



Следствие ведут цифровые гуманитарии: что на самом деле думали в средневековом Китае про котиков и других животных?



University of
Zurich^{UZH}

Asien-Orient-Institut

Кто я такая



University of
Zurich^{UZH}



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ





Чем я занимаюсь:

- Династия Тан (618–907)
- «Вещи» в классической китайской поэзии *ши*
- Поэзия *ши* «о вещах»



Что такое «классическая» поэзия

- Поэтическая форма, один из основных видов классической поэзии Китая
- «Золотой век» был в 7–10 вв.
- Для регулярной поэзии характерны строгие правила, регламентирующие:
 - Длину строки, количество слов в стихотворении
 - Фонетическую составляющую (последовательность тонов)
 - Типы рифм
 - Тематика строк
 - Параллелизм

平平仄仄仄平平，
仄仄平平仄仄平。
仄仄平平平仄仄，
平平仄仄仄平平。

○○●●●○○，
●●○○●●○。
●●○○○●●，
○○●●●○○。



Поэзия о вещах

- Главный объект описания – какая-либо вещь
- Понимание «вещей» отличается от нашего. В список могли включить:
 - животные
 - горы, реки
 - природные явления
 - людей (женщин чаще, чем мужчин)
- Содержание может:
 - быть детальным описанием конкретной вещи
 - метафорически передавать переживания автора



Корпус

- «Полное собрание стихов династии Тан» (Цюань Танши 全唐詩)
- Начало создания – 1705 год
- Поэзия между 618–907 гг.
- Содержит почти 50000 стихотворений
- Более 3,5 млн иероглифов
- Огромный объём данных, учитывая период
- Собрание «поэм о вещах» (Юньуши сюань 詠物詩選)
- Сборник выпущен в 1706
- Содержит около 14 000 стихотворений, династия Тан – 5751
- Стихотворения поделены по темам



**University of
Zurich** ^{UZH}

Asien-Orient-Institut

Как выживать в этом жестоком мире

У гуманитария есть несколько путей.

Использовать консервативные методы; пользоваться только бумажными книгами.

Искать информацию в оцифрованных текстах; не признавать никаких “точных методов” кроме подсчёта частот.

Использовать точные методы для того, чтобы исследовать корпус или подтверждать гипотезы.

Обучить нейронную сеть и надеяться на лучшее.





Как можно найти информацию?

- В самих поэтических текстах
- В энциклопедиях
 - Практические (об управлении, сельском хозяйстве)
 - Учебные (с самыми важными цитатами из классиков)
 - Рифм – важные сочетания и удачные примеры использования слов
- Литературная критика
- Дополнительно можно попробовать использовать компьютерные методы, чтобы найти то, что:
 - Забыли критики
 - Не хотели рассказать составители энциклопедий



**University of
Zurich** ^{UZH}

Asien-Orient-Institut

ЖИВОТНЫЕ



Выборка

- Список животных – из энциклопедии начала Тан, созданной для начального образования (Начало учения, Chixueji, 初學記) + тигр и кот
- кот, тигр, единорог, собака, лошадь, лиса, мышь, корова, заяц, слон, осёл, баран, свинья, олень, верблюд, лев, обезьяна

КОШКИ-МЫШКИ

Коты.

- Котов никто не любит и про них мало пишут
- Их едят и используют для изготовления одежды
- Приход кота – к бедности
- Дикие коты хорошо охотятся, а домашние – ленивы

Мыши и крысы.

- Съедают запасы
- При этом могут приносить счастье, если прогрызут верхнюю одежду
- В отличие от котов, возвращают долги



Тигры

- Одно из самых опасных животных
- Сильные, яростные, благородные
- «Царь животных»
- Пример подражания для военных
- Отгоняют злых духов



Единорог

- Один или несколько рогов, чешуя, тело коня, ноги оленя, хвост быка
- Появляется во времена праведных правителей
- Гуманен по отношению ко всем животным, кроток
- Образец добродетели и благородности





Тигр и устоявшиеся выражения

Коллокаты:

- 猛虎 Яростный тигр
- 龍虎 Дракон и тигр
- 豺虎 Шакал и тигр
- 白虎 Белый тигр
- 貔虎 Барс и тигр



Word2vec и вычисления

- Единорог:
 - Синонимы: феникс, украшенный, возвышающийся, парадная одежда, совершенствовать, критиковать, Нюйва (богиня)
 - Что такое единорог – «гуманность»? Волосы, шпилька, расчёска – т.е. орнамент
- Тигр:
 - Синонимы: леопард, олень, спячка, пещера, змея/дракон, хищник, свирепый и т.п.
 - Чем занимается тигр (тигр + движение)? Рычит, спит, (оказывается) связанным или застреленным, совершает преступления
- Мышь:
 - Синонимы: лузга, кот, лиса, кожа, расхищать, гомонить и т.п.
 - Если мышь маленькая, какой слон? Большой (на самом деле «большая балка»)



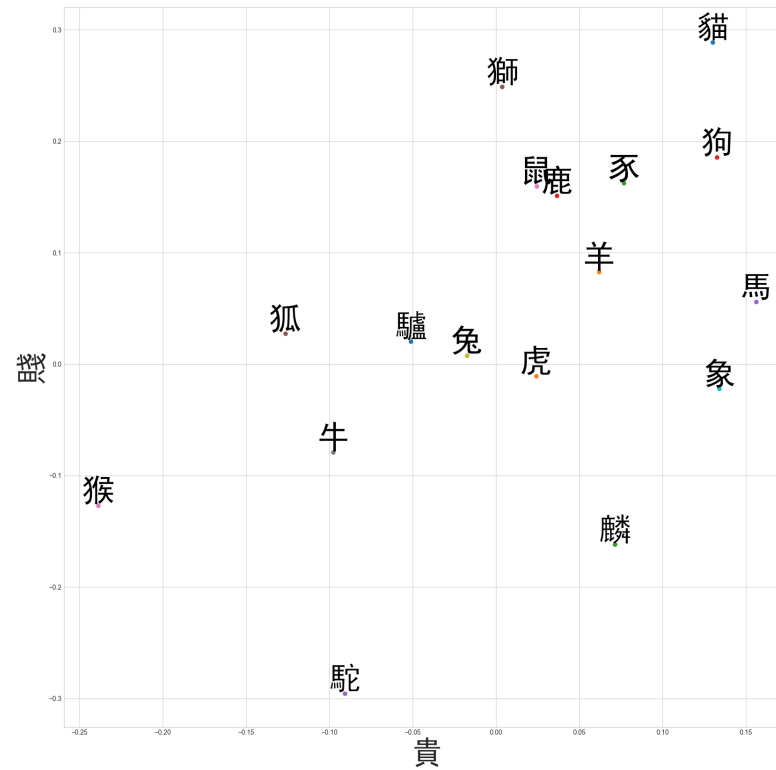
**University of
Zurich** ^{UZH}

Asien-Orient-Institut

Word2vec и неожиданное

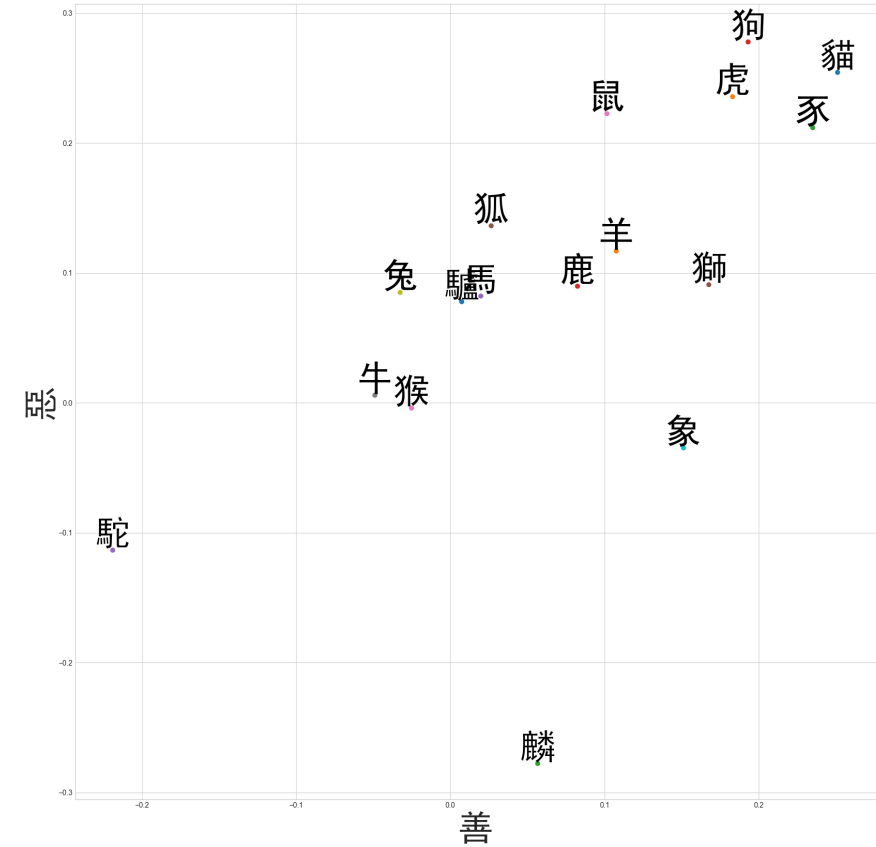
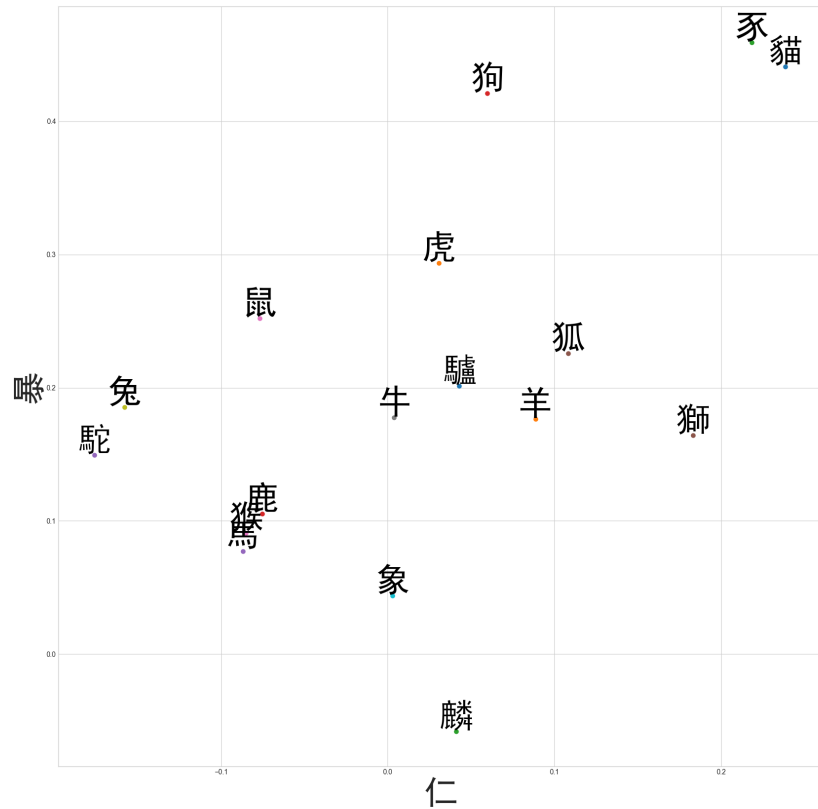


СКОЛЬКО СТОЯТ ЖИВОТНЫЕ?



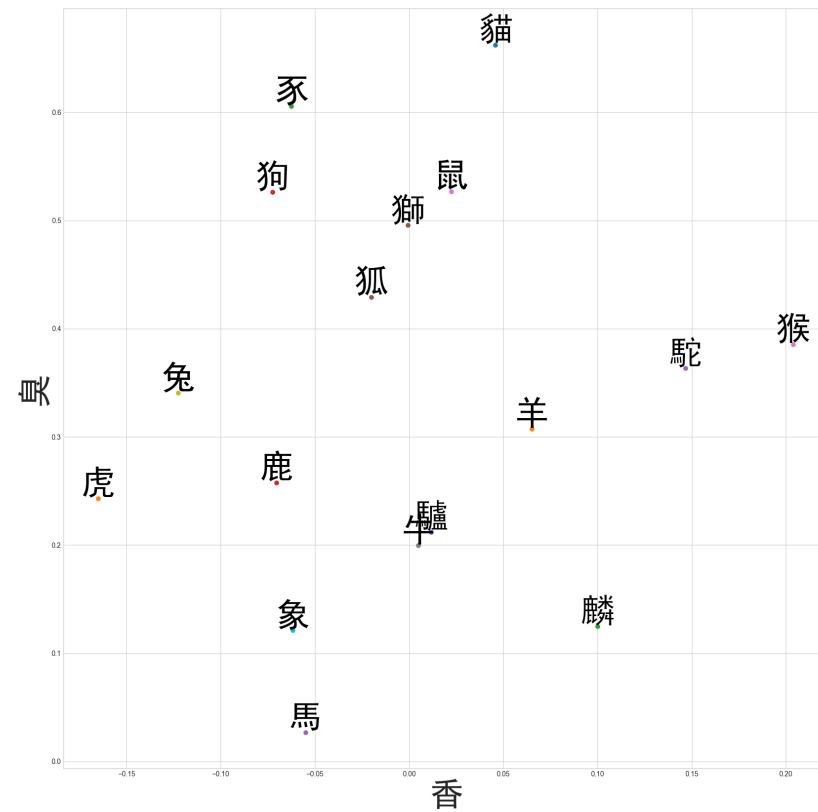


Воплощают ли животные добродетели?



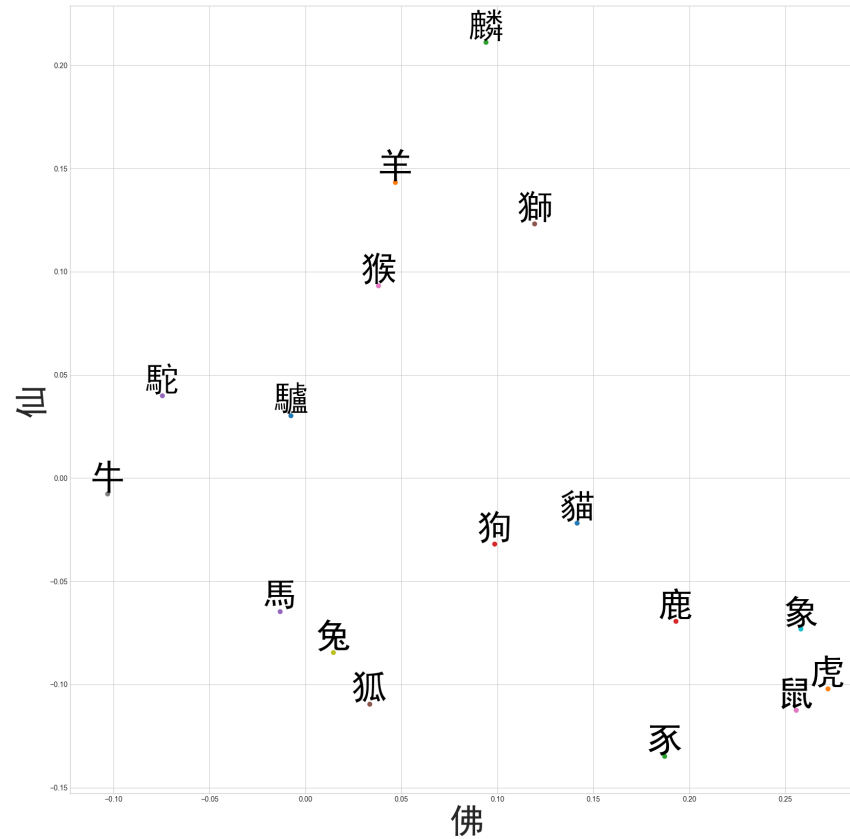


А кто хуже всего пахнет?





Животные и религия?



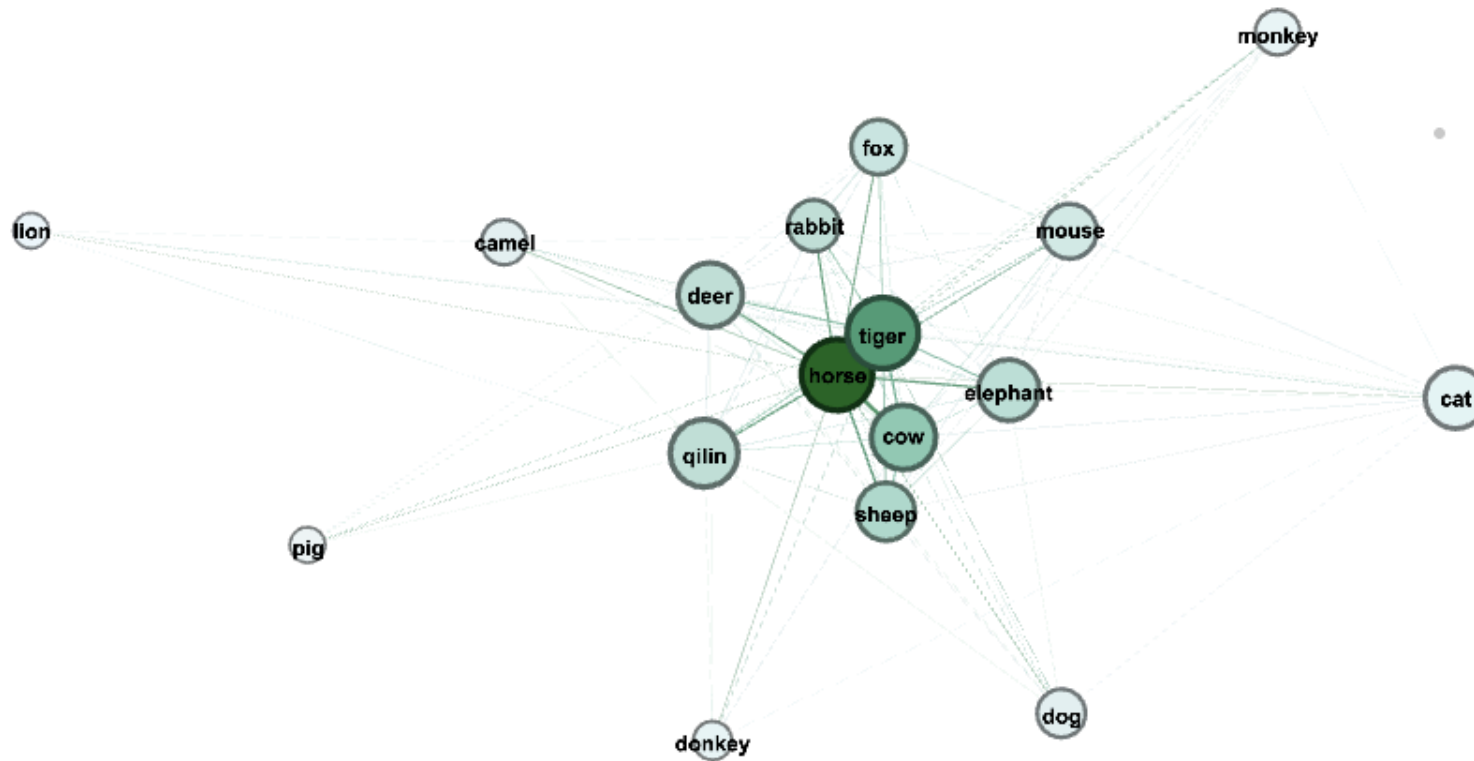


**University of
Zurich** ^{UZH}

Asien-Orient-Institut

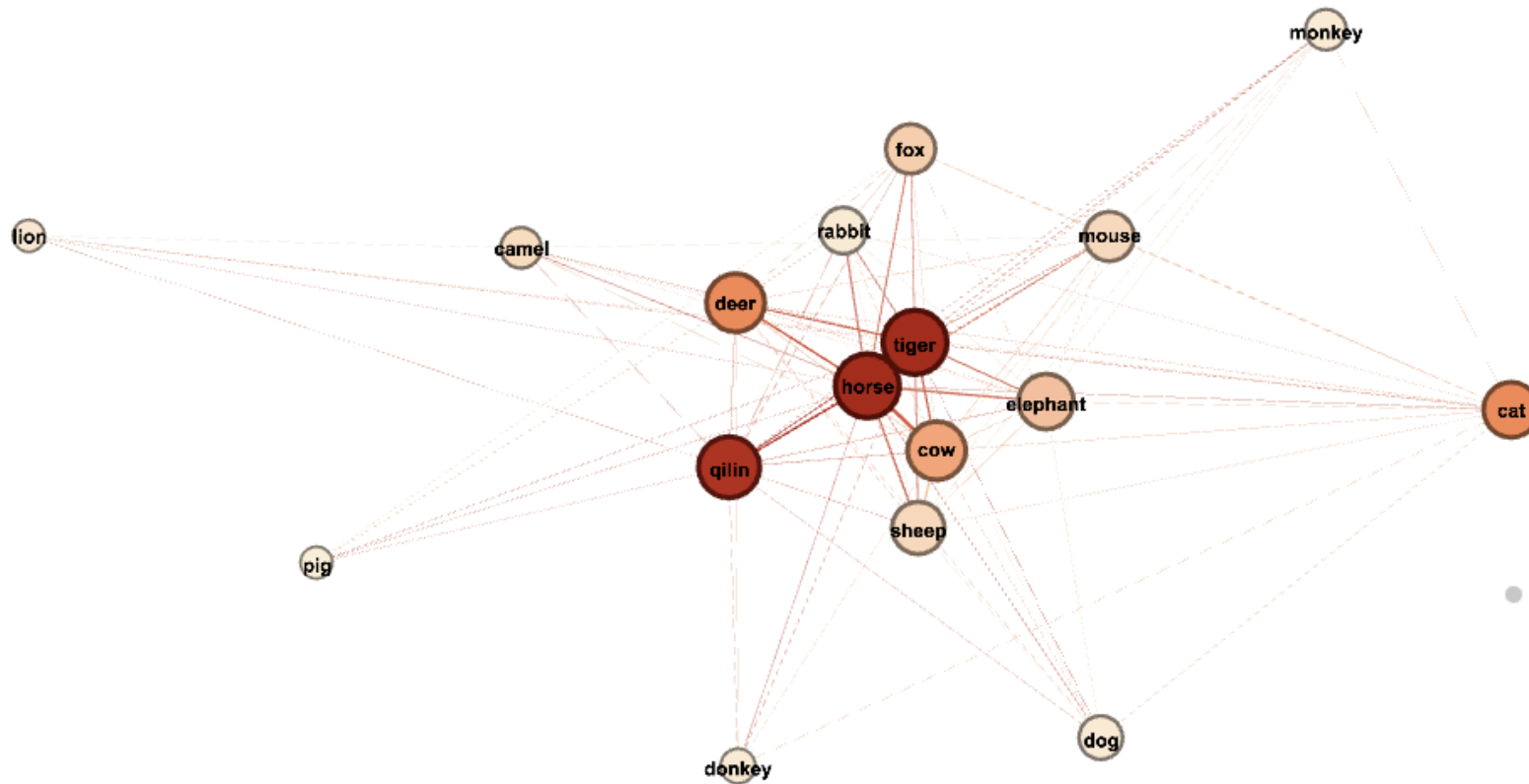
Сетевой анализ

Какое животное самое «социальное»?





Кто на свете всех центральной?





**University of
Zurich** ^{UZH}

Asien-Orient-Institut

Тематическое моделирование



Оно вообще работает?

- «Собрания поэм о вещах»
- Насколько осмысленной была классификация стихотворений по группам вещей?
- Есть ли какая-то принципиальная разница между стихотворениями о разных животных?
- Что будет, если искать гораздо меньше тем, чем у нас животных?



Весь сборник и разные параметры

1. Сельское хозяйство
2. Реки и озёра
3. Еда
4. Весна
5. Осень
6. Военная тематика
7. Бессметрные и даосизм
8. Бюрократия
9. Музыка
10. + разные типы животных – ПТИЦЫ и т. п.



Только 17 животных

1. Верблюды и международные отношения
2. Гиббоны и грусть по ночам
3. Собаки и гости (?)
4. Лошади, драконы и (военные) походы
5. Тигры и война
6. Олени, бараны, единороги и бессмертные
7. Зайцы, лисы и придворная жизнь
8. Кошки, мыши, свиньи и домашнее хозяйство
9. Путешествия и смысл жизни
10. Коровы и сельское хозяйство