

Русско-китайский параллельный корпус НКРЯ



Что такое параллельный корпус?

- Корпус:
 - Собрание текстов
 - Каждое слово в котором размечено (т.е. есть грамм. и другая информация)
 - По каждому слову можно вести поиск
- Параллельный:
 - Тексты собраны на двух языках и выровнены по предложениям

Cui bono?

- Изучающим иностранный язык (и преподавателям)
- Переводчикам
- Компьютерным лингвистам и программистам

Наш корпус

- Где:
 - https://linghub.ru/rnc_parallel_chinese/search
 - <http://www.ruscorpora.ru/new/search-para-zh.html> (хромает)
- Кто:
 - Институт русского языка РАН
 - ВШЭ (ФикЛ, ФКН и ДИЯ)
 - МГУ
 - РГГУ (Институт Конфуция)
 - МФТИ/Huawei
 - СПбГУ
 - Чжэцзянский университет
 - Хэйлунцзянский университет

Как это выглядит сейчас?

(INSERT CORPUS TITLE IN messages.po) RU | EN |  

Query


Word #1

Word:

Lemma:

Grammar:

Language/tier:




Full-text search:


☐ Precise match


Search sentences

Search words / lemmata



Select subcorpus





Search result: 2 occurrences, 2 sentence(s) found in approximately 2 document(s).

狂人日记 / Kuangren riji 鲁迅 / Lu Xun 1918

☒ 

Чтобы **справиться**, раскрыл книгу по истории, в книге не было дат, зато каждая страница изобилвала словами «гуманность», «справедливость», «мораль» и «добродетель».

我翻开历史一查, 这历史没有年代, 歪歪斜斜的每叶上都写着“仁义道德”几个字。

祝福 / Zhufu 鲁迅 / Lu Xun 1924

☒ 

На этот раз в дом дядюшки наняли поденщика, но он не сумел со всем **справиться**, и в помощь ему позвали старуху Лю. Однако резать кур и гусей она отказалась, потому что верила в Будду, постилась и не могла убивать живые существа. Сян-линь поддерживала огонь в печи и от нечего делать смотрела, как старуха моет жертвенные сосуды. Это было единственное, что та согласилась делать.

四叔家里这回须雇男短工, 还是忙不过来, 另叫柳妈做帮手, 杀鸡, 宰鹅; 然而柳妈是善女人, 吃素, 不杀生的, 只肯洗器皿。祥林嫂除烧火之外, 没有别的事, 却闲着, 坐着只看柳妈洗器皿。

- Языки:
 - Русский
 - Китайский (иероглифы)
 - Китайский (пиньинь)
- Поиск по:
 - Словоформе
 - Начальной форме (lemma)
 - Грамм.показателю
- Можно выбрать подкорпус

Размер корпуса

- 177 000 русских слов
 - Скоро будет + 450 000 слов
- Произведения XIX – нач. XX вв.
- 7 русских произведений, 3 китайских

Наша цель:

Reverso Context ПЕРЕВОД | СЛОВАРЬ | КОРРЕКТОР | СПРЯЖЕНИЕ | СИНОНИМЫ | ГРАММАТИКА

влиять на Русский Английский

Спрягать этот глагол

Перевод "влиять на" на английский

Глагол

influence affect influence on say in interfere with be influenced by sway have an impact on impact the have an effect on impacts on have a bearing on ...

Внезапно у меня появилась способность **влиять на** маленькие объекты. I was suddenly able to **influence** objects that were smaller and smaller.

Прекрасный способ приобрести друзей и **влиять на** людей. It's a good way to win friends and **influence** people.

Слушайте, не позволяйте другим **влиять на** ваш день. Listen, don't let the behavior of others **affect** your day today.

- Большой
- Разнообразный
- Удобный
- Популярный

Наши задачи (для китаистов)

- Выравнивать тексты (готовить их к формату корпуса)
- Искать новые ПАРАЛЛЕЛЬНЫЕ тексты
- Проверять качество алгоритмов:
 - Словоделения
 - Пиньиня
 - Выравнивания
- Поддерживать корреспонденцию с китайскими коллегами
- (в будущем) заниматься СММ

Наши задачи (для программистов)

- администрирование сайта на сервере
- фронт-энд: создать две новые версии сайта:
 - более простой интерфейс на русском языке для пользователей-нелингвистов + прикрутить нейронки
 - интерфейс на китайском языке
- адаптация других корпусов под наш формат
- встроить автоматическую систему присвоения транскрипции иероглифам с хорошей точностью
- делать любые другие исследовательские и технические работы с нашим корпусом. Инициатива приветствуется!

Формы кооптации

- Практика
- Волонтерство
- Ярмарка проектов
- НУГ (научно-учебная группа)
- Оплачиваемая «практика»
- Кооптация со стороны (классический аутсорс)

Спасибо за внимание!

Кирилл Семенов, ФикЛ:

kir.semenow@yandex.ru

Vk: cyrilsemenov

Tg: @cyr_sem