选题

# 2/22/2023

* 用Ian的工具，别的工具等找recommendations
* From a collection of papers / raw texts, discover cutting edge (new) topics that are worth digging / researching
  + If a corpus of literature / papers (realm of meta-analysis):
    - [TODO] A good paper on this approach!! <https://www.sciencedirect.com/science/article/pii/S2590159122000231>
    - Most common tags (if available)
    - Extract abstracts of these papers => raw-text dataset =>
      * Most common (research) topics
        + But how to determine? Should differentiate with just “tags”
        + [TODO] Would probably need a topic modelling / topic classification: <https://monkeylearn.com/blog/introduction-to-topic-modeling/>
    - Extract “future work” sections of these papers => raw-text dataset => make sense of the dataset
      * This strategy is good, but may not be useful to implement because usually the num of papers at this stage is small enough for human to process and should be processed by human too.
  + If individual author:
    - Language pattern(s) shift along the author’s timeline?
    - Most common words / phrases / imageries that are significantly different from other authors?
      * Needs comparison dataset
      * Needs a lib of literary words / phrases / imageries (meaningful)
    - Sentiment analysis etc.
    - …?
  + If a corpus of a genre (e.g. Gothic):
    - Language pattern(s) shift along the historical timeline?
      * Emerging trends?
    - Most common words words / phrases / imageries that are significantly different from other genres?
      * Needs comparison dataset
      * Needs a lib / pre-trained model of literary words / phrases / imageries (meaningful)
    - Continue writing for unfinished work using NLP DL models
      * Or suggestions for new writing patterns in the contemporary era

# 2/26/2023

决定topic 是爱伦坡vs陀氏

目前TODOs:

* 找相关综述，最好是精神分析 and / or 文本data分析方向
* 阅读并继续缩小研究问题范围
* （基本文本阅读）