

Начинаем работать с электронными терминологическими базами данных

Мариана Секеракова, Философский факультет ПУ, Прешов,
mariana.sekerakova@gmail.com

Ключевые слова: ТБД, память перевода, CAT-программы, глоссарий, импорт-экспорт, SDL Trados Multiterm, MemoQ

Key words: TB, translation memory, CAT-tools, glossary, import-export, SDL Trados Multiterm, MemoQ

В процессе перевода переводчик неоднократно обращается на переводческий словарь, терминологический словарь, а в не последнюю очередь и свои собственные глоссарии. Мы не раз встречались со ситуацией, когда перевод был почти закончен, но один термин не давал спать два дня. Такие ситуации не идут на пользу экономии рабочего процесса. Заказчики переводов требуют не только высококачественной работы, но главным образом, экспрессного выполнения заказа.

Ваши собственные, годами разработанные глоссарии, будут намного эффективнее, если будут переведены в электронную форму. Конечно, их переписывание займет не мало времени, но в будущем сохранит его намного больше. Основные способы разработки и сохранения глоссариев нам предлагают:

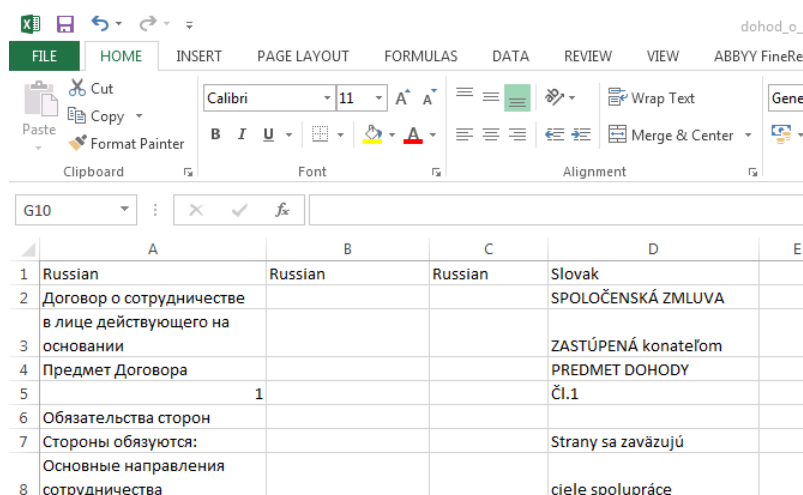
- Отдельные программы для создания терминологических баз (далее ТБД)
- Функции комплексных CAT-программ – обзор основных [здесь](http://cs.wikipedia.org/wiki/CAT_Porovn.C3.A1n.C3.AD_rozd.C3.ADln.C3.BDch_CAT_n.C3.A1stroj.C5.AF)http://cs.wikipedia.org/wiki/CAT_Porovn.C3.A1n.C3.AD_rozd.C3.ADln.C3.BDch_CAT_n.C3.A1stroj.C5.AF
- Табличные процессоры – информации [здесь](https://ru.wikipedia.org/wiki/%D0%AD%D0%BB%D0%B5%D0%BA%D1%82%D1%80%D0%BE%D0%BD%D0%BD%D0%B0%D1%8F_%D1%82%D0%B0%D0%B1%D0%BB%D0%B8%D1%86%D0%B0)https://ru.wikipedia.org/wiki/%D0%AD%D0%BB%D0%B5%D0%BA%D1%82%D1%80%D0%BE%D0%BD%D0%BD%D0%B0%D1%8F_%D1%82%D0%B0%D0%B1%D0%BB%D0%B8%D1%86%D0%B0

Рынок с переводческим софтвером предлагает платные и бесплатные, более и менее качественно разработанные программы. Существуют также возможности разработки, экспорта и импорта ТБД онлайн. Те, которые не желают покупать дорогое программное обеспечение или онлайн членство на специализированных вебах, могут начать работать именно с табличными процессорами. Для этого достаточно установить одну из программ:

- **Бесплатные** – OpenOffice Calc, LibreOffice Calc, NeoOffice, Google Sheets (онлайн в рамках Gmail аккаунта)
- **Платные** – Apple Numbers, Corel Quattro Pro, Lotus 1-2-3, Microsoft (Excel, Works)

Список бесплатных программ для установки в компьютере и онлайн программ для переводчиков можно найти (*J. Vedral, 2007*) [здесь](#).

Нашим требованиям послужит программа Microsoft Excel, которая установлена почти в каждом компьютере с операционной системой Microsoft Windows. Открываем программу и вводим данные, как показано на рисунке 1. Так мы сможем разработать двух- и даже многоязычный глоссарий. В первой строчке вводим не обязательную информацию о языке, которая служит только для большей наглядности. Таким способом разработанный глоссарий сохраняем в любом из этих форматов: *.xml, *.xlsx, *.xlsm, *.xls, *.xlt.



	A	B	C	D	E
1	Russian	Russian	Russian	Slovak	
2	Договор о сотрудничестве			SPOLOČENSKÁ ZMLUVA	
3	в лице действующего на основании			ZASTÚPENÁ konateľom	
4	Предмет Договора			PREDMET DOHODY	
5	1			Či.1	
6	Обязательства сторон				
7	Стороны обязуются:			Strany sa zaväzujú	
8	Основные направления сотрудничества			ciele spolupráce	

Рис. 1

Этот документ может служить в качестве электронного глоссария, в котором можно найти нужный термин при помощи функции «Найти Ctrl+F». Документ также можно разделить на листы по нужным темам или другим критериям. Если планируете в будущем использовать переводческие программы, то документ сохраненный в одном из универсальных форматов табличного процессора может быть импортирован в такую программу.

Со временем, когда вас работа с электронным глоссарием заинтересует настолько, что захотите создавать более разработанные и профессиональные терминологии, то предлагаются следующие бесплатные альтернативы:

Онлайн экстракторы терминологии:

Five Filters – бесплатный онлайн проект экстрактора терминологии. После вложения текста и нажатия кнопки «GetTerms», пользователь получает список наиболее релевантных терминов. Поддержка английского языка.

TerMine – онлайн экстрактор предназначенный для разработки терминов из области биомедицины. Позволяет вкладывать текст для анализа, файл с текстом в *.txt или *.pdf формате, а также URL-адрес. Поддержка английского языка.

Программы для экстракции терминологии:

Anchovy – бесплатная мультилингвальная программа для правки глоссариев и экстракции терминов.

ArayaBilingual Terminology Extraction – программа для автоматизированной экстракции терминов. Платная с возможностью скачать тест-версию.

Лучший способ использования разработанной, купленной или скачанной терминологии – импорт в САТ-программу. Каждая из таких программ предлагает возможность автоматически разыскивать нужный термин в импортированном **ТБ**, т.е. **TermBase** или терминологическая база данных. Большинство из программ «кошек» (рабочее название для САТ-программ) работает не только с авторскими ТБД, но и с т.н. **plug-in** или подключаемыми модулями как напр.:

EuroTermBank plugin – модуль подключающий доступные онлайн ТБД к САТ-программе. Глоссарии доступны и онлайн на сайте, который предлагает следующие функции:

Поиск в терминологии из разных ресурсов

Определение возможных терминов и их экстракция

Подборка терминов-кандидатов в различных ресурсах

Правка терминов и их перевод

Предоставление доступа к терминологии другим пользователям

Сотрудничество с друзьями и коллегами

Пользование терминологией в других программах

(наш перевод <http://www.eurotermbank.com/>)

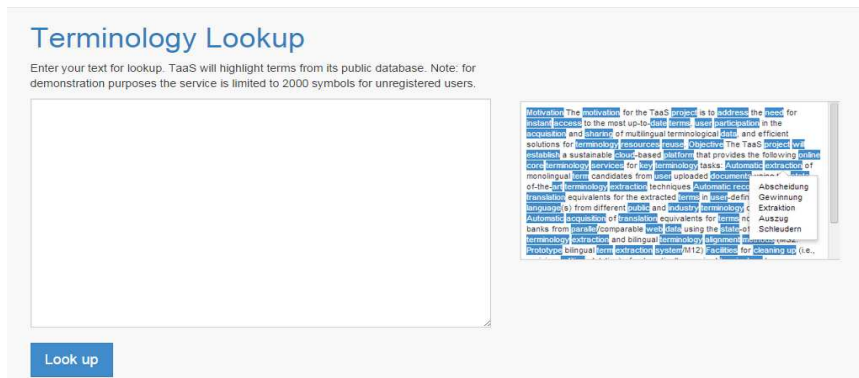


Рис. 2

Кроме этого, сайт предлагает онлайн экстракцию терминов. Рис. 2

EuroTermBank подключенный к программе, предлагает не только перевод термина, но и информации, как область использования и определение термина. Рис. 3

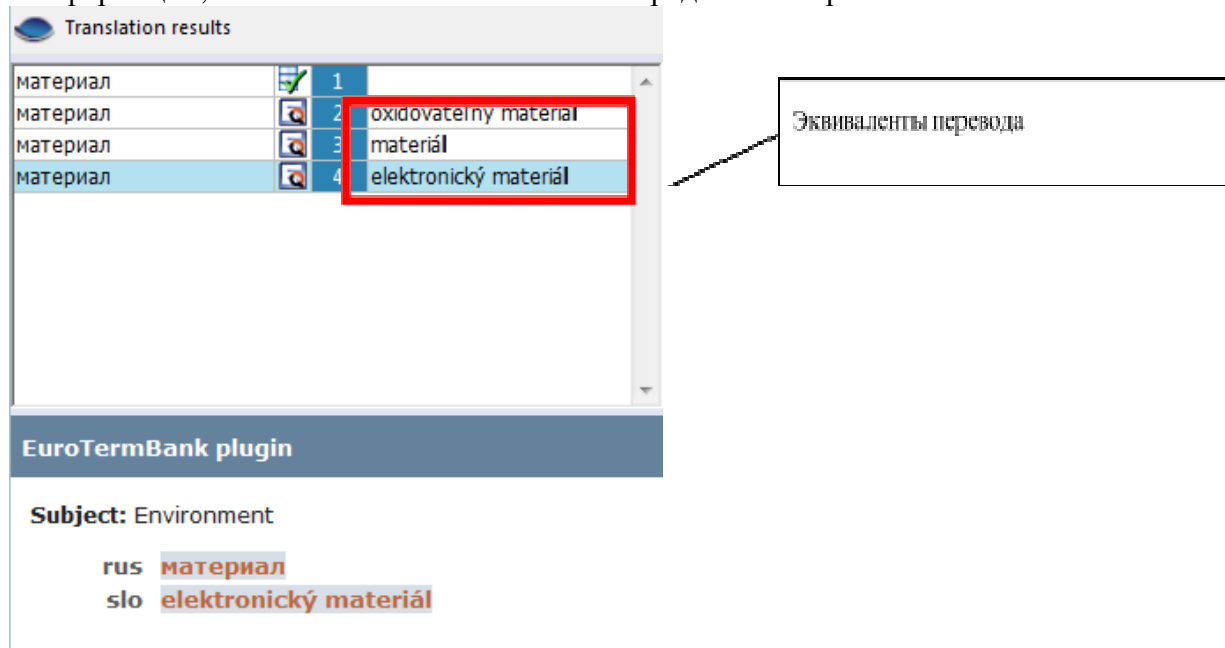


Рис. 3

Следующим модулем является TaaS plugin – платформа, основанная на принципе «облака» (Cloud) – подключенная автоматически к САТ-программе (напр. MemoQ) предлагающая:

«...удобный для пользователя, совместный и мультилингвальный терминологический сервис для переводчиков.»

(наш перевод <http://kilgray.com/news/taas-integrated-memoq>)

Подходящий способ разработки глоссариев предлагают и САТ-программы. Благодаря им возможно из готовых текстов экстрагировать термины (по настроенным критериям), исправлять их, переводить, добавлять дополнительные информации и в

конечном итоге, использовать для перевода проектов. Можно также сделать экспорт ТБД по запросу клиента. Решением всех терминологических проблем является менеджмент терминологии:

«Оставив терминологию без менеджмента, она может стать нелогичной и включать в себя противоречивые определения, что означает невозможность повторного использования перевода.»

(<http://www.translationzone.com/products/terminology-management/#tag4>)

Под фразой «терминологический менеджмент» понимаем – сбор существующих терминологических ресурсов, разработку и расширение терминологии, ревизию и обновление ТБД. Качественный менеджмент терминологии приносит ряд выгод как отмечает Яна Левицка:

- Уменьшение финансовых и временных утрат на перевод и разработку специализированных текстов (минимизация ошибок и неясностей в коммуникации)
- Эффективность при разработке терминологических продуктов
- Способ реализации – *terminology management system* (например, как части программ для автоматизированного перевода)

Доступным, очень простым и интуитивным для пользования является онлайн модуль от компании Kilgray (разработчик MemoQ) под названием **qTerm**. Заполнив формуляр на [странице компании](#), сможете получить доступ к онлайн менеджменту ваших ТБД, пользоваться доступными ТБД коллег и проводить импорт-экспорт ТБД в свои «кошки».

Рассмотрим основные и самые известные CAT-программы на сегодняшнее время: SDL Trados, MemoQ. Обе эти программы платные, но обе предлагают самую широкую палитру функций для работы с переводом, ТМ (памяти переводов, из англ. Translation Memory) и ТБД.

SDL Trados

SDL Trados – система автоматизированного перевода предлагает эти основные модули:

Translator's Workbench — модуль работы с базами данных памяти переводов ТМ (создания, обслуживания, импорта, экспорта). Создание документов перевода и их редактирование производятся в отдельных модулях.

Панель Trados в Microsoft Word — модуль для перевода документов в Microsoft Word.

TagEditor — модуль для перевода документов в формате PowerPoint, Excel, HTML, XML и т. п.

WinAlign — модуль для создания памяти переводов на основе ранее переведённых двуязычных текстов.

S-Tagger — модуль для перевода документов в формате FrameMaker и InterLeaf.

T-Window — модуль для перевода текста из буфера обмена.

MultiTerm — модуль для ведения глоссариев.

ExtraTerm — модуль для автоматического поиска терминологических кандидатов в тексте и создания глоссариев на их основе.

(<https://ru.wikipedia.org/wiki/Trados>)

MultiTerm работает как отдельная программа системы перевода и позволяет экстрагировать термины с молилингвальных и билингвальных документов, расширять и управлять ТБД так, как это требует работа с ТБД. Конечно, возможен и ввод терминологических пар вручную во время перевода текста. Интерфейс программы можете видеть на рис. 4.

Создаем новый проект перевода

Для перевода текстов разных форматов используем модуль SDL Trados Studio, которого интерфейс работает с английским, немецким, французским, японским, испанским и китайским языками.

Нажимаем кнопку «New project» → пропускаем первый шаг «Next» → вводим название проекта в строчке «Name» и продолжаем «Next» → в строчке «Source Language» выбираем исходный язык, а в строчке «Target language» один или больше языков перевода → продолжаем на следующий шаг «Next» → нажимаем на «Add Files» и выбираем файл или файлы для перевода → «Next» → нажимая «Add» добавим существующую ТМ (File-based ТМ), а нажимая «Create» создадим новую ТМ (File-based ТМ) → «Next» → в этом шаге можно добавить существующую ТБД разработанную в **MultiTerm** → следующих три шага пропускаем «Next» → в окне «Project Summary» можем проверить заданные параметры → «Finish». Проект создан и мы можем начать переводить текст открыв файл на панели «Files».

Добавляем термины из существующего текущего документа в существующую ТБД

Прежде чем начать добавлять термины из текущего перевода в ТБД, мы должны создать ее в модуле **MultiTerm**. В отличие от программы **MemoQ** модуль **MultiTerm** функционирует как самостоятельная программа и терминологический менеджмент. ТБД созданной в **MultiTerm** можно управлять, экспортировать и импортировать в **SDL Trados Studio**, как ТБД к конкретному проекту. Шаги к созданию новой ТБД:

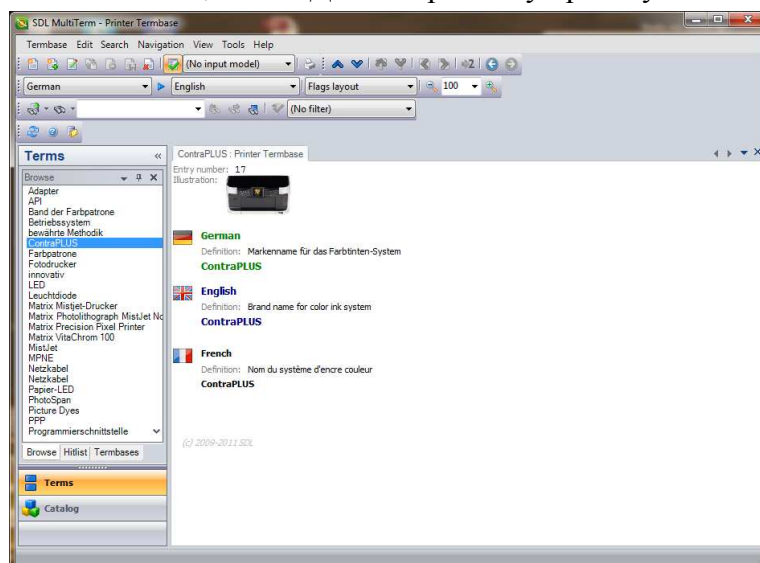


Рис. 4

Открываем модуль **MultiTerm** и на верхней панели в закладке «Termbase» нажимаем «Create Termbase» → сохраняем ТБД в компьютере на нужном месте → открывается окно «Termbase Wizard» нажимаем «Next» → шаг 1 – выбираем первую возможность и нажимаем «Next» → шаг 2 – вводим название ТБД и нажимаем «Next» → шаг 3 – выбираем все нужные языки, добавляем в список кнопкой «Add» и нажимаем «Next» → шаг 4 – вводим область

терминологии или пропускаем шаг и нажимаем «Next» → шаг 5 пропускаем → «Finish». Рисунок 4.

Каждый новый термин добавляем клавишей «F3». На рис. 4 вы можете видеть готовую мультилингвальную ТБД. В ней находятся не только языковые эквиваленты термина, но и его определение, информации о нем или даже иллюстрация.

Чтобы добавить в данную ТБД термины из текущего перевода, то открываем перевод, обозначаем термин и его перевод и клавишей «F2» вкладываем языковую пару в ТБД проект, добавляем нужные данные а при помощи клавиши «F12» сохраняем.

Импорт существующего глоссария в .xls формате происходит следующим образом:

Утилита **Multiterm Convert**: «New conversion session» → нажимаем «Next» → выбираем тип конвертируемого файла → указываем место расположения файла и место сохранения → указываем для обоих языков «Index Field» и меняем язык →

конвертируем и получаем промежуточный файл → выбираем в меню Multiterm «Termbase -> Create termbase» → задаем название новой ТБД и место ее сохранения → нажимаем «Next» → выбираем пункт «Load an existing termbase definition file», где указываем на полученный файл → в следующем шаге меняем название и описание → пропускаем следующие окна нажимая «Next» → открывается главное окно программы и нажимаем «View -> Catalog» или «Alt+Shift+C» → в появившемся окне нажимаем правой кнопкой по пункту «Import» и «Process» → выбираем файл для импорта — обычно, «Filename.mtf.xml» и нажимаем «Next» → Называем файл для неправильных записей и нажимаем «Next». Подробное описание целого процесса с рисунками по ссылке <http://bit.ly/1nqKZ99>

Экспорт разработанной ТБД

Для собственных потребностей, расширения информации, сохранения резервной копии или на требование заказчика, можем сделать экспорт ТБД. Multiterm предлагает экспорт документов разных форматов, с которых самыми выгодными являются:

- **Tab-delimited export** «*.tbx» – универсальный формат ТБД, который можно сконвертировать на другие форматы, перевести в таблицу или сделать импорт в другую CAT-программу
- **Word Dictionary export** «*.rtf» – текстовый файл в форме принтового словаря.

МемоQ

Создаем новый проект

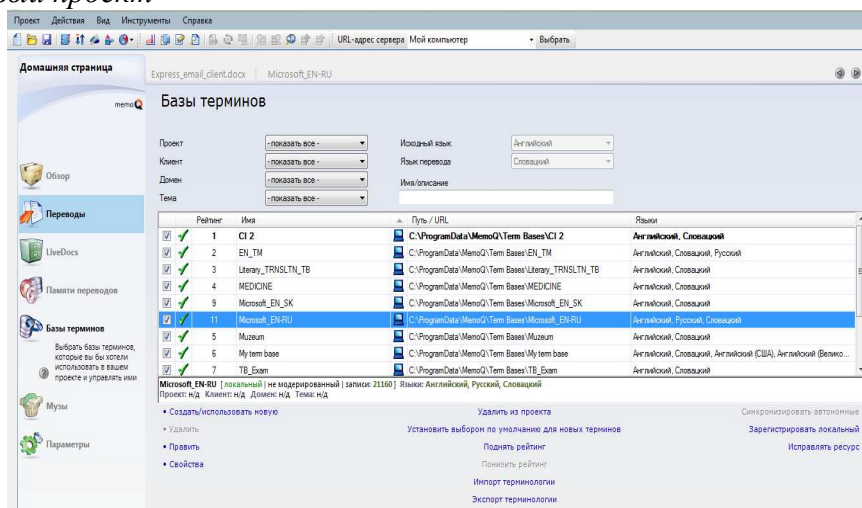


Рис. 5

МемоQ (рис. 5) представляет очень интуитивную и удобную для пользователя CAT-программу. Пользовательский интерфейс предлагает следующие языковые модификации: английский, русский, немецкий, испанский, французский, венгерский, японский, польский и португальский. Для создания нового проекта открываем программу:

«Создать новый проект» → ввести название проекта, исходный язык и язык перевода, нажимаем «Далее» → импорт файла для перевода и нажимаем «Далее» → выбираем существующую память переводов или создаем новую, нажимаем «Далее» (можно также сделать импорт экстренной ТМ) → создаем новую базу терминов (можно также использовать существующую или импортировать экстренную) → «Готово».

Если вы не создали ТМ и ТБД в этих шагах, то это возможно сделать позже на панели «Памяти переводов» и «Базы переводов» на ленте «Домашняя страница». Открываем наш документ и можем начинать переводить. С этого документа вы можете

извлечь терминологию с помощью функции «Извлечь термины» в закладке «Действия». В появившемся окне задаем название сессии и источник терминов – из активного документ или из памяти переводов. Вводим параметры будущих терминов или как их называет программа МЕМОQ «**термины-кандидаты**». Мы также можем создать список «стоп-слов» или слова, которые программа не должна учитывать, как кандидата. Программа проведет анализ и в новом окне открывается список «терминов-кандидатов», который предстоит перевести. Нужные термины подтверждаем клавишами «Ctrl+Enter» а ненужные забросим «Ctrl+D». После завершения работы экспортируем термины в ТБД «Export to term base». В закладке «Базы терминов» можно управлять нашей ТБД открыв ее кнопкой «Править». Здесь термины можем переводить, исправлять существующие переводы и добавлять данные: меру совпадения (для поиска в ТБД), пример употребления термина, грамматические данные (часть речи, род, число) а также ввести его определение. Ваш перевод и все будущие переводы могут черпать из созданных вами ТБД.

Добавляем термины из существующего текущего документа в существующую ТБД

Как мы уже упомянули выше, слова-термины можно добавлять в ТБД во время перевода текста. Для этого используем сочетание клавиш «Ctrl+E». Для добавления термина в ТБД:

Обозначаем слово в исходнике и его эквивалент в столбце перевода и нажимаем «Ctrl+E» → в появившемся окне добавляем все нужные характеристики терминов, или другие варианты перевода при помощи знака «+» → нажимаем «ОК».

*Импорт существующего глоссария в *.xls формате происходит следующим образом:*

Для импорта глоссария, скачанного из интернета или разработанного вами в любом табельном процессоре, в части «Базы терминов», создаем новую базу терминов:

«Создаем новую» → вводим название, в части «Языки» обозначаем флажком нужные → нажимаем «ОК» → открываем созданную ТБД для правки → на панели управления нажимаем на кнопку «Импорт» → находим в компьютере нужный глоссарий (в формате *.csv, *.txt, *.txv, *.tmx, *.xls, *.xlsx, *.xml, *.tbx) → в окне настроек «*.tmx файла» вводим языковую пару → нажимаем «Ок» и запускаем импорт. ТБД готова к использованию.

Экспорт разработанной ТБД

В случае, если заказчик требует от вас не только перевод, но и память перевода или ТБД перевода, то экспорт проводим одним из следующих способов:

1. Открываем ТБД для правки → на панели управления нажимаем «Экспорт» → выбираем формат файла и его место сохранения → убираем флажки из ненужных полей (изменения видны в окне просмотра CSV) → в качестве разделителя выбираем точку с запятой → нажимаем «Экспорт».

2. В закладке «Базы терминов» нажимаем на нужную ТБД → на нижней панели в среднем столбце находим приказ «Экспорт терминологии» → дальше как в пункте № 1.

Конечно, приучиться работе с «кошками» не просто а разработка ТБД займет не мало времени, но все это инвестиция в будущее. О том, что все это не зря, говорит и тот факт, что профессиональные переводчики, агентства и даже учреждения как Европейская комиссия в своем Департаменте перевода, пользуются всеми доступными электронными средствами для автоматизированного перевода а также базами для их хранения и распространения. Больше информации о работе Департамента и об упомянутых программах можно прочесть в публикации «Translation at the European Commission – a history», которая доступна в электронном виде [здесь](#).

Советуем попробовать разные способы разработки глоссариев и терминологических баз и найти самый подходящий для данного пользователя. Самым

простым и универсальным остается способ разработки терминов в табличном процессоре и хранение ТБД в таком виде. Существуют разные форумы и статьи, где вы сможете прочесть о пользе «кошек», ТМ и ТБД, но то насколько это упрощает процесс перевода, решать Вам.

Literatúra

LEVICKÁ, J.: Terminologické databázy ako nástroj terminologického manažmentu (online) [citované 04. 11. 2014]. Dostupné na internete :

<https://wiki.ufal.ms.mff.cuni.cz/_media/seminare:paha_prezentacia.pdf>

VEDRAL, J.: Software pro překladatele zdarma a legálně. (online) In: Z přednášky na veletrhu EXPOLINGUA 2007. [citované 03. 11. 2014]. Dostupné na internete :

<http://www.jtpunion.org/spip/article.php?id_article=137>

СОЛОВЬЕВА, А.: Профессиональный перевод с помощью компьютера. СПб.: Питер 2008. с. 160.

Internetové zdroje

Dostupné na internete:

<http://cs.wikipedia.org/wiki/CAT#Porovn.C3.A1n.C3.AD_rozd.C3.ADln.C3.BDch_CAT_n.C3.A1stroj.C5.AF>

Dostupné na internete: <<http://fourfeathers.by/2013/01/multiterm-терминологические-базы-данных/>>

Dostupné na internete: <<http://www.translationzone.com/products/terminology-management/#tag4>>

Dostupné na internete: <<https://ru.wikipedia.org/wiki/Trados>>

Dostupné na internete: <https://ru.wikipedia.org/wiki/Электронная_таблица>

Dostupné na internete: <<https://www.memoq.com/about-us/news/taas-integrated-with-memoq-2014>>

Dostupné na internete: <<https://www.memoq.com/translation-products/compatibility,-extensions/qterm-memoq-server-add-on>>

Summary

How to start with electronic terminological databases.

If you want to make your translation work faster and easier but still of high-quality then you should give a try to modern technology. CAT-tools provide you with various functions to simplify translation work and function of creating your own Term Base is one of them. Glossaries saved in excel tables are the first step into the world of electronic term bases of other formats. Let them help you to improve your translations and make your life easier.