

大选数据中的“秘密情报”

李熙爱

联合早报，2020年10月8日

大选 2020 的结果已成定局，但投票日前若能估计民意方向再调整策略，结果是否会不同？

本届大选的主要战场转向网络，数据分析公司 Analytix Labs 分析了面簿数据，从中找出有关民意的四大“情报”。Analytix Labs 首席数据分析师蔡振峰认为，各政党虽然在本届大选中更懂得制作内容和利用社交媒体，但忽略从社交媒体收集“情报”并即时调整策略。到底从这些“情报”能看出什么，让 zaobao.sg 带大家了解。

“情报”来自哪？

Analytix Labs 从面簿贴文收集了公开数据，通过搜寻媒体（联合早报、联合晚报、海峡时报、MotherShip 等）、政党（行动党、工人党、民主党）和政治人物（李显龙、毕丹星、陈清木、李显扬）的面簿页面，找出与各个大选关键词相关的面簿贴文，再从中计算总互动量（interactions）。在这里，面簿互动指的是与贴文进行的九种互动：分享、留言、点赞，以及六种表情（爱心、惊讶、笑脸、伤心、生气、关怀）。

虽然数据只源自面簿这一个社交媒体平台，但蔡振峰指出这是新加坡人最常接触新闻的社交媒体渠道。他强调，这些数据显示的趋势无法和最终的选举结果形成因果关系，它只能作为估计民意的一个参考指标。

1. 国人对疫情的关注度已降低 并非预计中的“疫情大选”



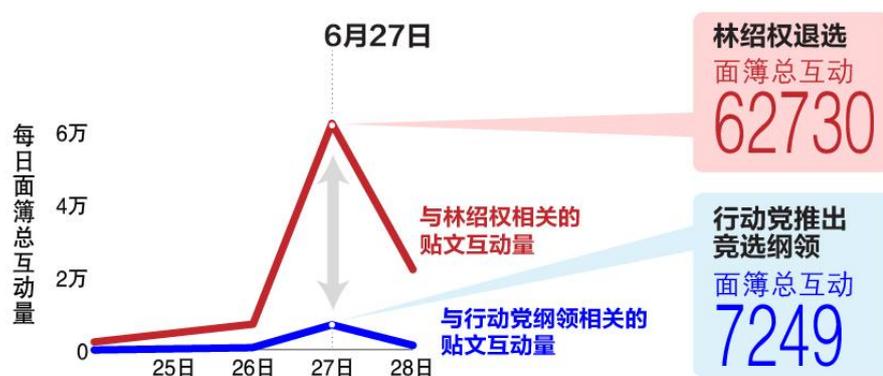
注 / 数据源自32,602个与冠病疫情相关的面簿贴文
资料来源 / Analytix Labs

在疫情期间举行大选，不少人普遍认为疫情带来的危机和不确定因素会使选民更保守，让他们投靠安全的选项。但根据疫情相关贴文的互动量，国人对疫情的关注度从4月开始已逐渐下跌，到了大选期间关注度更低。

蔡振峰认为，到了六七月大选期间时，政党就应该对国人会因疫情而投安全选项的假设质疑了。

所谓的“疫情大选”更多在于形式，例如得戴口罩和保持安全距离等，而非本质上因疫情而影响选举结果。

2. 退选的林绍权关注度比行动党政纲高 政纲讯息被淹没？



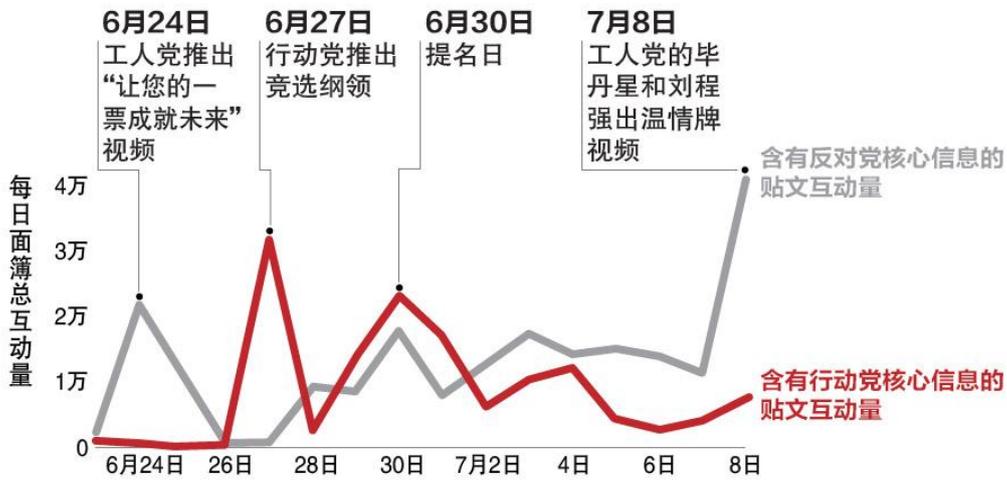
注 / 数据源自出现“行动党”+“林绍权”，以及“行动党”+“竞选纲领”字眼的89个面簿贴文
资料来源 / Analytix Labs

zaobao.sg

行动党正式介绍林绍权为候选人后，网上陆续有人分享与他不太愉快的经历，并质疑他是否合适参选。风波越闹越大，四天后林绍权在6月27日中午首度发表声明说会“坚持走下去”，结果当天晚上却突然宣布退选。短短一天之内发生戏剧性变化，让他占据了社交媒体和新闻页面。

不巧的是，行动党也在当天推出竞选纲领，由秘书长李显龙亲自分享。然而数据显示，与林绍权相关的面簿贴文所获得的反应是纲领的8.65倍。蔡振峰认为，行动党从中就应该推断纲领的“守护生命、保障工作、共创未来”信息没能获得网民的共鸣，网民很明显得对林绍权事件更有兴趣。

3. 行动党关注度过早达顶峰 工人党越战越勇



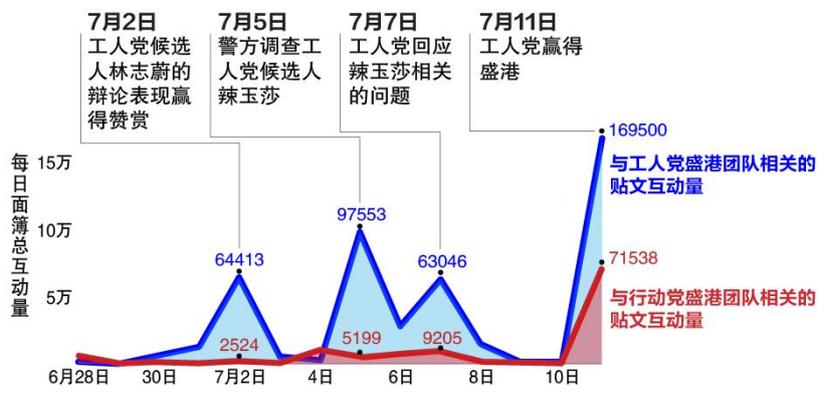
注 / 出现行动党和反对党所传达的讯息关键词的154个面簿贴文
资料来源 / Analytix Labs

zaobao.sg

虽然行动党的纲领关注度不如林绍权，但与行动党核心信息有关的贴文互动，在当天其实已达到高峰，随后与行动党要传达信息有关的贴文互动率就减少，后劲不足。李显龙总理7月6日在线上午餐群众大会上发表的重要演说稍微推高了行动党在面簿的关注，但效果不大。

相反的，反对党要传达的信息反而在后期表现得更好，工人党在竞选的最后一天更是出动了秘书长毕丹星和前秘书长刘程强录制感性视频，面簿互动达到高峰（4 万多个互动），让工人党在情绪高昂的气氛中结束九天的竞选活动。

4. 盛港之战：工人党候选人曝光率远远超越行动党



注 / 有盛港相关关键词（如候选人名字）的325个面簿贴文
资料来源 / Analytix Labs

zaobao.sg

盛港集选区被工人党攻下，对多数人或许是意外，但蔡振峰认为从面簿数据早就能看出端倪。从与盛港集选区相关的贴文来看，不难发现有关工人党候选人的贴文引来更多的互动和反应，当中林志蔚在辩论的表现和辣玉莎被警方调查这两起事件都引起了很大关注。

相对而言，盛港集选区行动党候选人的曝光率低了许多，在社交媒体上也无法吸引公众的关注和兴趣，就连李总理 7 月 4 日为盛港团队站台时也起不了太大作用。从头到尾，行动党盛港团队每天最多只有约 1 万次的互动，与工人党团队的差距甚远。

蔡振峰分享，美国政党在竞选期间往往会准备 Plan B 甚至 Plan C，根据民意调整策略或所要传达的信息。但新加坡政党的信息由始至终都相当一致，并没有根据情况而改变。他说，本地政党未来若能利用数据并正确解读这些迹象，在选举中可成为优势。

“情报”有多可靠？

当然，这样收集“情报”的方式有它的缺点，它只能粗糙地勾勒出民意的轮廓，不能直接预测选民的投票决定，毕竟选民在社交媒体上的互动不一定与他们投票的方式有直接关联。这就如去反对党群众大会的选民一般比行动党群众大会来得多，但大选结果出炉时多数人仍投票给执政党一样道理。

总计互动数量也是比较缺乏弹性的工具，总计互动数量也无法分辨不同的反应，在同一个贴文按赞的人，想法会和按“愤怒”的网民截然不同。就如行动党的林绍权和工人党的辣玉莎都在竞选期间得因个人事件引起高度关注，但结果却截然不同。

要如何在一堆数据中作出解读以搜集准确的“情报”，将是政党未来的挑战。