
Статистический машинный перевод

Статистический машинный перевод — это разновидность машинного перевода текста, основанная на сравнении больших объемов языковых пар. Языковые пары — тексты, содержащие предложения на одном языке и соответствующие им предложения на втором, могут быть как вариантами написания двух предложений человеком - носителем двух языков, так и переводом с исходного на язык перевода, выполненный человеком. Таким образом статистический машинный перевод обладает свойством «самообучения». Чем больше в распоряжении имеется языковых пар и чем точнее они соответствуют друг другу, тем лучший результат статистического машинного перевода.

Программа PROMT TerM — специализированное средство для автоматического извлечения и обработки терминологии, программный продукт компании ПРОМТ. PROMT TerM относится к группе программных средств, разработанных для корпоративных пользователей. Данный продукт предназначен для автоматического извлечения и управления терминологией PROMT TerM используется для расширения функциональности систем машинного перевода и для создания автоматизированной цепочки перевода.

Возможности системы:

- Подсчет частотности слов и словосочетаний в анализируемых документах;
- Составление терминологических списков;
- Управление терминологией (возможность установки фильтров и условий обработки; интеграция с системами машинного перевода и системами Translation Memory).

Описание основных этапов работы с PROMT TerM:

- Выбор файлов для анализа.
- Возможности по настройке системы. Так, например, пользователь может отличать слова, набранные в разном регистре, получать список терминологических кандидатов в канонической форме. Пользователь может указать слова, которые должны быть проигнорированы в процессе анализа (например, предлоги, союзы и т.п.). Или, наоборот, пользователь указывает только те слова или словосочетания, которые будут учитываться в процессе анализа. Остальные слова могут быть проигнорированы.
- Функции поиска перевода в словарях системы PROMT как пользовательских, так и специализированных. Может быть осуществлен поиск на предмет соответствия словарным фильтрам, извлечена информация для сбора статистики по словарям, а также получен машинный перевод терминологических кандидатов (ТК).
- Возможность выбора параметров для формирования словосочетаний. Например, выбор первого и последнего слова словосочетания; минимальное и максимальное количество слов, присутствующих в словосочетании, указание количества раз, которое слова должны встретиться в тексте для того, чтобы было сформировано словосочетание.
- Пользователю предоставляется возможность выбора Translation Memory и терминологического словаря для обработки терминологических кандидатов.

- Широкие возможности по предоставлению результирующего списка ТК. Список результатов может быть отсортирован по частоте нахождения ТК в документах, по алфавиту и в обратном порядке.

- Для каждого ТК может быть отображен контекст.

- Кроме того, пользователи системы МП PROMT могут пополнять словари непосредственно из списка ТК.

- Опыт использования программы PROMT TerM в процессе автоматического перевода.

- Краткое описание технологической цепочки перевода с применением PROMT TerM. Использование систем МП и ТМ.

- Взаимодействие PROMT TerM с системами МП и ТМ.

- Использование данных технологических решений для перевода документации компании SAP.

Выводы об эффективности применения интегрированного решения:

- Увеличение скорости всего цикла перевода и редактирования (до 50 %);

- Повышение эффективности процесса извлечения и управления терминологией (до 80%);

- Обеспечение единообразия терминологии и стиля при коллективной работе;

- Уменьшение объема печатной и ручной работы;

- Снижение общей стоимости переводческого проекта.

Сократ Персональный - система автоматического перевода с английского языка на русский и наоборот, предназначенная для работы со сравнительно небольшими текстами общей тематики.

Низкие системные требования, интуитивно понятные способы использования, удобный интерфейс и гибкие настройки - вот лишь небольшой перечень его достоинств. Различные варианты ввода текста. Сократ Персональный позволяет пользователям переводить текст набранный в своем собственном окне, находящийся в буфере обмена или файле. Перевод почтовых сообщений. Программа имеет plug-in для MS Outlook, позволяющий переводить сообщения непосредственно в этом почтовом клиенте. Перевод стандартных Help'ов других программ. Если иконка "Сократа Персонального" уже находится на панели задач Windows, то открыв справочную систему любой программы, Вы увидите на верхней панели кнопку или пункт меню "Перевести". Щелкнув по кнопке Вы получите перевод содержимого текущей страницы справочной системы. Доступ к сервису перевода из других приложений. Для того, чтобы перевести текст или слово из почти любого приложения MS Windows, достаточно выделить его и, используя сочетание 'горячих' клавиш, вызвать сервис перевода. Переведенный текст появится во всплывающем окне. Вызов интерактивного словаря. При желании выяснить значение какого-либо слова Вы можете просто щёлкнуть по иконке "Сократа

Персонального" правой кнопкой мыши. Набрав в окне ввода слово, Вы автоматически получаете его перевод.

Англо-русский словарь TransLite - удобный и быстрый словарь, который позволит Вам одним щелчком перевести слово или фразу в любой программе. Англо-русский словарь в демоверсии содержит более 94 000 терминов, Немецко-русский словарь - порядка 69 000 терминов. Вы получите перевод прямо под курсором. TransLite переведет фразу, даже если слова в ней написаны в различных склонениях, множественном числе и т.п. Мгновенный перевод под курсором. Перевод мгновенно появляется под курсором в виде всплывающей подсказки.

Режим Smart-Click. Данный режим позволяет при щелчке по слову, если это слово является началом фразы, имеющейся в словарной базе, предоставить перевод всей фразы.

Продуманный алгоритм перевода. Вам не надо заботиться о том, чтобы выделять отдельные слова. Вы получаете перевод всей фразы. TransLite постарается найти перевод слов во множественном числе, различных склонениях и падежах.

Автоматическое переключение раскладки (транслитерация). При ручном вводе текста вам не надо задумываться в какой раскладке вы это делаете. TransLite автоматически переключит раскладку и исправит краказябры.

Зарегистрировав программу, Вам станут доступны в полном объеме все словарные базы TransLite, которые на данный момент содержат более 3.500.000 терминов (более 200 тематик).