

Использование информационных технологий для оптимизации процесса перевода

презентацию подготовила студентка
группы № 33804/2,
Рудь Ксения

План презентации:

- ▣ **Состояние современного рынка переводов**
- ▣ **Виды программ оптимизации перевода**
- ▣ **Автоматизированный и машинный перевод, понятия и основное различие**
- ▣ **Автоматизированный перевод: CAT-системы и их ключевые функции**
- ▣ **Примеры CAT-систем**
- ▣ **Преимущества и недостатки CAT-систем**
- ▣ **Машинный перевод (МТ) и его основные виды:**
- ▣ **Преимущества и недостатки машинного перевода**

Современный рынок переводов



Программы оптимизации перевода:

- **Автоматизированный перевод (CAT-tools)**
- **Машинный перевод (MT)**
- **Программы распознавания PDF-файлов и изображений (ABBYY FineReader, Adobe Acrobat Pro)**
- **Программы для перевода чертежей (AutoCad)**
- **Программы создания памяти переводов (ABBYY Aligner)**
- **Программы проверки качества перевода (Verifica, xBench)**
- **Голосовой набор текста (Speechpad)**

Автоматизированный перевод

CAT-система (Computer Assisted (Aided) Translation)

– это программное обеспечение, используемое человеком-переводчиком в процессе перевода для повышения производительности труда.

Машинный перевод (Machine Translation, MT) – технология перевода текста с одного естественного языка на другой с использованием методов лексического, грамматического, синтаксического и семантического анализа.

Автоматизированный перевод ≠ Машинный перевод

Автоматизированный перевод осуществляется человеком, машинный не требует участия человека.

CAT-tools

Функции:

- Память переводов (Translation memory)
- Управление терминологией (глоссарными базами)
- Встроенные словари
- Автоматизированная проверка качества перевода
- Совместная работа нескольких специалистов над одним документом
- Распознавание изображений и PDF-файлов

Память переводов

Translation memory

- база данных, содержащая набор ранее переведенных сегментов текста.

Принцип работы:

- сопоставление новых переводческих данных с ранее переведенным материалом
- ПОИСК СООТВЕТСТВИЙ
- применение найденных соответствий в тексте перевода

- Нечеткое соответствие (fuzzy match) – хранящийся в базе сегмент не полностью идентичен новому
- Точное соответствие (exact match) – единица перевода, хранящаяся в базе, полностью совпадает с новой единицей перевода
- Контекстное соответствие (context match) – сегмент, хранящийся в базе, совпадает не только с новым сегментом, но и окружающим его контекстом

Управление терминологией

Терминологическая база – это словарь, содержащий слова и словосочетания, имеющие специфическое значение в определенном контексте

Этапы управления терминологией:

- ручное или автоматическое извлечение терминов и перевод (Monolingual extraction, Bilingual extraction)
- создание глоссария, интеграция в рабочий процесс
- автоматический контроль правильности перевода терминологии

Встроенные словари

- ABBY Lingvo
- Мультитран
- SlovoED
- Мультилекс
- Polyglossum

Автоматизированная проверка перевода

AQ (Quality Assurance) – проверка качества

Функции:

- проверка орфографии и опечаток
- специфические проверки

Совместная работа нескольких специалистов (**Collaboration**)

Специалисты, участвующие в переводе текста:

- Переводчик
- Редактор
- Корректор
- Эксперт

CAT-системы:

- SmartCAT
- SDL Trados
- Wordfast
- Memsource
- Déjà Vu
- OmegaT
- XTM
- Wordbee
- CASMACAT

SmartCAT

← Я smartcat.ai The mystery of why you can

< Домой The mystery of why you can > The mystery of why you can

SmartCAT

Прогресс

Все сегменты сохранены. Готово

Сегменты

Оригинал Перевод

№	Оригинал	Перевод	Статус
1	The mystery of why you can't remember being a baby?	Почему мы не помним, как были детьми	Готово
2	Babies are sponges for new information – so why does it take so long for us to form your first memory?	Дети – губки, постоянно впитывающие новую информацию; так почему же проходит так много времени, прежде чем у нас появятся первые воспоминания?	Готово
3	You're out to lunch with someone you've known for a few years.		Перевод
4	Together you've held parties, celebrated birthdays, visited parks and bonded over your mutual love of ice cream.		Перевод
5	You've even been on holiday together.		Перевод
6	In all, they've spent quite a lot of money on you – roughly £63,224.		Перевод
7	The thing is: you can't remember any of it.		Перевод

Комментарии к документу | Комментарии к сегменту | История | Проверка

№	Описание	Игнорировать
1	Оригинал и перевод оканчиваются на разные знаки пунктуации	<input type="checkbox"/>

Секция CAT

№	Оригинал	ТМ	Перевод
1	The mystery of why you can't remember being a baby?	102%	Почему мы не помним, как были детьми

Секция CAT info

Конкордантный поиск | Словари

Различия

The mystery of why you can't remember being a baby?

Языки: Английский – Русский

Создана: 04.12.2016 01:00:05

Изменена: 04.12.2016 01:00:05

Память переводов: The mystery of why you can

Файлы Статистика Предварительный перевод

+ Создать проект

ПРОЕКТЫ

РЕСУРСЫ

ПЛАТЁЖНЫЙ КАБИНЕТ

КЛИЕНТЫ

НАСТРОЙКИ

Прогресс

Дата создания 04.12.2016

Описание Sample project

Статус Создан

Документы

Поиск по имени... + Загрузить

<input type="checkbox"/>	Документ перевода	Тип	Прогресс	Статус	Язык ...	Автор	Дата со...	Пров...
<input type="checkbox"/>	001 (1)	jpg	<input type="text"/>	Создан	en	Ксения Рудь	04.12.2016	0
<input type="checkbox"/>	Научитесь работать в SmartCAT	docx	<input type="text"/>	Создан	en	Ксения Рудь	04.12.2016	0
<input type="checkbox"/>	сексизм	docx	<input type="text"/>	Создан	en	Ксения Рудь	04.12.2016	0

Память переводов

Сохранить Изменить

Название	Язык оригинала	Язык перевода	Порог	Запись
Sample project for Ксения Рудь	Русский	Английский	75 <input type="text"/>	<input checked="" type="checkbox"/>

Машинный перевод

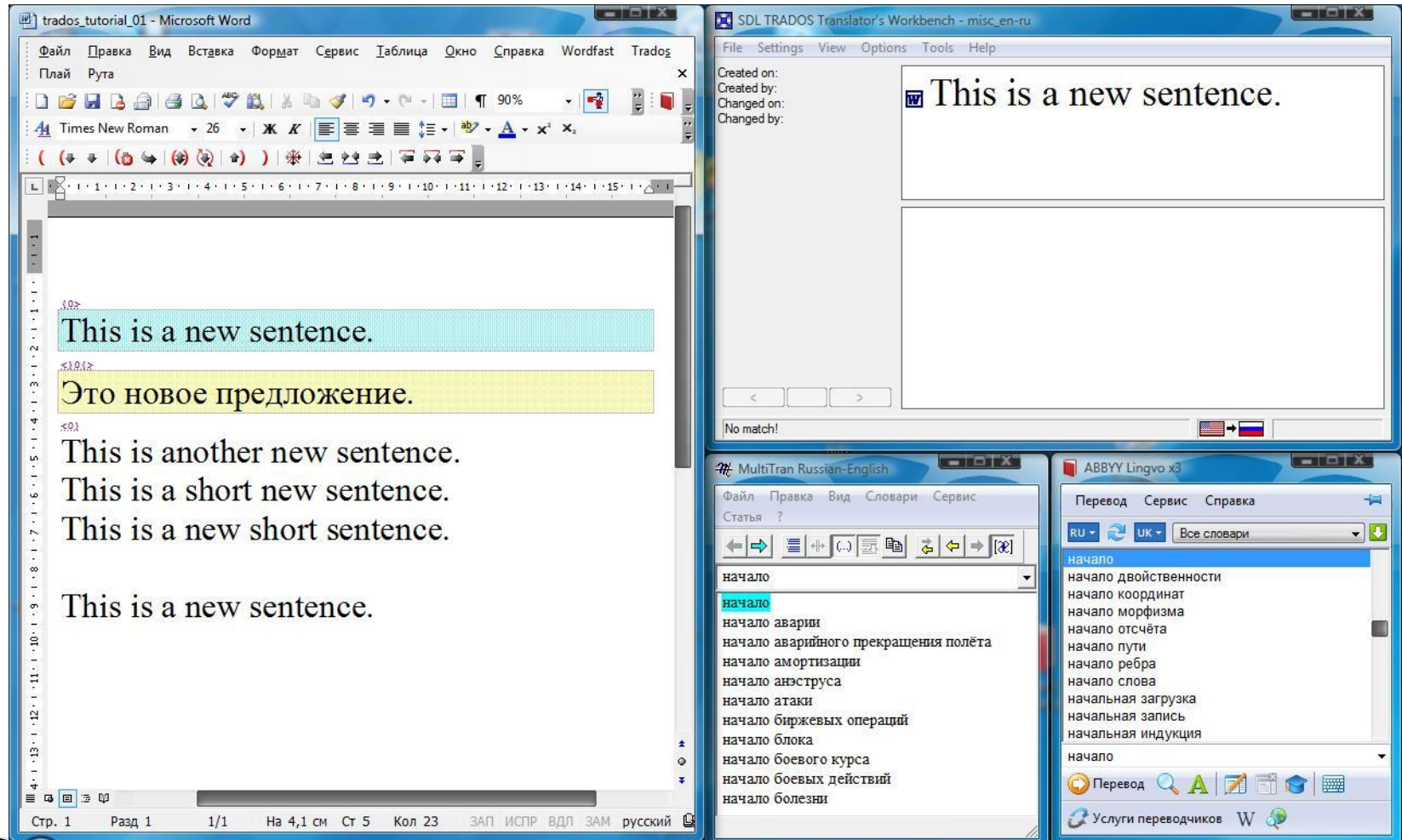
Сохранить и запустить

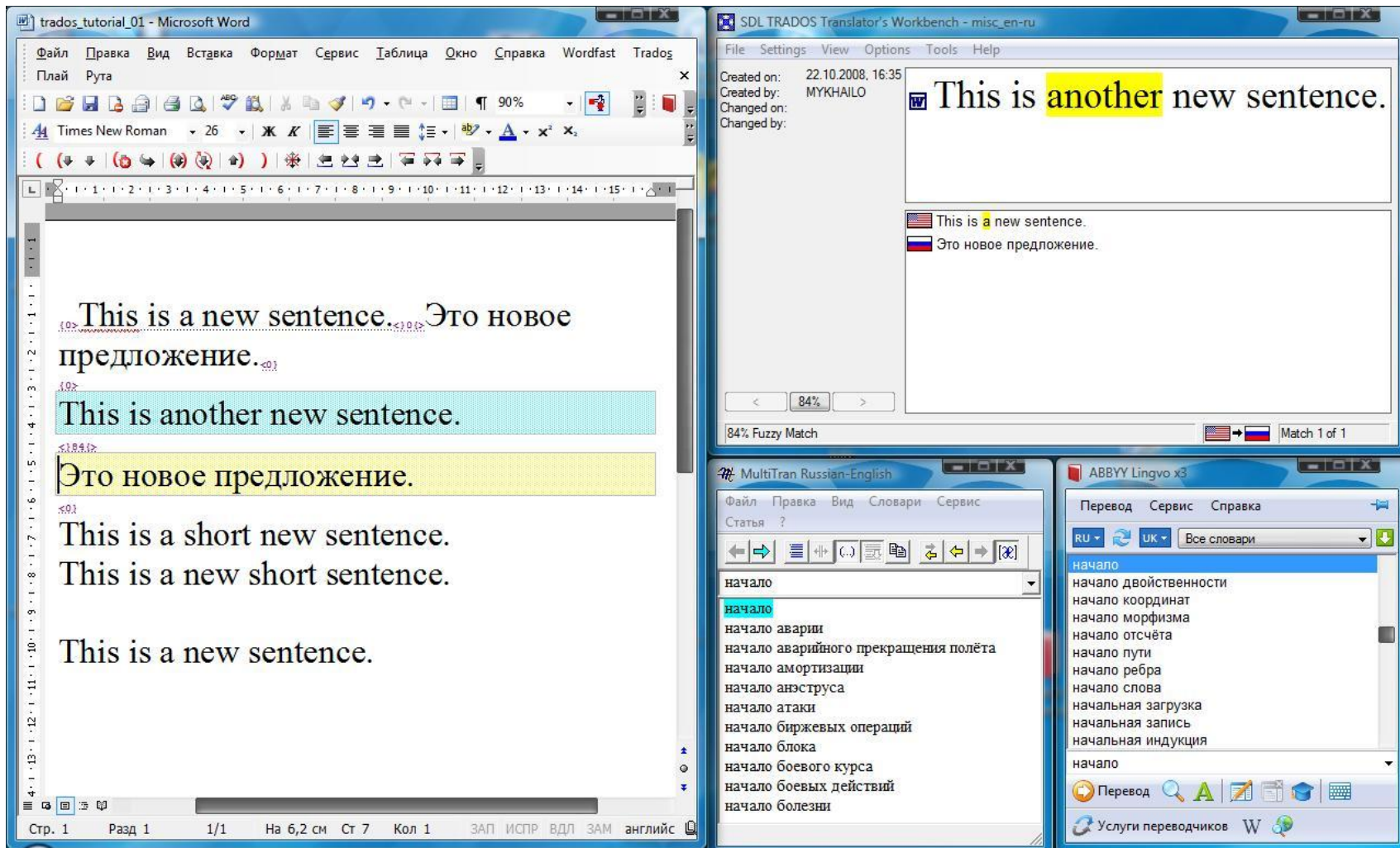
Использовать машинный перевод

Глоссарии

Сохранить

SDL Trados





Wordfast

The screenshot displays the Wordfast software interface. At the top is a menu bar with options: Wordfast Anywhere, File, Edit, View, Translation, Review, TMs & Glossaries, and Help. Below the menu is a toolbar with various icons for file operations and translation functions.

The main workspace is divided into several panels:

- TM Panel:** Located in the upper middle section, it shows a list of translation memory entries. The selected entry (row 6) is:

1	<1> Press release <2> Wordfast LLC <3>	LLC <1> del lanzamiento de prensa <2> <3> Wordfast	100
2	<1> - <2> < > Wordfast Brings Translation Memory Software to the iPad <3> TM <4> < > <5> < > <6>	<1> - <2> < > Wordfast trae software de la industria de la traducción al iPad <3> TM <4> < > <5> <6>	100
3	iPad Users Can Now Use Their Devices to Run The World's Leading Cloud-Based Translation Memory Software	los usuarios del iPad pueden ahora utilizar sus dispositivos para funcionar el software Nube-Basado de la memoria de la traducción del mundo que conduce	100
4	Paris and New York, August 29, 2010	París y Nueva York, 29 de agosto de 2010	100
5	News Facts	Hechos de las noticias	100
6	<1> Wordfast introduces the first professional Translation Memory software application that supports translation on the iPad.	<1> Wordfast introduce el primer uso profesional del software de la memoria de la traducción que apoya la traducción en el iPad.	100
- Document Panel:** Located in the lower middle section, it shows a list of document segments. The selected segment (row 6) is:

6	<1> Wordfast introduces the first professional Translation Memory software application that supports translation on the iPad.	<1> Wordfast introduce el primer uso profesional del software de la memoria de la traducción que apoya la traducción en el iPad.	100
---	---	--	-----
- Glossary Panel:** Located on the right side, it shows a list of terms and their translations:

application	aplicación	view
memory	memoria	view
software	programas y sistemas de programación	view
software	"software"	view
software	software	view
support	soporte	view
support	matriz	view
translation	traslación	view

At the bottom of the interface is a **Status Bar** showing the current document: 'ipad_sample.doc', 94-32%, 6/24, EN-ES.

Преимущества

- исключение повторного перевода
- единообразии перевода
- контроль правильности перевода



- повышение качества перевода
- сокращение временных затрат

Недостатки

- основная работа выполняется человеком
- временные затраты на изучение программного продукта
- высокая стоимость программ

Машинный перевод (МТ)

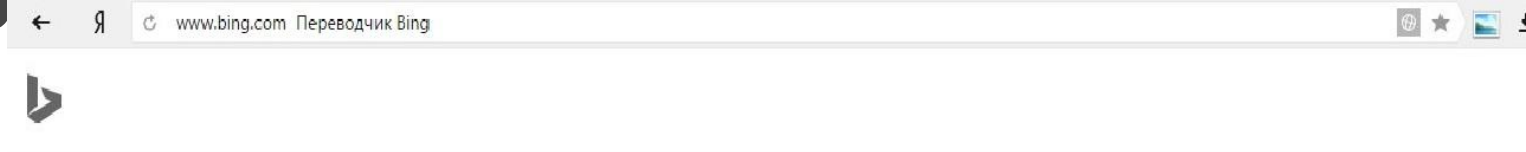
Виды МТ:

- ▣ **Статистические**(Statistical МТ, SMT)
- ▣ **Основанные на правилах**(Rule-based МТ, RBMT)
- ▣ **Гибридные**(Hybrid МТ, HMT)
- ▣ **Основанные на моделях**(Model-based МТ):

Статистический

- Основан на статистическом анализе корпусов параллельных текстов
- Примеры: Google, Bing, Яндекс

Bing



Испанский ▾ Русский Итальянский

Somos responsables por los que hemos domesticad

48/5000

🔊 ▾



Перевести

Русский ▾ Испанский

Мы ответственны за тех, кого приручили

🔊 ▾ 🔗 🔍 👍 👎

Основанный на правилах

- ▣ Основан на словарной информации и анализе грамматических правил конкретных языков
- ▣ Примеры: PROMT, Linguatex

PROMT

Загрузка файла или архива

Язык оригинала: Испанский

Язык перевода: Русский

Профиль перевода: Business Correspondence (IT)

Формат: Определить формат

Перевести с разметкой: Нет

Создать базу в формате TMX: Нет

Правила перевода XML: Не задано [Загрузить...](#)

Имя файла (архива): [Выбрать файлы](#) Файл не в

[Простой режим](#)

PROMT in PROMT PROMT\EvgeniyaF | Совет дня

Перевод Мои настройки PROMT Агент Администрирование [Справка](#)

Перевод/Перевод документов

[Загрузить документ](#)

Документ	Язык оригинала	Язык перевода	Профиль перевода	Состояние	Добавлен	Действия
DFI-1_2012-08-20.doc	Английский	Русский	Business Correspondence	Завершен	27.05.2013 14:53:31	
Draft Declaration of the High-level Meeting of the 67th Session.pdf	Английский	Русский	Business Correspondence	Завершен	27.05.2013 14:42:41	

Всего документов: 2 (занято 2 Мб из 256 Мб)

© ООО "ПРОМТ", 2003-2013. Все права защищены. [Мобильные приложения](#) | [Написать разработчику](#) | [PROMT.ru](#)

Гибридный

- Основан на совмещении методов SMT и RBMT
- Примеры: Systran, Asia Online

Systran

The screenshot displays the Systran web interface. On the left, a document viewer shows the text of a license agreement for 'SYSTRAN Desktop-Produkte und Anwendungspakete'. A blue callout bubble over the text reads: 'To translate and select'. On the right, the 'Document Metadata' panel shows 'Control Number: DE_19'. Below this, the 'SYSTRAN Queue' table lists translation requests. A blue callout bubble over the table reads: 'View the status of any translation request.' A yellow arrow points from this bubble to a specific row in the table. Another blue callout bubble with a yellow arrow pointing to the 'Message' column of the same row reads: 'SYSTRAN detects inconsistencies between actual vs. selected language'.

Document 1 of 10

Document Identification

- Control Number: DE_19
- Control Number End:
- Control Number Beg Attach:
- Control Number End Attach:

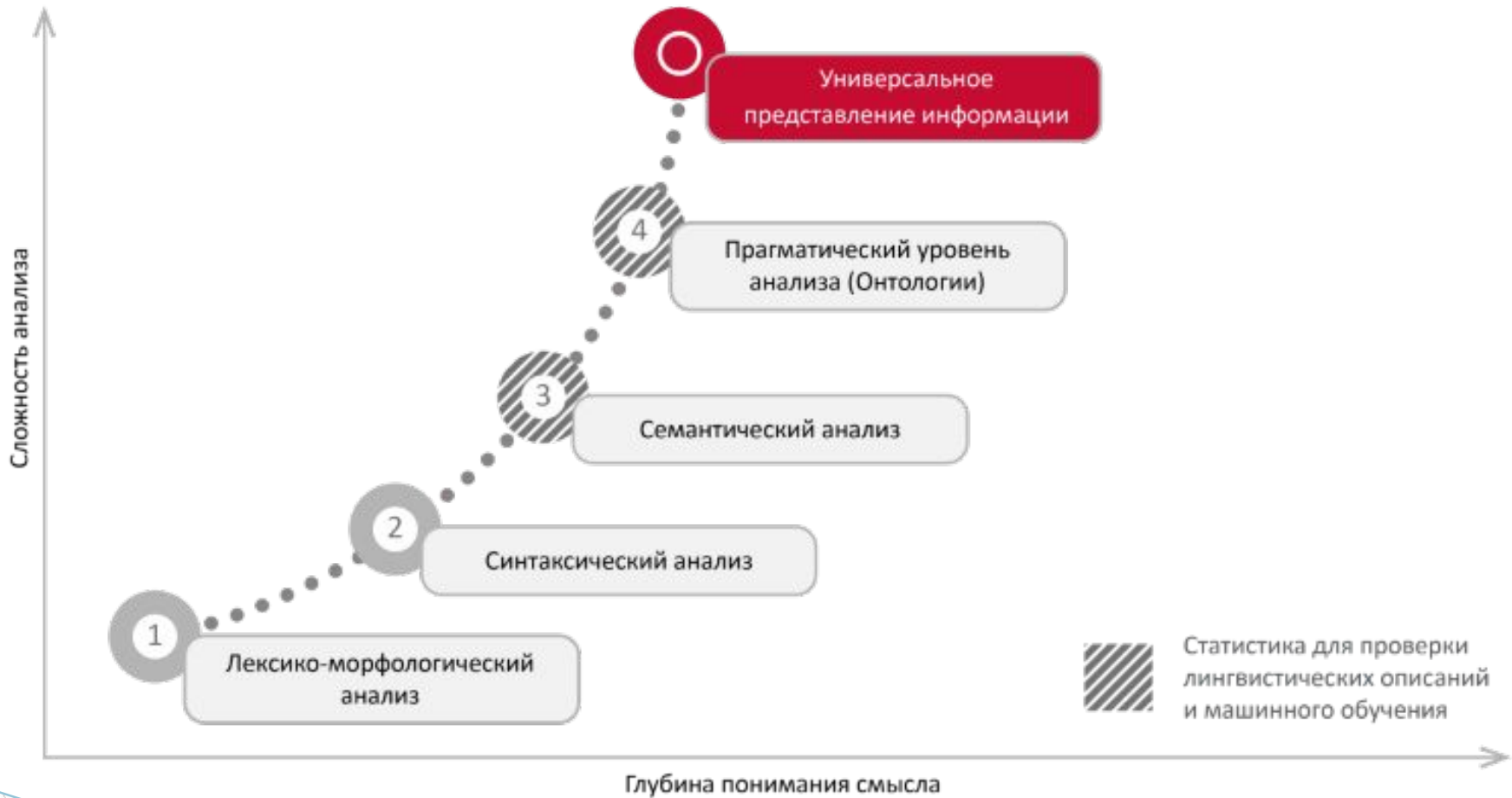
SYSTRAN Queue

Request Time	Completion Time	Control Number	Username	Status	Message
2014-09-24 14:59	2014-09-24 14:59	Multi_22	Relativity Admin	Success	Italian was selected for this document, however it seems most likely to be German.
				Success	French was selected for this document, however it seems most likely to be German.
				Success	Dutch was selected for this document, however it seems most likely to be French.
2014-09-24 14:55	2014-09-24 14:55	Multi_22	Relativity Admin	Success	
2014-09-24 14:42	2014-09-24 14:42	Multi_20	Aaron Spruill	Success	
2014-09-24 14:42	2014-09-24 14:42	Multi_21	Aaron Spruill	Success	
2014-09-24 14:42	2014-09-24 14:42	Multi_22	Aaron Spruill	Success	
2014-09-24 14:30	2014-09-24 14:30	FR_13	Aaron Spruill	Success	

Основанный на моделях

- Основан на универсальной семантической иерархии
- Примеры: ABBYY Compeno

ABBY Compreno



Официальный **курс** доллара установленный Центробанком России на 25 марта упал на 67 копеек до 58,77

курс: курс
RATE_AS_PRICE_OF_CURRENCY
PRICE
PRICE_CHAR
CH_PRICE_AND_SUMS

На **курсе** повышения квалификации слушатели смогут углубить свои знания по эффективному управлению финансами

курс: SET_OF_LESSONS
LESSON_SEMINAR
ARRANGEMENTS
EVENT

Выручка X5 во втором квартале выросла до 155,1 млрд руб. а в первом полугодии выросла до 299,1 млрд руб

Переговоры по иранской ядерной программе, продолжавшиеся в Багдаде последние два дня, завершились ничем

Преимущества

- высокая продуктивность
- низкая стоимость
- единообразиие перевода

Недостатки

- низкое качество



- большое количество ошибок
- невозможность применения при переводе разрозненного текста
- неприменимость в случае высоких требований к стилю
- юридическая ответственность

Список использованной литературы

Спасибо за внимание!