

# 五、考核方式

## 考试形式

- 期末报告： 大数据收集可视化、文本分析、机器学习，选择1~2类进行研究详细设计或初步成果报告。PPT+Word
- Presentation: 15分钟, 5.20/5.27/6.3
- Word: 2500字以上, 6.30日之前发到邮箱chenzhuo@nju.edu.cn
- 评论: 评论别人的报告两次



# 研究设计/成果报告形式

- 1. 研究题目

必须体现出社会科学的问题意识与计算方法的结合

- 2. 研究背景

- 3. 研究现状-评述（详细）

找准目前研究的gap，如何运用大数据和计算方法创新或者弥补

- 4. 研究问题

- 5. 研究数据（详细）

详细展示数据形式、结构、收集方式、大小、可得性

- 6. 研究方法（详细）

详细展示分析步骤、方法、工具、指标操作化…

- 7. 研究重点和难点

- 9. 研究局限

- 10. 参考文献





- 1. **研究类型：**描述性研究（大数据/大时空）；解释性研究；预测性研究，预测+解释整合研究
- 2. **大数据的来源：**搜索大数据（百度）；图书大数据（谷歌图书）；法律大数据（裁判文书）；媒体大数据（自媒体平台）；科学大数据（论文）；地理大数据；政策大数据…
- 3. **文本分析方法：**词典法、实体关系抽取、文本向量编码、相似度计算、主题挖掘、文本聚类、文本分类、大语言模型…
- 4. **社会预测方法：**有监督学习，可解释性机器学习，因果机器学习…



- 1. 研究亮点
- 2. 研究问题：研究问题是否真的存在，大数据和计算方法如何发挥长处？
- 3. 数据的合理性：数据是否可获得？数据包含的信息能支撑指标操作化？数据的代表性？有没有其他可能的数据源？
- 4. 方法与技术匹配度：是否合理的选择了合适的方法？分析流程是否清晰？方法存在什么局限？有没有其他更合适的方法？
- 5. 结构的合理性：研究设计的逻辑是否通顺，结构是否合理？

