

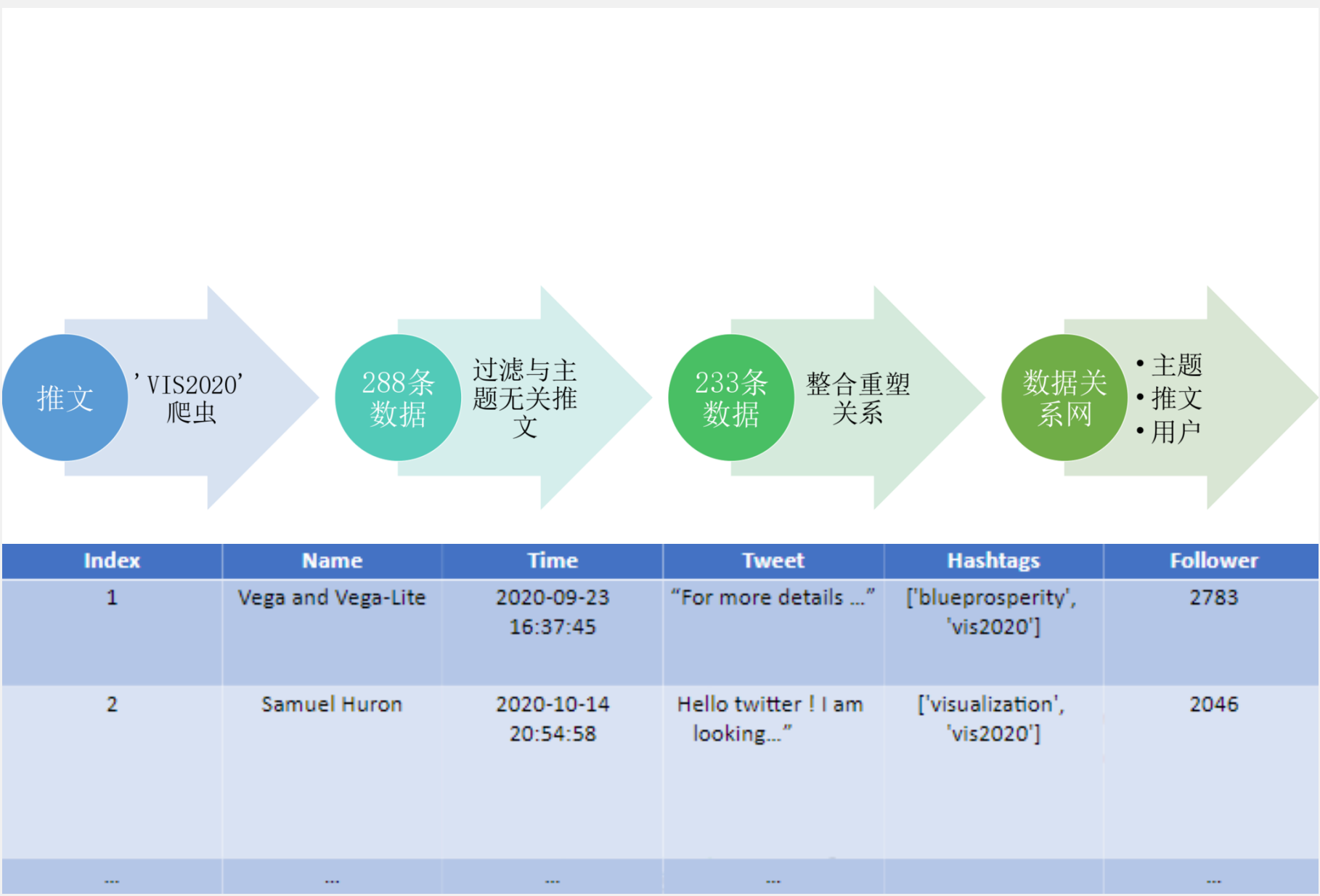


VIS2020学术会议推特数据可视化

黄孝谨，田野，周子淼，李钰

简介

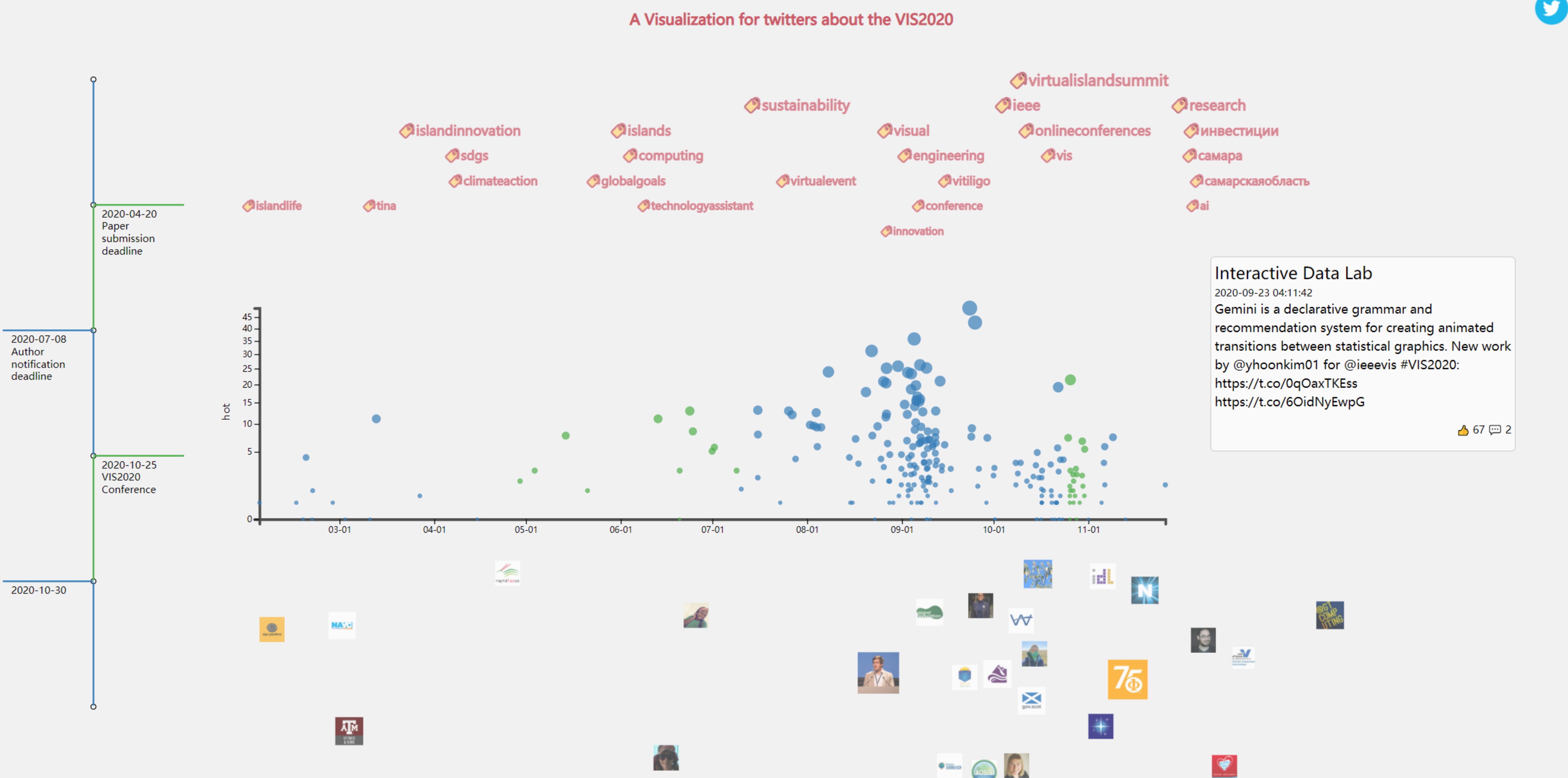
本可视化项目旨在分析可视化学术会议"IEEE VIS2020"相关的推文数据，从中得出会议进程、发送用户、标签等等的相关关系。我们选取了包含"VIS2020"标签的所有推文，尝试从推文的时间、



数据

数据分为两部分，一部分是推特数据，一部分是用户数据。推特数据的属性有：创建时间、发推用户名、推特内容、点赞数、回复数、转发数、相关标签。我们对点赞数、回复数和转发数这三个属性做了加权统一，得到一个新的属性，为推特的热度。用户数据的属性有：用户名、粉丝数量、用户头像。在进行数据处理后，我们得到了用户发推时间平均，以及不同tag的热度，从而获得系统设计的各个参数。

系统总览



1) 标签

标签选择爬取数据中相关推特数最多的25个标签进行分布，提到数量越多的标签分布越靠上字体越大，横坐标为该标签相关推特的横坐标均值。单击标签，可以在图二中高亮出带有该标签的所有推特并自动对时间轴进行缩放，以及在图三中放大发推的所有作者。

2) 推特内容

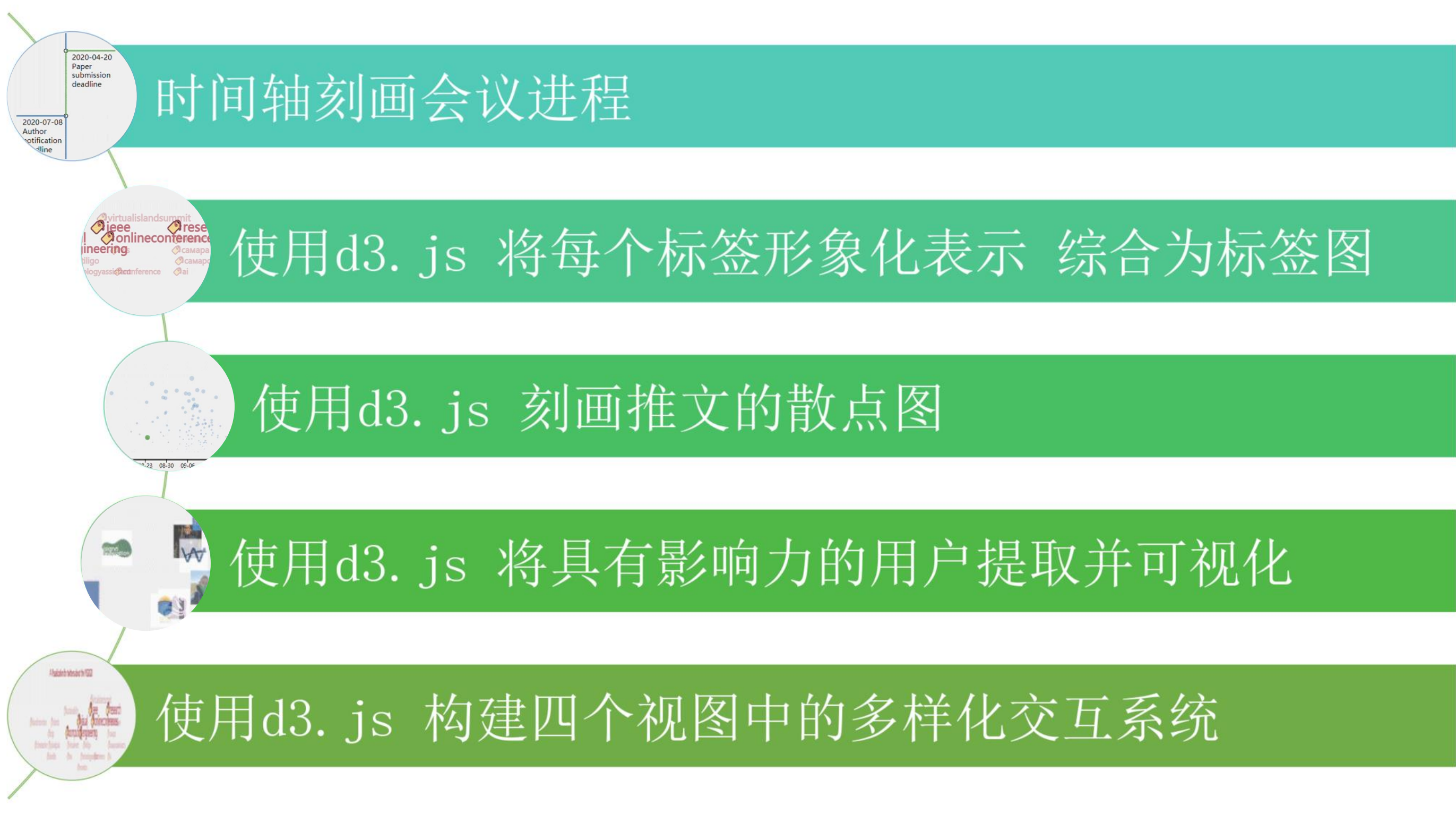
图二为以发推时间为横轴，热度为纵轴的散点图，颜色对应屏幕左侧会议进程的不同阶段。单击可以在屏幕右侧显示该推文的具体内容，并在图一放大所带标签，在图三放大其作者。为了使点均匀分布，纵坐标轴使用非线性。横坐标轴可以根据感兴趣的时间节点进行缩放。

3) 用户

用户以头像的方式呈现。其中关注者越多的用户认为影响力更大，在图中表现为更大的头像面积。它的主要作用是和以上两部分交互，在选取标签、推文时能够浮现对应的作者便于观察。

4) 会议时间线

时间线描述了会议的几个重要时间点。其中不同时间段分别用不同颜色来描绘，对应推文散点图上可以快速筛选出一定时间段内的推文。



结果分析

- 在7.25-10.25日，即作者通知截止时间到会议开幕时，推文数量最多。其中的主要内容是对与自己论文相关议题的预热和邀请讨论。
- Interactive Data Lab 所发布的关于 Gemini 和 Boba 可视分析方法以及其应用的可视分析工作获得了最高的热度。在可视化前沿工作中可以加以参考。
- 发推最多的用户是James Ellsmoor，他是可持续发展和可再生能源中的佼佼者，是太阳能国家首脑的联合创始人和董事。