

Домашнее задание 5

Вне зависимости от того, какой вариант вы выберете, обязательно запишитесь, в этом документе. Это важно, чтобы темы не повторялись.

Вариант 1, презентация статьи

Выберите одну из статей опубликованных в материалах воркшопа “Natural Language Processing meets Journalism”, прочитайте ее и сделайте по ней презентацию. Ваша презентация должна быть рассчитана на выступление длительностью 5 минут.

Материалы

- Сборник 2016
<https://drive.google.com/file/d/0BwmHO3smKNIpNFVncFFVU2FoejQ/view>
- Сборник 2017
<http://www.aclweb.org/anthology/W/W17/W17-42.pdf>
- Сборник 2018
http://lrec-conf.org/workshops/lrec2018/W13/pdf/book_of_proceedings.pdf

Вариант 2, применение сетевого анализа

Это задание можно выполнять в одиночку или в группе из двух человек.

Придумайте себе тему мини-исследования, это может быть анализ отношений героев в некоем произведении, анализ упоминаний персон в новостных текстах, извлечение ключевых слов из текстов, с которыми вы работали ранее, когда выполняли домашнее задание 3 и др.

Постройте граф и проанализируйте его. Опишите цель, ход работы и полученные результаты в презентации. Ваша презентация должна быть рассчитана на выступление длительностью 5 минут. Для построения графа воспользуйтесь программой Gephi.

Материалы

1. Коротко про работу в Gephi
<https://ancatmara.gitbooks.io/digital-literacy-for-sfl/content/seminar-11.html>
2. Тьюториал по Gephi
https://gephi.org/tutorials/gephi-tutorial-quick_start.pdf
3. Тьюториал на китайском
<https://www.udemy.com/gephi/>
4. Плагин Sigma для создания интерактивных графов
 - [Скачать плагин](#)
 - В статье из п.1 в конце описано, как воспользоваться плагином и выложить граф на GitHub.
5. Датасеты, которые можно использовать (не берите, пожалуйста, датасеты, для которых уже выложена презентация и сделан какой-то анализ)
<https://hum.hse.ru/digital/data>

