пока их продукты не попадут на рассмотрение в таможенный департамент такой страны. Поскольку систематизация местных правил может оказаться весьма далекой от четкой эмпирической классификации, решение о «целесообразности» может зависеть от личной интерпретации смысла регулирования контента конкретным правительственным чиновником. Несмотря на защиту свободы слова, обеспечиваемую Первой поправкой, даже Соединенные Штаты Америки сохраняют такие

защитные правовые инструменты, как концепция так называемого «предварительного запрета». Предварительный запрет представляет собой запрет правительства, который может быть использован для цензуры выпуска материалов, если они, например, могут представлять угрозу национальной безопасности или усиливать напряженность в обществе. В таких случаях из-за решительного запрета у контента нет никаких шансов на обсуждение.

Ситуация с недостатком ясности в регулировании контента, пусть медленно и лишь в некоторых секторах, но все же меняется в положительную сторону. Фильмы были формой возрастного ценза на протяжении десятилетий, и многие люди знакомы с американской рейтинговой системой G, PG, PG-13, R и NC-17. В индустрии видеоигр во многих странах существуют подобные рейтинговые организации, которые отвечают за специальную проверку игр на допустимость контента. В Северной Америке есть рейтинг Entertainment Software Ratings Board, который функционирует как добровольная, но повсеместно используемая система (фактически, это требование к продаже игр через любого розничного торговца в Соединенных Штатах Америки, но оно не является действующей правовой нормой). В других странах есть аналогичные рейтинговые организации: рейтинговая система Pan European Game Information в Европе, Computer Entertainment Rating Organization в Японии и Game Rating Board в Южной Корее.

Рейтинговые управления приносят настоящую пользу разработчикам кон-

тента, поскольку они формируют определенную систему принципов для определения приемлемого и неприемлемого контента на местном уровне. Одним из их самых больших недостатков является то, что они часто нацелены на «большую тройку» проблем контента: ненормативную лексику, секс и насилие.

Вряд ли они будут вникать в более глубокие вопросы, связанные с культурными аллегориями, политическим цинизмом, религиозной символикой и так далее. Я узнавала в некоторых рейтинговых организациях, будут ли они делать это в будущем, но большинство из них слишком плохо финансируется, чтобы усилить свой контроль. Создателю контента остается только либо заблаговременно распознавать эти риски за пределами «большой тройки», либо рискнуть в вопросе уязвимости на местном уровне.

Выходя за пределы действующего законодательства и добровольных (но, фактически, обязательных) оценок, мы попадаем в царство социальных норм и культурных практик, которые зачастую воспринимаются местным населением как само собой разумеющиеся вещи или, по крайней мере, считаются совершенно очевидными для посторонних. Однако многие из этих само собой разумеющихся вещей находятся вне досягаемости чужаков, если только они не читали об особенностях местной культуры или не провели какое-то время в стране. Мало у кого из разработчиков контента есть возможность сделать это ради углубления своих знаний, так что это еще одна область, где компании часто испытывают потрясение, когда у них возникают про-

блемы с выпуском контента на определенном рынке.

С этим связано одно юридическое понятие под названием «богохульство». Богохульство используется в религиозном контексте, таком как новые резолюции Организации Объединенных Наций о «диффамации» религий. Уровни законов о богохульстве в различных странах отличаются, но в тех из них, где культурные особенности связаны скорее с религиозными, чем со светскими принципами, они носят более непосредственный характер. Например, культура Саудовской Аравии сильно структурирована в соответствии с принципами ислама и, таким образом, корпоративные практики и контент соответствуют этим религиозным нормам. Великий муфтий возглавляет Совет высших улемов, который является собранием богословов, принимающих решения в связи с проблемами контента и его пригодности для страны. Их деятельность координирует Министерство по делам ислама, которое часто оказывает влияние на издательские компании с целью мониторинга создаваемых материалов.

Кроме законов о контенте, касающихся религии, действует также ряд других правовых норм, которые затрагивают различные виды контента. Существуют законы, известные как «соглашения о наследии», призванные предотвратить, например, продажу изображения Эйфелевой башни без разрешения французского правительства. Правовая категория морального права, в чем-то схожего с авторским правом, должна предотвратить продажу изображения стеклянной пирамиды, расположенной у входа в

Лувр, которую создал Ю. М. Бэй. Существует даже закон о товарных знаках, который может быть применен к различным видам флагов для предотвращения злоупотребления или имитирования.

В конце концов, поскольку изобилие законов о защите контента и сложностей с потенциальными судебными тяжбами уже начинает бить через край, возможно, это заставит нас вспомнить о том, что в отношении любви-ненависти, сложив-

шиеся у многих людей с юридической профессией, не мешало бы добавить немного больше любви!

Кэйт Эдвардс — географ и главный консультант компании Englobe в Сизтле, предоставляющей консультационные услуги в области геокультурной аналитики и стратегии разработки контента. Эдвардс проработала 13 лет в корпорации Майкрософт на должности географа и ведущего специалиста по геополитике.

Турботестирование машинного перевода в компании Cisco

Лори Тике (Lori Thicke)

Cisco Systems, Inc. — мировой лидер в сфере разработки сетевого оборудования для Интернета. На сегодня годовой доход корпорации Cisco, основанной в 1984 году группой ИТ-специалистов из Стэнфордского университета, составляет около 40 млрд долларов США, а количество ее сотрудников превышает 70 000 человек. Неудивительно, что Cisco является лидером в области машинного перевода (МП). Машинный перевод в Cisco начали применять в 2000 году. Это первая компания, опубликовавшая адаптированный, но необработанный машинный контент в Интернете в небывалых размерах. Для управления революционной функцией поддержки клиентов в корпорации Cisco используется сочетание перевода, выполненного человеком, и машинного перевода. Руководитель программ Cisco Services Пабло Васкес (Pablo Vazquez) является архитектором решений МП для поддержки клиентов Cisco и работает в Сан-Франциско, Калифорния.

Тике: Не будет преувеличением сказать, что Вы один из немногих со- временных специалистов, изучивших машинный перевод вдоль и поперек.

Над какими инновациями вы работаете в Cisco?

Васкес: Я изучаю сочетание машинного перевода и работы службы поддержки, краудсорсинга и машинного перевода и их совместное использование, анализ исходного текста, методы обучения системы для того или иного проекта, выбираю наилучшую систему для конкретного проекта и решаю другие вопросы, связанные с этими направлениями.

Тике: Корпорация Cisco — признанный лидер в области инициатив машинного перевода, начиная со знаменитого развертывания SYSTRAN в 2000 году. Расскажите, пожалуйста, о своей работе в Cisco в качестве архитектора решений машинного перевода для поддержки клиентов.

Васкес: Я работаю с локализованными версиями сайта технической поддержки. Что касается машинного перевода в частности, моя задача — непрерывно улучшать качество результата. На этом мне хотелось бы остановиться и кое-что объяснить. Важно понимать, что я не пытаюсь усовершенствовать систему машинного перевода; я не инженер. Я работаю над улучшением результата. Можно сказать, что я выступаю в роли опытного пользователя машинного перевода на данной должности: обучаю систему, анализирую результаты их работы с последующим определением лучших способов их обучения и анализа результатов в будущем. Сотрудники корпорации Cisco умеют рационально использовать машинный перевод. Это дает нам возможность осваивать новые языки и новые потоки операций, эконо-

мя деньги, улучшая качество перевода и увеличивая объем переводов без повышения издержек.

Тике: Для какого рода контента у вас используется машинный перевод?

Васкес: Я осуществляю поддержку и другой деятельности компании, но в основном машинный перевод используется для контента базы знаний на сайте службы поддержки клиентов. В будущем мы планируем расширить сферу применения.

Тике: Корпорация Cisco создала огромный сайт для поддержки клиентов, где используется перевод, выполненный человеком и компьютером. Сегодня многие компании стремятся следовать вашему примеру. Что изначально побудило корпорацию Cisco стать пионером в области автоматизации перевода для поддержки клиентов?

Васкес: Cisco — это международная компания, которая предоставляет тысячи продуктов и привлекает клиентов со всех континентов и почти со всех стран. Поэтому нам нужно предоставлять техническую поддержку на многих языках. Наша модель поддержки многогранна, как и у большинства ИТ-компаний. Одним из основных подходов является самостоятельное получение пользователями поддержки — клиенты хотят решать свои проблемы сами с помощью веб-сайта вместо того, чтобы звонить инженерам службы технической поддержки. Это дает нам возможность предоставлять для поддержки контент и знания, которые мы получаем из разных ресурсов, в многоязыковой веб-среде. Самый лучший и эффективный способ перевести на множество языков наши

материалы для поддержки — воспользоваться машинным переводом.

Тике: Вы фиксируете количество вопросов, которые получается решить с помощью веб-сайта?

Васкес: Мы фиксируем все. Для нас сайт — это способ обеспечения быстрой поддержки клиентам на основе знаний, собранных из всех наших технических ресурсов и накопленных в ходе работы с пользователями. У нас эффективная, быстрая и легкая в использовании модель поддержки, с помощью которой можно решать распространенные и типичные проблемы. В то же время наши сотрудники могут сосредоточиться на более сложных, новых вопросах. Наш коэффициент снижения количества обращений по поводу старых повторяющихся проблем значительно выше среднего по отрасли. Почти 80 % всех проблем, с которыми сталкиваются клиенты службы поддержки, решаются в Интернете, в среде самообслуживания. Чтобы дать вам представление о трафике: у нас более двух миллионов посетителей каждый месяц и около 18 миллионов просмотров страниц.

Тике: У корпорации Cisco огромный объем материалов для поддержки клиентов. Как вы определяете приоритеты для контента?

Васкес: Мы обращаем внимание на то, что просматривают наши клиенты. Мы знаем, какие документы не читает никто, а к каким обращаются сотни тысяч раз каждый день. Некоторые документы могут просматривать миллион раз в день, а другие — два или три в квартал. Разумеется, мы расставляем приоритеты для работы соответствующим

образом. Еще мы используем другие показатели, дающие возможность прогнозировать использование; например, новые продукты, которые должны стать популярными, или документы, для которых требуется значительный трафик.

Мы обращаем внимание на то, что просматривают наши клиенты... Разумеется, мы расставляем приоритеты для работы соответствующим образом.

Тике: Какие качественные показатели вы используете?

Васкес: Как я уже говорил, мы измеряем все. У нас различные качественные показатели. Хороший критерий качества — насколько объемная память переводов (ПП). Независимо от системы, мы всегда сначала выполняем анализ совпадений с памятью переводов, а затем уже осуществляем машинный перевод. Таким образом, мы знаем, какой процент документа переведен человеком, а какой — компьютером. В каждом документе сочетаются оба варианта. Весь перевод выполняется компьютером, но в нем может быть до 80 % из памяти переводов. Кроме того, при тестировании различных параметров мы автоматически оцениваем перевод. Другой критерий — удовлетворенность: решили ли мы проблему клиента. Каждый документ могут оценивать пользователи и эксперты. Что касается оценок экспертов, у нас много высококвалифицированных инженеров, говорящих на языке целевых стран, поэтому мы можем использовать такие ресурсы. Мы даем им прочитать документ и задаем ряд вопросов касательно понятности, качества перевода и т. д. Мы просим их

оценить не отдельные фразы, а документ целиком. Наша цель — добиться понятности. Это как в ресторане. Важнее всего — объяснить, какие блюда вы хотите заказать, как их нужно приправить и т. д., а официант должен подать их на стол. Больше сообщать ничего не нужно, и вы довольны.

Тике: С какими языками вам удастся работать с помощью машинного перевода?

Васкес: Cisco — международная компания, поэтому мы работаем со многими языками. На веб-сайте их около 90. Наша команда, специализирующаяся на машинном переводе, работает с языками, которые для нас стратегически важны: испанским, португальским, русским, японским, китайским и французским; в будущем мы планируем расширить этот перечень.

Тике: Корпорация Cisco известна своей длительной работой с SYSTRAN. Вы используете другие системы?

Васкес: Я использую все. В производстве задействовано много систем. Это SYSTRAN и PROMT; на этапе тестирования находятся Moses и несколько прототипов, которые мы настроили. Но основные системы на сегодня именно эти.

Тике: Как происходит выбор системы?

Васкес: Корпорация Cisco купила первый пакет лицензий SYSTRAN приблизительно в 2000 г.; с тех пор эти системы используются для нескольких языков. Moses применяется больше для экспериментов в различных опытных средах. SYSTRAN используется для европейских и азиатских языков; для русского используется PROMT.

Тестированию различных подходов способствует то, что корпорация Cisco является крупным предприятием, с большими интеллектуальными и другими ресурсами, памятью переводов и вычислительными мощностями. Например, на компьютерах Quad корпорации Cisco могут работать несколько виртуальных машин, поэтому можно тестировать множество параметров статистических систем перевода. За четыре часа работы я могу подготовить систему, затем клонировать ее десять раз и тестировать один и тот же корпус с десятью параметрами одновременно.

Все десять параметров можно протестировать за выходные, потому что у нас есть необходимая инфраструктура для измерений.

Такую же вычислительную мощность мы можем использовать для работы SYSTRAN, поскольку очень сложно исключить параметры, ухудшающие перевод. Я не знаю, кто еще это делает.

Тике: Вы сказали, что система Moses используется в качестве эксперимента?

Васкес: Как и остальные участники отрасли, мы надеемся объединить силы двух подходов (на основе правил и статистического), получив своего рода гибрид, или найти параметр, который даст нам возможность обойти возникающие трудности. Например, при использовании статистического машинного перевода (СМП) компьютер обрабатывает контент, а не анализирует, и это создает массу проблем, особенно с непереводаемыми разделами и текстовыми форматами. Поэтому развертывание системы СМП настолько проблематично. СМП может исказить смысл документа, а для

нашего контента это опасно. Нужно внимательно отбирать данные, которые поступают в систему СМП, поскольку программа унифицирует предложения и игнорирует уникальные значения. Простой пример: фраза «ошибка номер пять» переводится как «ошибка номер два» лишь потому, что система обнаруживает больше фраз «ошибка номер два».

Помимо этой конкретной проблемы часто возникают и другие, касающиеся форматирования. Все это затрудняет применение СМП в среде, похожей на нашу. У нас мало файлов, состоящих только из текста (пользовательский контент — еще одна проблема), используется много изображений, т. е. контент смешанный. Можно сказать, что это выходные данные консоли вперемешку с текстом. Система Moses отказывается декодировать такие данные и изменяет значения.

Тике: Похоже, в ней нет встроеной функции определения важности отдельных концепций, например неизменяемости чисел. Такое часто бывает?

Васкес: Не всегда, но достаточно часто, чтобы нам приходилось это проверять. А если абсолютно все нужно проверять и исправлять, машинный перевод теряет свою привлекательность. Скажем, у вас 10 000 документов, и на исправление каждого уходит пять минут — это 20 недель работы. Обеспечить работу подобным образом невозможно. Я ожидаю, что заработает гибридный подход, но все равно приходится настраивать параметры. Это одно из направлений наших экспериментов.

Тике: Расскажите о вашем процессе улучшения качества.

Васкес: Перевод, настройка системы, перевод и снова настройка. Наш контент улучшается постоянно, даже после публикации. Мы всегда переводим его повторно. Частота повторного перевода зависит от текста и колеблется в диапазоне от одной недели до четырех–шести месяцев. Если текст новый (например, новый продукт с новой терминологией), мы делаем это чаще, поскольку единственный способ выловить ошибки — пересмотреть результат. Например, у нас есть продукт Smart Care. Поначалу наша система не понимала, что это название продукта, и переводила его как «разумная забота» — небольшая, но очевидная ошибка, которую требовалось устранить перед выпуском. Еще один сетевой термин переводился обработчиком как «салат ромэн». Компьютерный термин «пингование» переводился как «свист пули». У нас подобрана группа высококвалифицированных сотрудников — терминологов, лексикографов и пользователей машинного перевода, — работающих над двумя различными процессами. Во-первых, они определяют, где система ошибается. В основном для этого выполняются перевод и проверка. Мы можем использовать извлечение терминов, извлечения с учетом статистики или отбор документов, но все равно первым шагом является чтение результатов машинного перевода и обозначение ошибок. Затем начинается второй этап: исправление ошибок. Выявленные ошибки классифицируются и устраняются. После завершения этих двух этапов мы повторяем процесс, чтобы посмотреть, все ли внесенные на втором этапе исправления пройдут контроль качества первого этапа.

Тике: Редактируете ли вы машинный перевод?

Васкес: В нашем рабочем процессе предусмотрено последующее редактирование, но всегда с целью улучшения качества машинного перевода.

Тике: Вы упомянули краудсорсинг. Как пользователи помогают в улучшении систем?

Васкес: Как и любая крупная компания, Cisco рассматривает краудсорсинг как средство накопления знаний за счет работы с сообществами и пользовательскими группами. Проблема в том, как объединить краудсорсинг с машинным переводом и понять, чем машинный перевод может помочь сообществу. Мы осуществляем несколько опытных проектов — не только для того, чтобы люди привыкли пользоваться машинным переводом, но и для предложения лучших вариантов перевода. Моя цель — привлечь постоянных пользователей в качестве лексикографов без их ведома. Мы работаем с высококвалифицированными техническими инженерами, владеющими не одним языком, для которых язык перевода является основным, и предлагаем им пользоваться функцией машинного перевода. Некоторые инструменты позволяют видеть машинный перевод и предлагать альтернативные варианты. Мы получаем информацию и определяем автора текста, таким образом мы узнаем, насколько можно доверять предложенному варианту. Вначале процесс выполнялся вручную, но сейчас пользователи могут настраивать свои параметры. Ре-

зультат действительно неплох, и к нам поступает значительный объем информации для улучшения систем от инженеров технической поддержки, которые работают непосредственно с клиентами. Это также способствует ускоренному внедрению новых систем. За год до запланированного развертывания новой системы — скажем, для тайского языка, — мы создаем базовую программу и предлагаем инженерам, говорящим по-тайски, использовать ее и делиться своими впечатлениями. Им выгодно пользоваться лучшим машинным переводом, который они помогают создавать, а в системе используются результаты лексикографической работы, проделанной специалистами, разбирающимися в терминологии.

Тике: Какой у вас запланирован следующий этап?

Васкес: Системы нового поколения, умеющие распознавать части речи и воспроизводить форматирование исходного текста в переводе, если форматирование соответствующего текста из памяти переводов не будет отвечать необходимому. Это будет переворот. Когда системы машинного перевода смогут выполнять такие действия, они создадут новую среду, способную заменить системы автоматизированного перевода, поскольку будут делать то же самое, но более разумно. Этому дня я жду с нетерпением.

Лори Тике — соучредитель и генеральный директор компании Lexcelera, соучредитель ассоциации «Переводчики без границ» и член редакционного совета MultiLingual.

Повторите, пожалуйста, или DITA для начинающих

Тони Селф (Tony Self)

Объяснить непосвященному, что такое DITA, непросто. Слушатель предполагает, что это продукт или технология, а на самом деле это стандарт и методика — в этом и заключается сложность. DITA — это средство написания технической документации с применением идеальных компонентов: модульности, централизации и многократного использования контента.

Эффективность документов с модульной структурой

Правда было бы неплохо создавать материалы по отдельным компонентам, а затем составлять из них разные документы в соответствии с требованиями? Такой модульный подход к документации эффективен в самых разных сценариях. Это очень практичный метод, поскольку каждый информационный блок нужно написать всего один раз. Изменение в одном компоненте отражается во всех документах, в которые входит этот компонент. Такую возможность часто называют многократным использованием.

Концепция модульной документации не нова. Ее применяли даже до компьютеризации, в эпоху пишущих машинок. Методика написания текстов называлась по-английски STOP (Sequential Thematic

Organisation of Publications — организация публикаций в виде тематической последовательности). В ней компоненты документов являлись «тематическими блоками дискурса»; сегодня такие компоненты обычно называются «темами».

Справочные системы всегда строились на основе тем. По мере совершенствования средств написания в справочных материалах стали интенсивнее использовать возможности модульной структуры. Интернет тоже построен по принципу тем, а неограниченные возможности связывания тем между собой позволяют и в этом случае говорить о концепции модульности.

Тем не менее, в параллельном мире печатных документов модульность не прижилась. В таких форматах документов, как Microsoft Word, основной единицей является документ, а не тема.

Методика DITA включает формат документов и разработана специально для документов модульной структуры. Другими словами, DITA делает модульность действительно простой для всех видов предоставления документов, включая Интернет, справку и печать.

В стандарте DITA используются два основных типа информационных структур: темы и схемы. Схемы — это простые спецификации документа с пере-

числением тем, организованных в иерархию и расположенных в порядке их отображения в документе.

Удобная работа с документами за счет выделения типов информации

В ряде методик составления документации используется идея о сегментации информации в соответствии с ее типом (и предназначением). В основе такой идеи лежит представление о том, что люди читают руководства для получения конкретных необходимых сведений. Иногда нужно найти инструкцию по выполнению той или иной процедуры либо узнать принцип работы устройства. Бывают случаи, когда необходимо посмотреть код для ввода. Редко кто открывает руководство для утоления читательского голода.

Для удовлетворения определенных требований читателя информация делится на три типа: инструкции, сведения о принципах работы и общая информация. В схематическом представлении информации, разработанном в 60-е гг. XX века Робертом Хорном (Robert Horn), использовался ряд типов информации: основные сведения, процессы, процедуры, концепции и структуры. Через несколько лет корпорация Майкрософт применяла в своей документации семь типов: концепции, вопросы и ответы, глоссарии, процедуры, справку, сведения об устранении неисправностей и учебные материалы.

В методике DITA информация делится на три основных типа: задачи (инструкции), концепции (принцип ра-

боты) и справку (общую информацию). Как ни странно, контент большинства руководств и справочных систем легко разбивается на эти три простые категории. Тем не менее, если тот или иной контент не получается отнести к какому-либо из этих трех типов, стандарт DITA дает возможность создавать новые типы информации. В большинстве случаев новые типы разрабатываются в масштабах отрасли или области применения, но их можно создать и самостоятельно.

Разделение информации на типы в методике DITA также помогает согласовывать контент с использованием перекрестных методик написания технической документации. Это возможно благодаря созданию документов по определенным правилам. Например, при документировании задачи нужно включить как минимум один шаг. Если его не включить, тема не сохранится! В свою очередь, принудительное применение правил написания стало возможным потому, что DITA — формат документов на основе языка XML, предназначавшегося для этой цели. Проверка документа на соответствие семантическим правилам в XML называется валидацией.

Создание качественной документации в рамках методике DITA по-прежнему во многом зависит от квалификации автора, но типы информации и валидация каждый раз упрощают достижение правильного результата.

Централизация с помощью семантической разметки

Термин «централизация» можно понимать по-разному. По сути, большин-

ство людей согласится с тем, что он означает использование одного источника контента для создания разных продуктов. Это расширенная идея о документах модульной структуры с включением различных режимов предоставления; один и тот же контент может появиться не только в разных публикациях, но и на совершенно разных носителях. Инструкции размещают в печатном руководстве пользователя, в разделе справки, на веб-странице и в файле ePub. В кратком руководстве они могут находиться на втором уровне, а в руководстве администратора — на пятом.

Для упрощения централизации в контент нельзя добавлять инструкции по форматированию. Бесполезно определять название темы как «Заголовок 2», если в другой публикации может понадобиться указать ее как «Заголовок 4». Кегль шрифта нельзя определять равным 12 пунктов, если текст будет отображаться на экране мобильного телефона, где 12 пунктов — слишком много. DITA дает возможность обходить такие потенциальные препятствия и эффективно осуществлять централизацию с помощью семантической разметки. Разметка текста осуществляется не на основе форматирования, а на основе вида текста. Названия размечаются как названия. Необходимые условия определяются как необходимые условия. Шаги — как шаги и т. д.

Семантическая разметка позволяет разделять контент и форму. Форма (или стиль) добавляется на основе правил семантической разметки схемы со стилями представления во время процесса публикации. Процесс публикации ав-

томатизирован; в определенном смысле он выполняется одним нажатием кнопки мыши после создания правил схемы для организации. Нужен PDF? Выбираем PDF. Нужен ePub? Выбираем ePub. Нужна справка Eclipse? Выбираем справку Eclipse.

Одной из проблем для специалистов по составлению технической документации является перенесение внимания с формы на контент. Возможно, это звучит просто, но для перехода с написания по стилям на написание с семантической разметкой нужно уметь переключаться. Преимуществ много: за счет автоматизации форматирования можно уделять больше времени подбору слов и фраз, а не шрифтам, выравниванию и нумерации! Таким образом, повышается качество и улучшается согласованность текста.

DITA принадлежит сообществу

Вы в восторге от стандарта DITA? Он экономит время, помогает повысить качество выпускаемых документов, высвобождает время для написания и облегчает профессиональную деятельность. Так где же купить этот стандарт?

Это еще один удивительный момент. DITA — стандарт, но не продукт. Более того, это стандарт с открытым исходным кодом. А значит, его можно получить бесплатно — то есть совершенно даром и безвозмездно. Стандарты с открытым кодом создаются и поддерживаются сообществами, а не отдельной компанией. Стандарт DITA «принадлежит» (если это слово здесь уместно) сообществу составителей технической документации. Им

управляют некоммерческий комитет по стандарту под названием OASIS и группа добровольцев из технического комитета OASIS DITA.

Для внедрения стандарта DITA нужны средства написания текстов с его поддержкой. Поскольку DITA — стандарт с открытым кодом, можно выбирать из десятков таких средств, включая FrameMaker, Arbortext, XMetaL, oXygen, Serna, XHE, DITA Storm, Xorus и многие другие. При желании можно даже использовать поочередно два редактора в процессе работы над темой! Но обилие вариантов может сбить с толку кого угодно, особенно новичка. Если стандарт DITA бесплатный, то коммерческие средства написания текстов с поддержкой DITA — нет. Цены на них разные, от менее 100 до более 1000 долларов США.

Разделение контента и формы в стандарте DITA влечет за собой использование разных типов средств для написания и публикации материала. Для потоков операций с применением стандарта DITA может понадобиться не одно средство, а два или три.

Множественное использование

Один из моих любимых принципов написания текстов по стандарту DITA — ХОР, или хватит одного раза. Почти все функциональные возможности DITA предназначены снизить рабочую нагрузку, что можно сделать, например, устранив повторные действия. После написания тех или иных фраз, блоков или тем (будь это название продукта или компании, предупреждение, последовательность шагов, тема или глава)

больше не нужно писать их снова. DITA предусматривает огромное количество механизмов многократного использования контента. У некоторых из них интересные названия, например «виртуальное включение» (transclusion) и «косвенное обращение» (indirection), но за неизвестными терминами скрывается гениальная простота функций. В среде для написания текстов встречаются переменные, но функции многократного использования в стандарте DITA — это как переменные на стероидах!

Можно многократно использовать фрагмент контента DITA любого типа: абзацы, примечания, фразы, термины, схемы, элементы индекса и целые темы. Это означает, что идея о документации модульной структуры шире, чем просто многократное использование тем в различных публикациях — это многократное использование чего угодно, что пришлось бы повторно набирать или копировать. Многократное использование значительно облегчает обновление контента, поскольку изменения нужно вносить всего лишь раз.

Необходимость в системе управления контентом компонентов

После внедрения документов модульной структуры, интенсивной централизации и многократного использования — неотъемлемых составляющих стандарта DITA — могут возникнуть проблемы с управлением контентом. Как узнать, что кто-то из вашей команды уже написал похожую тему? Где хранятся переменные с названием продукта? Кто

из вашей команды написал ту или иную тему? В какие публикации она вошла?

Такие задачи управления решаются с помощью программного средства, а именно системы управления контентом компонентов. Данная система умеет работать с компонентами документов модульной структуры.

Сложный ли стандарт DITA?

Стандарт DITA часто называют запутанным и слишком сложным. Современный стандарт DITA содержит свыше 500 семантических элементов... неужели можно запомнить, для чего они все предназначены? DITA во многом отличается от более ранних подходов к созданию документации. Эта разница препятствует внедрению.

Для использования преимуществ DITA нужно переосмыслить подход к документации. Необходимо понять принципы разделения контента и формы и быть готовыми забыть о форме! Нужно писать, следуя стандартам сверхорганизации, и принимать идеалы открытого исходного кода, отказаться от философии «одно средство для всего» и работать с набором соответствующих средств, изучить предназначение небольшого количества элементов семантической разметки (гораздо меньше 500) и случаи их применения. Требуется рассматривать проект написания документации как часть библиотеки, а не как отдельную публикацию, работать на более высоком уровне детализации и понимать, каким образом повторное использование элементов такого уровня упрощает жизнь.

Иногда даже нужно изучить основы XML, но в действительности все зависит от выбранных инструментов и вашего интереса к XML.

Некоторые специалисты считают, что методика DITA ограничивает, поскольку в ней множество правил, стандартов и проверок; такие люди уверяют, что DITA душит творчество. На мой взгляд, все как раз наоборот. Возьмем поэзию хокку. Здесь тоже множество правил, касающихся слогового ударения, фраз и размера. Но разве можно сказать, что хокку душит творчество? DITA, как и хокку, ему способствует.

Что такое DITA?

DITA — это методика и открытый стандарт на основе XML, поддерживаемый сообществом составителей технической документации. DITA дает возможность применять оптимальные методы написания текстов, например модульность, централизацию и повторное использование контента, главным образом за счет разделения контента и формы.

Методика используется для автоматизации процесса публикации и уменьшает рабочую нагрузку на составителей документации. Хотя стандарт DITA бесплатный, средства для авторов и издателей нет. Для работы со стандартом DITA могут понадобиться разные инструменты. Вначале все может показаться сложным, но после усвоения базовых понятий работать неожиданно становится проще. Наконец, использование стандарта DITA в соответствии с первоначальным за-

мыслом повышает качество текста и снижает издержки.

Кстати, забыл сказать. DITA — это аббревиатура от английского Darwin Information Typing Architecture (архитектура типизированной информации Darwin). Но знать это вовсе необязательно!

Тони Селф на протяжении 30 лет работает в сфере разработки документации. В 1993 г. он основал компанию HyperWrite, которая специализируется на консультациях по гипертексту. Тони Селф также читает лекции по обмену технической информацией и цифровой грамотности в Университете Суинберна, г. Мельбурн, Австралия.

Семь стратегий для судебных лингвистов

Дена Бугель-Шунра (Dena Bugel-Shunra)

Срочность и важность юридического перевода достигают пика, когда устных и письменных переводчиков вызывают для оказания помощи при рассмотрении дела в суде. К лингвистам часто обращаются после ходатайства о предоставлении документов, когда дело доходит до разбора вороха рукописных заметок, которые уже не помещаются в комнате, или архива электронной почты на языках, на которых адвокаты не в состоянии их прочесть. И это как раз то доказательство, от которого зависит все разбирательство, именно на нем будут сосредоточены показания при посредничестве квалифицированного переводчика.

Правильный выбор лингвистов для таких случаев весьма проблематичен из-за спонтанного характера таких команд, поскольку адвокаты разбираются

в вопросах права и стратегии, а не в лингвистических аспектах. Задача менеджеров проектов, которые должны подобрать лингвистов в команду юристов, куда сложнее, чем обычно. Существует семь стратегий, позволяющих добиться отличного взаимопонимания между лингвистом и юристом и сохранить прозрачность в языковых вопросах.

Во-первых, начните с поиска во всех подходящих местах. Чтобы определить хорошую кандидатуру для своей команды, рассмотрите несколько кандидатов. Будет полезно обратиться к спискам таких ресурсов, как базы данных Американской ассоциации переводчиков (ATA) и *MultiLingual Resource Directory*, но когда речь идет о работе в судах, рекомендуется также включить в круг поиска Национальную ассоциацию судебных переводчиков (NAJIT). Реестр федеральных сер-

тифицированных лингвистов доступен только для персонала суда, но звонок в офис секретаря часто дает возможность получить несколько рекомендаций. Дополнительные ресурсы см. во врезке.

Во-вторых, задайте насущные вопросы. Сравнивая квалификацию нескольких лингвистов, найдите время, чтобы с ними побеседовать: если возможно — лично, если нет — по телефону.

Вы же не хотите, чтобы ваша дача показаний оказалась дебютом ваших устных переводчиков. Расспросите о предыдущем опыте, в том числе обо всех аспектах, которые, вероятно, возникнут в данном случае. Могут ли они работать с рукописными материалами? А с каракулями в скорописи? А как насчет аудиозаписей, записей телефонных разговоров и видеоматериалов? Насколько знакома переводчику тематика дела?

Полезно спросить, есть ли у лингвиста напарник по устным или письменным переводам для таких случаев. Это важно знать как для проверки опыта лингвиста, так и для определения его способности работать в команде. Для выполнения обсуждаемого здесь вида судебной работы требуется сочетание умений и профессиональных навыков: в области устного (разговорная речь) и письменного (письменные материалы) перевода, а также в разного рода смежных областях. Весьма полезно уточнить наличие опыта в каждой из этих сфер.

В-третьих, проверьте квалификацию и рекомендации. Конечно, все мы просим предоставить рекомендации. Правильно? Увы, это делается не всегда. Насколько бесед с удовлетворенными быв-

шими клиентами помогут «убедиться в квалификации».

Профессиональную квалификацию проверить легко. Судебная администрация в любом штате может, по требованию, выдать сертификат устного судебного переводчика, который подтверждает профессиональную репутацию. Если сертификат был выдан через ATA (на многих языках) или NAJIT (только на испанском языке), за репутацию лингвиста будут ручаться эти организации.

В-четвертых, предоставьте данные об обстоятельствах дела. Перед тем как лингвисты приступят к самому первому переводу, убедитесь, что они совершенно четко представляют себе, что именно они могут разглашать на законных основаниях по сути дела. Лучше всего разобраться с этим до того, как будет сделан окончательный выбор кандидатуры для вашей команды, и это не просто вопрос тонкостей профессионального жаргона. Язык компьютерных наук, например, очень отличается от языка архитектуры и разработки приложений. Если дело связано с алгоритмами, будет использоваться другой словарный запас и формулировки, чем в деле о разработке программного обеспечения и маркетинге. Профессиональные лингвисты будут изучать материал уже с конкретными данными на руках. Иногда они составляют глоссарий терминов, а иногда готовят ресурсы, чтобы собрать примеры использования. Обеспечивая максимум возможностей сделать все так, как нужно, предоставьте максимальное время для подготовки.

В-пятых, обратитесь к книге *The Checklist Manifesto: How To Get Things Right* («Манифест о контрольном списке:

как сделать так, чтобы было, как нужно») хирурга Атула Гаванде (Atul Gawande), чтобы узнать о недавней революции в области безопасности хирургической практики. Автор выдвигает и излагает теорию о том, почему квалифицированные люди в критической ситуации допускают дорогостоящие ошибки. Более тщательная подготовка не избавляет от этих ошибок, объясняет он. Однако использование контрольного списка позволяет выявить и резко сократить их количество.

Когда этот принцип был применен на практике, Гаванде сам был удивлен, обнаружив, что в хирургии контрольные списки помогли операционным бригадам снизить послеоперационную смертность на двузначное число в каждой больнице, где они были введены. Аналогично, при юридическом переводе использование контрольных списков в процессе изучения дает возможность избежать ситуации, когда срочность может начать преобладать над проверкой материалов в ущерб конечному результату.

При отборе переводчиков создайте контрольные списки для проверки этапов по всем типам материалов. Если лингвист говорит вам, что это не нужно, вспомните о словах Гаванде о двузначном показателе снижения уровня смертности в послеоперационный период и найдите другого лингвиста.

Контрольный список, подходящий для вашей команды и ситуации, так же уникален, как и сама команда. Вот несколько примеров контрольных списков для двух типичных ситуаций, связанных с устным и письменным переводом.

Дача показаний с помощью видеоконференции

- Введение и форма адреса заполнены (в том числе судебным секретарем).
- Переводчик выполняет устный перевод и знаком с языком и диалектом.
- Справочные документы получены и прочитаны.
- Цель дачи показаний ясна всем членам команды.
- Готова стратегия консультаций не для протокола.
- Эффективность переводчика максимальна (языковые блоки легко усваиваются).
- Проведена проверка оборудования.
- Закрыты все дополнительные открытые вопросы.

Изучение документов

- Введение.
- Переводчик знаком с языком, диалектом и носителями информации (рукописные данные / микрофиши / программное обеспечение / аудио- и видеозаписи).
- Типы требуемой информации понятны всем членам команды.
- Основная предметная область относится к компетенции переводчика.
- Доступны глоссарии, аналогичные документы и специалисты по терминологии.
- Тактика расследования ясна (контекст, который может быть разглашен, отсутствует, косвенно упомянут, процитирован с указанием источника).
- Масштаб, график и сроки.
- Закрыты все дополнительные открытые вопросы.

Крайне важен также завершающий этап — убедиться, что были получены ответы на все дополнительные вопросы. Он должен быть включен во все предварительные контрольные списки проекта. Как ни удивительно, есть вещи, о которых люди хотят знать, но стесняются спросить напрямую.

В-шестых, избавьтесь от прозрачности. Считать переводчика «прозрачным» так заманчиво. Устные переводчики особенно хороши, когда они «незаметны», и передают поступающую к ним информацию без каких-либо собственных комментариев. Во многих случаях это свидетельствует об эффективном использовании навыков лингвиста, но в условиях судебного разбирательства межкультурный опыт переводчика может дать критически важную оценку. Как можно получить такую оценку?

В этом случае также будет полезным опыт медицины. В книге *Checklist Manifesto* Гаванде рассматриваются шаги, предпринимаемые бригадами перед входом в операционную. Один из них — предоставление слова каждому члену бригады, включая медсестер, врачей и других членов команды. Каждый должен назвать свое имя и рассказать о своих ожиданиях в связи с операцией. На этом предоперационном этапе каждый из членов бригады получает право высказать свое мнение. Гаванде поражает та информация, которая обнаруживается в таких разговорах, информация, которая осталась бы неизвестной, если бы медсестры «знали свое место» и промолчали бы, потому что у врачей должны быть ответы на все вопросы.

Лингвисты обучаются «невидимости», и для отказа от нее требуется как раз-

решение, так и подходящая обстановка, позволяющая высказывать свое мнение всем членам команды. Платой за такой подход вполне может стать та информация, которая и определит исход дела.

Наконец, организуйте возможность получать отзывы на месте. Команда работает вместе — лингвисты делают свое дело среди юристов и их помощников. Участвуют все. Помимо перевода и изучения документов команда интересуется мнением лингвиста и получает это мнение. Будет ли она работать идеально? Нет, если в этом участвуют люди.

Разница между эффективной и неспособной системой становится очевидной сразу же, когда возникают проблемы. Как команда справляется с ошибками и недоразумениями? Как может наиболее эффективно учиться под давлением?

Наличие предопределенного процесса для сглаживания неизбежной кривой обучения способствует активной командной работе. Обсуждение в конце каждого заседания помогает прояснить ожидания и способствует повышению лингвистической точности и культурной ясности. Альтернатива команде, которая должна учиться на ошибках, — команда, которая никогда не учится на ошибках. Если лингвистическая работа выполняется в режиме реального времени и для критически важных задач, лучше подготовить почву для обучения.

Заблуждение о вавилонской рыбке

В популярной научно-фантастической трилогии *Автостопом по Галак-*

тике автор Дуглас Адамс навсегда испортил мир для всех лингвистов. Он написал о рыбке, которую можно поместить в ухо и которая будет безупречно переводить с любого языка. По его словам, она питается мозговыми волнами, поглощая смысл и намерения на одном языке и испражняясь ими непосредственно в мозг на другом. Неплохо было бы иметь такую рыбку в ухе?

Ни один человек-переводчик не сможет быть настолько эффективным. Чтобы приблизиться к идеалу и понять все, что можно найти в суде, — от рукописных заметок, записей телефонных разговоров и банковских выписок в кассовой книге до нюансов языка дачи показаний — нам нужна хорошая специально подобранная команда. К сожалению, вымышленная рыбка не может выполнить эту работу. А вот удачно подобранная команда, в которую входят уважаемые профессиональные лингвисты, будет работать великолепно.

Ресурсы и справочники

1. Реестр программ сертификации для судебных переводчиков доступен в

NAJIT по адресу: www.najit.org/certification/certification.php.

2. Американская ассоциация переводчиков предлагает индексированный реестр письменных и устных переводчиков по адресу: www.atanet.org/onlinedirectories.
3. Список ресурсов LLRX устного и письменного перевода для системы уголовного правосудия доступен здесь: www.llrx.com/features/interpreters.htm.
4. Данные о консорциуме Consortium for Language Access in the Courts вместе со ссылками на координаторов устных штатных переводчиков во всех штатах можно найти здесь: www.ncsconline.org/D_Research/CIconsortContactspage.html.
5. Каталог *MultiLingual Resource Directory* доступен на странице www.multilingual.com/resourceDirectory.php.

Дена Бугель-Шунра работает переводчиком и писателем с 1989 г. Она является зарегистрированным судебным устным переводчиком для пары языков «иврит – английский» в Вашингтоне, и была членом как хороших, так и плохих судебных переводческих команд.

Многоязычная Европа: проблемы языковых технологий

Джордж Рехм (Georg Rehm) и Ганс Ускорайт (Hans Uszkoreit)

Владение одним родным языком, будь то латвийский, венгерский или португальский, не должно стать социальной или экономической преградой в объединенном европейском сетевом информационном обществе двадцать первого века. Многие европейские языки рискуют оказаться жертвами цифрового века из-за того, что они недостаточно представлены и не обеспечены надлежащими ресурсами. Из-за языковых барьеров огромные возможности региональных рынков по-прежнему остаются нереализованными.

Возможным решением этой крайне важной проблемы могут стать языковые технологии при условии их надежности, рентабельности и доступности для всех европейских языков и всех жителей Европы. Однако для достижения этих целей необходимо ускорить научно-исследовательские работы, дав им мощный целенаправленный толчок.

Ключом к защите необычайно разнородной группы из более чем 60 европейских языков и содействию ее сохранению являются языковые технологии, хотя интенсивность их использования для разных языков в Европе существен-

но отличается. Исследования в этой области за последние несколько лет дали значительные результаты, но, к сожалению, текущий темп научно-технического прогресса слишком медленный, чтобы отразиться в реальных программных продуктах, которые могли бы обеспечить большие успехи в развитии коммуникаций в многоязычной среде в течение последующих 10–20 лет. Эти широко используемые базовые технологии обычно одноязычные и доступны лишь для некоторых языков. Сетевые службы, такие как Переводчик Google или Bing Translator, удобны, если необходимо получить общее представление о содержании документа, однако эти приложения создают массу сложностей. Такие приложения, как языковые или голосовые пользовательские интерфейсы или диалоговые системы, используются только в узкоспециализированных областях, а их технические возможности весьма ограничены. А для технологий, которые потенциально могут быть связаны со спасением жизни, например технических средств, используемых для спасательных операций, и робототехники в области здравоохранения, точность перевода вообще сложно переоценить.

Общие сведения о META-NET

META-NET — это европейская сеть передовых технологий, состоящая из 44 расположенных в 31 стране центров исследования языковых технологий и организованная Технологическим альянсом многоязычной Европы (META) благодаря согласованным усилиям по созданию крепкого и сильного европейского общества, заинтересованного в языковых технологиях, которые смогут его объединить. Ее цель — подготовка базы для приложений на множестве различных языков, которая даст возможность выполнять автоматический перевод, управлять информацией и знаниями, в том числе локализацией, а также производить контент и приложения в связанных областях на всех европейских языках. Задача сети META-NET, деятельность которой началась 1 февраля 2010 г., — активизировать исследования языковых технологий как средства реализации концепции Европы, объединенной в единый цифровой рынок и информационное пространство. Выполнение этих задач будет осуществляться по трем направлениям: META-VISION (разработка концепции), META-SHARE (общий доступ) и META-RESEARCH (исследования).

Направление META-VISION связано с задачей, которая не только важна, но и стратегически необходима для общего успеха инициативы. Эта задача — формирование слаженного и однородного общества путем объединения представителей сильно раздробленных и разнородных групп заинтересованных лиц. По нашим оценкам, уже сейчас нам удалось привлечь свыше 2500 спе-

циалистов в области языковых технологий и изложить им задачи META-NET. Например, на конференции Translingual Europe 2010 (7 июня, Берлин) исследователи обсуждали текущие проблемы и концепции с представителями компаний из различных отраслей (такими, как корпорация Майкрософт, Asia Online и ProMT), а также с пользователями языковых технологий и машинного перевода (European Patent Office, Symantec, EC DGT). Основным мероприятием сети META-NET в первый год ее существования стала конференция META-FORUM 2010 (17–18 ноября, Брюссель, Бельгия). Более чем 250 участникам META-FORUM были продемонстрированы первые результаты процесса формирования концепции.

Вторая важная задача META-VISION — это коллективная подготовка, формирование и популяризация (в пределах сообщества и его силами) текущих исследовательских стратегических задач, которые должны будут стать долгосрочным инструментом, объединяющим под одной крышей промышленные и академические исследования и разработки в период до 2020 г. Это предполагает составление наиболее общих рекомендаций и предложений в отношении совместных действий, которые должны быть представлены Европейской комиссии, государственным и региональным органам и финансовым учреждениям.

Сеть META-NET также занимается созданием METASHARE, надежной одноранговой сети хранилищ языковых данных, инструментов и веб-служб, документированных с помощью высококачественных метаданных и объеди-

ненных в реестры, обеспечивающие единообразный поиск и доступ к ресурсам. Доступ к данным и инструментам может быть как открытым, так и ограниченным, платным или бесплатным. Сеть METASHARE будет охватывать не только существующие, но и новые, и возникающие языковые данные, инструменты и системы, необходимые для создания и оценки новых технологий, товаров и услуг. В связи с этим решающую роль играет повторное использование, сочетание, видоизменение и реструктуризация языковых данных и средств.

Полностью рабочий прототип META-SHARE был представлен на первой ежегодной конференции META-FORUM 2010. К важным и актуальным компонентам инфраструктуры относится универсальная схема метаданных для описания языковых ресурсов и языковых технологий. В данный момент ведется подготовка такой схемы метаданных и ее обсуждение в рабочей группе, состоящей из специалистов, которые участвуют в реализации данной инициативы, и нескольких других европейских экспертов. В то же время тщательно изучаются перспективы лицензирования языковых ресурсов, и при поддержке специалистов в области права уже был принят и подготовлен первый пакет шаблонов лицензирования. META-SHARE способствует развитию движения за открытые данные и инструменты с открытым исходным кодом и подстраивается под это движение, особенно в рамках инициативы Creative Commons.

Основная задача META-RESEARCH — учет дополнительных семантических аспектов в машинном переводе, оптимиза-

ция и распределение труда при гибридном машинном переводе, подготовка эмпирической базы для машинного перевода и использование контекста при выполнении автоматического перевода с помощью вычислительных мощностей. С этой целью в сети META-NET проводятся исследования, которые дают возможность установить связь с другими областями и дисциплинами, такими как обучение работе с компьютерами и сообщество Semantic Web. Направление METARESEARCH связано со сбором данных, подготовкой наборов данных и языковых ресурсов для проведения оценки, составлением реестров инструментов и методов, а также организацией семинаров и проведением курсов по повышению квалификации для персонала. К основным достижениям в этом направлении относятся определение проблем машинного перевода, семантических аспектов, которые могут оказать положительное влияние на уровень науки и техники, формирование рекомендаций по решению проблемы внедрения семантических данных в машинный перевод и составление списка инструментов и ресурсов, которые можно было бы использовать для этой цели. Сейчас проводятся заключительные работы по созданию нового языкового ресурса для машинного перевода, Annotated Hybrid Sample MT Corpus. Он будет предоставлять данные для языковых пар английский — немецкий, английский — испанский и английский — чешский. Третий важный результат — это программное обеспечение для сбора сводов скрытых веб-документов на различных языках. Средство объединяет новостные статьи на различных языках, в которых обсуждаются сходные вопро-

сы или события, и создает кластеры страниц, которые распознаны как взаимный перевод. Проводимые в этом направлении исследования призваны добиться существенного повышения уровня научно-технического прогресса в области машинного перевода.

Расширение и влияние

В рамках данной инициативы создан учредительный консорциум, состоящий из 13 партнеров в десяти странах. Однако META-NET функционирует на европейском уровне. Именно поэтому сеть была расширена в ноябре 2010 г. В расширенную сеть входит 44 партнера из 31 страны. Большинство новых членов участвует в трех проектах, финансируемых ЕС. Основная их задача заключается в содействии достижению целей META-NET путем систематического сбора языковых ресурсов и технологий, их отбора и описания с помощью записей метаданных и обеспечения их доступности с помощью META-SHARE, привлечения общественности в соответствующих странах и организации ознакомительных мероприятий общего характера. Реализация трех проектов началась 1 февраля 2011 г., спустя ровно год с момента начала функционирования сети META-NET.

Цель META-VISION заключается в предоставлении долгосрочного плана по выполнению основной задачи — созданию по-настоящему многоязычной Европы. Создав большое и разнородное сообщество для реализации поставленных целей, ассоциация META способствует достижению уровня критической массы, необходимого для существенного изме-

нения подходов к применению языковых технологий для обеспечения и сохранения всего многообразия языков в будущем Европы. Для исследователей, технологов, профессионалов и администраторов, разрабатывающих, предоставляющих или использующих языковые технологии, META предлагает уникальную возможность оставаться в курсе последних событий, выдвигать идеи или отстаивать потребности определенного языка, участвуя в дискуссиях специалистов, рабочих группах и планировании деятельности, от которой будет зависеть лингвистическое будущее Европы. Сеть META-SHARE, обеспечивающая доступ к пулу языковых ресурсов и технологий, что способствует планированию и проверке услуг META, приглашает к сотрудничеству авторов исследовательских проектов и проектов технологического развития. Также META предлагает коммерческим предприятиям поддержать ее концепции в отношении товаров и услуг, принять участие в процессе планирования и использовать META для налаживания выгодного партнерства. Учебные заведения и педагоги, журналисты и СМИ, политики и общественные учреждения и организации также приглашены к участию в открытых обсуждениях концепции и подходов к созданию по-настоящему многоязычного информационного общества. Чтобы получить дополнительные сведения, посетите веб-сайт www.meta-net.eu.

Джордж Рехм работает главным консультантом в компании DFKI и является менеджером сети и проектов META-NET.

Ганс Ускорайт — научный директор лаборатории LT Lab при DFKI и профессор вычислительной лингвистики в Саарском университете.

Происхождение (отрывок)

Диана Абу-Джабер



Посвящается Скотти

Глава 1

На четвертом этаже я выхожу из лифта и сразу замечаю посетительницу: она сидит напротив кабинета на складном металлическом стуле.

Блестящие каштановые волосы выбиваются из пучка, кожа испещрена веснушками, лицо сосредоточенно-серьезное.

За моей спиной дверцы лифта закрываются.

Потерпевшие для меня существуют в другом измерении — в области теории. Потерпевшими занимается полиция, а мы — исследованиями. Имей я желание встречаться с потерпевшими, не стала бы дактилоскопистом.

Старательно пряча глаза, я проскальзываю в кабинет. Элис, моя начальница, многозначительно смотрит на меня.

— Привет, Лена...

Женщина пробирается в кабинет, идет между столами — высокая, бледная, она пугает внутренним напряжением, которое, в моем понимании, возникает от большого горя. Не успев добраться до рабочего места, слышу ее голос:

— Вы Лена? Лена Досон?

Я вздрагиваю.

Элис встает из-за стола. Незнакомке она от силы по плечо, зато буквально из-

лучает целеустремленность и воинственный пыл.

— Мисс, послушайте! Я не знаю, как вы сюда попали, но в нашу лабораторию посетители не допускаются. Вам ведь уже объясняли...

Женщина подходит слишком близко — лицо бледнее полотна, голос дрожит, — и сначала я с трудом разбираю, что она бормочет. Отступаю к столу, но незнакомку не останавливает и это — она шагает следом.

— Меня зовут Эрин Коган. Мой малыш... Он умер пять недель назад, а полиция не сделала ничего. Абсолютно ничего... — Незнакомка частит, готовая к тому, что ее выгонят, и хватается мою руку. Высокий голос электрическим эхом вибрирует в ушах. — Пожалуйста, Лена... Мисс Досон... Говорят, вы можете...

Еще одна моя коллега, властная Марго, врывается в кабинет вместе с Эдом Уэлмором, который, вероятно, уже собрался идти домой после ночной смены — верхняя пуговица на форме расстегнута, подмышками темнеют полукружья пота.

— Ну, все! — решительно говорит он, войдя в комнату. — Миссис Коган, вам пора.

Эрин Коган выпускает мою руку, но взгляда не отводит.

— Пожалуйста, мисс Досон, пожалуйста...

Эд подходит к незваной гостье. Роста он небольшого, зато крепко сбитый. Упершись руками в бока, смотрит на меня поверх головы Эрин.

— Выходите немедленно!

Она поворачивается сначала к нему, потом ко мне. В голосе звучит столько страданий и ужаса, что я не выдерживаю. Эту женщину я вижу впервые, однако хорошо понимаю, что она чувствует. Ее отчаяние вызывает страх, почти такой же сильный, как жалость. Руки Эрин судорожно стиснуты, даже суставы побелели...

— Ну ладно, ладно. — Пытаясь привести в порядок дыхание, я упираюсь в свободный участок рабочего стола. — Мисс... миссис Коган, так? Пойдемте, я вас провожу.

В кабине лифта Эд многозначительно смотрит в угол — ему не нравится, что я покинула лабораторию. Элис тоже с нами едет: руки скрещены на груди, очки подняты на лоб, глаза буравят несчастную Эрин. Понятно, меня ждет нагоняй — нечего поощрять идиоток, нужно стать тверже, научиться говорить «нет», и так далее, и тому подобное.

Эрин Коган заламывает руки и смотрит только на меня.

— Я ждала у кабинета с шести утра. Упросила вахтера... Извините, я просто не знаю, что еще предпринять. Никто не желает говорить со мной о деле Мэтью! Я схожу с ума. Мой малыш, мой Мэтью умер, и никто не желает со мной говорить!

— Лена, к ним посылали следователей и двух экспертов... — раздражен-

но напоминает Эд, глядя на меня через плечо Эрин. Ясно, еще немного, и он сорвется.

Лифт останавливается, но выходим мы не сразу.

— Я не поняла... — Мой голос похож на карканье, и я откашливаюсь. — Мисс Коган, я не совсем поняла, что вы от меня хотите.

Придерживая дверцу лифта, Эд выводит нас из кабины. Эрин явно испугана, ее взгляд затравленно мечется между мной и Элис.

— Вы же эксперт, верно? Вы можете найти улики... По крайней мере, так говорят. Дескать, вы даже лучше полиции!

Элис закатывает глаза.

— Ну, это вряд ли! — Я категорично качаю головой. — В подобных делах улик, как правило, не бывает, естественно, в зависимости от причины смерти... — Осекшись, я с тревогой поворачиваюсь к Элис. Та, устремив взгляд куда-то вдаль, потирает высокую скулу. — От чего, по мнению патологоанатомов, наступила смерть?

— От синдрома внезапной детской смерти, — с горечью отвечает Элис. — Другими словами, следователи не знают, в чем дело; формулировка обтекаемая. — Она многозначительно поглядывает на Эда.

— Миссис Коган, — говорит он, — лаборатория — не полицейское управление, вам нельзя здесь находиться! Пора домой.

— Уже полчаса назад было пора, — уточняет Элис.

Увы, Эрин Коган видит и слышит только меня.

— Пожалуйста! Можете мне не верить, считайте сумасшедшей! Но даже если так, выслушайте! Я знаю, моего малыша убили. — Она наклоняется вперед. — Я только прошу... Нет, умоляю... Умоляю, ознакомьтесь с нашим делом...

Ненавидя себя за слабость, спрашиваю:

— У вас есть доказательства, что это не внезапная детская смерть?

Эд задумчиво чешет затылок. Эрин доверчиво опускает голову, и ее лицо из-за отблесков уличного света и возбужденного сияния глаз принимает то ли дикое, то ли безумное выражение. Словно облаком горячего пара, она опалает меня своим объяснением.

— В доме кто-то был! Я смотрела сериал на первом этаже, как вдруг над головой закричали половицы. Кто-то проник в дом и убил малыша. Мэтью спал наверху, и я внезапно услышала этот звук! Услышала и решила, что показалось. Я так устала... С ребенком порой устаешь, и хочется немного отдохнуть. Мне ведь никто не помогает... Муж целыми днями на работе, поэтому... — Она не договаривает. Ее взгляд становится рассеянным, устремившись сначала на пол, потом на меня. — У вас есть дети?

Элис с шумом выдыхает.

— Нет, — отвечаю я.

Миссис Коган хлопает ресницами, словно я отвесила ей пощечину, и шепчет:

— Простите...

Ей на плечо ложится твердая рука Эда.

— Мадам, отдел медицинской экспертизы обязательно примет к сведению вашу версию. Не сомневайтесь,

они сделают все возможное. — Доброта в голосе Эда смешивается с сильным раздражением.

Эрин подходит ко мне слишком близко, и теперь ее возбуждение действует подобно статическому электричеству. Позади нас большие окна фойе, за которыми белой пеленой висит снегопад.

— Нам обеим известно, — твердит Эрин, словно безумная. — Нам обеим известно... — Она с излишним рвением смахивает пыль с рукава пальто, и я впервые замечаю: оно дорогое, вероятно, кашемировое. — Округу это неинтересно, полиции — тем более. Я для них ничто, истеричная мать, а это еще хуже, чем ничто, ведь правда, правда?

Элис с Эдом делают каменные лица.

Эрин поворачивается ко мне, и ее голос звенит от напряжения.

— Мой муж, Клей, работает инженером-строителем и знаком со многими чиновниками, даже с Робом Каммингзом, они вместе играют в гольф в загородном клубе Онондага. После той... трагедии мы сначала ждали результатов от полиции. Не получив ровным счетом ничего, Клей начал наводить справки и каждый вечер дома говорил: «Лена Досон... Лена Досон». Дескать, она наверняка что-нибудь найдет, особенно в нашем деле, ведь ребенок погиб! Лена с уликами чуда творит... — Она снова пронзает меня горящим взглядом. — Патологоанатом твердит: подобное изредка случается. Мол, грудные дети слабые и порой умирают во сне... Но ведь такое случается далеко не всегда, верно? Мэтью было шесть месяцев. Совершенно здоровый, красивый... очень красивый. А теперь он убит, а тот, кто это совершил, до сих пор жив... — Эрин показывает на

дверь. — Он свободно разгуливает где-то там... Представляете, что я чувствую? Что чувствую, думая об этом? — Миссис Коган снова хватается меня за руки, выворачивает пальцы, и я едва не морщусь. Мертвенно-бледное лицо близко, слишком близко...

Эд силой оттаскивает ее назад.

— Довольно! — Он начинает выпроваживать Эрин из фойе, но она, громко вскрикнув, отталкивает охранника и впивается мне в запястья. Я не могу шевельнуться от ужаса, в крови бурлит адреналин.

Зрачки Эрин сжимаются, а потом между нами втискивается Элис и истошно кричит:

— Отпустите! Вы же делаете ей больно!

Эрин громко всхлипывает. Бессильно опустившись на корточки, она продолжает цепляться за мою кисть, и ее массивное обручальное кольцо царапает мне пальцы. Судорожно глотая воздух, я стараюсь вырваться.

— Все, довольно! — не унимается Элис.

Эрин сама разжимает пальцы. Голова опущена, руки спрятаны за спину.

— Простите, простите, простите, — скороговоркой лепечет она.

В главную дверь кто-то входит, и у меня появляется надежда: наверное, Чарли пришел меня спасать!.. Нет, это Келлер Дьюски, детектив из полицейского управления, что находится рядом с лабораторией. Едва переступив порог, он оглядывается.

— Здесь все в порядке?

— Да, Кел, я справлюсь, — уверяет Эд, а я киваю.

— Простите, простите, простите, — беспрестанно повторяет Эрин. С каждым «простите» она слабеет, будто превращаясь в невидимку, с которой слетают оболочки-слова. Мне больше всего хочется, чтобы скороговорка закончилась, только как оборвать безумный лепет?

— Пожалуйста, я ничем... — сбивчиво бормочу я.

Слышится жалобный всхлип Эрин. Материнское горе подобно тени или мрачной ауре: от него щемит сердце, и в глубине души пробуждается неутоленная грусть.

— Миссис Коган, я просто...

Договорить не удастся: я понимаю, что не смогу от нее отмахнуться.

Эрин снова поднимает глаза, на этот раз затуманенные, покрытые пеленой.

— Я не увижу, как он растет, — страшным мертвым голосом произносит она. — Не устрою ему день рождения, не постригу его, не познакомлюсь с его девушкой...

Она продолжает говорить, и ее голос звенит во мне, меняет форму, становится материальным, вызывает ассоциации: кажется, когда-то давно я знала эту женщину. Подобные открытия у меня случаются редко и вызывают тревогу, словно пробуждение в компании призрака.

— Боже мой... Ладно, дайте мне подумать.

...

Меня зовут Лена. В лабораторию я пришла, потому что администрация предлагала бесплатное обучение. Именно так написали в опубликованном «Геральд джорнал» объявлении: «Технический специалист в лабораторию судмедэкспертизы. Годичный

курс дактилоскопии в Заочной школе ФБР, двухгодичный вечерний спецкурс в колледже, практическое обучение делопроизводству».

Я работаю в Центре судебных исследований Уорделла, футуристическое прямоугольное здание которого построили в 1989, то есть за год до моего поступления на службу. В том же здании находится Лаборатория токсикологии при Управлении окружной медицинской службы, Банк донорской ткани при Красном кресте и Городская лаборатория судмедэкспертизы. Управление полиции расположилось по соседству. Пол лаборатории выложен блестящей синей плиткой, которую при определенном освещении можно принять за воду, а стены выкрашены в казенный блекло-бирюзовый цвет.

...

Естественно, после ЧП с Эрин Коган рабочего настроения в лаборатории как не бывало. Я нахожусь в прострации, словно стала свидетелем аварии. Кабинет накрывает мертвенная, как после катастрофы, тишина: мои коллеги сидят за столами, точно погрузившись в транс.

Сосредоточиться на отпечатках пальцев, с которыми работала вчера, не удастся. Пару минут бездумно гляжу в окно: свет распадается на крылатых насекомых и ящериц, а потом опять сливается в преломляющийся о стекло луч. Открываю файл с другим делом, заставляя себя вчитаться в полицейский отчет, но в конце концов, уступив внутреннему порыву, подхожу к высокому шкафу со стороны, где стоят материалы о «Несовершеннолетних-2002», и достаю заветный файл Коганов. В том отделении

еще две папки, которые я внимательно просматриваю. И так, еще две смерти: жертвы того же возраста, погибли примерно в то же время и в том же районе. Хм, подозрительно...

Элис расхаживает взад-вперед по кабинету, буравя меня свирепым взглядом.

— Ну что?

Она морщится.

— Сама знаешь! Поверить не могу...

— Поверить во что? — Я задыхаюсь от раздражения и полной беспомощности, а перед глазами на миг встает дикое затравленное лицо Эрин Коган.

— Та женщина... Неужели обязательно было с ней разговаривать?!

— Элис, а как, по-вашему, мне следовало поступить?

Неодобрительно поцокав языком, она уходит, а другая коллега, Сильви, сидящая у самой двери, сочувственно смотрит из-под пряди белокурых волос. Марго вздыхает, откидывается на спинку стула и кладет на лоб влажную салфетку.

— Что это было? — вопрошает она. — Что случилось?

Мы маемся без дела, пока кому-то не приходит в голову пораньше отправиться на ленч. Вчетвером усаживаемся за столик в «бассейне» — так у нас называют зону отдыха из-за выложенных белым кафелем стен, натянутых тонкой проволочной сеткой окон и люминесцентных ламп. Марго разворачивает ко мне стул, и я чувствую ее взгляд, хотя сама, спрятав сэндвич за папку, просматриваю файл.

Марго пришла в отдел криминалистики пять лет назад, и в свои двадцать девять — самая младшая из сотрудников

и единственная из нас четверых, у кого есть дети. Начала она со следов поджогов, а сейчас учится делать анализ ДНК — эта область кажется ей самой интересной — и скоро уйдет в новое отделение на одном из нижних этажей.

— Это и есть файл Коганов? — спрашивает она.

Я показываю корешок папки.

— Ну, что думаешь?

Я вожу пальцем по строчкам отчета патологоанатома.

— Мать — курильщица, ребенок спал на животе — по крайней мере, в таком положении его нашла бригада скорой помощи. — Я качаю головой, подпираю рукой подбородок и бормочу в ладонь: — Не знаю, очень похоже на синдром внезапной детской смерти.

Лицо Элис суровеет.

— Если возникли подозрения, ей следовало обратиться в полицию. Зачем в лабораторию идти?

— Умер ее ребенок, — напоминает Марго. — На месте Эрин любая мать поступила бы так же. Только троньте моих малышей, увидите, что случится!

— Знаете, из какой семьи эта Эрин? — вмешивается Сильви. — Из богатой. Я в истории болезни смотрела. Ее отец — Питер Биллингз, слышали про школу Биллингза при Сиракьюском университете?

— Нельзя допускать, чтобы люди вот так запросто к нам заявлялись. — Элис скрещивает руки на столе и опирается на локти. — Меня не волнует ни кто они, ни чьи они матери или дочери! Здесь проводятся серьезные исследования, и всем сотрудникам, включая Лену, необходимы элементарные условия для работы!

Сильви смотрит на меня, Марго прячет глаза.

— Сколько случаев внезапной детской смерти зарегистрировано в последнее время? — барабаня по крышке стола, спрашивает Элис.

Не решаясь встретиться с ней взглядом, я верчу в руках сэндвич с тунцом.

— Точное число назвать трудно: обычно подобные случаи происходят три-четыре раза за год, — отвечает Марго. — Но за последние два месяца «смерть в колыбели» была зарегистрирована дважды, не считая гибели ребенка той женщины... По-моему, смерть Мэтью Когана к этому разряду не отнесли...

— По каким признакам определяют «смерть в колыбели»? — интересуется Сильви.

— Понятия не имею, — говорит Марго со вздохом. — Официально зарегистрировано два случая, — тихо добавляет она.

— Знаете, я еще заметила, что Коганы живут в Люциусе. — Сильви обхватывает чашку ладонями. — Там в водопроводной воде кадмий нашли!

— Точнее, эту сплетню пустили студенты-хиппи из местного колледжа, — ворчит Элис.

Я смотрю на сэндвич: в гастрономическом великолепии майонеза и маринованных огурчиков скрываются кусочки некогда живого существа. Нетрудно представить, как тунец мелькает в морской воде, его блестящую чешую, острый ум... Нет, сейчас не нужно об этом думать, сейчас нужно «просто есть» — именно так всегда велит Пия, когда, сложив руки на груди, смотрит куда угодно, только не на стол.

Сильви задумчиво потирает лоб.

— На мой взгляд, это ненормально, даже как-то странно.

— В отчетах упоминаются кровавые пятна или отпечатки пальцев? — спрашивает Марго. Ее сыновья, Амаль и Фарид, еще малы. Марго носит в бумажнике их фотографии, а иногда после дача старший, Амаль, дожидается ее в коридоре. Фрэнк, директор лаборатории, дает мальчишке цветные мелки или красный фломастер и разрешает рисовать в своем блокноте. Амаль такой спокойный: сидит по-турецки на полу и с головой уходит в рисование...

— Все чисто, — отвечает Элис, — ну, то есть, насколько мне известно. В отчетах о вскрытии ничего необычного.

Ни один следователь не отошлет материалы о «смерти в колыбели» на экспертизу, разве только выехавшая на место происшествия группа заметит нечто подозрительное. Подозрительным может считаться странная реакция одного из членов семьи, резкий запах в детской или просто желание организовать дополнительную проверку. Внезапная детская смерть считается подозрительной, но материалы по таким делам добавляют лишь несколько капель в море вещественных доказательств, которые ежедневно присылают нам на экспертизу. В судебной лаборатории к подобным вещам привыкаешь, а синдром внезапной детской смерти редкостью не назовешь. Под эту классификацию попадают все необъяснимые смерти детей до года. Получив материалы, ничего кроме стандартной проверки, я не проводила, просто анализировала папиллярные линии рук матерей и слабые отпечатки паль-

цев погибших детей. Четких младенцы не оставляют: лежат себе на спинках в своих кроватках, практически не касаясь окружающих предметов.

— Окажись я на ее месте, — тихо, но очень четко и обдуманно начинает Марго, — и имей хоть малейшие подозрения, обязательно наняла бы профессионала.

Элис хрипло вздыхает и мрачно смотрит на салат, который ей ежедневно доставляют из студенческой столовой. Моей начальнице уже за пятьдесят (сколько именно, никто не знает). Она занимается судебной химией и двадцать лет назад, в пору, когда судмедэкспертизу для города и округа проводили разные организации, участвовала в открытии лаборатории. Шестью годами позже под ее руководством начала работать Сильви, специалист по анализу следов. Сильви тридцать шесть, и она клянется, что если в этом году не найдет мужа, обязательно обратится в банк спермы.

Наша четверка работает вместе — в одном кабинете и смотровой — уже много лет. Мы довольно странная команда. Элис перевели из отдела токсикологии в качестве временного начальника отдела анализа следов, и она до сих пор занимает эту должность. Порой кажется, между нами все гладко, но зачастую возникают трения. Элис с Марго особенно любят метать друг в друга молнии. Марго намекает, мол, Элис пора вернуться в «свой» отдел и оставить анализ следов «тем, кто в нем разбирается». Элис в ответ говорит, что руководить отделом лучше человеку со стороны: такие, как правило, более объективны, а потом дает понять, дескать, у Марго комплекс звезды и она портит своих мальчиков. Марго

на это заявляет, что женщина, не выносившая и не родившая ребенка, не может рассуждать о воспитании детей.

Сильви склоняется над разваливающимся сэндвичем с вареной колбасой.

— Неужели? Хочешь сказать, что наняла бы частного детектива?

— Копа на час? Да они же сплошь пенсионеры! Разве такие помогут? — вопрошает Элис. — Кстати, ты знаешь, сколько дерут эти ребята?

— Если бы речь шла о моих мальчиках? — вскидывается Марго. — Я бы ни перед чем не остановилась! Дом бы заложила, банк ограбила — какая разница! Наверняка бы удостоверилась, что сделала все возможное.

— Думаю, поэтому Эрин Коган и пришла к Лене, — парирует Элис.

— Пожалуй, я бы даже киллера наняла! — Марго вздыхает и смотрит на меня.

Я прячу глаза. Есть тунца не хочется, хотя сэндвич сегодня утром готовила сама. Я кладу его обратно в вощеную бумагу, аккуратно заворачиваю и перехватываю взгляд Марго. Радужка у нее почти такая же темная, как зрачок.

— Лена!

Сэндвич отправляется в бумажный пакет.

— Лена! — тихо зовет Марго. — А ты насчет этого дела... Ничего не чувствуешь?

Я качаю головой. Сейчас я вижу в ней мать, выжидающую, как тигрица в клетке.

— Нет...

Иногда обстоятельства и мотивы преступлений представляются мне так четко, что потом несколько дней я толь-

ко о них и думаю. Смотрю на побуревший от крови носовой платок, и в голове, откуда ни возмись, появляются мотивы: она уже давно хотела убить своего мужа; или: он всегда боялся своих одноклассников; или: она больше ни секунды не могла выносить жуткий шум в доме.

Однажды я нашла на месте преступления пожелтевшую страничку из блокнота, которую, как выяснилось, окропили слезами. Подняла и через секунду увидела: мужчина, сделавший записку, знал, что за ним придет убийца.

— По-моему, ничего необычного здесь нет, — растянув губы в улыбке, отвечаю я Марго.

— А вот мне так не кажется, — возражает она.

Судебные исследования всегда прямолинейны: появляющимся в ходе расследования вопросам противопоставляются научные принципы. Одни правила противоречат другим. Например, специалист изучает соскоб из-под ногтей убитого, частички кожи и волос — сначала под сильной лупой, потом под микроскопом, — дожидаясь, когда появится некое подобие улики. Естественно, хочется надеяться, что чем дольше и внимательнее смотришь, тем больше увидишь. Но иногда полезнее просто откинуться на спинку стула, расслабиться и закрыть глаза. Главное — не торопиться.

Рука Марго бессознательно тянется к сумочке, где лежат фотографии ее мальчиков. Порой в лаборатории мы перебираемся ничего не значащими фразами, только чтобы подбодрить друг друга. Сейчас Марго решает мне поверить. Ей прекрасно известно: под одним углом вещественное доказательство может вы-

глядеть так, а под другим — эдак. Марго кивает и сжимает мою ладонь.

— У меня просто паранойя, — тихо говорит она.

— А чего ты ждала? — Элис машет рукой, будто мы находимся у врат ада. — С нашей-то работой?

— Мы ежедневно сталкиваемся с темной стороной человеческой души, — вмешивается Сильви. — Совсем как студенты-медики, которые находят у себя все изучаемые болезни!

— Всех преступлений нам все равно не раскрыть! — добавляет Элис.

Марго улыбается, но не перестает бурчать меня взглядом. Мои глаза краснеют и чешутся, будто их рентгеном просвечивают.

Элис, Марго и Сильви переглядываются, и в «бассейне» раздается смех. Звук испуганными серебристыми волнами накатывает на меня, покрываясь кругами беспокойства.

Из-за своеобразной манеры держать ся и сардонических шуток детективы считают экспертов чудаковатыми, но уличные копы догадываются: без чувства юмора в нашем деле никуда. Мы занимается четко определенной частью расследования. Именно это меня привлекает — возможность в одиночку, без постороннего вмешательства выполнять собственное задание.

— Хорошо, что у тебя нет детей, — произносит Марго. — Лена, солнышко, ты такая счастливая!

Глава 2

Официально моя должность называется «специалист по дактилоскопиче-

скому анализу и лаборант», но неофициально мне чаще других достаются дела, в которых фигурируют пропавшие, неполноценные или получившие телесное повреждение дети. В мире судебных исследований бездетная женщина считается наименее отягощенной эмоциональным багажом. Вот вам квинтэссенция коповской философии: холодный ум и беспристрастность ценятся до такой степени, что будь на то воля начальника полиции, он запретил бы подчиненным вступать в брак и плодиться.

В общем, в мою картотеку попадают самые неприятные дела. Каждый файл, каждая папка — не что иное, как ящик Пандоры, в них периодически мелькают школьные фотографии, миниатюрные отпечатки пальцев, а у младенцев — отпечатки стоп.

Эти материалы окрашивают мир в цвета преступной халатности, равнодушия и насилия. Потом я выбрасываю страшные образы из головы и перехожу к следующему досье.

Однако по-настоящему все началось с дела Хаверстро, которым полиция занималась в 1997 году, почти пять лет назад. Маленького Троя Хаверстро убили в его собственной постели, и у детективов не оказалось ни единой зацепки. Когда я обнаружила необходимую улику, в отдельных детективах проснулся скепсис, граничащий с возмущением. Я ведь была лаборанткой — позиция на одну ступень выше секретарской, — вот детективы и сочли меня физически неспособной найти то, что пропустили сами. Некоторые восприняли это как личное оскорбление.

Директор лаборатории Фрэнк Визо после успешного окончания дела Хавер-

стро предложил мне повышение. Он боится, что, окрыленная успехом, я сбегу из лаборатории в ФБР. Меня не интересовало ни то, ни другое. Новая должность значила общение: безрадостную работу на месте преступления, непосредственное участие в расследовании со страдающими паранойей детективами, бесконечные часы в суде. Терпеть не могу, когда меня трогают. Не хочу ни готовить, ни ходить на танцы, ни присматривать за детьми, ни водить машины, ни сажать цветы, ни тратить время на йогу или другие занятия, которые мне настойчиво рекомендуют. Я люблю тишину и покой: аккуратным, ювелирно-точным движением отрегулируешь предметное стекло, настроишь кратность увеличения, и окружающий мир перестает существовать, а лежащий на стекле предмет, образно говоря, растворяется, превращаясь из вещественного доказательства в частицу правды. Он раскрывается перед моим пытливым взором, ноздри расширяются, рот наполняется слюной, а сердце бешено бьется. Тогда в сознании и возникает догадка: здесь преступник допустил ошибку.

...

Каждый день я окунаюсь в свой лазурный мир. Тонированные окна, сияющая плитка пола... Каждый день я упиваюсь сладкой меланхолией. Я где-то вычитала: Исаак Ньютон признавался, что ощущал себя ребенком, собирающим ракушки на берегу океана правды. Глядя на стеклянные межкомнатные стены и столы, уставленные современным оборудованием, я хорошо понимаю его чувства.

У нас новые компьютеры, микроскопы, усовершенствованные химика-

ты — флуоресцирующий порошок и реактивы, — возможность проводить анализ ДНК, но за пределами лаборатории люди продолжают совершать ужасные по своей предсказуемости поступки. Преступления идут бесконечной чередой, а с ними и убийцы, прикасающиеся к запястью жертвы, вазе, клочку бумаги, оставляющие тончайший след пота или жира, соли или крови, обозначающий их местопребывание. Вряд ли мне помешало бы предвзятое мнение, но чем меньше знаю о том, кому принадлежат отпечатки, тем легче их читать. Скрытые отпечатки состоят из пота и пыли, папиллярные линии стоп и ладоней остаются повсюду. Во всем мире не найдется двух одинаковых отпечатков пальцев: у каждого человека они свои, уникальные до малейшего завитка, зато не меняются от рождения до самой смерти.

Мой стол завален отпечатками пальцев, собранных либо на липкую ленту, либо сразу на карточку, и каждый нужно идентифицировать, то есть сопоставить с имеющимся в базе данных образцом.

— Лена, Лена! — Элис подбегает сзади, когда я возвращаюсь из уборной в кабинет.

Мой палец апатично чертит линии на запотевшем окне коридора. Холодные закорючки помогают сосредоточиться, но сегодня вечером наша уборщица Дейзи снова будет клясть меня за то, что залапала окна... Не говоря уже о целой горе бумажек, носовых платков и крошек, которую я опять оставила на рабочем месте. Да, да, за собой нужно убирать, только вот постоянно забываю...

— Эй, как ты? — нагнав меня, спрашивает начальница.

— Лучше не бывает.

— Нам не следует снова поднимать дело Коганов. — Элис рассеянно накручивает на палец длинную серебристо-пепельную прядь.

В облике моей начальницы присутствует что-то неутолимо грустное — эдакая вечная дева. Марго язвит, мол, Элис похожа на учительницу, зато Сильви настаивает, что именно это внушает спокойствие. Тем не менее, я знаю, как нетерпима Элис к людям и просьбам, которые кажутся ей «несерьезными».

— Но ведь вы об этом думали? Ну, чтобы копнуть поглубже?

— Нет. — Она сжимает губы, старательно изображая суровость.

— Ну, тогда, сами понимаете... — Я бессильно развожу руками.

— Хочешь сказать, что послушаешься? Я улыбаюсь.

— В таком случае, пожалуйста, отведь, мне любопытно, — просит Элис. — Та женщина, думаешь, в ее словах что-то есть? Нет, история, конечно, печальная, впрочем, как и большинство остальных, которыми мы занимаемся. Только я заметила еще там, в фойе.... Похоже, тебя проняло. Потому что она сумасшедшая, или есть иная причина? — Рот Элис снова превращается в тонкую полоску.

Замявшись, я выдавливаю из себя:

— Кажется... кажется, я ее знаю.

— Ты ее знаешь? Знаешь Эрин Коган?

Я затравленно прижимаю руки к груди.

— Нет, не так. Просто чувствую, что была знакома с Эрин, близко знакома, а

потом напрочь о ней забыла. Видите в этом смысл?

Элис смотрит на меня и хмурится.

Внезапно от стен доносится сдавленный стон — в коридоре не все в порядке с отопительной системой, — и я испуганно подпрыгиваю. Скрестив по-девичьи тонкие руки, Элис пронзает меня внимательным взглядом.

— Когда ты в последний раз говорила со своей приемной матерью?

— С Пией? А в чем дело?

— Не знаю, может, пора?

— Поговорить с Пией? Ну, спасибо за совет!

— Вдруг она сообщит нечто ценное? Например, является ли та женщина частью твоего прошлого.

— Как раз об этом Пия и словом не обмолвится.

— И тем не менее. — Пригладив тонкие седые волосы, Элис перебрасывает их через плечо. — Когда мир рушится, общение с семьей порой оказывается весьма полезным.

Я смотрю на нее во все глаза. Элис прекрасно понимает, что тема болезненная: Макуильямсы не удочерили меня, а лишь взяли на воспитание. Мне то хочется слиться с ними, то, наоборот, отречься. Я уже давно не звонила им и не заезжала в гости.

— М-м-м... прости...

Перевод Аллы Ахмеровој.

Diana Abu-Jaber, «Origin», Jordan-USA.

Источник: www.bakanov.org (отрывок публикуется с разрешения владельца ресурса www.bakanov.org).