

Поэзия, написанная детьми: дистантное чтение

Третья Московско-Тартуская школа Digital Humanities

Ход работы

2 группы:

- 1) Пристальное чтение, выделение признаков
- 2) Дистантное чтение, формализация и анализ признаков

Материал

3 текстовые коллекции
современной англоязычной
поэзии ~ 62 - 70 тыс. слов:

- 1) Детская поэзия
- 2) Взрослая непрофессиональная поэзия
- 3) Взрослая профессиональная поэзия

Кто автор?

Wonder

When the snow turns into water
Spring will emerge from its slumber.
And the birds will sing,
and the sun will rise.
The world is full of wonder.

old

each tear to my skin
each fade to my face
each crack to my bones
my story is long over
yours can begin

The Red Wheelbarrow

so much depends
upon

a red wheel
barrow

glazed with rain
water

beside the white
chickens.

Кто автор?

Wonder

When the snow turns into water
Spring will emerge from its slumber.
And the birds will sing,
and the sun will rise.
The world is full of wonder.

Автор - ребенок

old

each tear to my skin
each fade to my face
each crack to my bones
my story is long over
yours can begin

Автор -
непрофессиональный
поэт

The Red Wheelbarrow

so much depends
upon

a red wheel
barrow

glazed with rain
water

beside the white
chickens.

Автор - модернист,
William Carlos
Williams

"ЗАНИМАТЬСЯ ТВОРЧЕСТВОМ С ДЕТЬМИ-НЕВЕРОЯТНО ЗДОРОВО!"



ОЖИДАНИЯ



РЕАЛЬНОСТЬ

Детская поэзия, ожидания и реальность :

Спонтанность самовыражения

“Свежесть”

“Поэтический” взгляд на мир

Изображение счастливого
детства

Классы creative writing

Подстройка под аудиторию

Отбор стихотворений для
публикации

Исследователи: учителя (bias)

Особый пул поэзии “для” детей

Creative writing classes

In essence the Method consists of giving youngsters a topic - the city, the sea, the weather, night and the moon, morning, people, animals - then asking them to make a list of words suggested by the topic, and then telling them to write lines using their most descriptive words

- Описательность
- “Личный” подход
- Форма (хайку)
- Использование WOW-слов
- Задания, связанные со школьной повседневностью
- Ориентировка на аудиторию

Гипотезы

Детская поэзия:

- Набор тем
- *Влияние creative writing classes*
- Ограниченный лексикон(?)
- Специфический лексикон (dragon, unicorn)
- *Назидательность/Дидактичность*
- Эмоциональный заряд (happy poems)
- Специфическая форма (параллелизм, противопоставление, повторы)
- Использование прилагательных

Гипотезы

“Любительская” поэзия

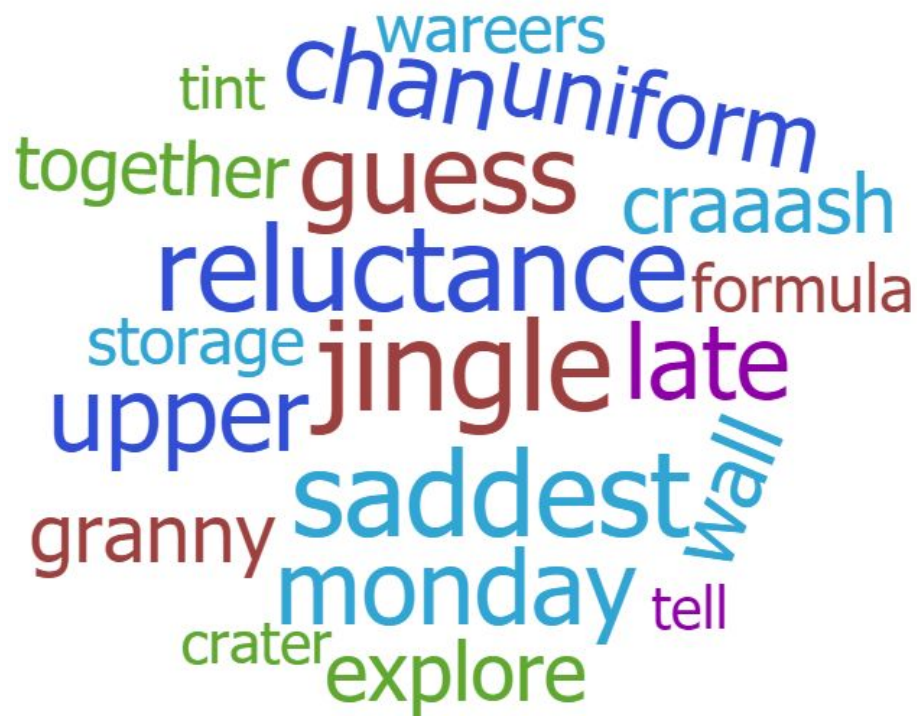
- Абстрактная/“психологическая” лексика
- Клишированные сочетания слов и коллокации (life - death etc)
- Простые рифмы
- Назидательность
- Исследование темы amateur
- Актуальные времени темы (отклик на события)
- Депрессивная тональность
- Сюжетность

“Профессиональная” поэзия

- “Редкие” слова
- БОльшая длина строк при сохранении рифмы
- Выдержанная форма
- Проявление поэтической “моды”
- Словотворчество

Уникальные слова

Детская поэзия



Уникальные слова

Наивная поэзия

emptiness medical
revelation wall
imminently
variety cun regale
retain
wart unanswered
guess profound late
willingly quill idiocy
uniform wrinkle
decipher

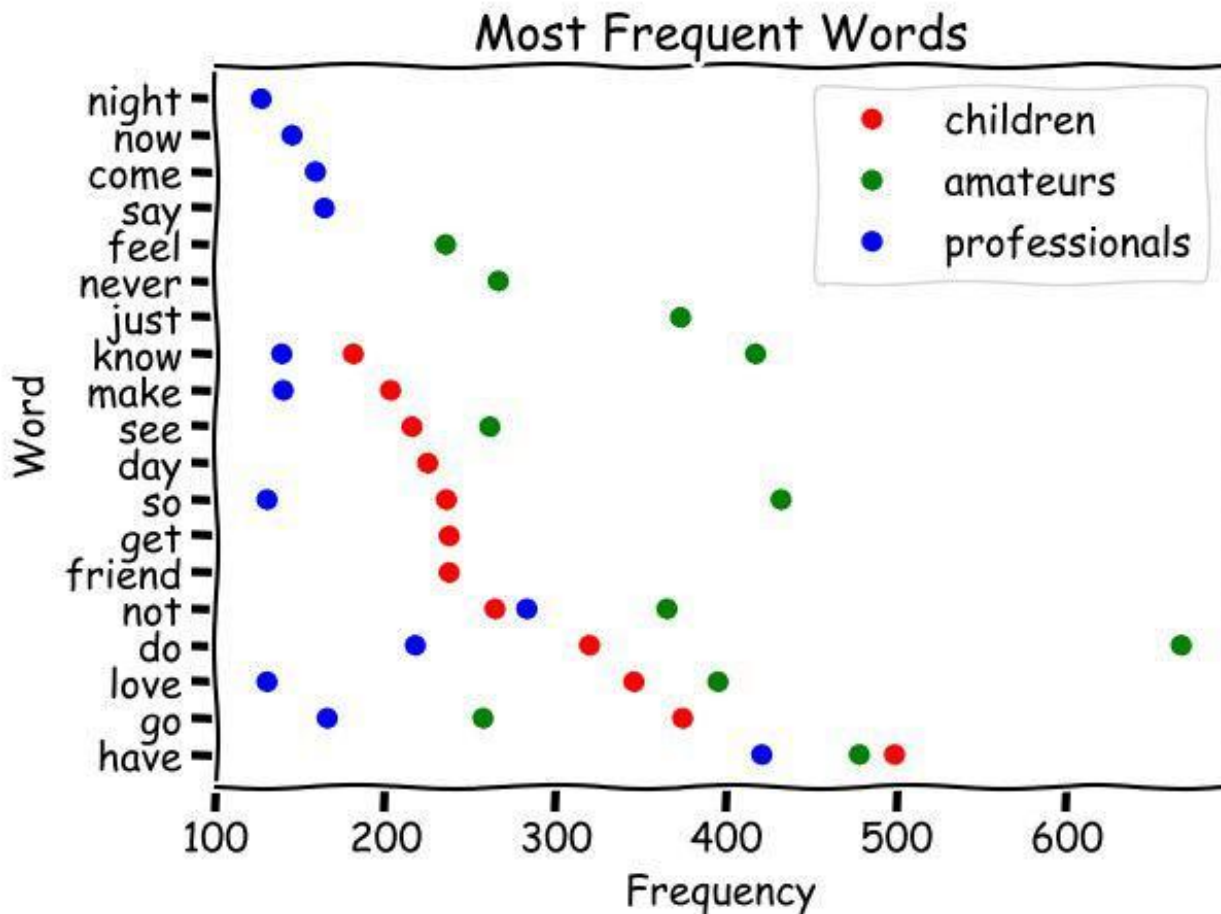
Уникальные слова

Профессиональная
поэзия

A word cloud of unique words from professional poetry. The words are arranged in a roughly triangular shape, pointing upwards. The words are in various colors (blue, green, purple, red, brown) and sizes. The words include: mounting, moon, dust, monday, crane, wing, marenma, blizzard, quaintly, you, oust, flautist, sky-floor, infamy, wherefore, emptiness, coronado, plainly, february, incetown, and concede.

mounting
moon
dust
monday
crane
wing
marenma
blizzard
quaintly
you
oust
flautist
sky-floor
infamy
wherefore
emptiness
coronado
plainly
february
incetown
concede

Самые частотные слова



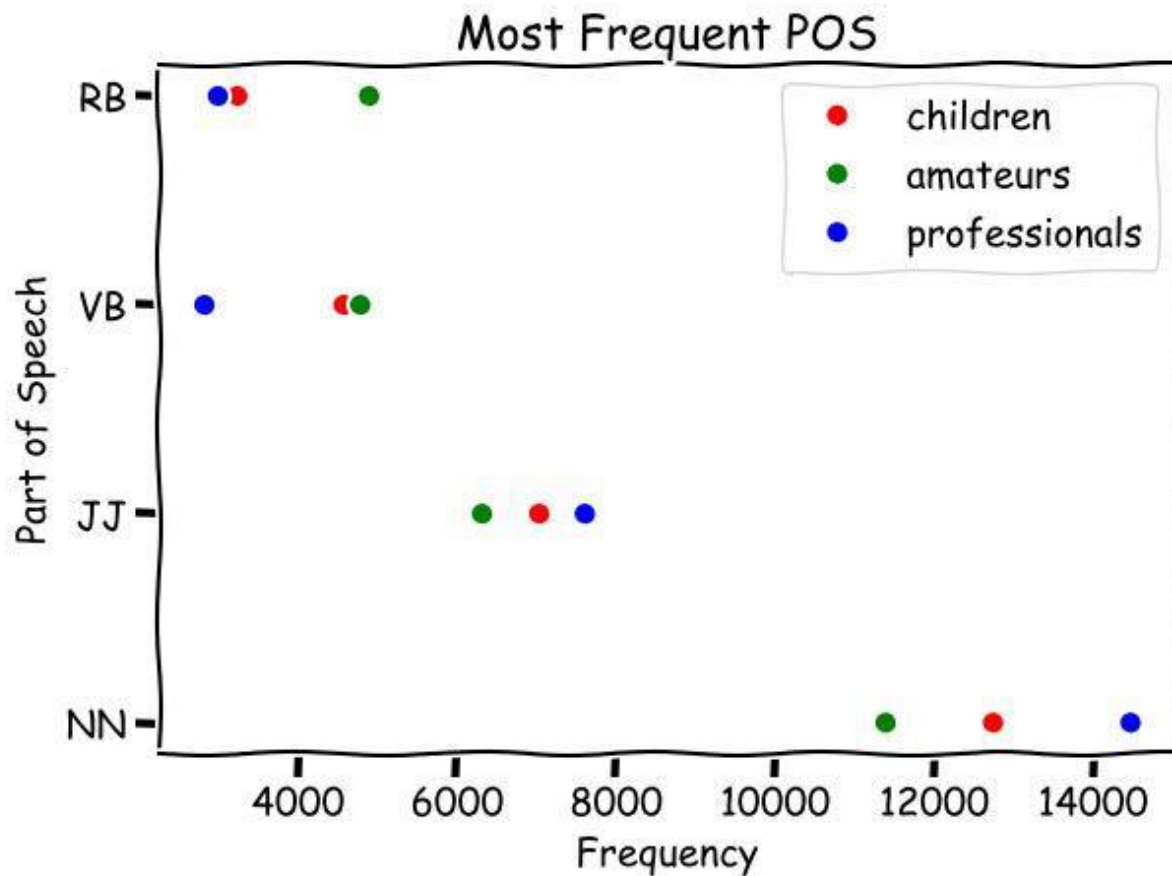
Распределение частей речи

JJ: adjective or numeral,
ordinal

VB: verb, base form

NN: noun, common,
singular or mass

RB: adverb



Распределение частей речи

TO: "to" as preposition or
infinitive marker

UH: interjection

FW: foreign word

CC: conjunction, coordinating

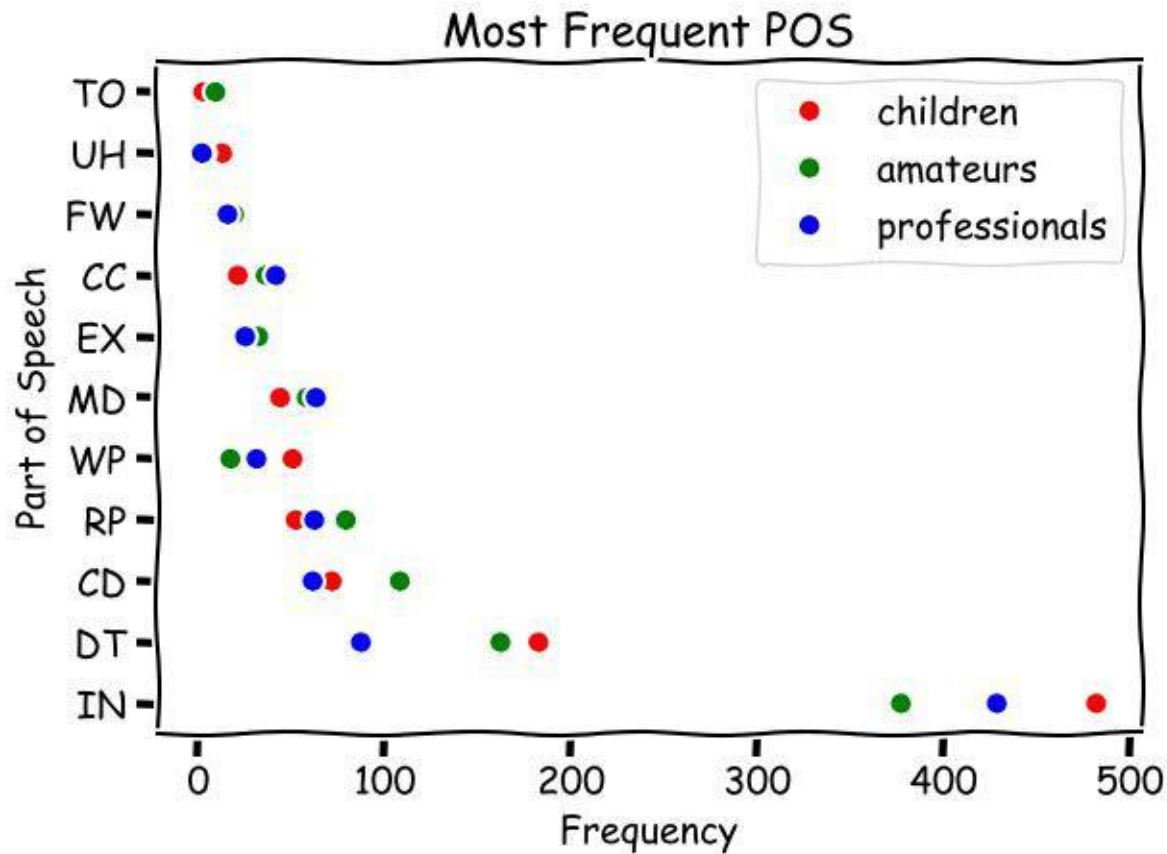
EX: existential there

MD: modal auxiliary

RP: particle

DT: determiner

IN: preposition or conjunction,
subordinating



Распределение частей речи

VBD: verb, past tense

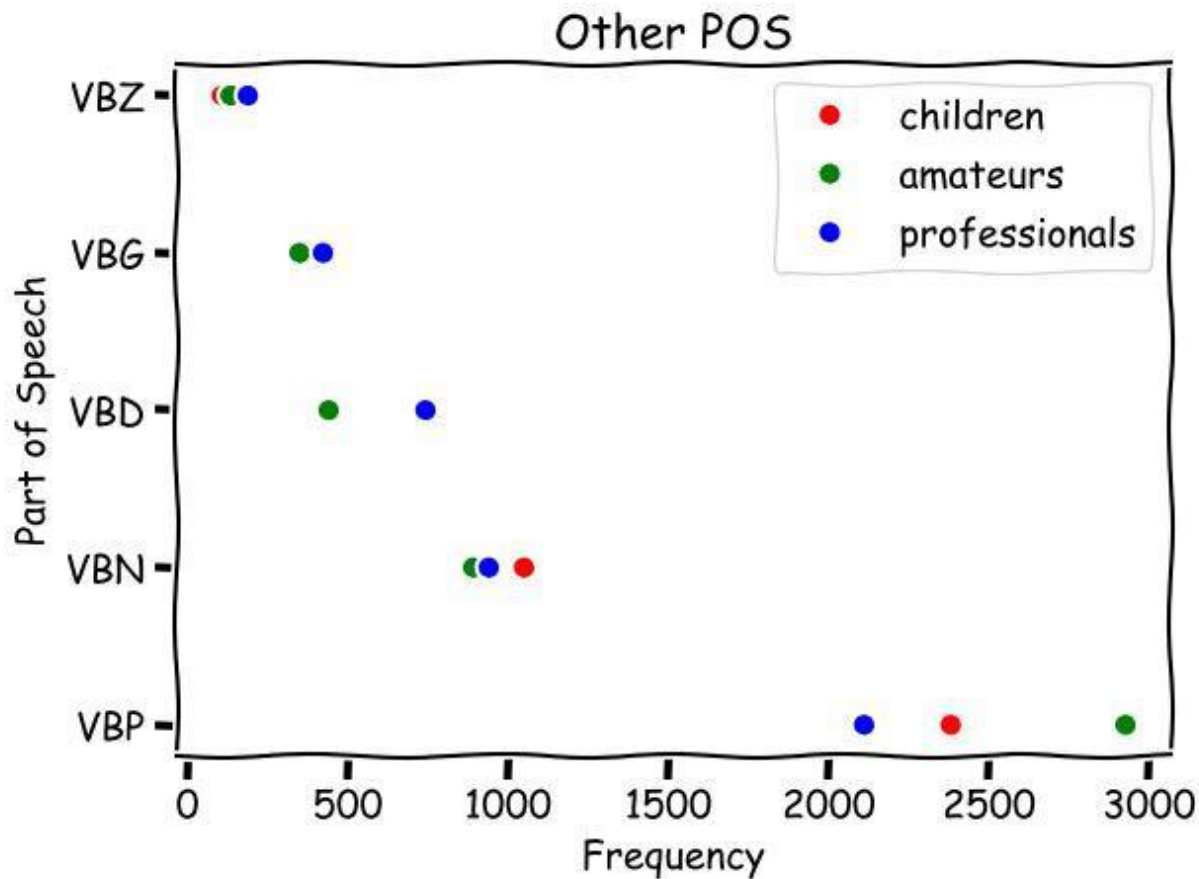
VBG: verb, present

participle or gerund

VCN: verb, past participle

VBP: verb, present tense,
not 3rd person singular

VBZ: verb, present tense,
3rd person singular



Распределени частей речи

WP\$: possessive wh-pronoun
(whose)

PDT: pre-determiner

WRB: Wh-adverb

RBS: adverb, superlative

WDT: WH-determiner

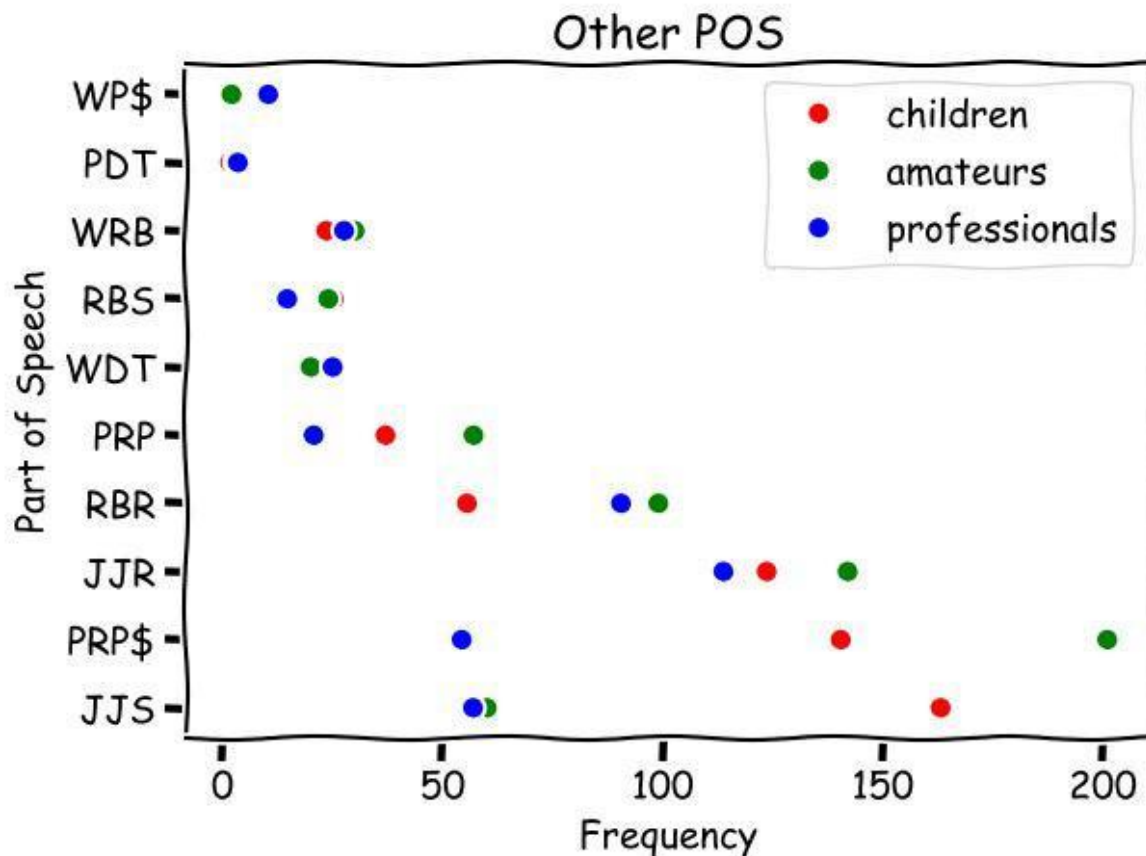
JJR: adjective, comparative

PRP: pronoun, personal

RBR: adverb, comparative

PRP\$: pronoun, possessive

JJS: adjective, superlative



Tf-idf

Детская поэзия



Tf-idf

Наивная поэзия



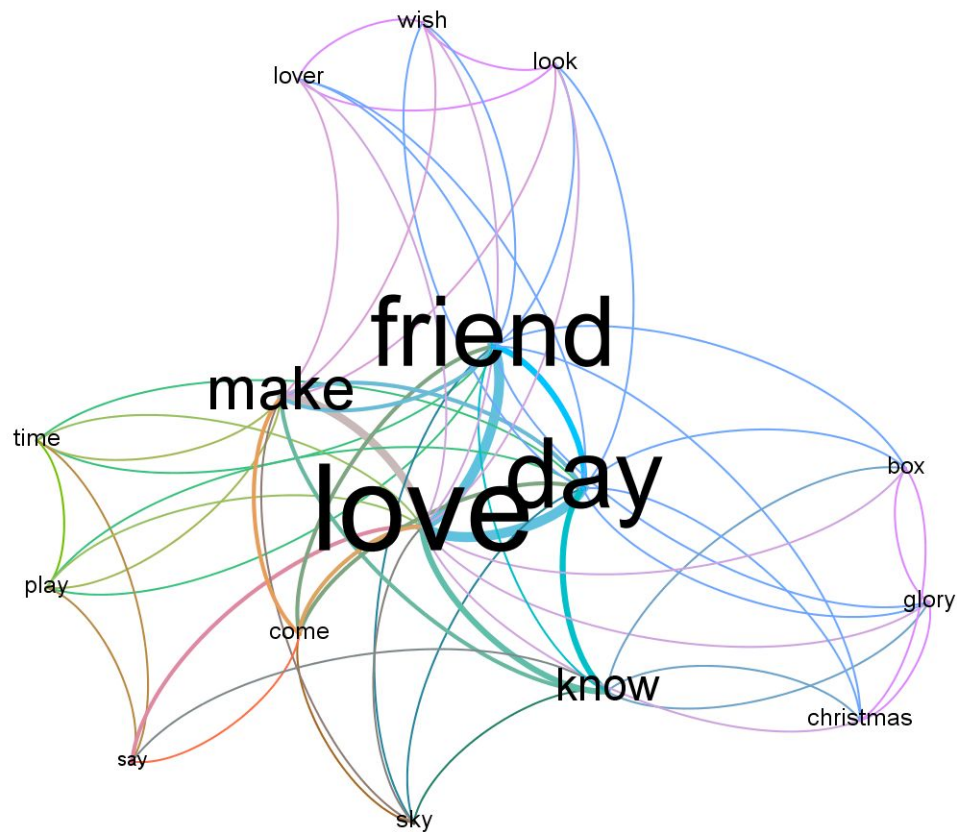
Tf-idf

Профессиональная
поэзия



Topic Modeling

LDA
Children



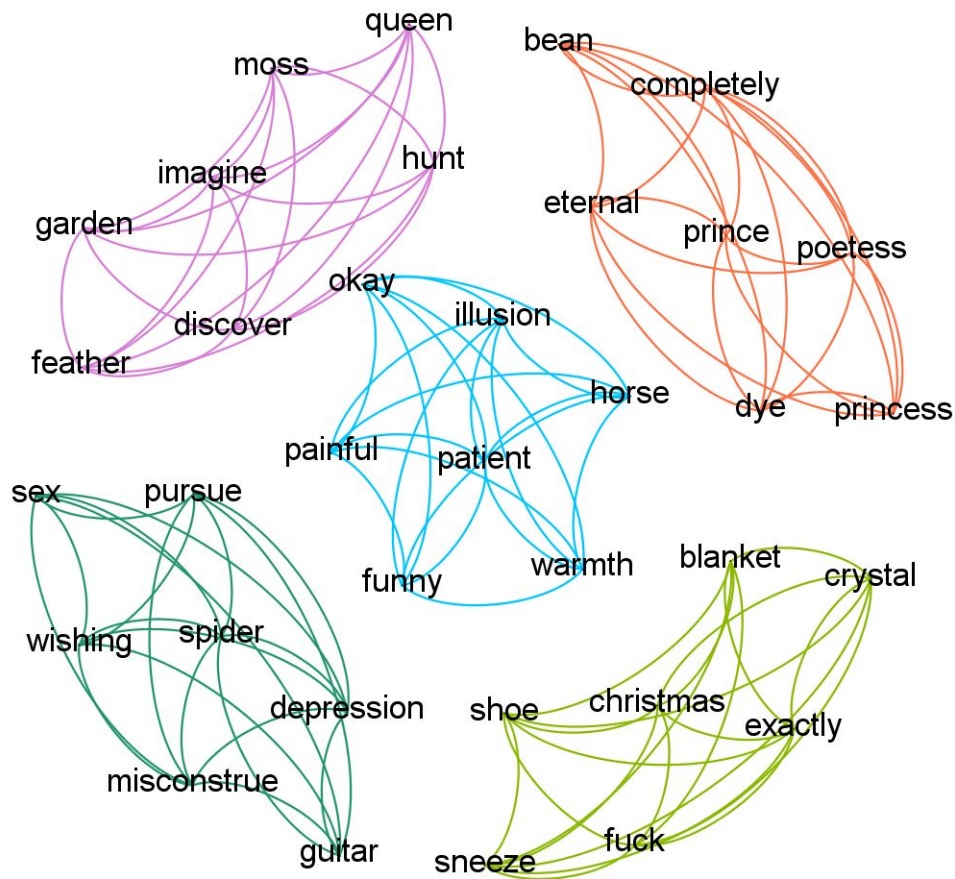
Topic Modeling

NMF
Children



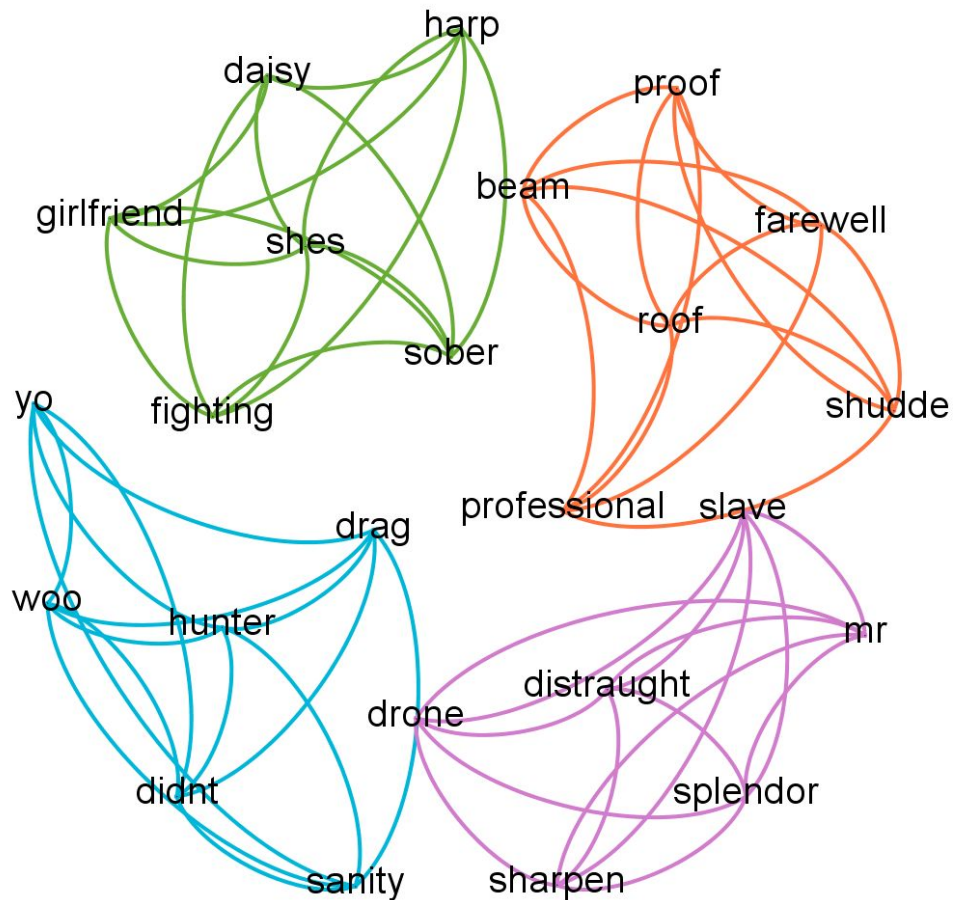
Topic Modeling

LDA
Amateurs



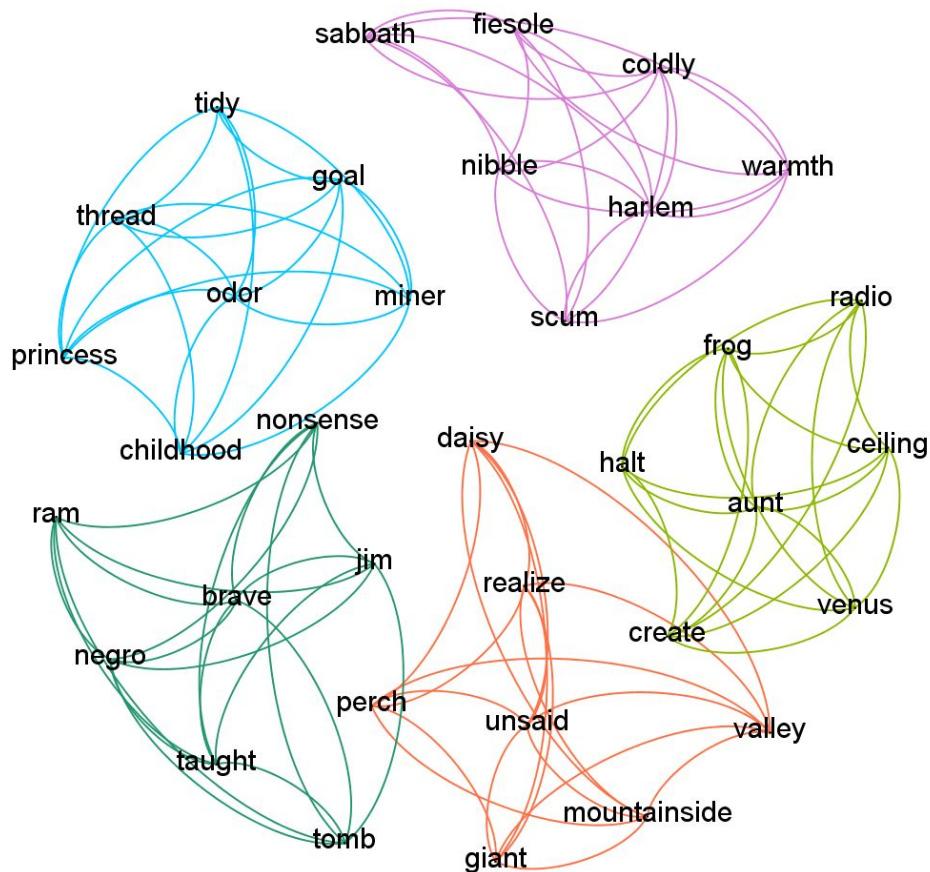
Topic Modeling

NMF
Amateurs



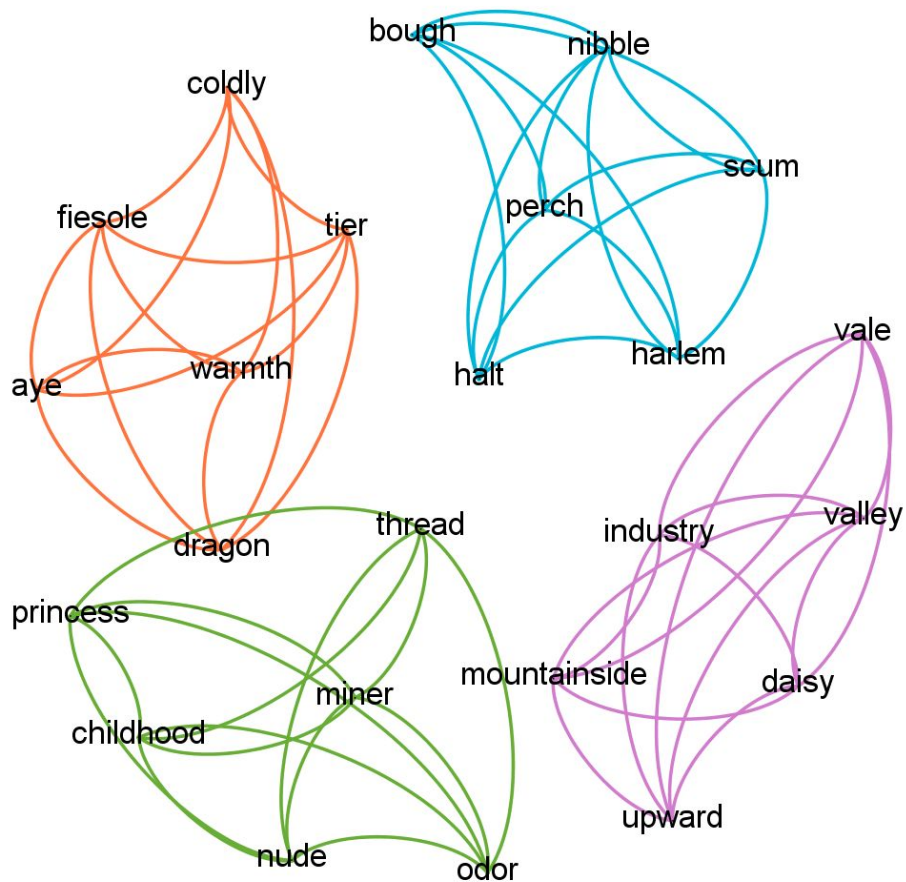
Topic Modeling

LDA
Professionals



Topic Modeling

NMF
Professionals



Классификатор

	precision	recall	f1-score	support
child	0.90	0.91	0.91	138
lay	0.86	0.80	0.83	124
prof	0.83	0.87	0.85	127
avg / total	0.86	0.86	0.86	389

	precision	recall	f1-score	support
child	0.91	0.93	0.92	148
lay	0.82	0.78	0.80	106
prof	0.82	0.84	0.83	135
avg / total	0.86	0.86	0.86	389

	precision	recall	f1-score	support
child	0.92	0.96	0.94	136
lay	0.85	0.82	0.84	114
prof	0.87	0.86	0.86	139
avg / total	0.88	0.88	0.88	389

Multinomial Naive Bayes

Cross validation (sklearn KFold, 3 splits)