

# Meeting Agenda

8-7-2021

# 近期目标

- 面向一般用户的报告

- 以PPT的形式，报告近期工作进展（两周后）

- 简历分析

- 标签分类与查询

- 职业社交网络

- 简历标准化：根据分词与预处理结果

- 对格式不规范的简历进行标准化（将多个经历拆分，地名对齐）

- 对错误格式（如错误时间，地名无法判断或名分词歧义）的经历进行标注

只需对比“处理前”与“处理后”的文本信息

- 交互式演示（八月底、九月初）

- 职业社交关系查询功能（个人轨迹展示）

- 文字交互即可（不一定需要可视化）

- 标签查询（支持模糊语义）

- 例如查找双一流大学毕业、三年以上法律方向工作经验的正科级干部（对应“双一流大学”，“政法”，以及当前职级为正科级的标签）

- (optional) 通过annotator标注数据，由用户增添新标签，并更新分类模型

- (optional) 标签学习的其他应用。如判断一个领导班子是否符合标注。例如要求一个选定的五人小组中两位来自基层，一位处级干部，一位具有教育背景)

# To Dos

- 解决分词问题
  - NER 分词模型优化（待分配）
  - case-by-case NER分词结果的后处理 (基本完成)
- 解决OC Tree与职业网络问题
  - bug fixes（优先确保共事网络的正确性，争取上下级关系使用职级信息）— dexu
  - 查询职业网络的API实现 — dexu
  - 查询职业网络的互动脚本 — kiki, yongchi
- 开发迭代流程：everyone 下周一讨论
- 在线学习与知识图谱标签学习部分的融入：
  - 手工标注数据 — everyone
  - 与当前repo的融入 — zhiwei