# 大数据可视化技术考试题目1

可视化你最喜欢的一本小说的网络结构

步骤与要求：

1. 选择你本你喜欢的图书，如果是大部头多卷作品的，选其中一部即可，如《冰与火之歌》共发表5部，可选第1部《权力的游戏》；
2. 下载该图书txt文档；
3. 设计算法并用Python实现，提取出文字中的复杂网络模型，可以进行分词，虚词过滤，同义名词预处理等操作，以保证网络建模更加有效；
4. 调用NetworkX包，将建模的网络生成.gexf格式的网络数据，采用Gephi或NetworkX计算中心性、社团；
5. 导出中心性数据，使用Excel进行中心性分布可视化；
6. 导出社团数据，使用Excel进行网络中节点在不同社团比重的可视化；
7. 使用Gephi，对网络结构进行可视化；
8. 通过可视化和统计数据，配图说明两个问题：1）哪一个你曾经对这本书中一个问题的感性认知，在可视化和统计分析过程中，找到了证据？2）哪一个你曾经对这本书的一个感性认知，在可视化和统计分析过程中，被推翻?

------------------------------------分割线------------------------------

上交材料要求：

1. 打印版纸质的实验报告；首页、目录单面打印，报告正文双面打印；
2. 实验报告文档电子版，pdf或word文件均可；目录名：学号姓名
3. 实验代码和数据电子版，.py和.ipynb均可，实验代码与所需数据放在一个子目录中，这个子目录放置在实验报告电子档的目录中；
4. 自评报告纸质版：一张打印，按照评分要求对自己上交的材料进行评分，分析优势和扣分原因；

------------------------------------完结线------------------------------