



中美关系可视化

杨昊翔，吴宇恒，刘云蛰，易普

数据集

设计目标

- 收集近百年的中美关系数据
- 呈现中美关系随时间的变化

数据集

- 对9份报纸逐一网络爬虫、筛选
- 向纽约时报邮件申请数据

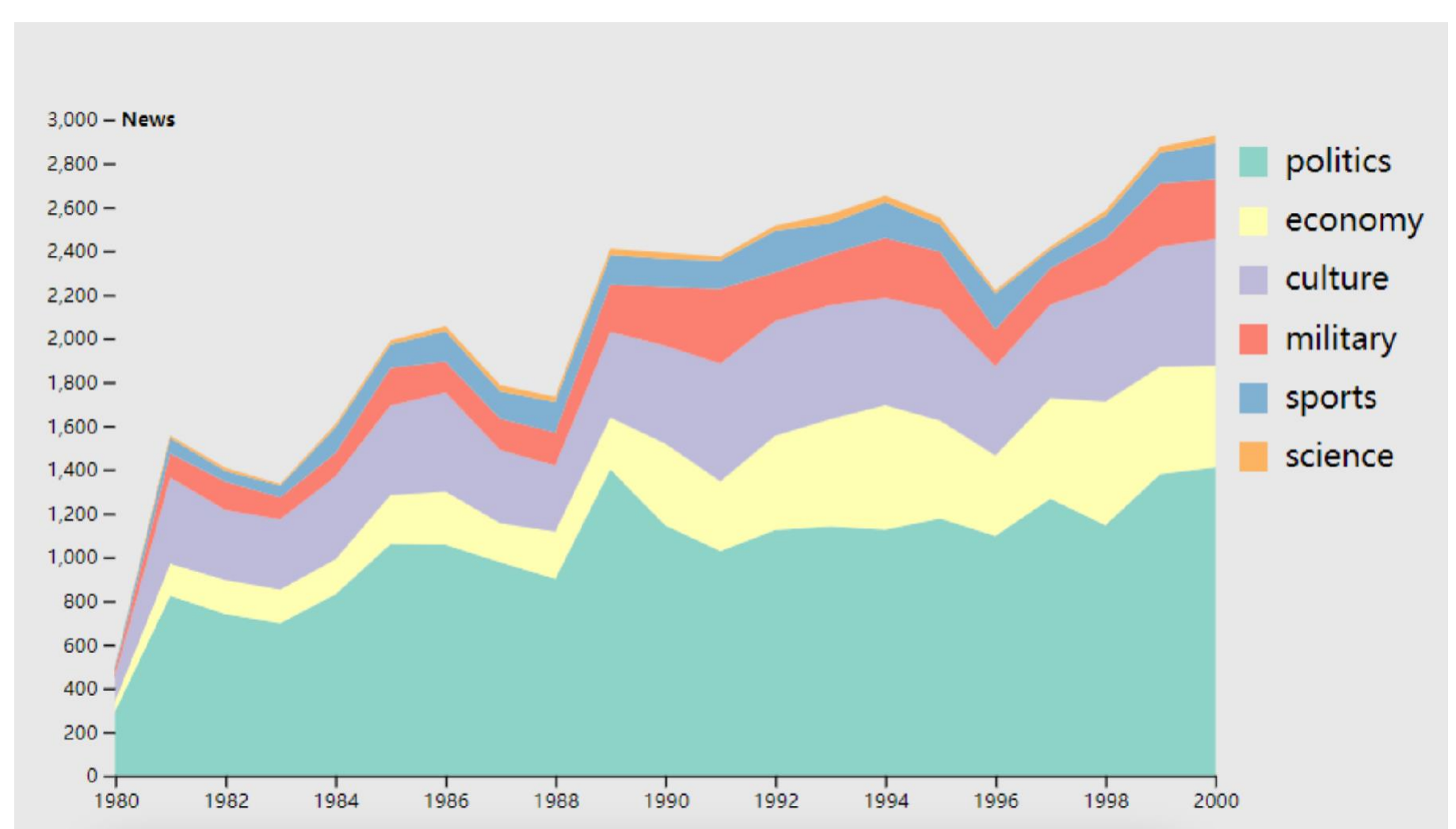
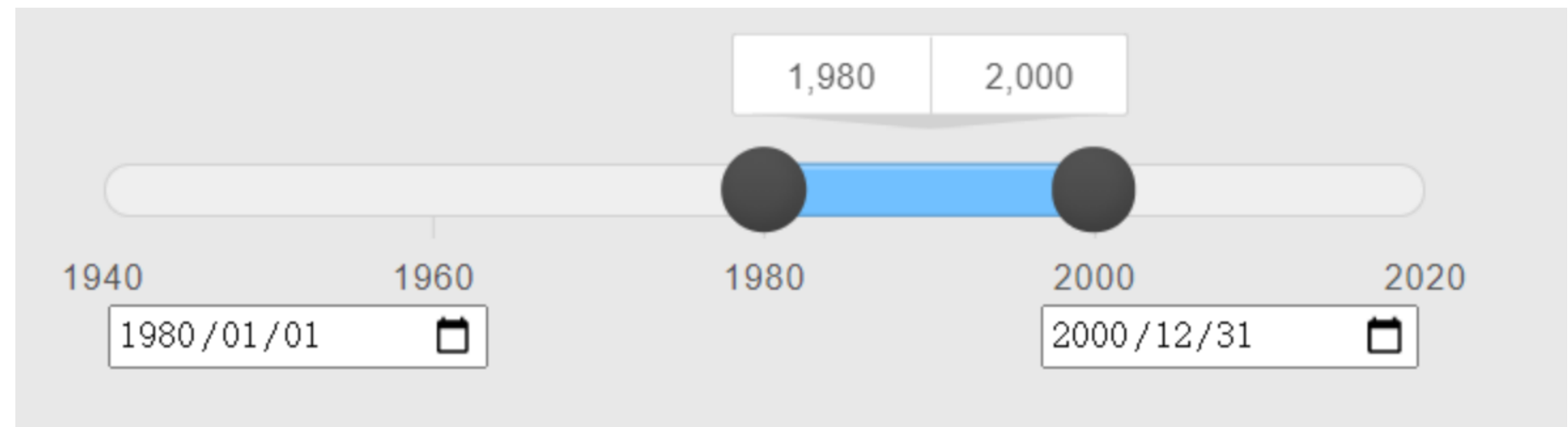
数据格式

- 标题、摘要、时间、超链接
- 共约30W条数据
- 每条数据超过10个单词

报纸数据来源列表

- New York Times
- Christian Science Monitor
- Los Angeles Times
- Scientific America
- The Sun
- Wall Street Journal
- China Daily
- Global Times

系统设计



概览：时间滚轮、面积堆叠图

- 时间滚轮选定欲查看的事件区间，日期模块精确到天
- 面积堆叠图展示选定时间区间政治、经济、文化等六个分类的新闻数量变化
- 由于数据无标签，采集六个类别的常用词，将每条新闻指定为出现词频最高的那一类

筛选：时间线与词云

- 从时间线看关键信息
- 时间线可左右平滑切换
- 从词云可筛选关键词
- 词频使用TF-IDF统计



Relevant News

- Four Instrumental Virtuosos From China; Chinese Virtuosos Performing ...
- STRATEGIES FOR GRADUATES IN THE NEW CHINA\$
- New Wariness Over China Deals; New Wariness Develops Over China C...
- China Rejects U.S. Protest And Hints at More Purges
- Work With China, Without Illusions
- Letters to the Editor; Wild Dogs Sadden Seattle? More on Movies Aloft...
- China's Paragon of Corruption; Meet Mr. Chu, a Hero to Some, an Emb...

细节：相关新闻

- 多选关键词，右侧文章哪找关键词出现的词频从高到低排序显示
- 点击链接访问对应新闻

设计原则：

- 先概览，按需展示

挑战：

- 数据量偏大
- 多模块交互设计

解决方案：

- 分模块进行多层采样

结果分析

历史趋势

- 中美关系总体上升，略有波澜
- 经济文化交流占比上升
- 可视化结果与重要历史事件对应
- 1971年乒乓外交与2012年贸易竞争影响颇大

趣事逸闻

- 苏联解体前常有 soviet communist
- 64年我国原子弹爆炸，军事新闻大幅增长
- 越战前后 vietnam 频频出现
- 美国总统成榜上常客