

一带一路相关新闻全球媒体监测与分析系统

杨立 马佳佳 方明哲 孔维一
软件发展研究部

联系方式: 杨立 yangli2017@iscas.ac.cn 13301228953

系统介绍

背景: “一带一路”倡议自提出以来,得到了世界上越来越多的国家和国际组织的认同,其相关报道与日俱增。

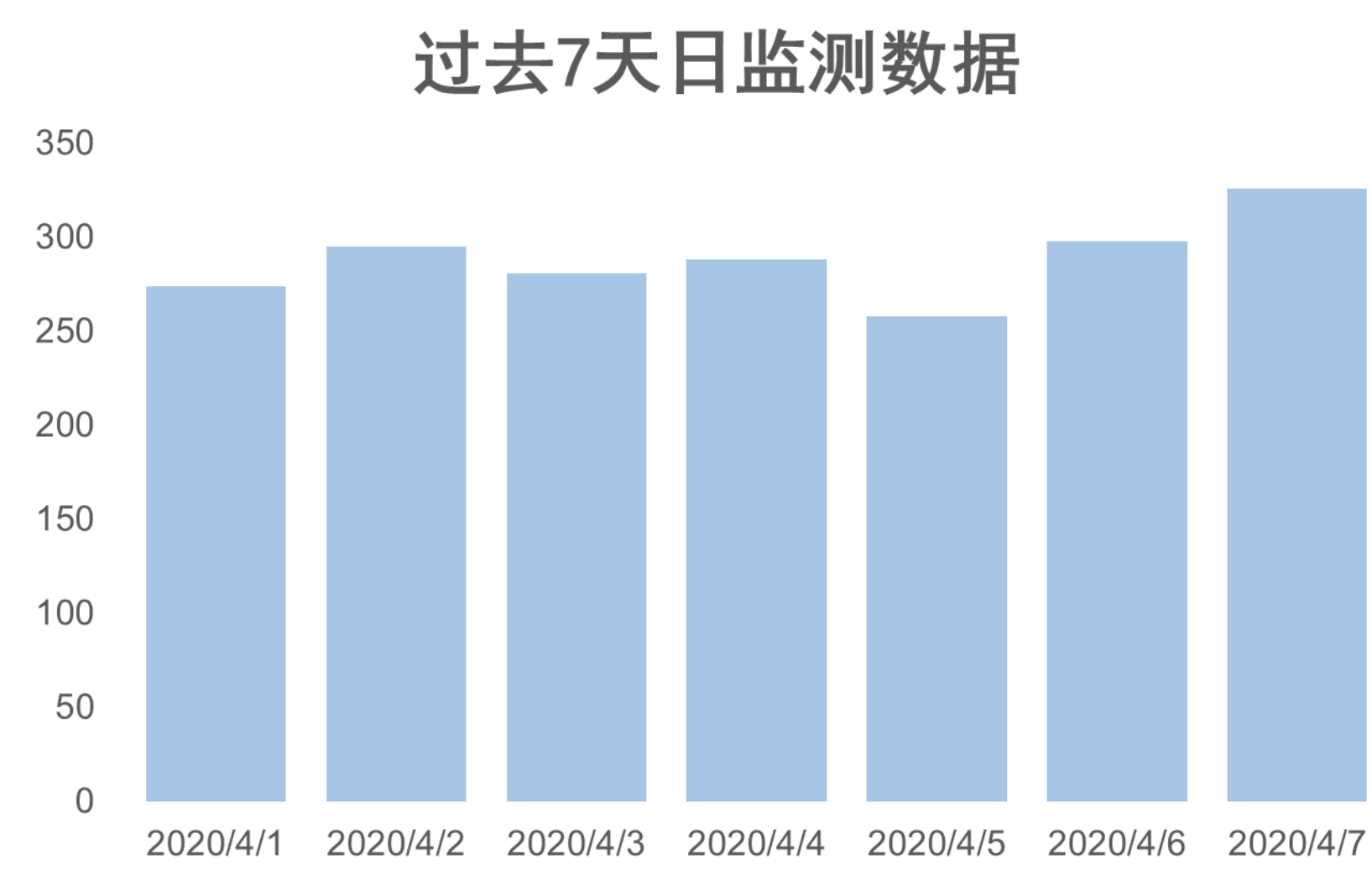
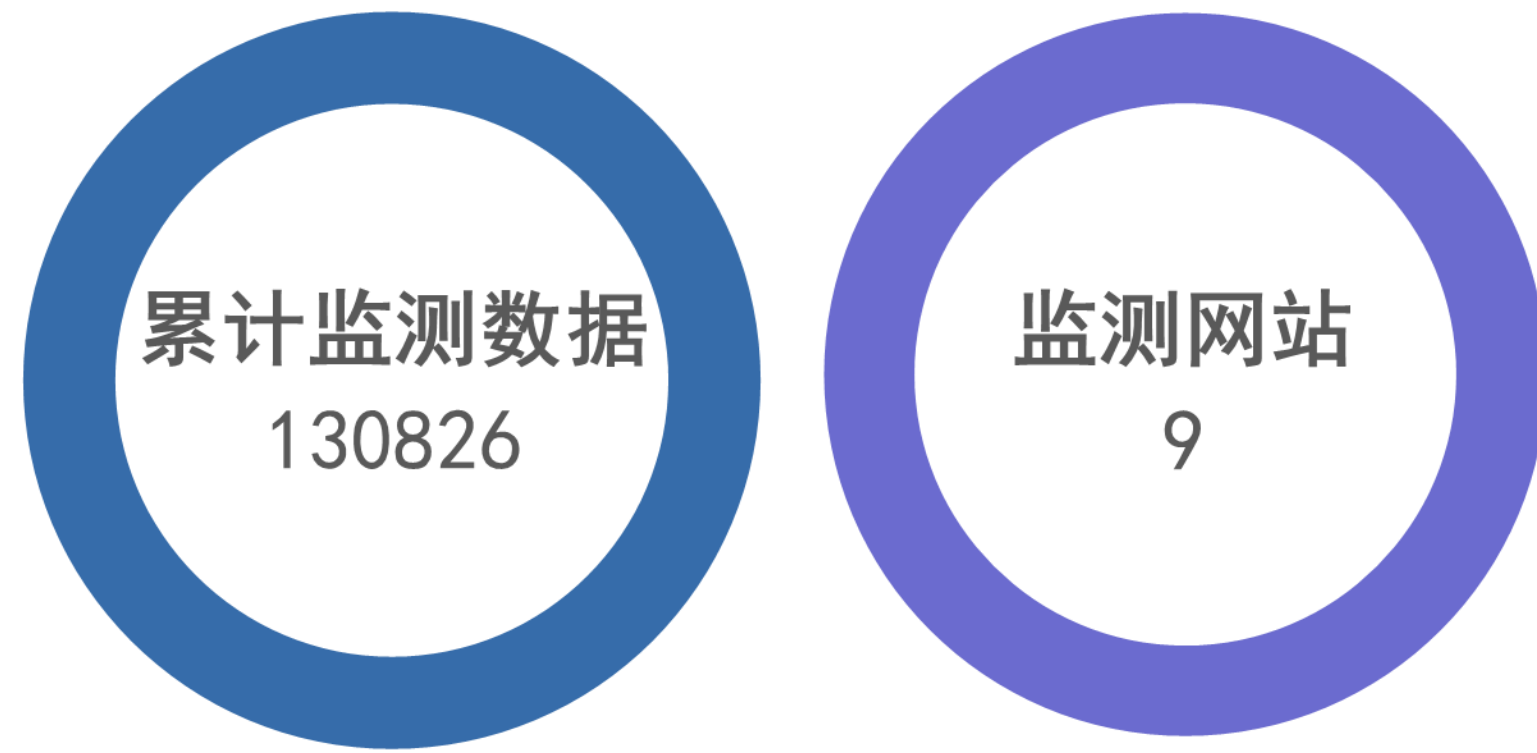
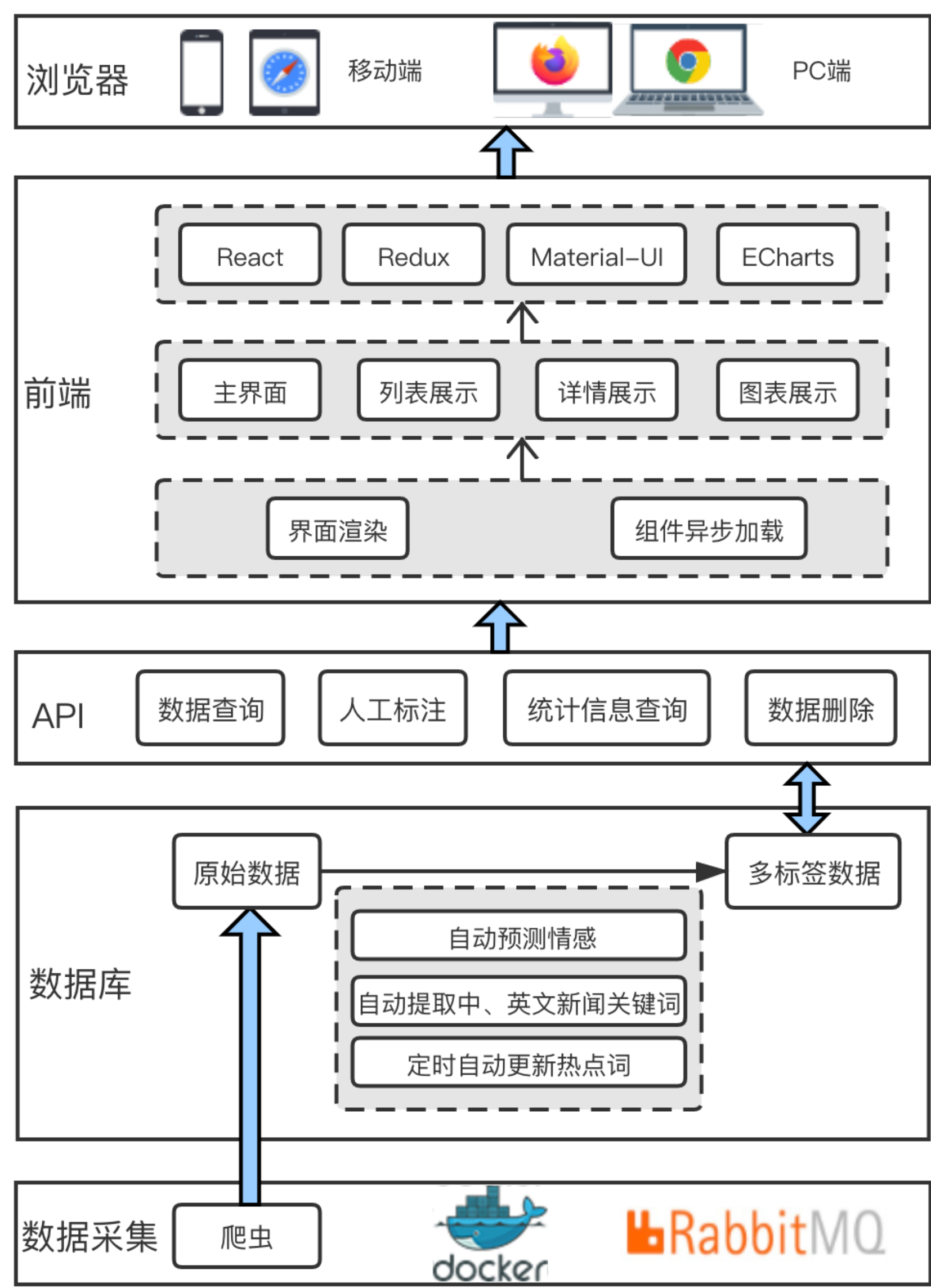
挑战: 如何在海量新闻报道中快速全面地监测“一带一路”相关报道?

如何准确地自动提取“一带一路”相关报道的重要信息?

如何智能监测“一带一路”相关报道的情感态度?

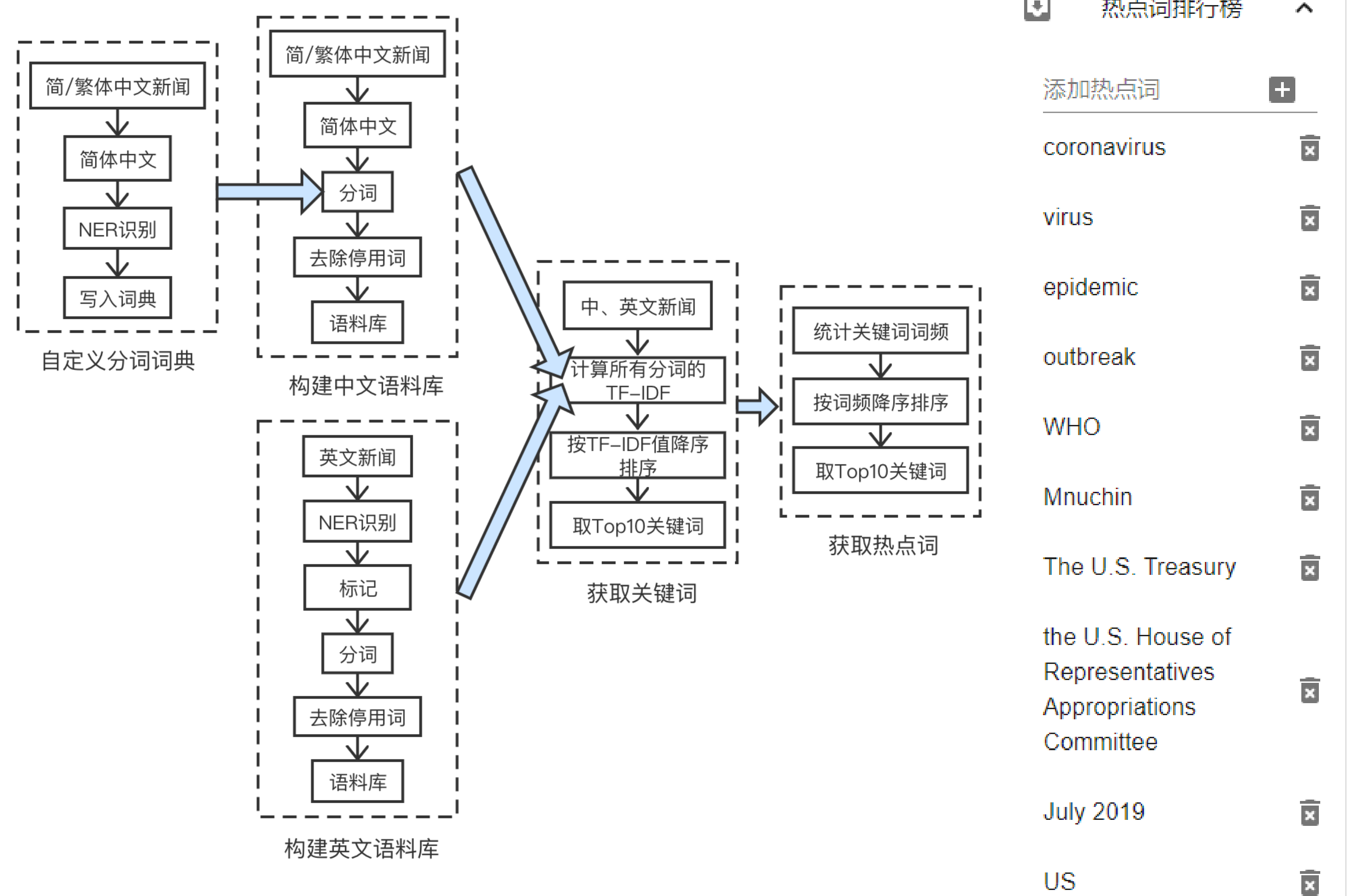
工作内容: 本系统通过对媒体覆盖监测可快速获知“一带一路”最新信息(采集周期最快秒级,最慢几小时/次),依托人工智能技术进行智能分析(挖掘热点词、负面新闻等),同时提供可视化界面辅助用户决策。

系统架构



热点词分析

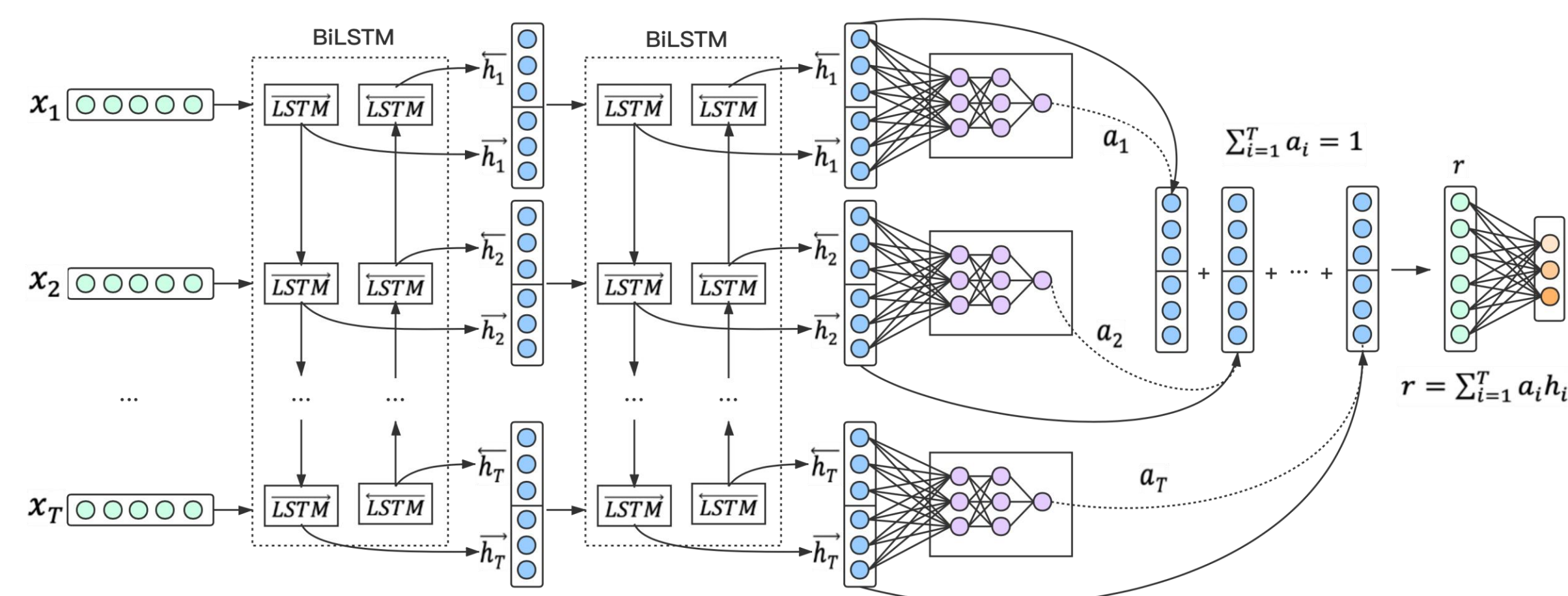
关键步骤: 1) 分别对中、英文新闻进行关键词提取; 2) 根据关键词的词频确定排名高低。



情感分析

任务: 预测新闻情感为正面、中立或负面

模型: 双向LSTM + Attention



Twitter识别效果举例:

Text snippet: "@vaiala If you read the thread she appears to have gotten the info from other tweets. Most of the replies to her, as in all her other tweets about China, are extremely racist and she doesn't discourage them. She's super opposed to belt and road initiative & written extensively about it"

Text classification: positive (green), negative (red)

Model prediction: 预测情感 negative

Annotation: 内容中包含关于中国和“一带一路”倡议的负面信息

时间线提取

任务: 针对某关键词,归纳涉及该关键词的新闻,识别新闻中的时间,并按照时间降序排列。

以“coronavirus”为例,部分时间线如下所示:

The timeline shows news items for 'coronavirus' from 2020-03-24 to 2020-04-02. To the right, two donut charts provide statistics: '新闻来源统计图' (News source statistics) showing counts for bbc, beltandroad, reuters, and twitter; and '预测情感极性统计图' (Predicted sentiment polarity statistics) showing percentages for negative (51.81%), neutral (27.11%), and positive (21.08%) sentiments. A list of detected countries is also provided.

系统应用

该系统已应用在中国科学院A类战略性先导科技专项“泛第三极环境变化与绿色丝绸之路建设”中,并可扩展到新闻采编、舆情监控等应用领域。