Машинный перевод как инструмент переводчика

Выгодский Олег

Центр переводов «Аргонавт»

О компании «Аргонавт»

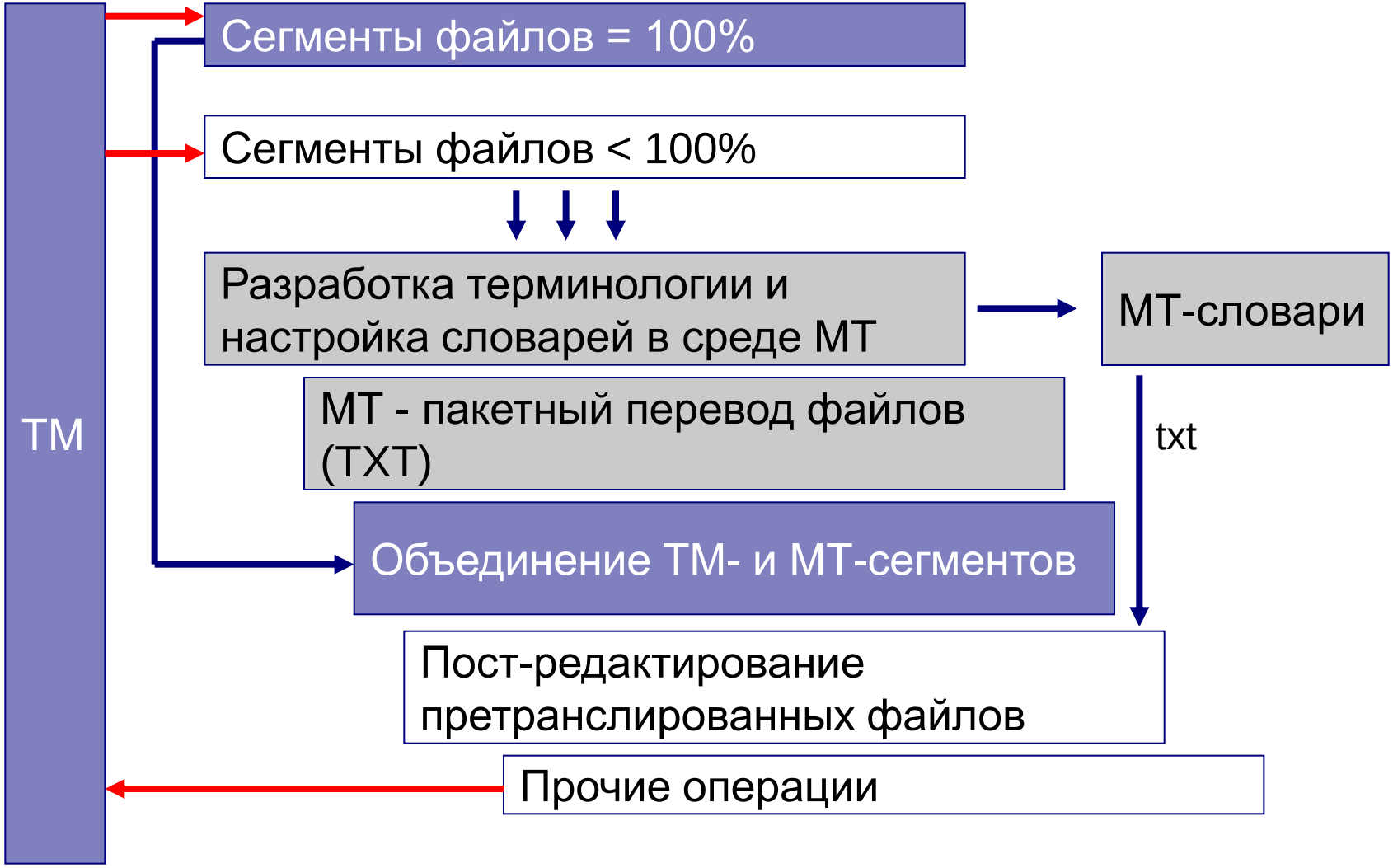
- перевод документации в сфере Telecom + IT, с английского на русский, с 1992 г.
- основные заказчики: AT&T, Italtel, Samsung, Ericsson, ZTE, Siemens
- применение системы машинного перевода (MT) с 1993 г.
- применение систем Translation Memory (TM) с 1997 г.

Структура презентации

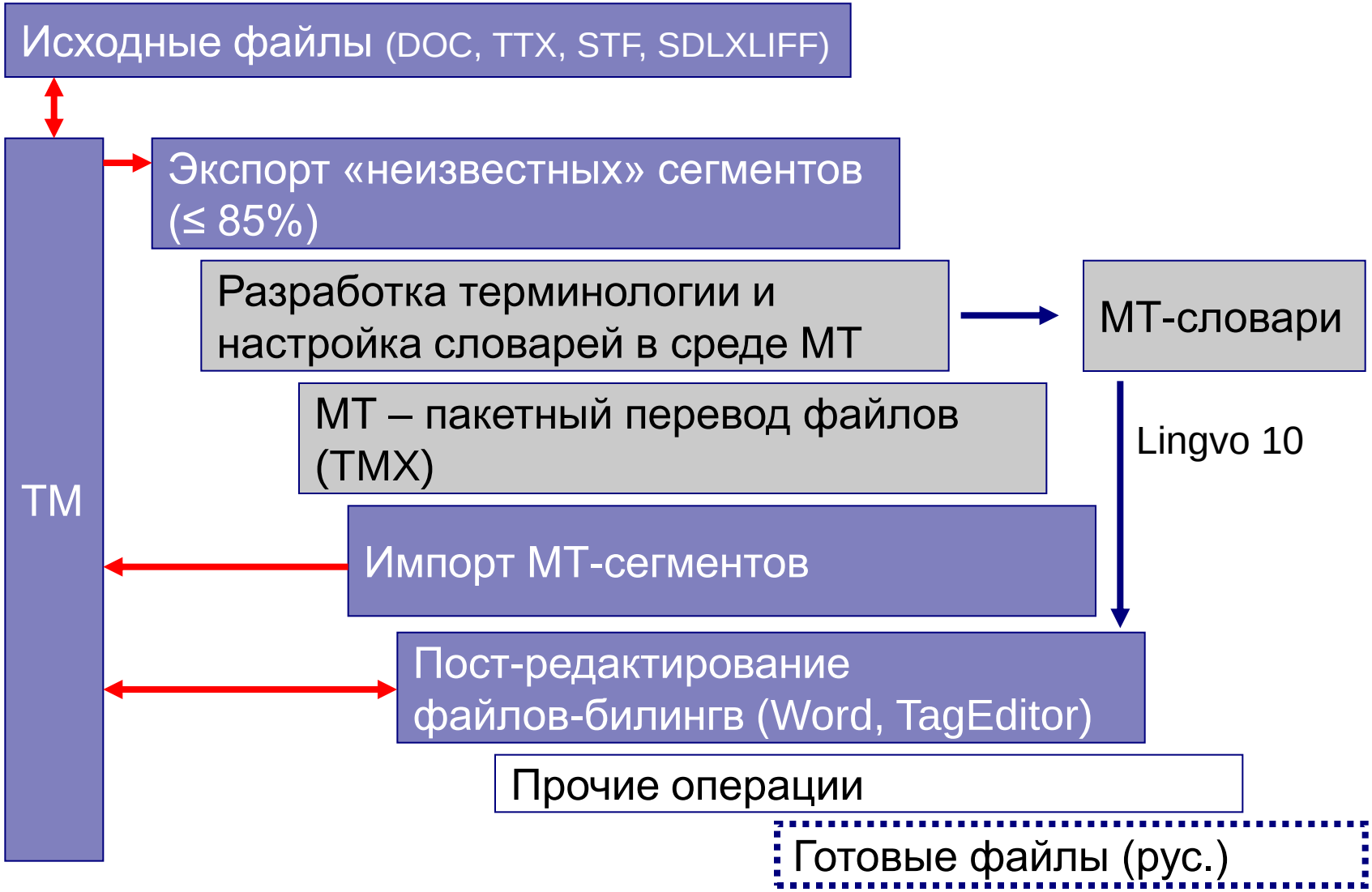
1. Интеграция системы МТ в технологическом процессе компании
2. Разработка и сопровождение словарей для МТ
3. Управление МТ-профилями и стилем перевода
4. Преимущества применения МТ

1993 г. Stylus (DOS) for Software

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

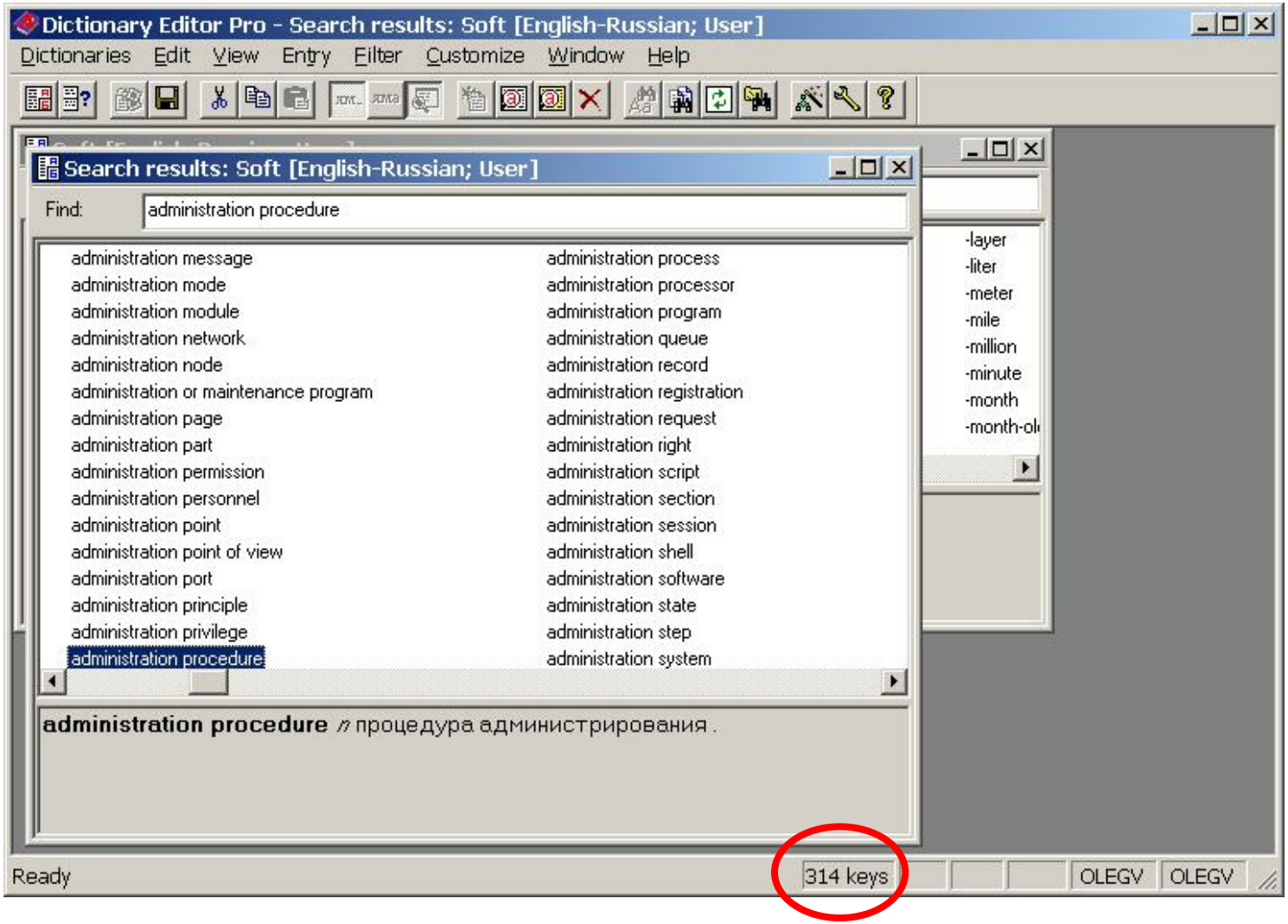
3 Управление МТ-профилями и стилем перевода

4 Преимущества применения МТ

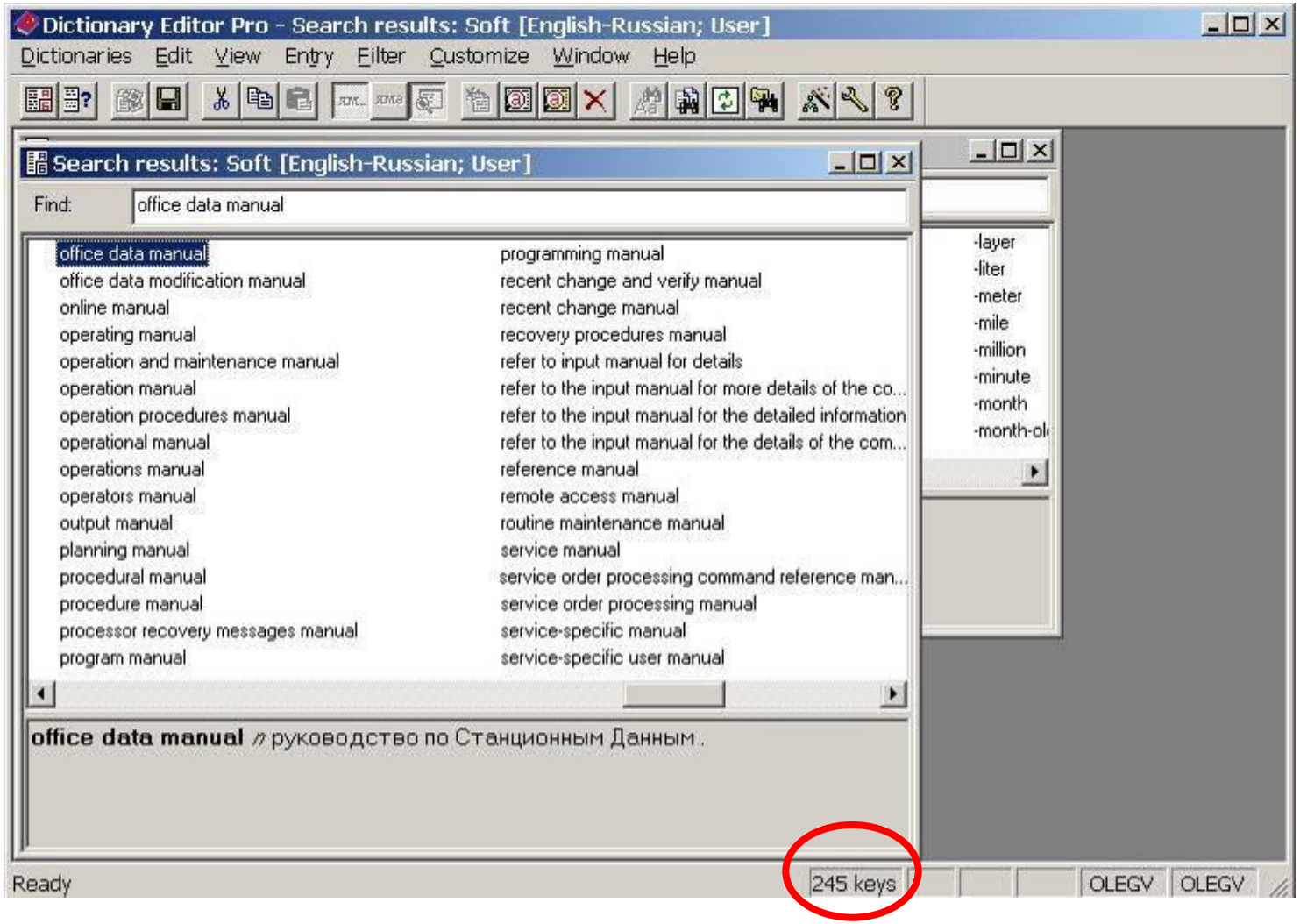
Источники лексики и уровни терминологии для МТ

- 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

Dictionary Editor Pro - Search results: Soft [English-Russian; User]

Dictionaries Edit View Entry Filter Customize Window Help

Search results: Soft [English-Russian; User]

Find: Q3 event log

Q3 event log	alarm log	archived redo log file
Q3 log	alarm log administration tool	archivelog
Q3 log file	alarm log file	archivelog file
Windows event log	alarm log record	audit log
accounting log	alarm log retrieval	audit log file
accounting log file	alarm/error log	backup log
accounting log record	alert log	backup log file
action log	alert log file	batch job log file
activity log	application event log	call log
activity log file	application log	command and response log
activity log group	application log file	command log
agent log	archive log	command log file
agent operation log	archive log file	communication log
alarm and event log	archive redo log	configuration change log
alarm event log	archive redo log file	configuration log
alarm event log file	archived redo log	console log

Q3 event log // журнал регистрации Q3-событий .

Ready 216 keys OLEGV OLEGV

Разрывные глаголы

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

The FA:MB then removes the appropriate card from operation, i.e. it switches to a standby board.

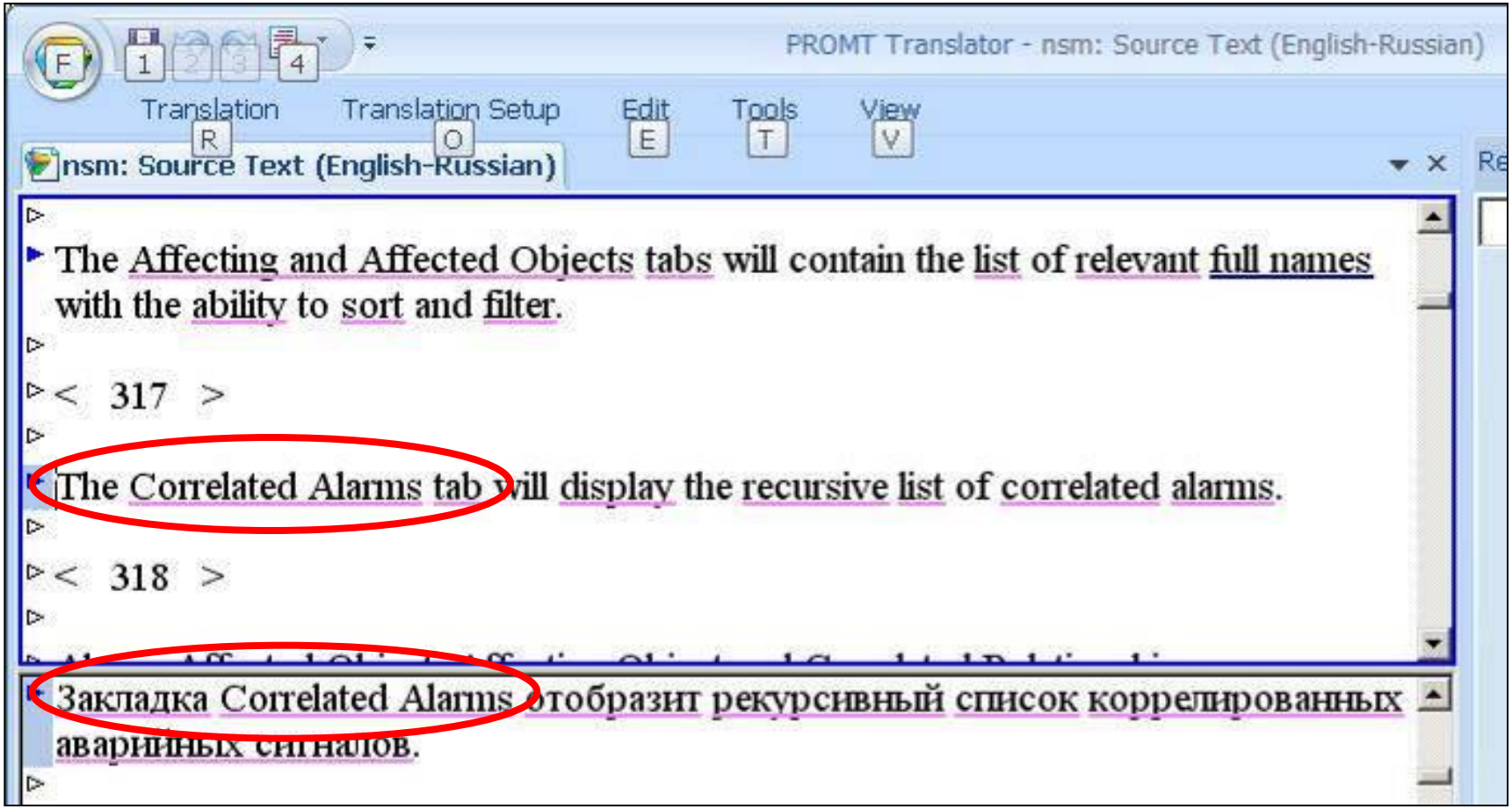
FA:MB затем выводит соответствующую плату из обслуживания, то есть это переключается на резервную плату.

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


PROMT Translator - nsm: Source Text (English-Russian)
Translation
Translation Setup
Edit
Tools
View
nsm: Source Text (English-Russian)
The Affecting and Affected Objects tabs will contain the list of relevant full names with the ability to sort and filter.
< 317 >
The Correlated Alarms tab will display the recursive list of correlated alarms.
< 318 >
Закладка Correlated Alarms отобразит рекурсивный список коррелированных аварийных сигналов.

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

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

Translation | Translation Setup | Edit | Tools | View

English-Russian
EWSD
General Settings

Translate
Translation

syd: Source Text (English-Russian)

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.

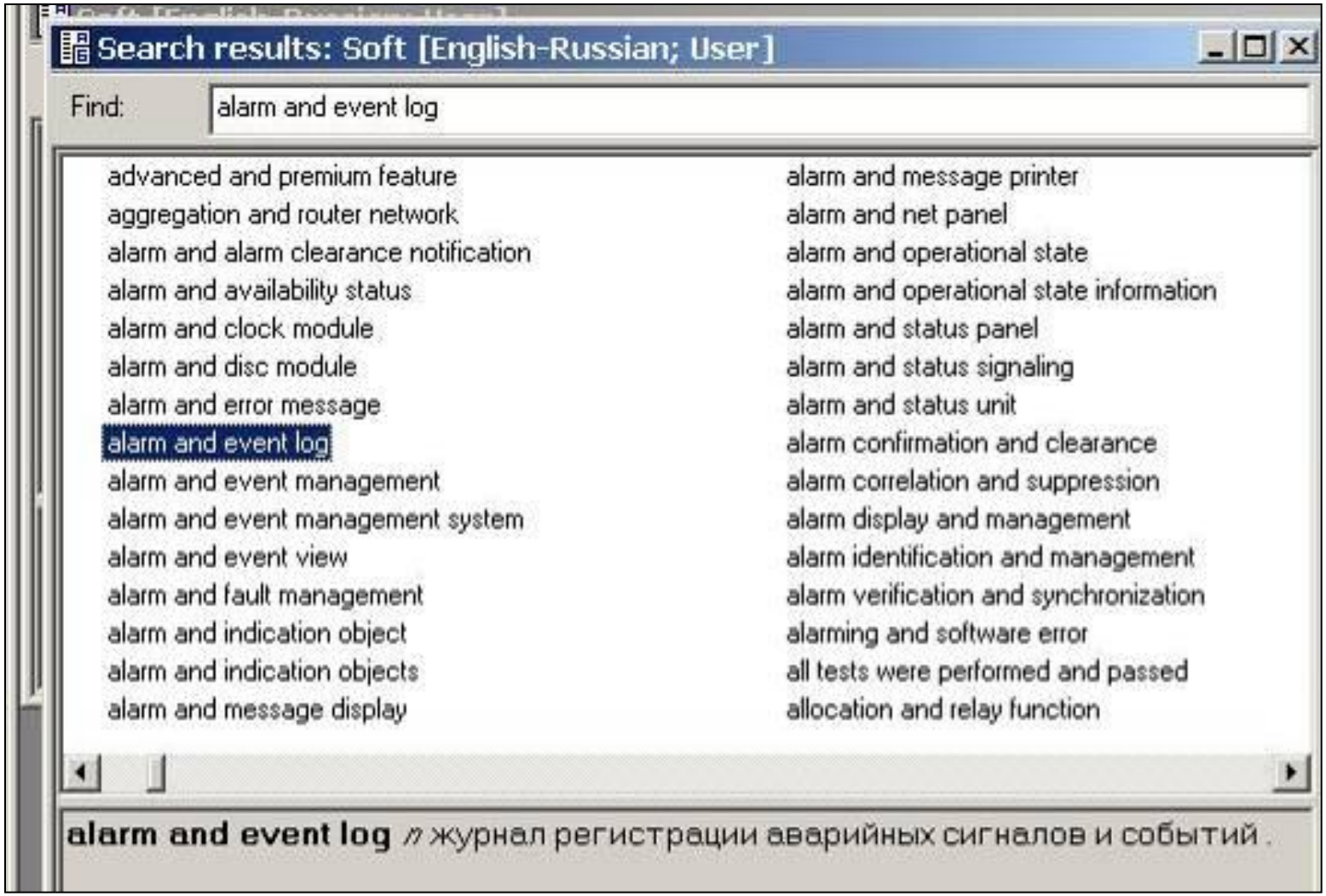
References

Foreign OS

PROMT Dictionaries

foreign OS, *Noun*
"Soft", user dictionary
1. внешняя система 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 (ATM) является технологией, разработанной для высокоскоростной передачи речи, видео, и данных через сети общего пользования и частные сети, использующие технологию ретрансляции ячеек.

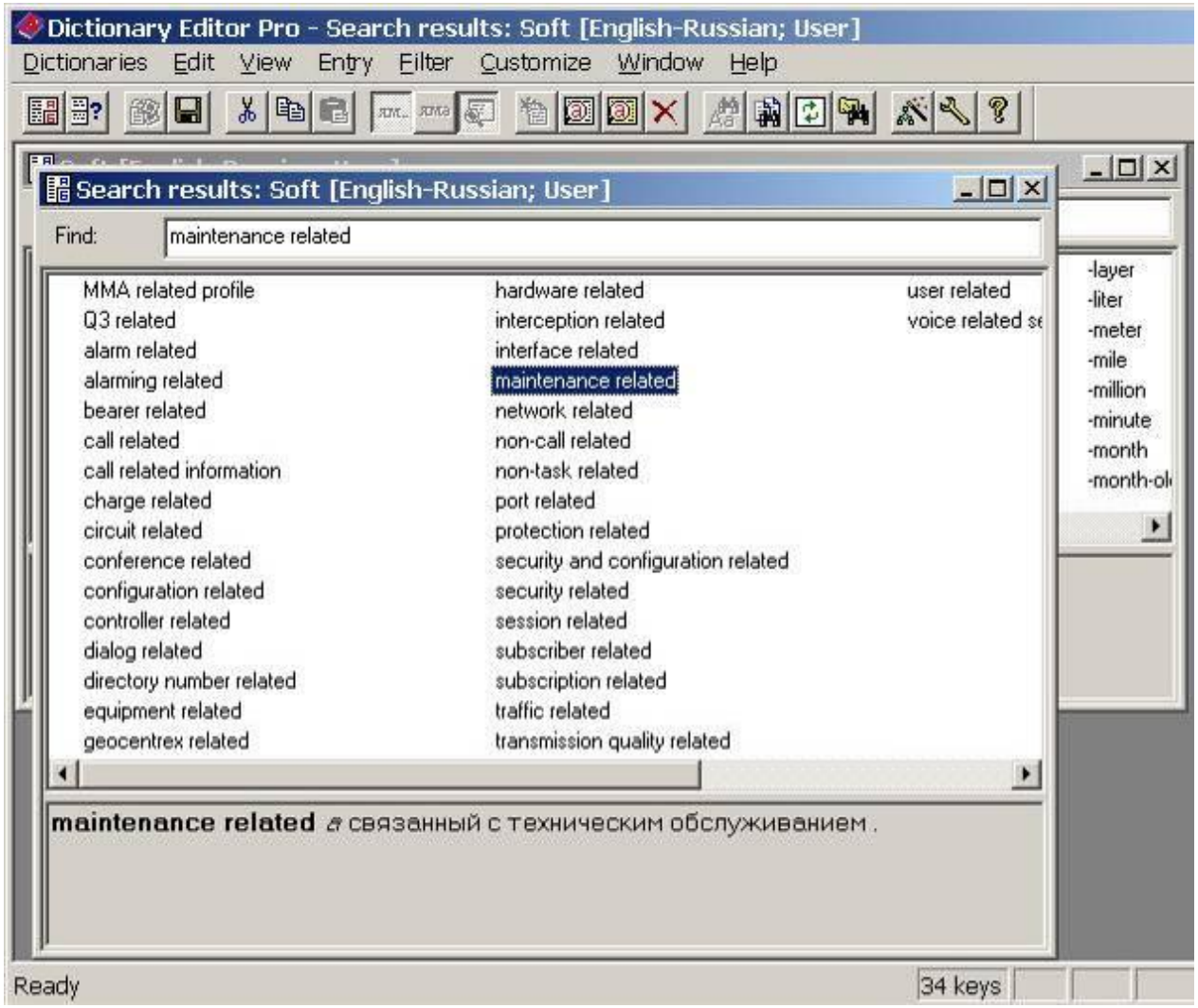
Analysis of Translation

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

Dictionari

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



Пример Security related

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

Translation | Translation Setup | Edit | Tools | View

English-Russian
EWSD
General Settings

Translate
Translation

atm_text_plain: Source Text (English-Russian)

ATM Basics, page 1-2

•

Traffic Contracts and Service Categories, page 1-12

Creation of security related parameters (e.g., IPSec, SSH, SFTP), which you may assign to a network element interface.

Создание связанных с защитой параметров (например, IPSec, SSH, SFTP), который можно назначить интерфейсу сетевого элемента.

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

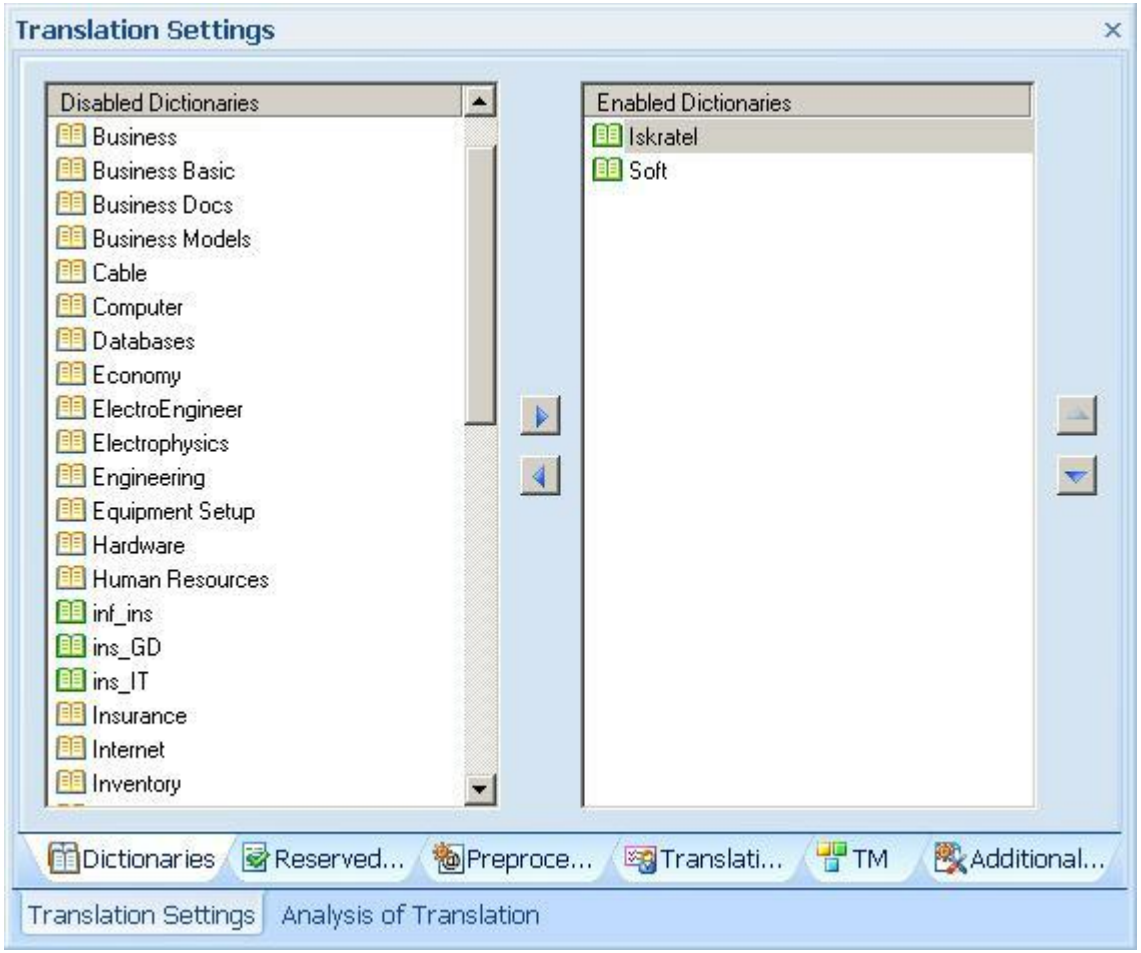
3 Управление МТ-профилями и стилем перевода

4 Преимущества применения МТ

Профили (шаблоны)

Профиль

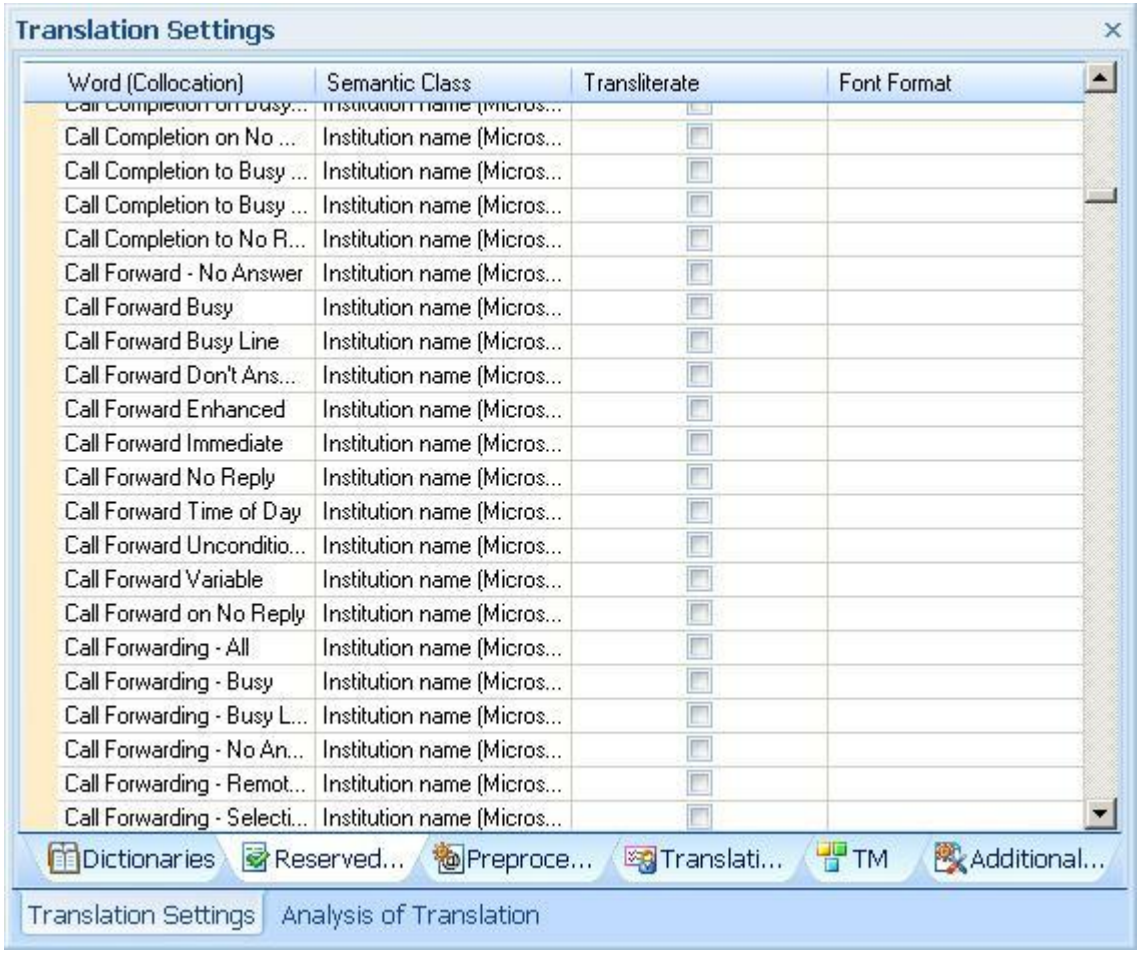
- Словари
- Список Not to Translate
- Правила перевода



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 NT4.0
Microsoft Windows Server 2003

- названия услуг

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

- Всего свыше 3 500 терминов (строк)

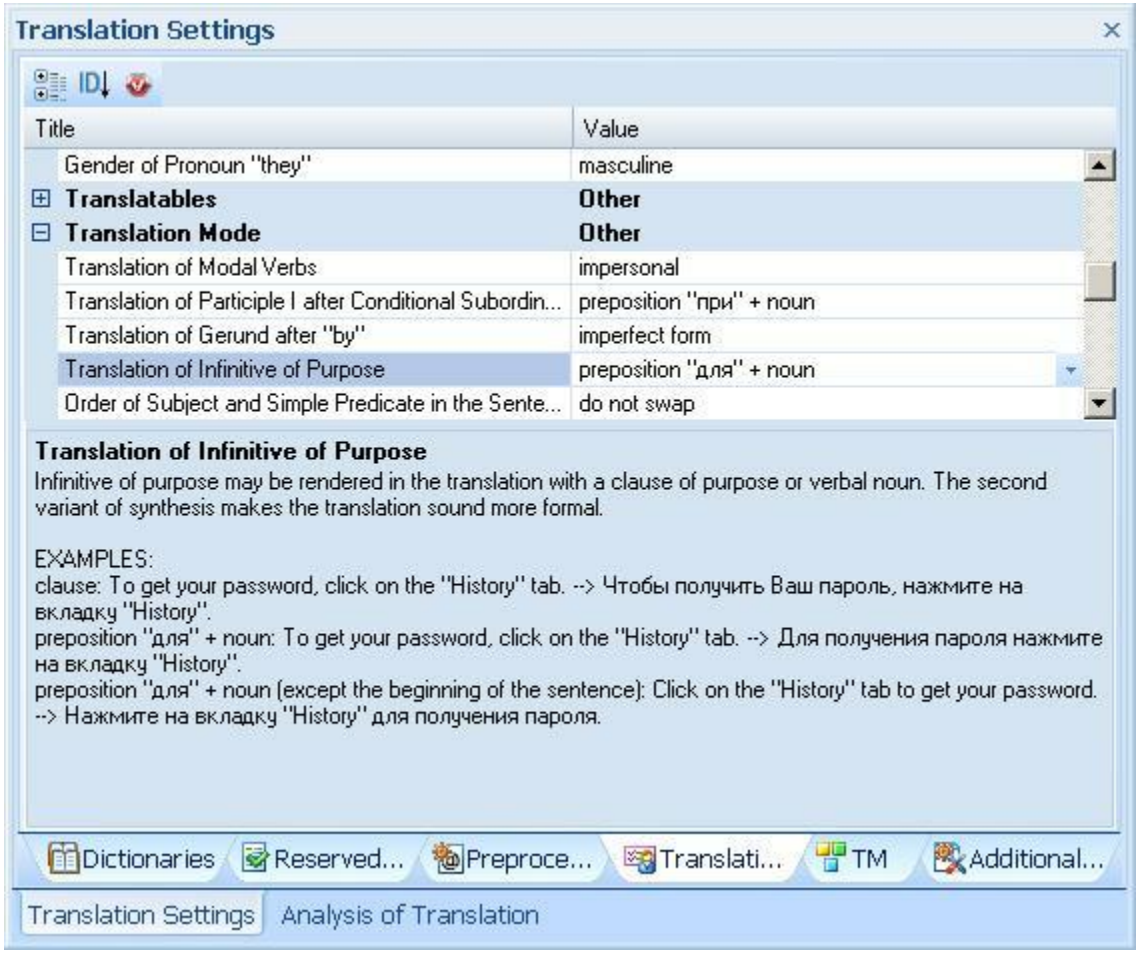
Профили (шаблоны)

Профиль

Словари

Список Not to Translate

Правила перевода



Пример Правила перевода

Statistics can also be collected on-demand by pressing the Collect button on the right of the Statistics window.

< 1509 >

Статистические данные могут также быть собраны по запросу путем нажатия на кнопку Collect (Собрать) справа от окна Statistics (Статистика).

Analysis of Translation

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

Properties of Fragment | 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	

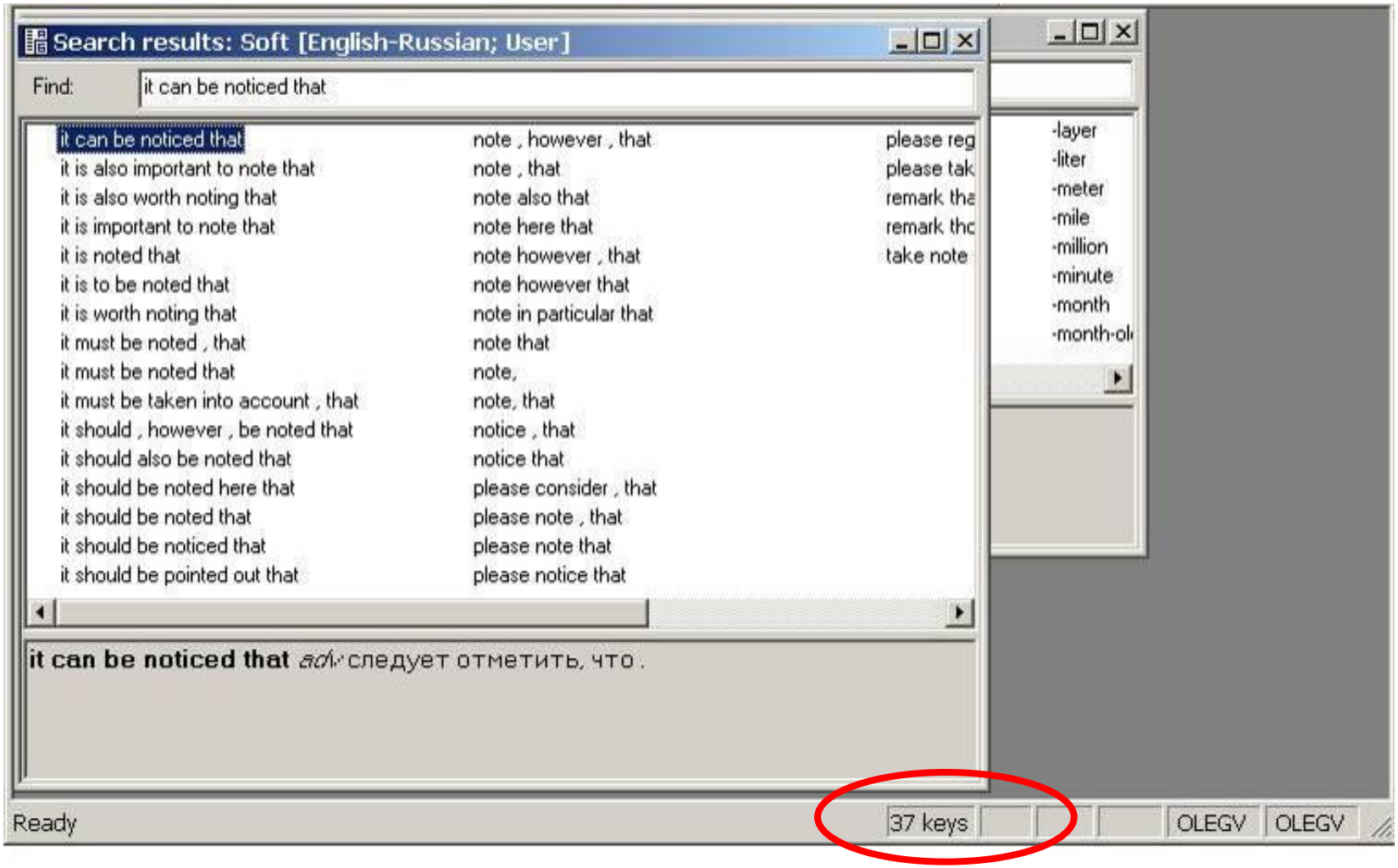
Properties of Fragment | Unknown Words

Управление единообразием на уровне стиля

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

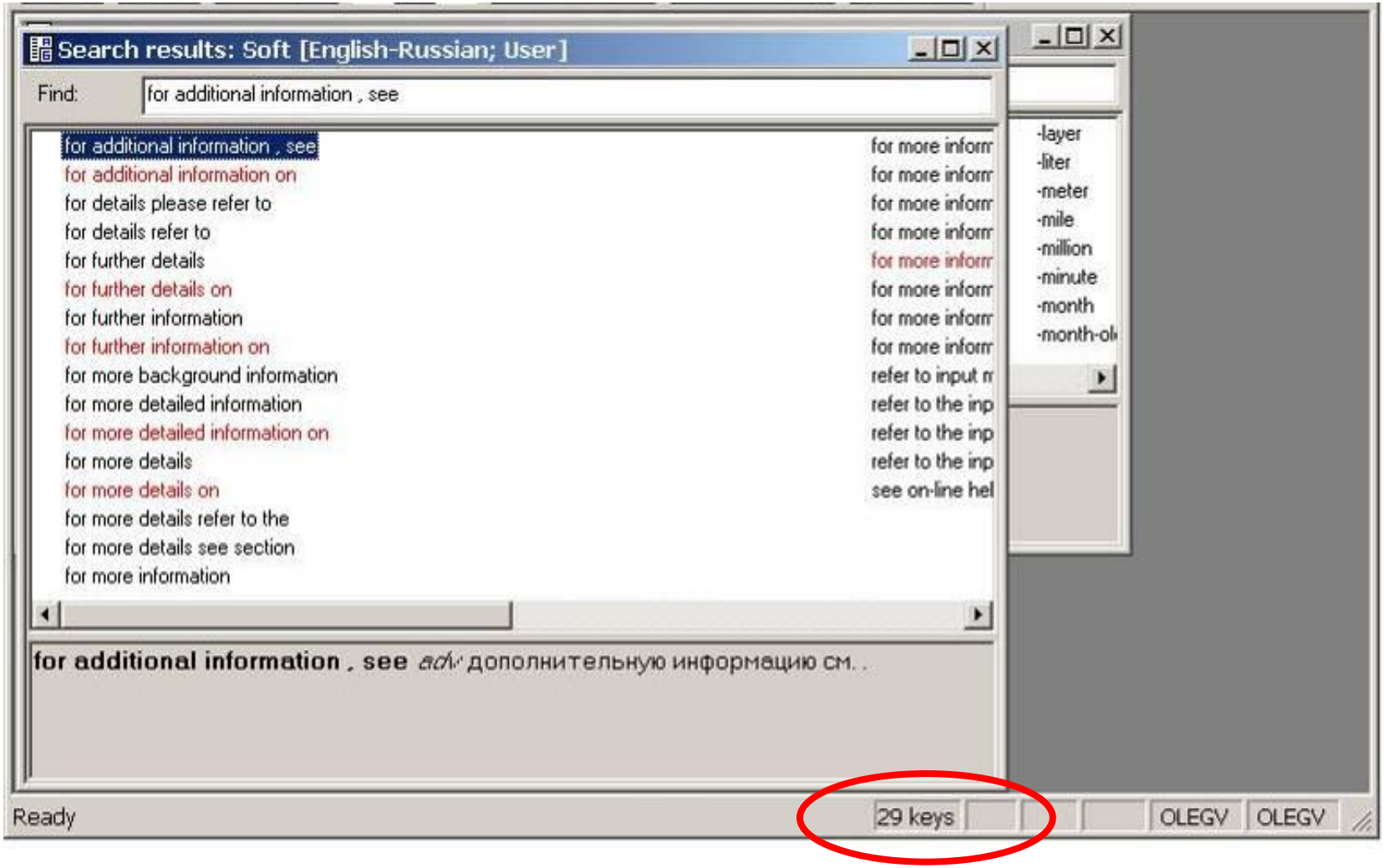
this chapter describes	в этой главе приводится описание (кого-чего)
this command deletes	эта команда позволяет удалить (кого-что)
this manual covers	в этом руководстве приводится описание (кого-чего)

Пример Следует отметить, что



Пример

Дополнительную информацию см.



Пример This chapter describes

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

Translation | Translation Setup | Edit | Tools | View

English-Russian
EWSD*
General Settings

Translate
Translation

manage: Source Text (English-Russian)

Management/Portal Server Administration and Maintenance

< 1 >

This chapter describes important administration and maintenance tasks that you can carry out after successfully installing servers.

< 2 >

В этой главе приводится описание важных задач администрирования и технического обслуживания, которые можно выполнить после успешной инсталляции серверов.

References

For further information , see

PROMT Dictionaries

For further information , see, *Preposition*
"Soft", user dictionary
• 1. **дополнительную информацию**
"P5_tmp", user dictionary
• 1. **дополнительную информацию**

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

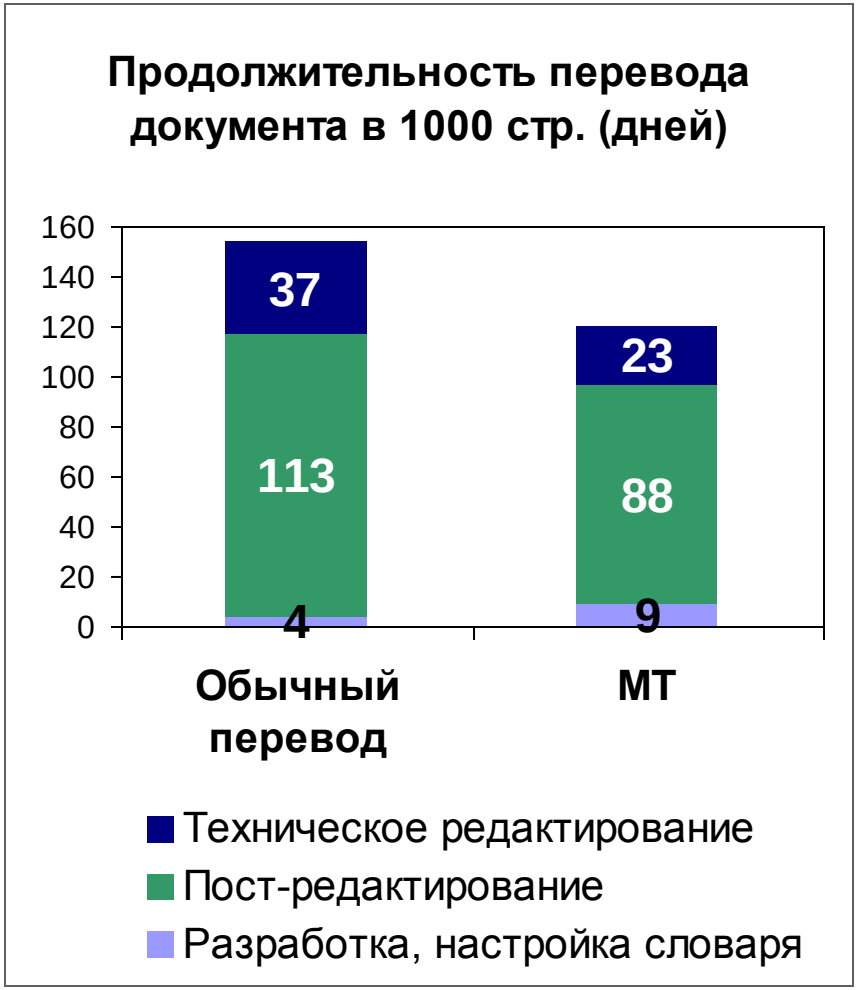
- **Автоматизация.** Словарь – не просто справочник, а инструмент автоматической замены терминов с исходного языка на целевой.
- **Производительность.** Минимизация рутинной работы по набору текста на клавиатуре. Повышение скорости работы переводчиков (в разумных пределах).
- **Качество.** Повышение 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

  oleg.vigodsky@gmail.com

- Бесплатный пробный MT-перевод файлов (по тематике Telecom + IT)
  - до 50 стр.
  - TXT, RTF, TTX