# Задание 2. Поиск похожего текста

## Задание

Совместно в группах решить задачу поиска похожего текста с использованием того корпуса текстов, который вы собрали в Задании 1. Необходимо выбрать один из корпусов и выполнить следующие задания.

* Построить облако слов по корпусу текстов.
* Создать векторизатор по корпусу текстов.
  + Необходимо выбрать минимум 3 комбинации параметров min\_df и max\_df.
* Определить меру сходства текстов, которую вы будете использовать, и выполнить поиск ближайших текстов.
  + При необходимости искусственно дополнить корпус самыми коротким и самыми длинными текстами с целью дальнейшей нормировки текстов;
  + Выполнить поиск похожих текстов ДО нормирования и ПОСЛЕ нормирования;
  + Вывести минимум 5 текстов в отсортированном порядке.

## Ожидаемый результат

1. Файл jupyter-notebook, который содержит решение все перечисленных выше задач с поэтапным выводом результатов (как было в примере в лекции).
2. Текстовый документ с интерпретацией результатов - словесное описание результатов (дало ли результат изменение параметров векторизатора, нормирование данных).