

Научный семинар Школы лингвистики НИУ ВШЭ

с участием лингвистической лаборатории по корпусным
технологиям и лаборатории языков Кавказа

Даниил Скоринкин

преподаватель и аспирант Школы лингвистики НИУ ВШЭ

Лев Толстой и сетевой анализ

Методы теории сетей (network theory/network science/network analysis) применяются в диапазоне от политических наук до эпидемиологии — везде, где уместно моделировать объект исследования в виде графа. Не осталось в стороне и литературоведение: и художественных персонажей, и прочие именованные сущности в текстах, да и сами тексты можно представить в виде узлов и связей между ними. Серьезных преимуществ тут два:

1. Из всех формальных абстракций сеть (граф) — одна из самых наглядных и интуитивно понятных. Небольшие сети легко визуализировать и анализировать невооруженным глазом, в то же время существует множество математических метрик для оценки структурных особенностей сети в целом и «значимости» отдельных узлов, разработано немало алгоритмов выделения сообществ в сетях.
2. При наличии подходящих инструментов NLP или достаточно «умной» разметки текстов можно производить количественные исследования (в духе [distant reading](#)) на больших корпусах художественных произведений. Здесь открываются возможности для исследования типологии произведений, анализа [эволюции](#) определенного типа/жанра текстов и т.д.

Я расскажу о применении сетевого анализа для исследования текстов Л.Н. Толстого. Первый и основной сюжет моего рассказа — извлечение и анализ сетей в романе

Старая Басманная 21/4,
аудитория 518
28 апреля (пятница) в 10:30

Сайт семинара:
ling.hse.ru/SciSeminar

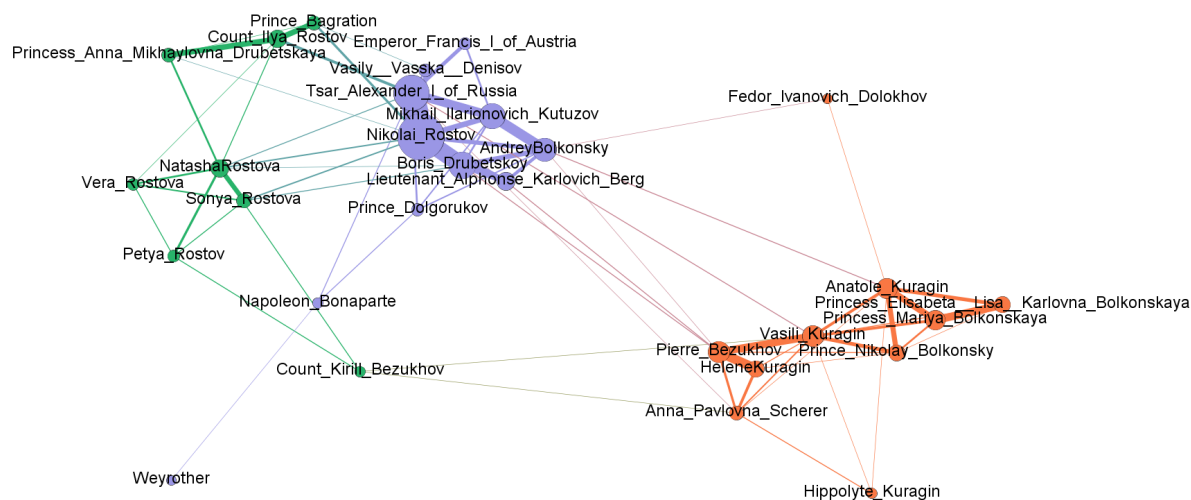




ШКОЛА
ЛИНГВИСТИКИ
НИУ ВШЭ

«Война и мир». Я покажу, как некоторые особенности сетей связаны с антитезой «мирного» и «военного» в романе, как с помощью метрик центральности можно отследить переключения «точки зрения» между героями, какие группы персонажей получается выделить стандартными методами кластеризации графов и на какие литературные параллели всё это наталкивает.

Второй (побочный и маленький) сюжет связан с недавно созданным под руководством Бориса Орехова прекрасным [цифровым указателем](#) к 90-томному собранию сочинений Л.Н. Толстого. Среди прочего этот ресурс позволяет работать с указателем имен собственных к 90-томнику как с графом, связь в котором устанавливается на основе совместного упоминания. Я попытался использовать возможности цифрового указателя для анализа художественной части наследия Толстого.



Старая Басманная 21/4,
аудитория 518
28 апреля (пятница) в 10:30

Сайт семинара:
ling.hse.ru/SciSeminar

