

# Машинный перевод как инструмент переводчика

**Выгодский Олег**

**Центр переводов «Аргонавт»**

# О компании «Аргонавт»

- перевод документации в сфере Telecom + IT, с английского на русский, с 1992 г.
- основные заказчики: AT&T, Italtel, Samsung, Ericsson, ZTE, Siemens
- применение системы машинного перевода (MT) с 1993 г.
- применение систем Translation Memory (TM) с 1997 г.

# Структура презентации

**1**

Интеграция системы МТ в технологическом процессе компании

**2**

Разработка и сопровождение словарей для МТ

**3**

Управление МТ-профилями и стилем перевода

**4**

Преимущества применения МТ

# 1993 г. Stylus (DOS) for Software

Исходные файлы  
(англ.)

Фильтрация (TROFF, RTF, MIF,  
HTML/SGML/XML, Ventura, PageMaker, ...)

Разработка терминологии и  
настройка словарей в среде MT

MT-словари

MT - пакетный перевод файлов

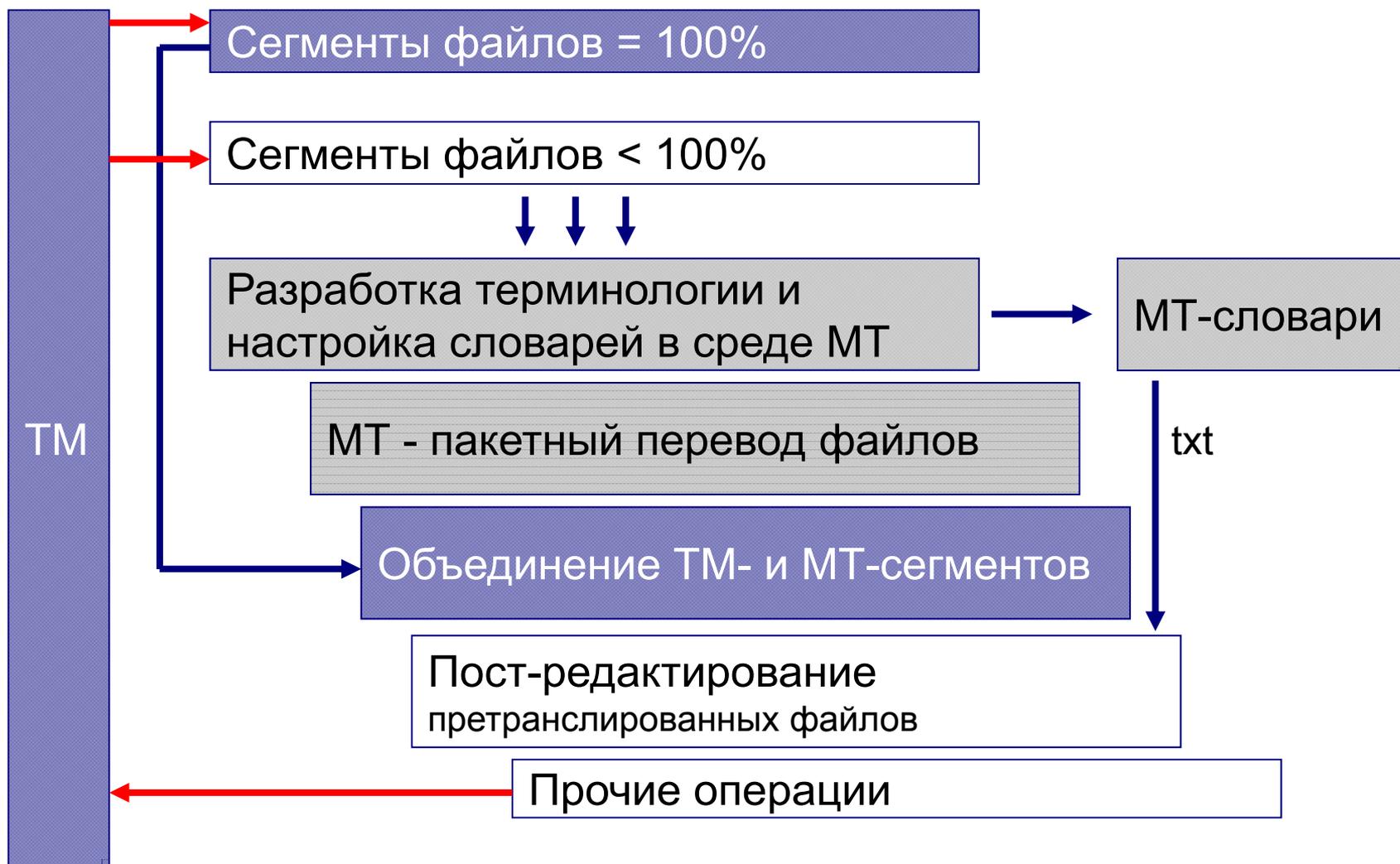
txt

Пост-редактирование MT-файлов во  
внешнем текстовом редакторе

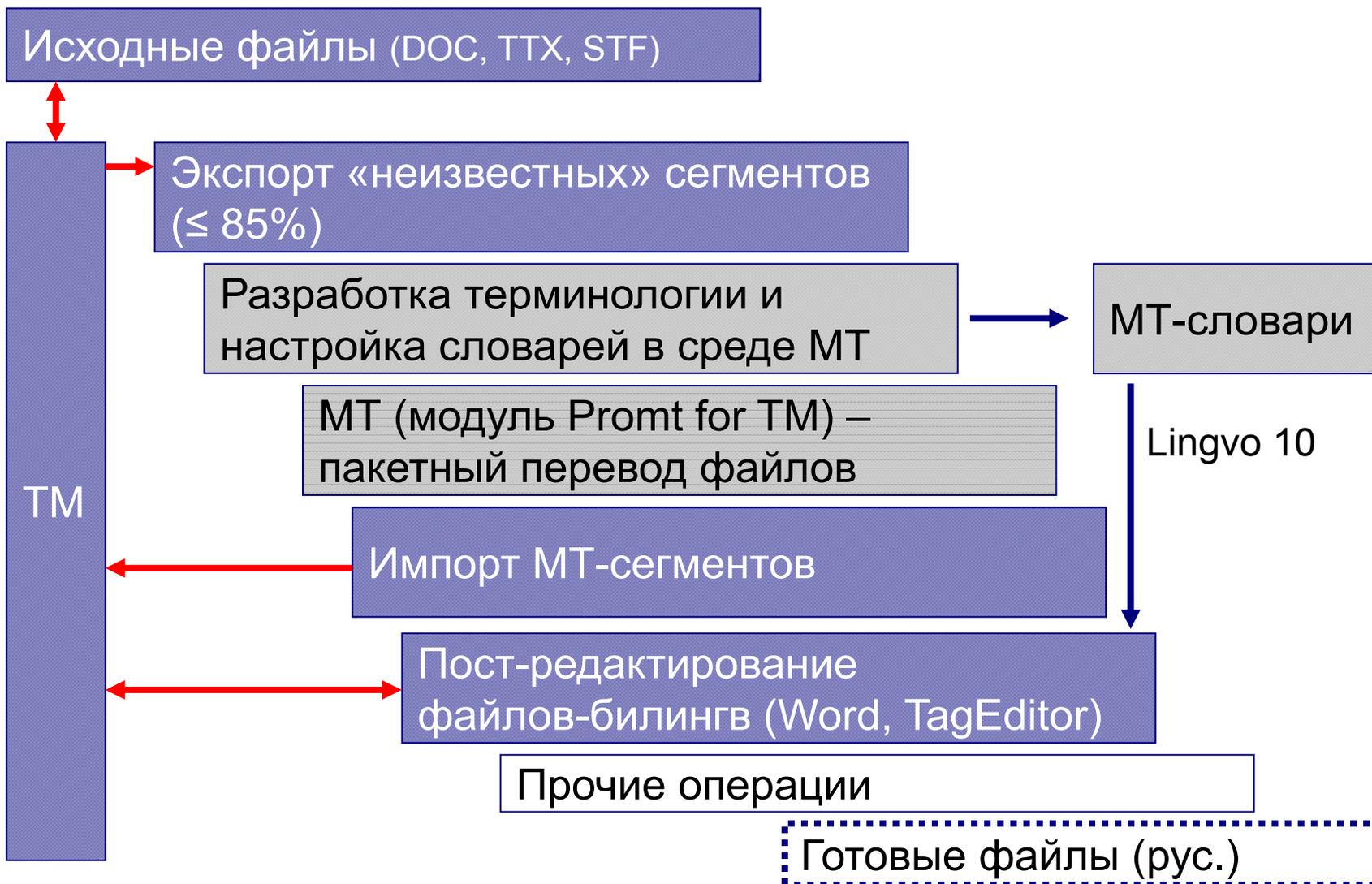
Прочие операции (техническое  
редактирование, обратная  
фильтрация, верстка, корректура)

Готовые файлы  
(рус.)

# 1997 г. TM (Transit) + MT (Promt)



# Сегодня: TM (SDL Trados) + MT (Promt)



1

Интеграция системы МТ в технологическом процессе компании

2

Разработка и сопровождение словарей для МТ

3

Управление МТ-профилями и стилем перевода

4

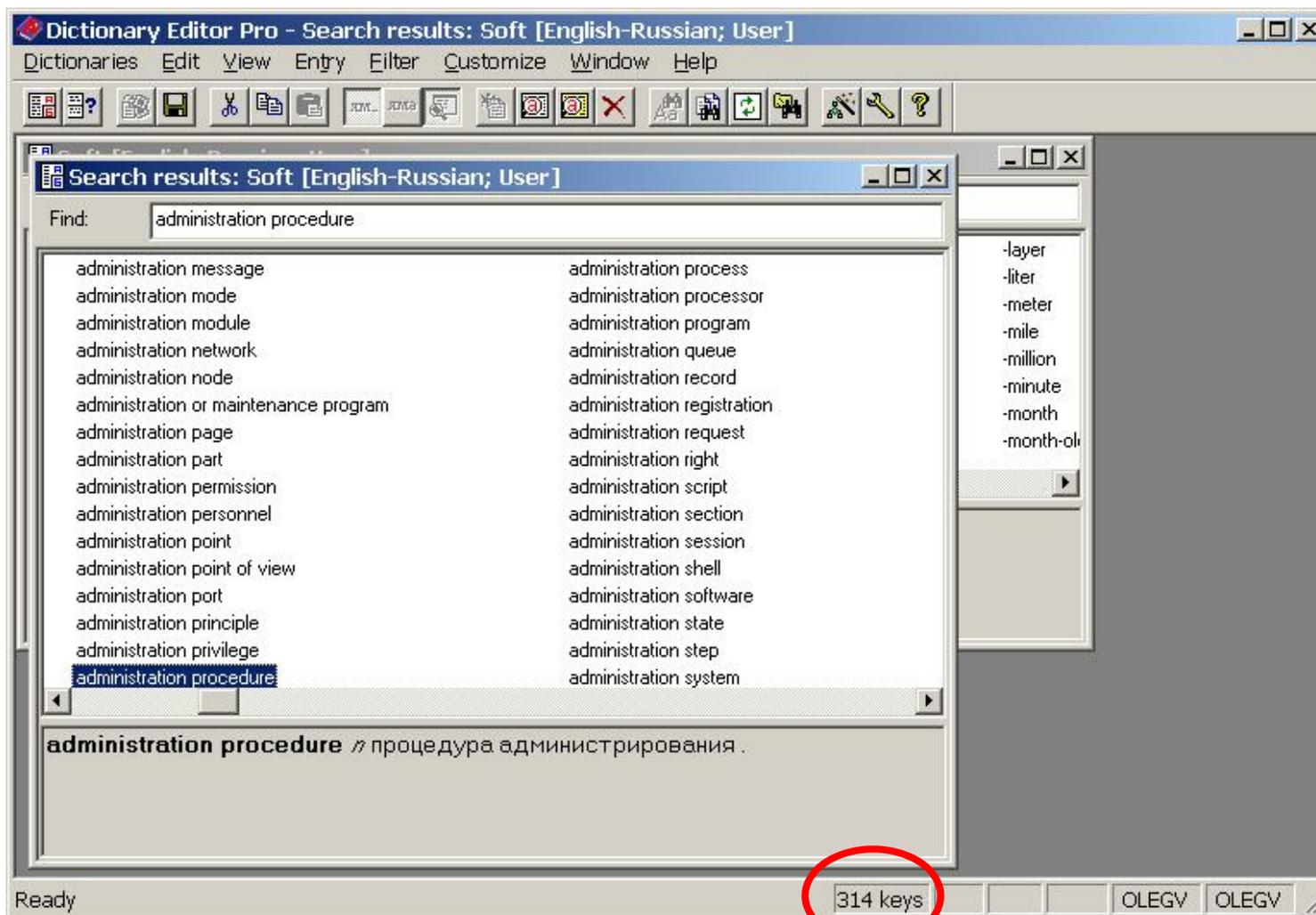
Преимущества применения МТ

# Источники лексики и уровни терминологии для МТ

- **Источники лексики:** переводимая документация, международные стандарты, "полезная" документация в Интернете (разных фирм), общедоступные словари и т. д.
- **Уровни терминологии:**
  - **Базовая лексика:** типы сетей и систем (IN, ISDN, ATM), применяемые технологии (SDH, MPLS, GSM), типы сигнализации, защита, обеспечение безопасности, учет стоимости, электропитание, основные услуги связи и т. д.
  - **"Вторичная" лексика (специфическая для фирмы-производителя):** системы, блоки, устройства, функции, режимы, состояния, руководства и справочники, приложения, решения, услуги и т. д.
  - **Типовые выражения и обороты:** "this chapter describes", "for more information see" и т. д.

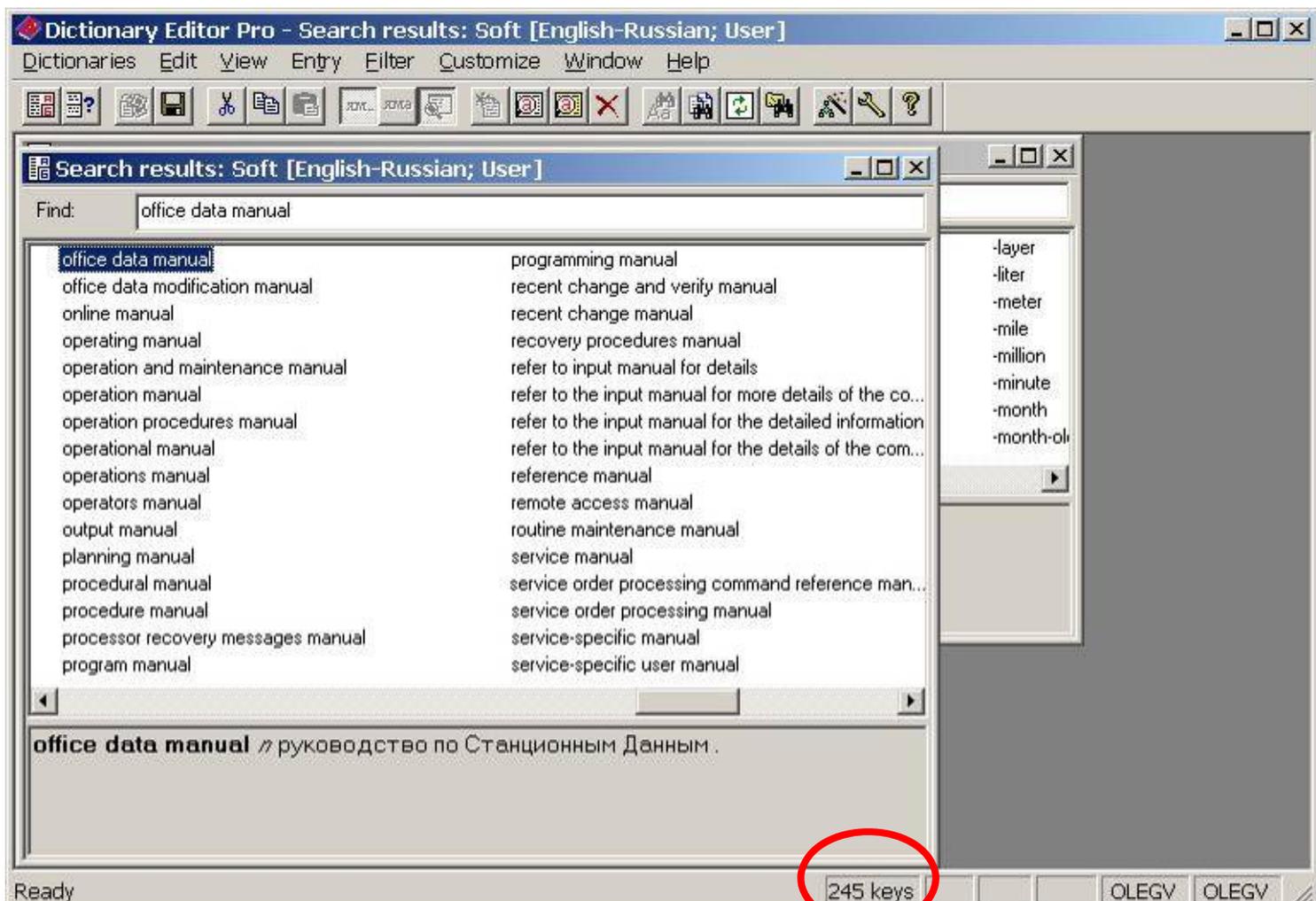
# Пример

# Administration



# Пример

# Manual (руководство)



# Особенности словарей для МТ

- термины в единственном и множественном числе
- комбинации с корневыми терминами
- разрывные глаголы
- термины-команды
- "расшифровка" сокращений
- термины с союзами *and (и) / or (или)*
- конструкции с прилагательным, например, "*... related*" (*связанный с ...*)

# Термины в единственном и множественном числе

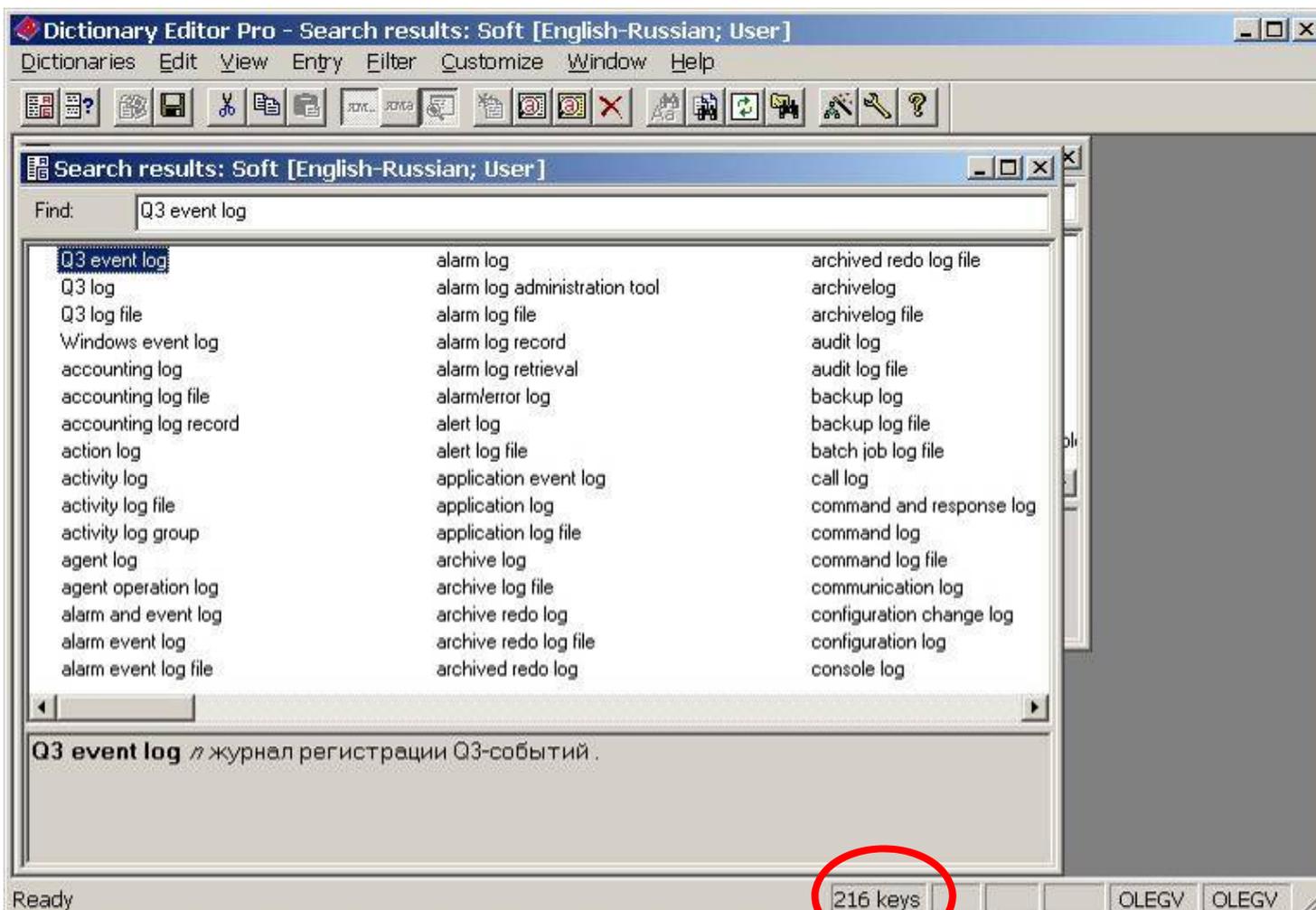
IP address	IP-адрес
port name	имя порта
port names	именА портОВ
parameter value	значение параметра
parameter values	значениЯ параметрОВ

# Комбинации с корневыми терминами

Корневой термин	Пример словарной статьи	Перевод	Кол-во словарных статей
... list	event list	список событий	411
... table	status table	таблица состояний	295
... range	value range	диапазон значений	96
... monitoring	transaction monitoring	текущий контроль транзакций	227
... management	fault management	обработка отказов	592
... log	error log	журнал регистрации ошибок	216

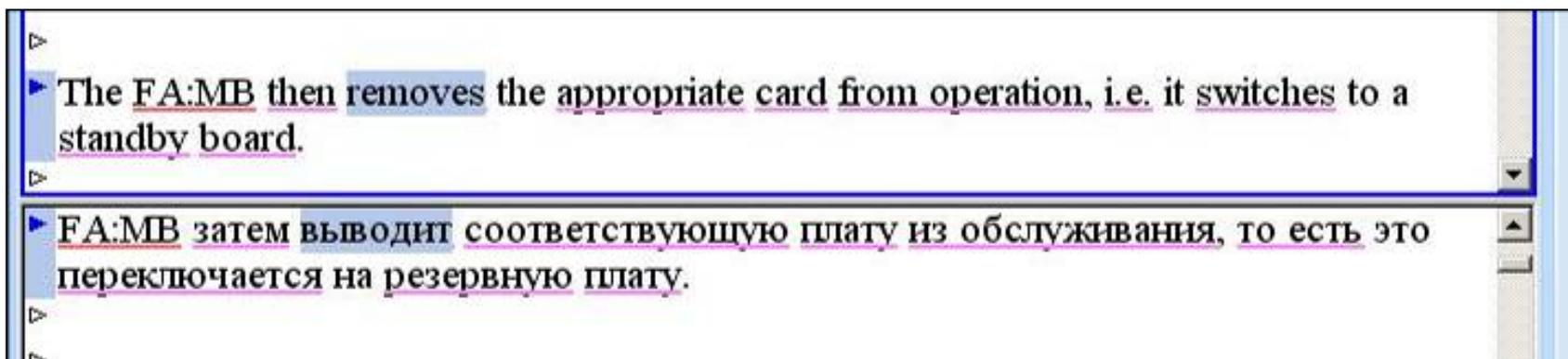
# Пример

# Log



# Разрывные глаголы

- to clear = очищать
- to clear ... alarm = сбрасывать (НЕКИЙ) аварийный сигнал
- to remove = удалять
- to remove ... from service = выводить (ЧТО-ТО) из обслуживания



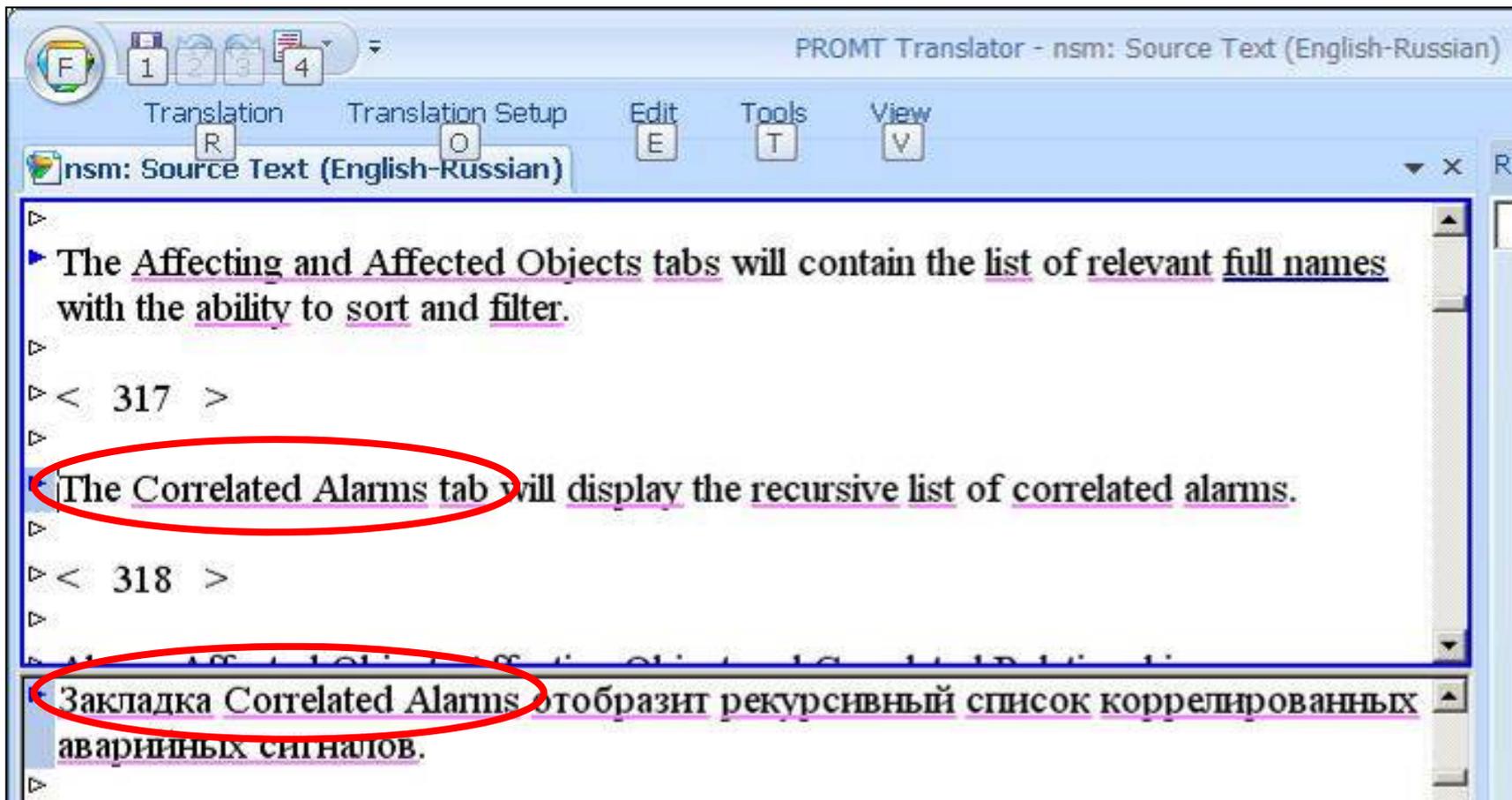
# Термины–команды

- Menu, menu item, window, screen, tab, button, option, check box ...
- Текст перед такими "командами" воспринимается как элемент GUI. Этот текст можно оставить в оригинале, перевести и т.д. (5 вариантов настройки).

MT-вход	MT-выход
the Edit menu	меню Edit
the "Save As" menu item	пункт меню "Save As"
the "Advanced Properties" dialog window	диалоговое окно "Advanced Properties"

# Пример

# Tab



# Некоторые сокращения

OSs	системы OS
backup OS	резервная система OS
export OS	экспортирующая система OS
foreign OS	внешняя система OS
higher OS	система OS более высокого порядка

# Пример

# Foreign OS

The screenshot displays the PROMT Translator interface. The main window shows a source text in English: "Alternatively to IP connection configuration set up CP database for dialog and FTAM over a X.25 DCN for EWSD Classic network elements." Below this, there are two numbered sections: "136" and "137". Section 137 contains the text: "Foreign OS gateway for NE alarm forwarding and remote operation" and "For further information refer to the Foreign OS Administration Manual." The Russian translation of this text is shown below: "Дополнительную информацию см. в Руководстве по администрированию Внешней системы OS." Red circles highlight the terms "Foreign OS" in both the English and Russian versions. On the right side, a "References" pane shows a dictionary entry for "foreign OS" as a noun, defined as "foreign OS" and "foreign OS", with the Russian translation "внешняя система OS".

PROMT Translator - syd: Source Text (English-Russian)

Translation Setup | Edit | Tools | View

English-Russian | EWSD | Translate

General Settings | Translation

syd: Source Text (English-Russian)

References

Foreign OS

PROMT Dictionaries

**foreign OS**, Noun  
"5 of", user dictionary  
• 1. **внешняя система OS**

▶ Alternatively to IP connection configuration set up CP database for dialog and FTAM over a X.25 DCN for EWSD Classic network elements.

< 136 >

▶ Foreign OS gateway for NE alarm forwarding and remote operation

< 137 >

▶ For further information refer to the Foreign OS Administration Manual.

▶ Дополнительную информацию см. в Руководстве по администрированию Внешней системы OS.

# Термины с союзами *and* (*и*) / *or* (*или*)



# Пример

# Public and private networks

What is ATM?  
Asynchronous Transfer Mode (ATM) is a technology designed for the high-speed transfer of voice, video, and data through public and private networks using cell relay technology.

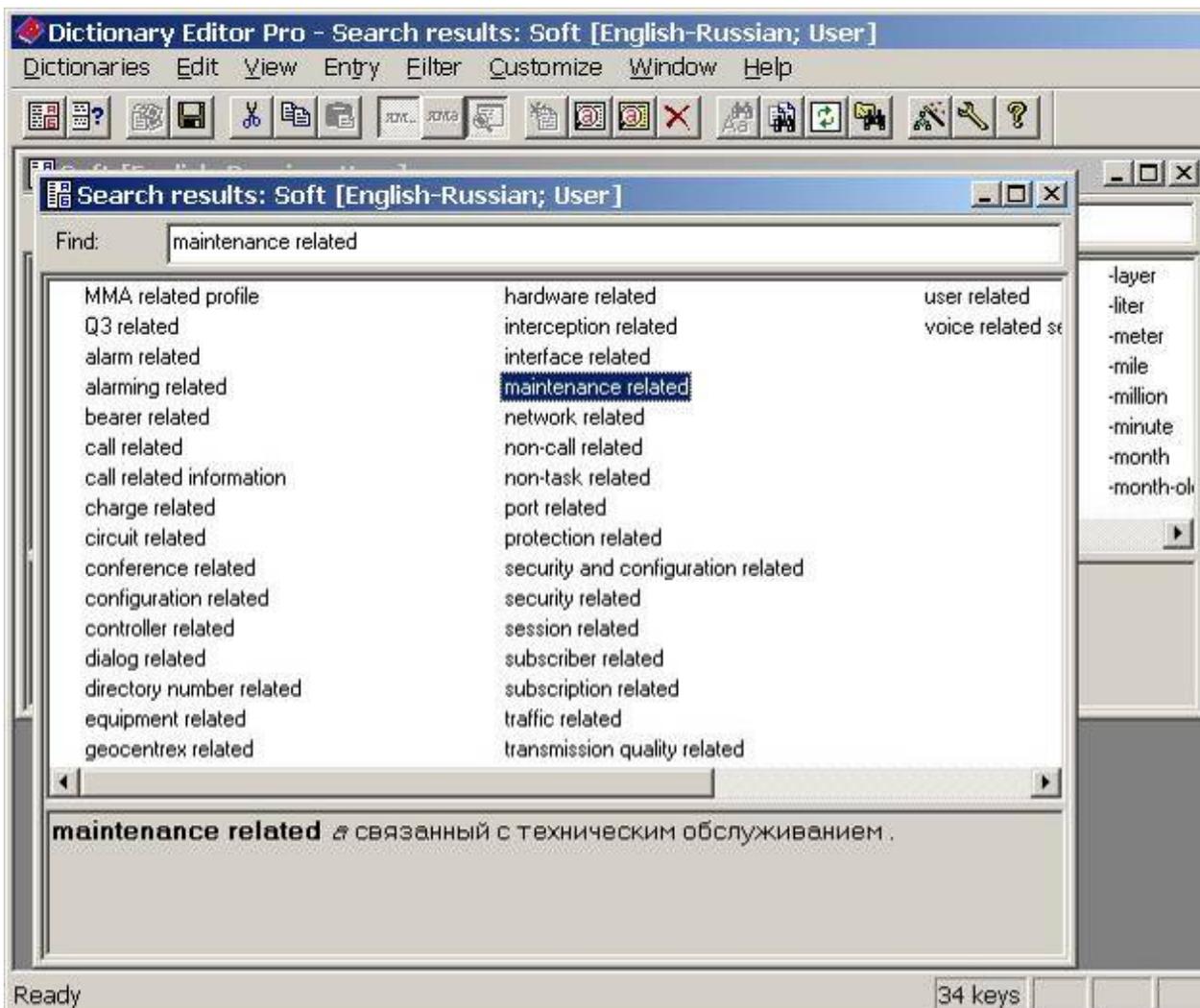
Режим асинхронной передачи (АТМ) является технологией разработанной для высокоскоростной передачи речи, видео, и данных через сети общего пользования и частные сети, использующие технологию ретрансляции ячеек.

Analysis of Translation	
In Source Text	public and private networks
In Translated Text	сети общего пользования и частные сети
Type of the Fragment	Word from a user dictionary

Dictionary

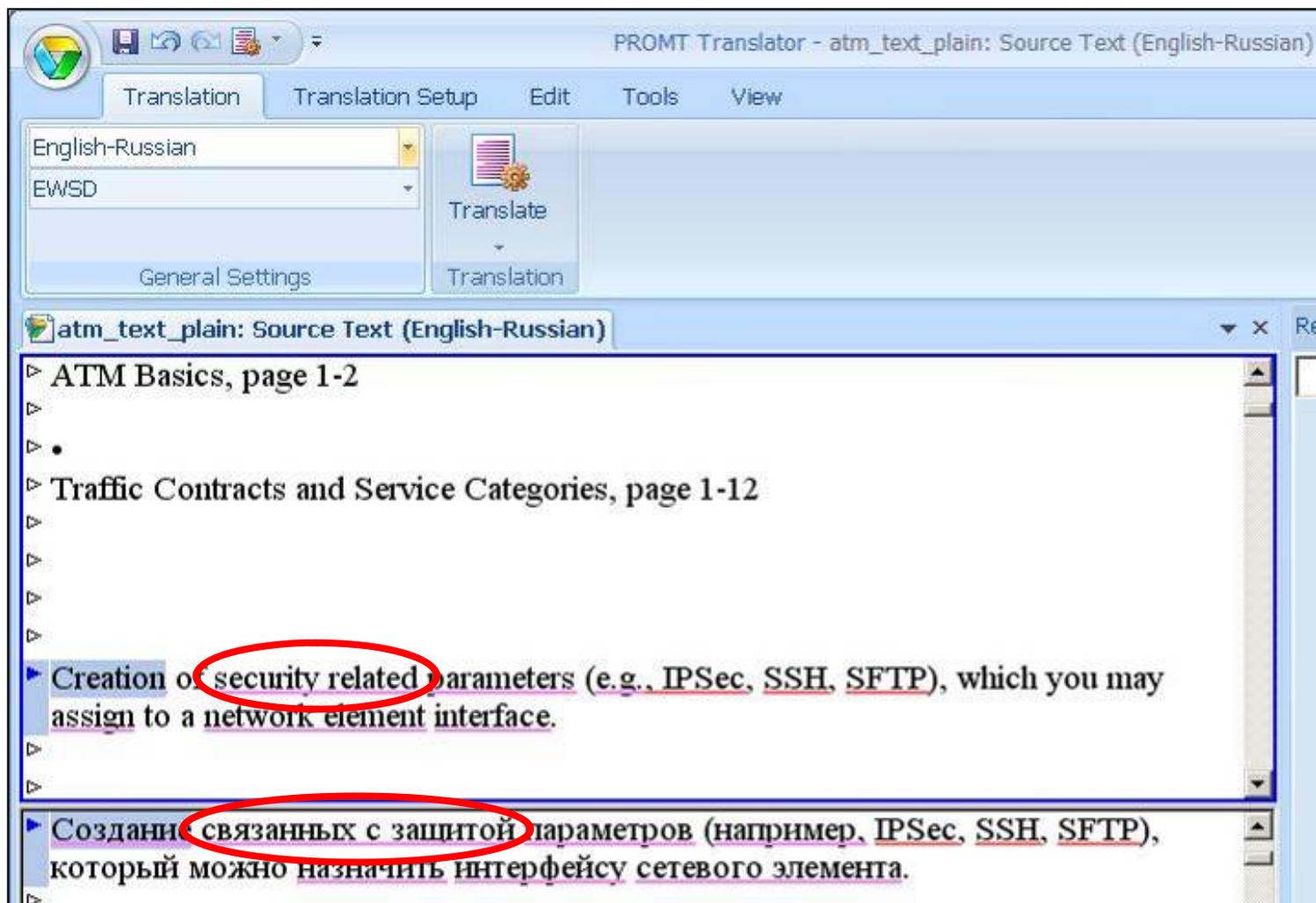
Properties of Fragment Unknown Words

# Конструкции с прилагательным, например, *related* (связанный с)



# Пример

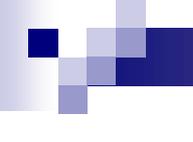
# Security related



# В пользовательском словаре свыше 65 000 словарных статей

Число слов в словарной статье	Число словарных статей	%
1	14 377	21.7
2	34 813	52.7
3	12 726	19.2
4	3 245	4.9
5	616	<1
> 5	280	<1

Часть речи	Число словарных статей	%
Существительное	57 208	88.0
Глагол	2 811	< 1
Прилагательное	3 914	< 1
Наречие	1 925	< 1
Предлог	169	< 1



1

Интеграция системы МТ в технологическом процессе компании

2

Разработка и сопровождение словарей для МТ

3

Управление МТ-профилями и стилем перевода

4

Преимущества применения МТ

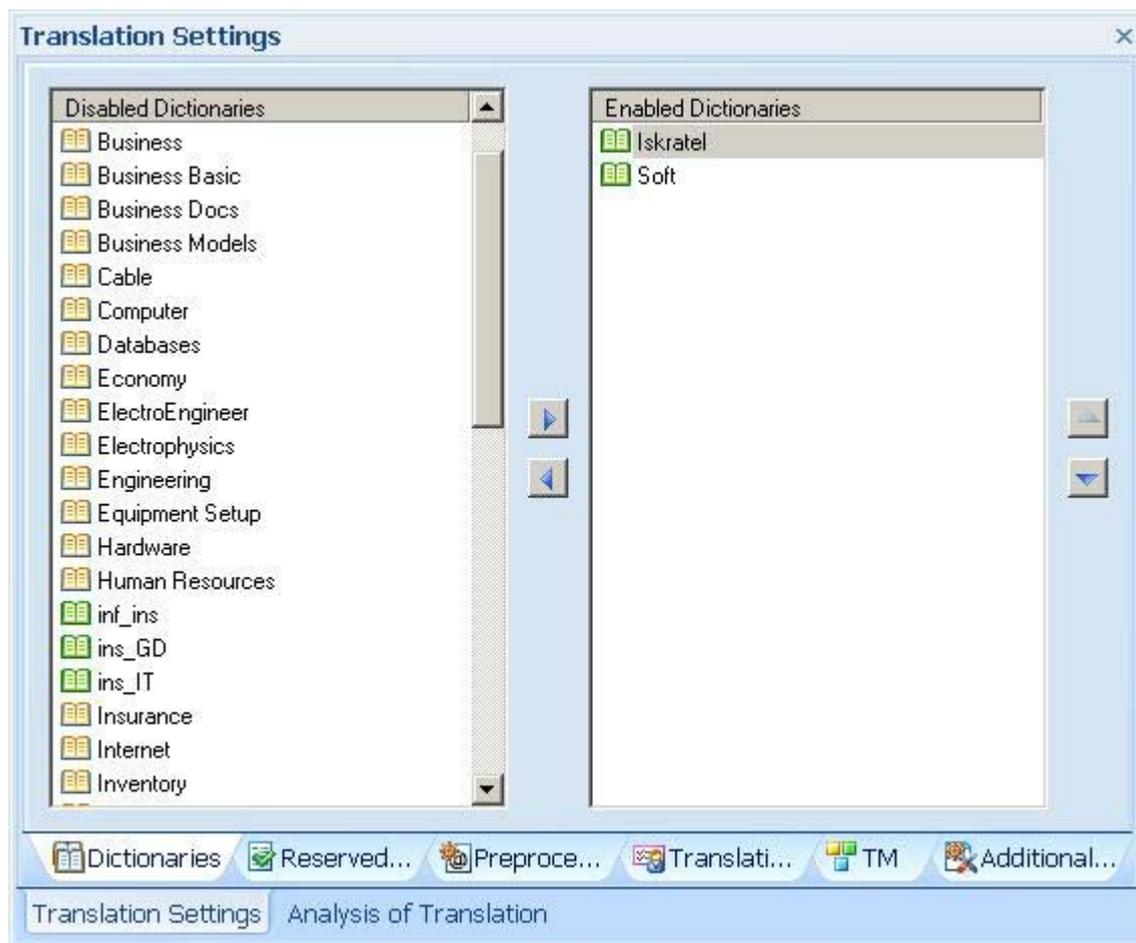
# Профили (шаблоны)

Профиль

Словари

Список Not to Translate

Правила перевода



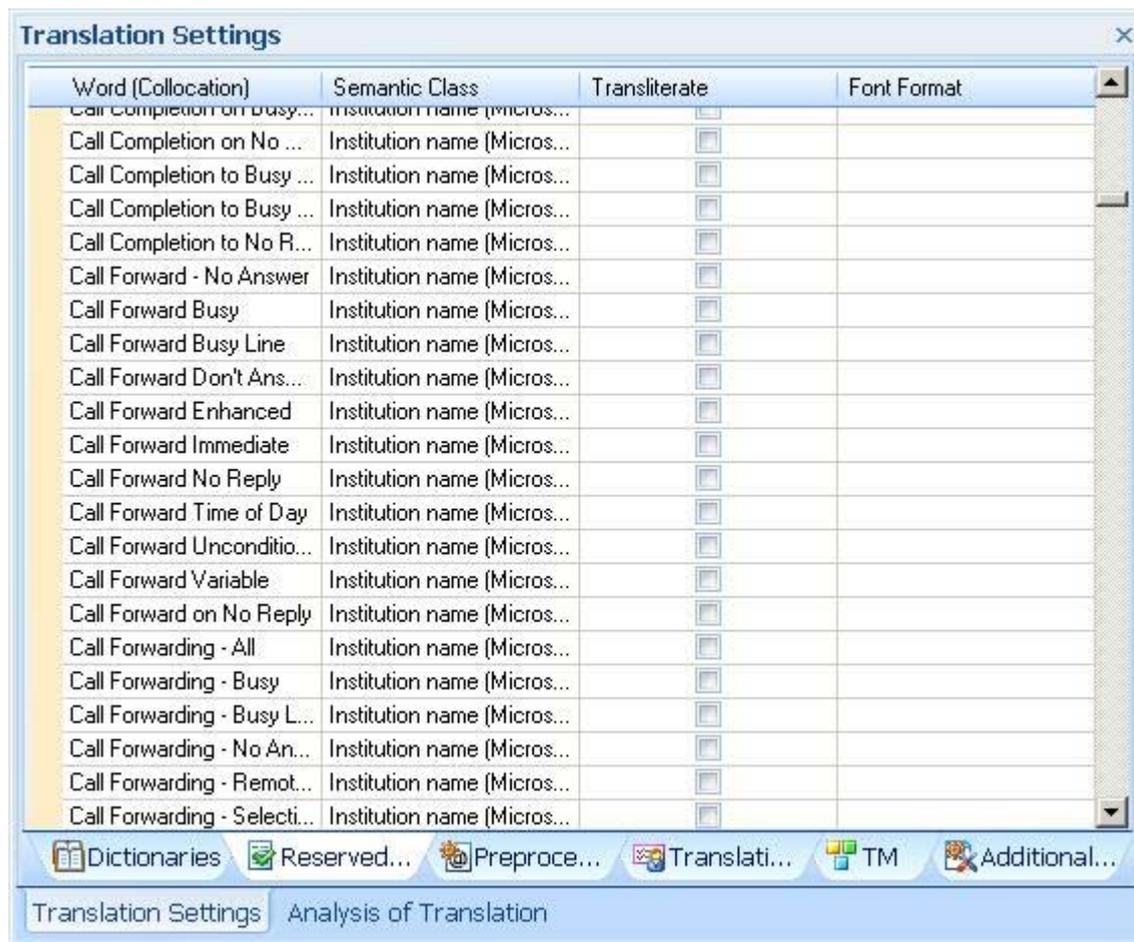
# Профили (шаблоны)

Профиль

Словари

Список Not to Translate

Правила перевода



# Список зарезервированных слов (Not to Translate)

- команды, названия систем, сообщения

Access Integrator Monitor
Access- Request
Access-Accept

- СОСТОЯНИЯ

Availability Status= degraded
Availability Status=failed / degraded

- программные продукты и модули

Microsoft Windows NT 4.0
Microsoft Windows Server 2003

- названия услуг

Call Completion to No Reply
Call Forward - No Answer
Call Forward Busy

- 
- Всего свыше 3 500 терминов (строк)

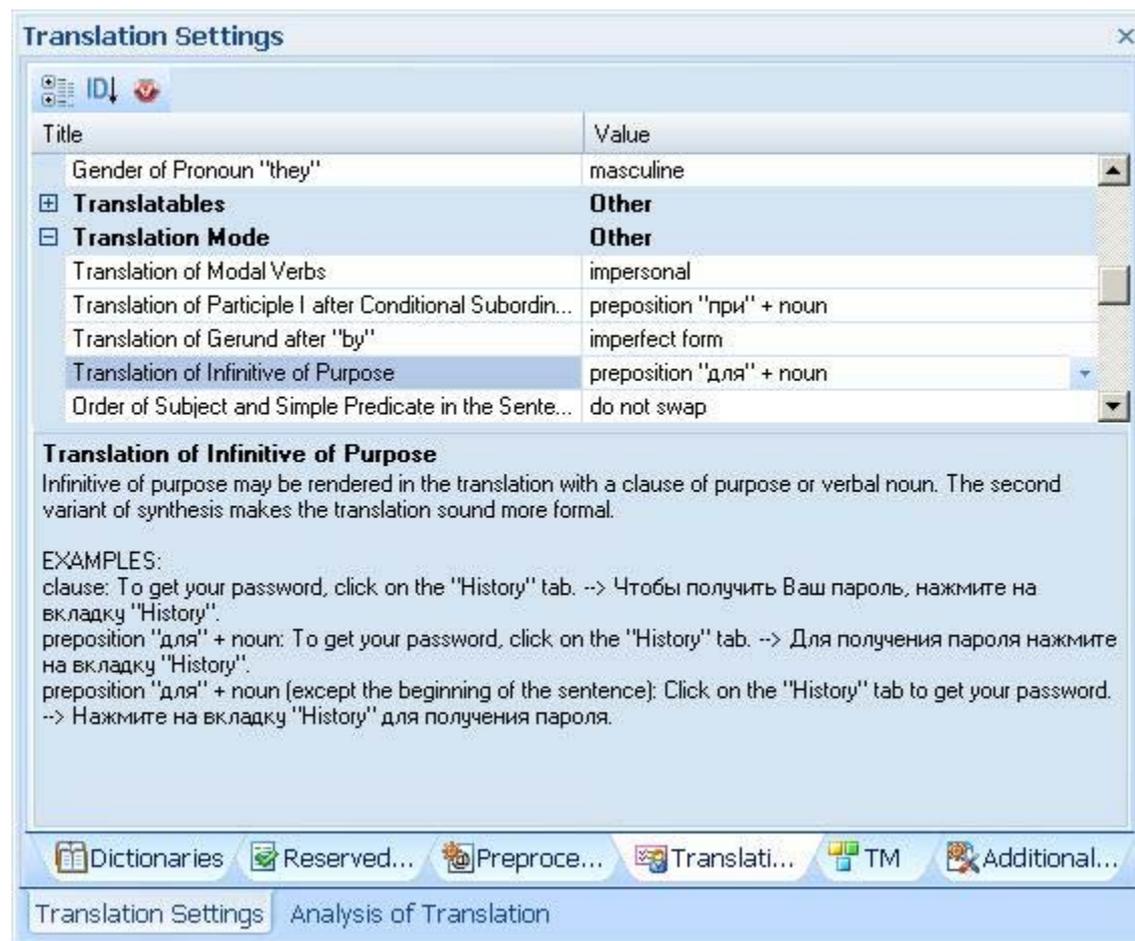
# Профили (шаблоны)

Профиль

Словари

Список Not to Translate

Правила перевода



Title	Value
Gender of Pronoun "they"	masculine
<b>Translatables</b>	<b>Other</b>
<b>Translation Mode</b>	<b>Other</b>
Translation of Modal Verbs	impersonal
Translation of Participle I after Conditional Subordin...	preposition "при" + noun
Translation of Gerund after "by"	imperfect form
<b>Translation of Infinitive of Purpose</b>	preposition "для" + noun
Order of Subject and Simple Predicate in the Sente...	do not swap

**Translation of Infinitive of Purpose**  
Infinitive of purpose may be rendered in the translation with a clause of purpose or verbal noun. The second variant of synthesis makes the translation sound more formal.

EXAMPLES:  
clause: To get your password, click on the "History" tab. --> Чтобы получить Ваш пароль, нажмите на вкладку "History".  
preposition "для" + noun: To get your password, click on the "History" tab. --> Для получения пароля нажмите на вкладку "History".  
preposition "для" + noun (except the beginning of the sentence): Click on the "History" tab to get your password. --> Нажмите на вкладку "History" для получения пароля.

Dictionary | Reserved... | Preproce... | Translati... | TM | Additional...

Translation Settings | Analysis of Translation

# Пример

# Правила перевода

The screenshot displays a translation rule in a software interface. The rule text is shown in both English and Russian. In the English text, "Statistics can also be collected on-demand by pressing the Collect button on the right of the Statistics window." The words "collected", "Collect button", and "Statistics window" are circled in red. The Russian text is "Статистические данные могут также быть собраны по запросу путем нажатия на кнопку Collect (Собрать) справа от окна Statistics (Статистика).", with "Собрать", "Collect (Собрать)", and "Statistics (Статистика)" circled in red. Below the text is an "Analysis of Translation" table.

Analysis of Translation	
In Source Text	collected
In Translated Text	, собранных
Type of the Fragment	Word from a user dictionary
Dictionary	

At the bottom of the interface, there are tabs for "Properties of Fragment" and "Unknown Words".

# Пример

# Правила перевода

In order to perform a more complex sorting, right click on the list heading and choose Show Sorting menu item from the contextual menu.

< 1515 >

Для того, чтобы выполнить более сложную сортировку, щелкните правой кнопкой мыши на заголовке списка и выберите пункт меню Show Sorting из контекстного меню.

Analysis of Translation	
In Source Text	In order to
In Translated Text	Для того, чтобы
Type of the Fragment	Word from a user dictionary

Dictionary

Properties of Fragment Unknown Words

# Управление единообразием на уровне стиля

- Типовые выражения (например, “*it should be noted that*”)
- Часто встречающиеся обороты

<i>this chapter describes</i>	в этой главе приводится описание (кого-чего)
<i>this command deletes</i>	эта команда позволяет удалить (кого-что)
<i>this manual covers</i>	в этом руководстве приводится описание (кого-чего)

# Пример

# Следует отметить, что

Search results: Soft [English-Russian; User]

Find: it can be noticed that

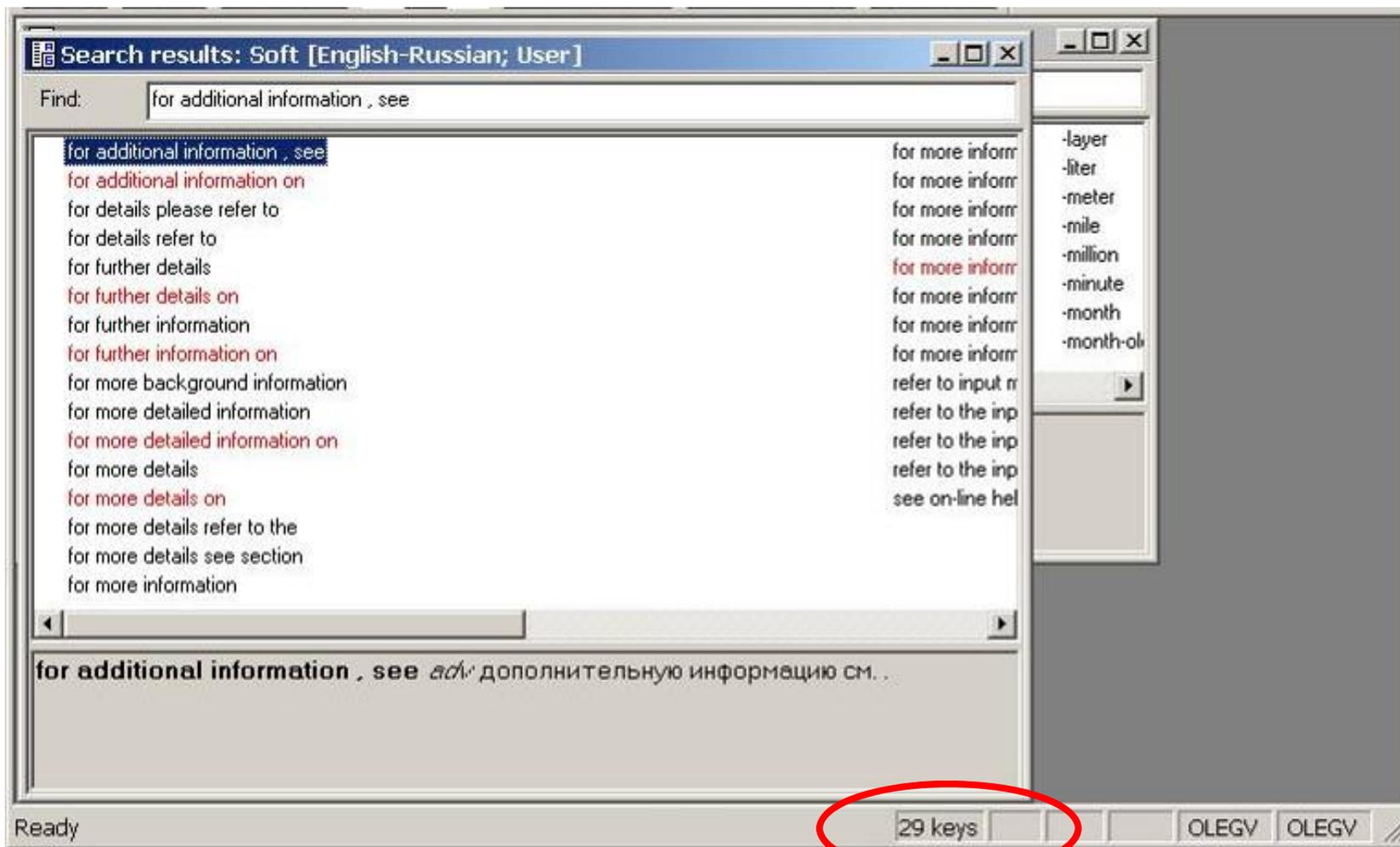
it can be noticed that	note , however , that	please reg
it is also important to note that	note , that	please tak
it is also worth noting that	note also that	remark the
it is important to note that	note here that	remark thc
it is noted that	note however , that	take note
it is to be noted that	note however that	
it is worth noting that	note in particular that	
it must be noted , that	note that	
it must be noted that	note,	
it must be taken into account , that	note, that	
it should , however , be noted that	notice , that	
it should also be noted that	notice that	
it should be noted here that	please consider , that	
it should be noted that	please note , that	
it should be noticed that	please note that	
it should be pointed out that	please notice that	

it can be noticed that *следует отметить, что.*

Ready 37 keys OLEGV OLEGV

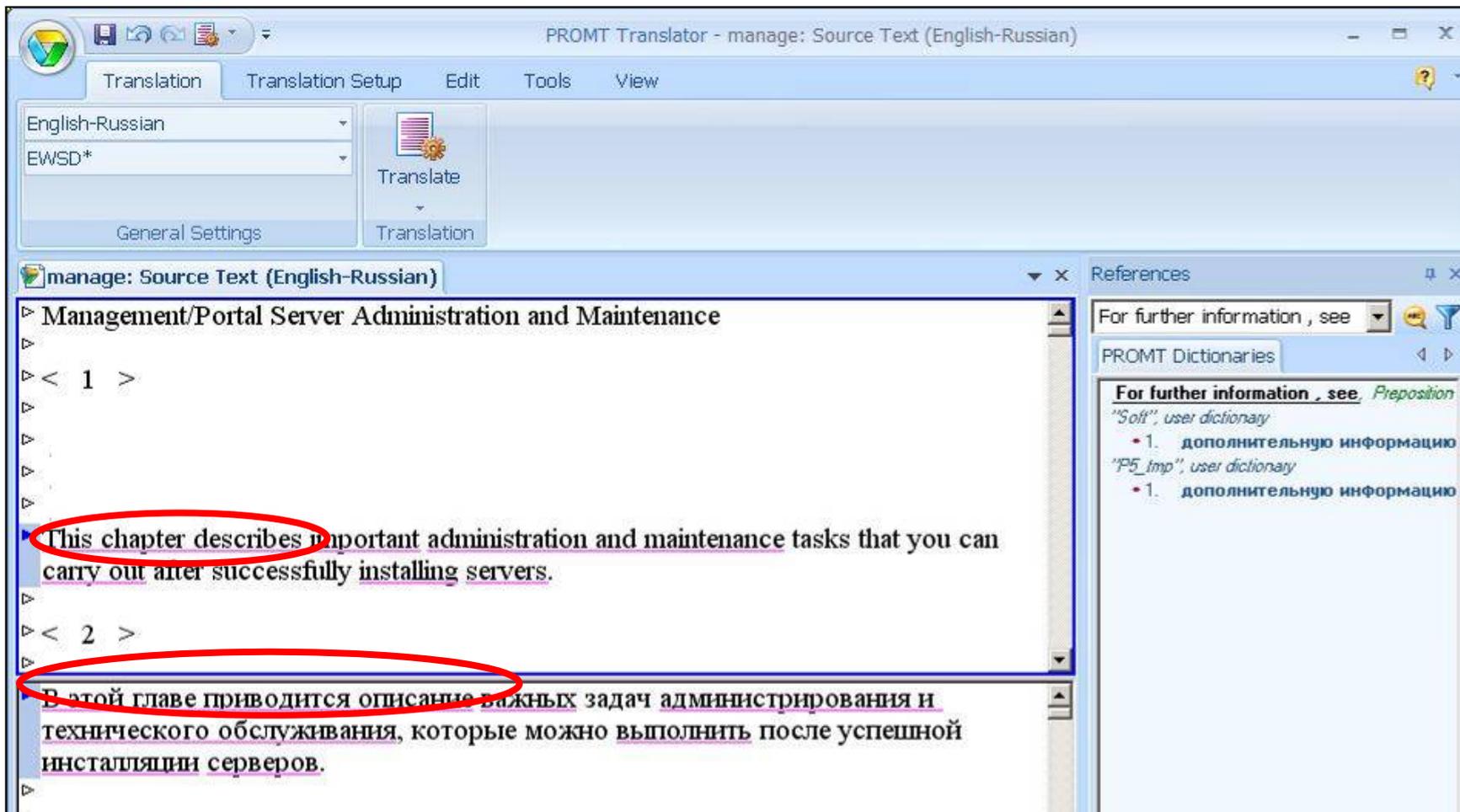
# Пример

## Дополнительную информацию см.



# Пример

# This chapter describes



1

Интеграция системы МТ в технологическом процессе компании

2

Разработка и сопровождение словарей для МТ

3

Управление МТ-профилями и стилем перевода

4

Преимущества применения МТ

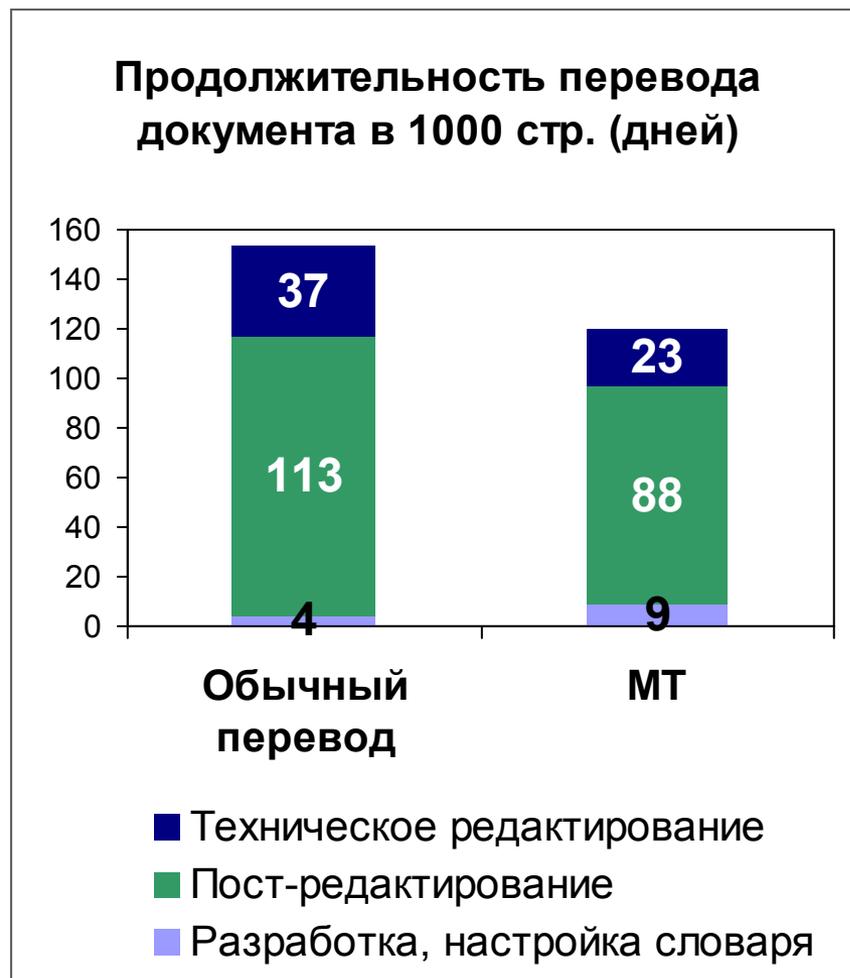
# Преимущества настроенного МТ

- **Автоматизация.** Словарь – не просто справочник, а инструмент автоматической замены терминов с исходного языка на целевой.
  - **Производительность.** Минимизация рутинной работы по набору текста на клавиатуре. Повышение скорости работы переводчиков (в разумных пределах).
  - **Качество.** Повышение качества перевода благодаря управлению единообразием терминологии и элементов стиля (уже на начальном этапе процесса перевода).
  - **Обучение персонала.** Новые сотрудники быстрее осваивают особенности терминологии и стиля.
  - **Диверсификация переводческих услуг.** Возможность предоставления качественного чернового МТ-перевода (без последующего редактирования) – понятно, оперативно, недорого.
- 
- **Повышение эффективности бизнеса переводческой компании за счет оптимизации технологических процессов**

# Эффективность применения МТ

- Разработка и настройка словаря занимает в 2 раза больше времени
- Производительность пост-редактирования увеличивается на 25-30%, технического редактирования – на 60-67%
- **Пример:** перевод документа в 1000 страниц при использовании МТ занимает на 34 дня меньше (22%). При постоянной загрузке за этот период можно получить **дополнительную выручку 142 тыс. руб.\***

\*при стоимости страницы для заказчика 500 руб.



# Спасибо за внимание!

- Для контактов:

[olegv@argonaut.spb.ru](mailto:olegv@argonaut.spb.ru)

[oleg.vigodsky@gmail.com](mailto:oleg.vigodsky@gmail.com)

- Бесплатный пробный МТ-перевод файлов (по тематике Telecom + IT)
  - до 50 стр.
  - ТХТ, RTF, ТТХ