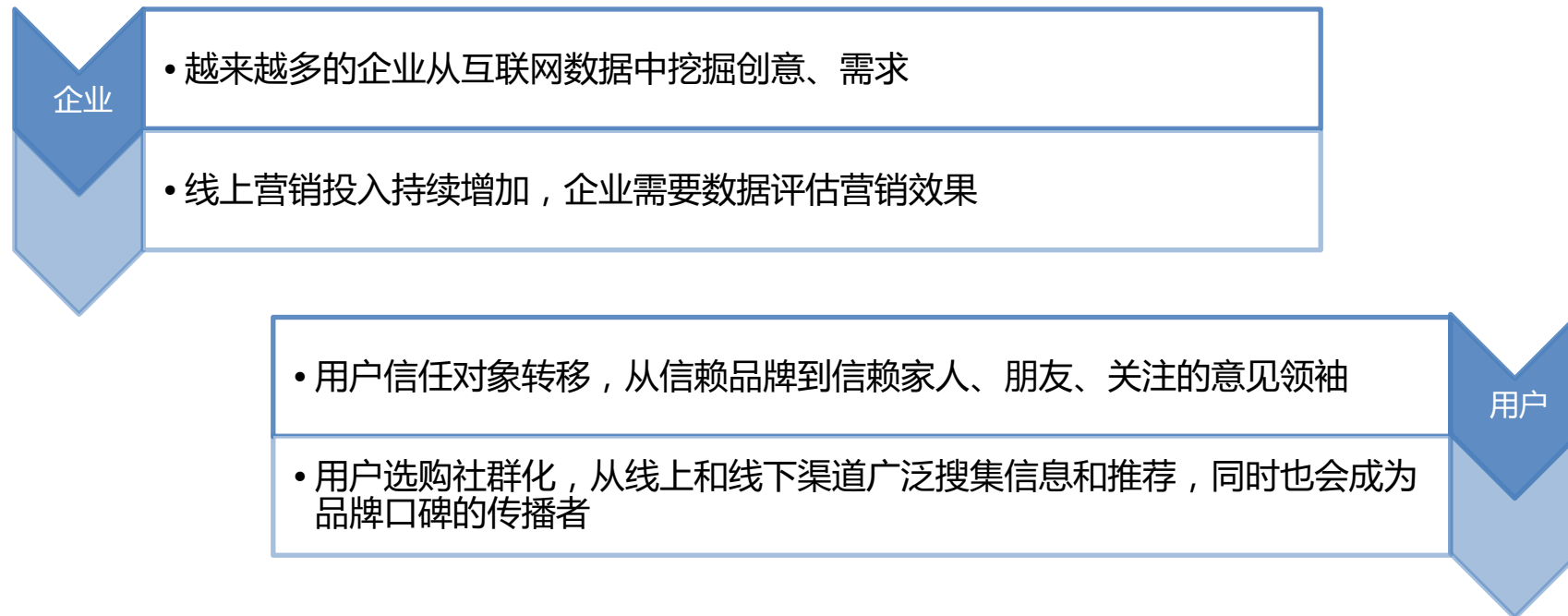


# 数说聚合



- ◆ 所属类型：平台
- ◆ 参选类别：年度最佳移动营销服务系统/工具/平台

当下越来越多的用户在互联网上表达自己态度、分享观点和见解，维护自己的社交关系网，对品牌企业来说，互联网是品牌聆听用户声音的重要渠道，越来越多的企业需要探索如何低成本快速构建互联网大数据洞察能力，以驱动企业数字化营销转型。



## 互联网大数据的应用困境

### 获取数据

#### 自建采集



耗时/耗人/耗钱；  
反爬/改版频繁，维护力不从心

#### 外部采购



全不全？稳不稳？贵不贵？

#### 数据清洗



文本数据如何结构化处理；  
结构化标签体系如何快速构建

### 分析数据

#### 分析工具



各类BI工具价格高、上手难；  
人员招聘、培训成本高

#### 周期/成本



常规一般不少于5个人天；  
价格不低于5万

#### 分析方法论



如何构建靠谱的分析方法论

一站式获取全网高质量数据

无门槛开启专业的数据分析



数说聚合-互联网大数据采集与分析平台

## 数说聚合核心功能与价值



### 数据采集

指定关键词/KOL/单帖等方式获取数据，轻松构建行业数据库

- 90%声量覆盖
- 多维度采集
- 内置丰富ETL
- 多种方式输出数据
- 边采集边推送



### 实时仪表盘

全网声量实时查询，仪表盘多维度下钻归因，探索式分析

- 全网数据实时洞察
- 峰值事件挖掘
- 双微互动指标获取热帖
- 热度趋势预判
- 热门词云/话题词云洞察



### 分析市场

内置多种商业分析模型，一键生成报告和解读

- 微博/B站/抖音等多平台人群研究
- 事件分析
- 口碑分析
- 微博热门评论分析等

全面覆盖各类数据源，千亿数据量积累



## 核心数据优势

- 每日入库数据量超过**1亿**，累积数据超**千亿**
- 微博**亿级**账号监控，全面集成**商业API**
- 抖音**千万级**账号监控
- 小红书/微信**百万级**账号监控
- 双微自带**T+3互动量**
- 持续集成**弹幕、直播、电商社区**等数据源

数说故事站点和KOL覆盖行业有：科技、数码、母婴、汽车、美妆、健康、家居、房产、影音娱乐、食品、金融、旅游、体育、游戏、电商资讯等

## 多维度全网数据获取

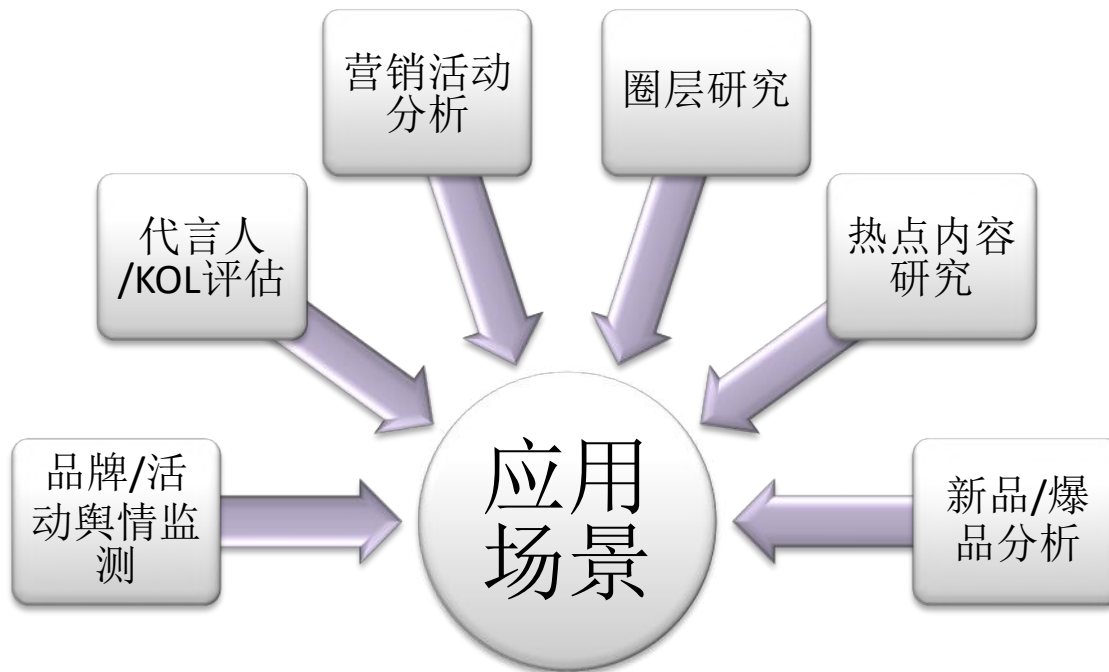
✓ 关键词/过滤词

✓ KOL

✓ 话题/标签

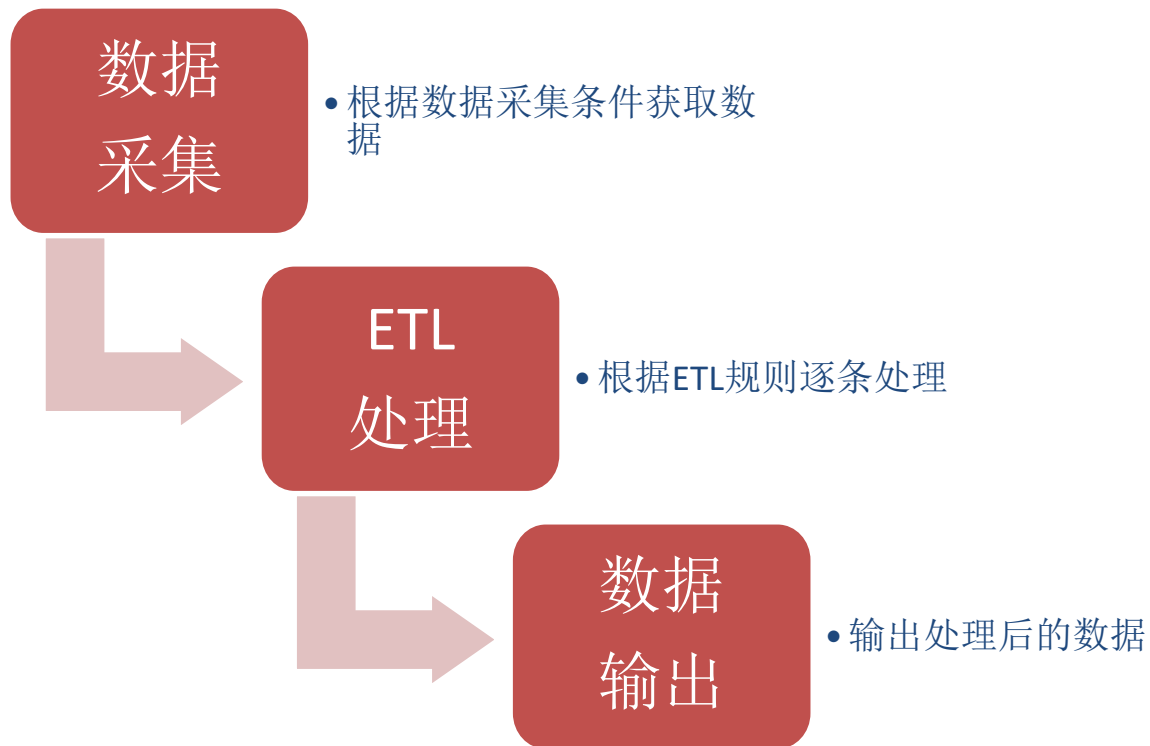
✓ 单帖/URL

✓ 店铺ID/品类URL



抖音、小红书、B站等平台使用标签采集数据，能剔除SEM搜索引擎优化，使采集数据更加精准，排除噪音影响

## 应用前沿的NLP处理技术构建ETL平台



- 标签抽取
- 分词提取
- 口碑提取
- 情感判断
- 知识图谱
- 广告杂音判断
- 新词发现
- 事件聚类



## 多种方式导出数据

### 数说立方

- 无缝衔接数说立方，自由构建分析模型和dashboard，文本分析更高效
- 支持多任务追加
- 千万级数据秒级查询响应
- 支持通过API获取数据
- .....

### ES集群

- 自定义集群，资源独立，稳定性更强
- 支持多任务追加
- 支持通过API获取数据
- .....

## 多种方式 导出数据

### FTP/SFTP

- 文件服务器一键配置，通过CSV文件推送数据
- 按目录组织项目
- 定期从服务器扫描文件，架构更安全
- 其他
- .....

### Excel/CSV

- 少量数据一键下载
- 其他
- 其他
- .....

## 千亿数据探索式实时分析



关键词: 口红

2020-05-17 00:00:00 ~ 2020-06-16 23:31:04

正面  中性  负面

新闻 46852 | 论坛 21184 | 微博 1161453 | 微信 304030 | 小红书 191805 | 短视频 22515 | 仪表盘 1747839

2019-06-12 00:00:00 ~ 2020-06-11 15:49:52 | 请选择情景 | 数据源 | 关键词: 雅诗兰黛 | 过滤词: 无 | 保存

### 声量趋势

月

平台	2019-06	2019-08	2019-10	2019-12	2020-02	2020-04	2020-06
总数	~50k	~100k	~200k	~150k	~100k	~50k	~20k
新闻	~5k	~10k	~20k	~15k	~10k	~5k	~2k
论坛	~5k	~10k	~20k	~15k	~10k	~5k	~2k
微博	~5k	~10k	~20k	~15k	~10k	~5k	~2k
微信	~5k	~10k	~20k	~15k	~10k	~5k	~2k
小红书	~5k	~10k	~20k	~15k	~10k	~5k	~2k
短视频	~5k	~10k	~20k	~15k	~10k	~5k	~2k

### 数据分布

情感分布

情感	占比
正面	27.4%
中性	67.82%
负面	4.68%

### 内容词云

普通 | 微博话题

肖战代言小鹿茶 肖战代言小鹿茶 肖战代言小鹿茶 肖战代言小鹿茶 肖战代言小鹿茶

肖战代言小鹿茶 肖战代言小鹿茶 肖战代言小鹿茶 肖战代言小鹿茶 肖战代言小鹿茶

肖战代言小鹿茶 肖战代言小鹿茶 肖战代言小鹿茶 肖战代言小鹿茶 肖战代言小鹿茶

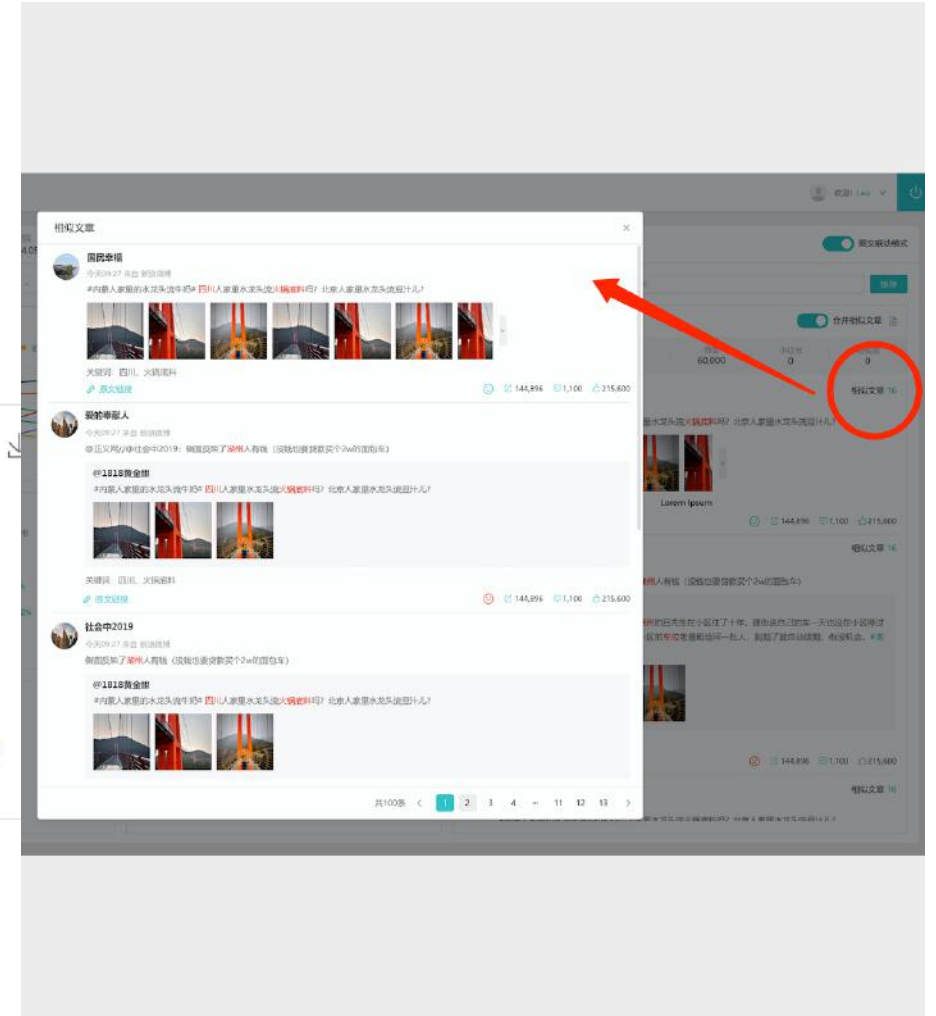
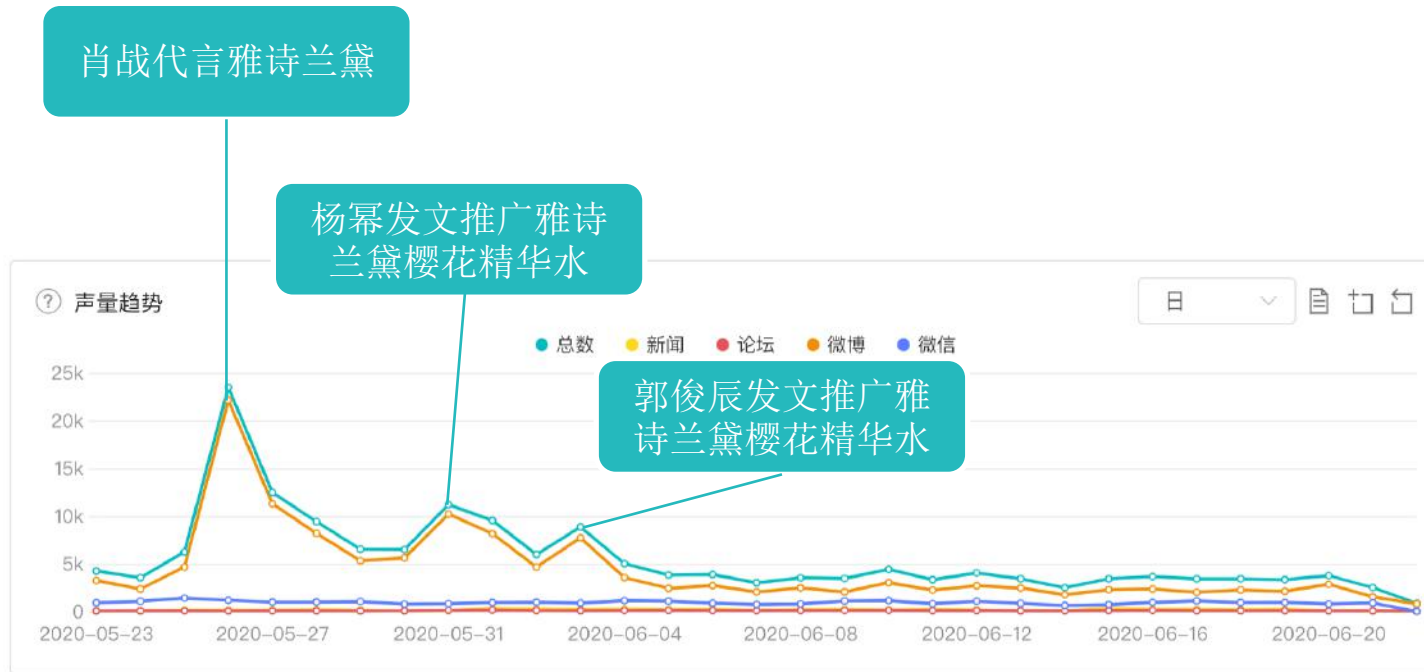
### 多维度下钻探索

原文内容

胖精精阿  
2020-06-11 14:40  
...瞬间觉得用完的CPB好瘪阿... 洗脸宗旨: 痘肌洗干净就好! 所以我挑的基本都是清...  
关键词: 乳白色, 山茶花, 洁面, 洗面奶, 香奈儿

sue  
2020-06-11 14:04  
50ml海蓝之谜精华1xxx就慢了  
50ml海蓝之谜精华1xxx就慢了 现在不真真的对不起自己的脸[笑哭R][捂脸R] !!!海蓝之...  
关键词: lamer, 海蓝之谜, 粉身碎骨, 精华, 雅诗兰黛

通过声量趋势峰值点，基于实时的聚类结果，快速挖掘峰值事件的驱动因素



注：搜索关键词“雅诗兰黛”

## 寻找最高互动帖子

	发布时间	转发数	评论数	点赞数	情感类型
@雅诗兰黛	2020-06-15 20:11	1,000,000	694	27,726	正面
@薇娅viyaaa	2020-06-15 20:11	49,796	43,089	29,235	正面
	发布时间	阅读数	在看数	发布位置	情感类型
@丁香医生	2020-05-20 12:15	10万+	781	次条	正面
@李佳琦Austin	2020-05-26 17:50	10万+	435	头条	正面



为了帮助大家了解自己的皮肤屏障状态，丁香医生携手雅诗兰黛共同发起《2020 皮肤屏障调研》，快点击阅读原文填写问卷，只需 2~3 分钟就能填完。



注：搜索关键词“雅诗兰黛”，使用“总互动量”排序（也可按照转发、评论、点赞单项指标排序）

获取搜索词的相关词时，可选择隐藏跟研究目的不相关词汇或搜索关键词本身；点击词云联动原文，快速聆听用户声音




注：搜索关键词“雅诗兰黛+白金粉底液”



## 分析市场核心价值

### 分钟级生成分析报告

- 数据预处理，人群指标定期更新即取即用
- 根据分析模型，迭代存储结构
- 独立服务器，保障资源提升效率

### 多种专业的分析模型

- CCI种草机
- 事件传播模型
- 智能口碑模型

### 前沿算法模型

- NLP自然语言处理
- 机器学习
- 多元统计分析

### 灵活贴心的功能特性

- 报告带解读，可灵活编辑
- 词云可编辑
- 报告可下载为多种文件形式

## 人群研究

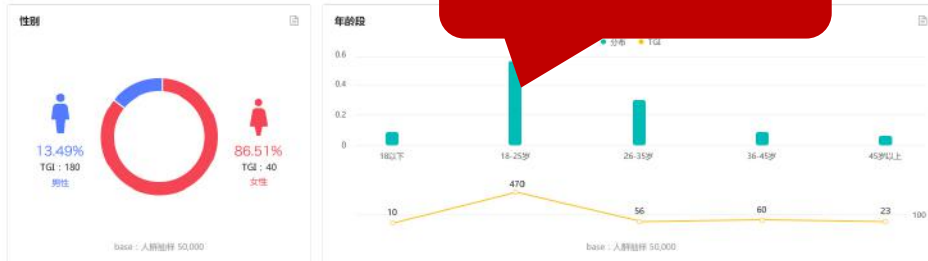
研究目的：全面深入了解圈层的人群属性、媒体偏好和内容偏好，为制定具有针对性的沟通策略提供数据基础，实现用圈层关注的内容选择合适的KOL进行触达，以增强品牌与目标人群之间的有效沟通及营销效果。



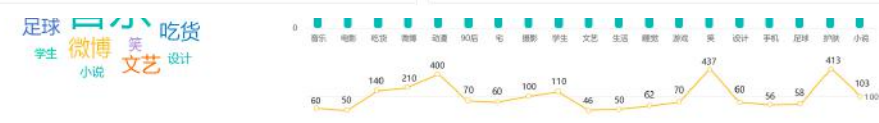
## 人群属性

该人群中(18-25岁女性)的占比最高。

多维度描绘人群特征



TA们集中在(一线城市, 二线城市), 其中(海外地区)的占比显著高于随机人群。



TA们较关注的话题是(“易烱千玺朋友请听好”, “易烱千玺阿玛尼男士护肤代言人”)



关注的明星、KOL

## 共同关注

根据他们的关注情况, 以下博主具有较大影响力。

明星/KOL	粉丝数	关注比例	TGI
Mr_Honor易烱千玺	84,513,538	8.14%	167
Mr_Honor易烱千玺	84,513,538	8.14%	167
Mr_Honor易烱千玺	84,513,538	8.14%	167
Mr_Honor易烱千玺	84,513,538	8.14%	167
Mr_Honor易烱千玺	84,513,538	8.14%	167
Mr_Honor易烱千玺	84,513,538	8.14%	167
Mr_Honor易烱千玺	84,513,538	8.14%	167
名字很长很长很长很长很长很长...	84,513,538	8.14%	167
名字很长很长很长很长很长很长很长...	84,513,538	8.14%	167



## 事件分析

研究目的：基于社交媒体数据对**舆情事件或营销事件**的演化和传播进行分析，感知用户情绪，挖掘事件发展的多个阶段和各个阶段中的观点或需求变化，为企业相关部门及时回应热点、评估传播效果等提供客观的大数据分析依据



通过关键词圈定多平台数据源

1

### 事件传播概况

声量趋势、互动量趋势、情感分布、阵地分布  
关键词词云、微博话题词云、微博话题阅读数排行

2

### 渠道分析

渠道热门KOL排行  
渠道热门帖子排行

3

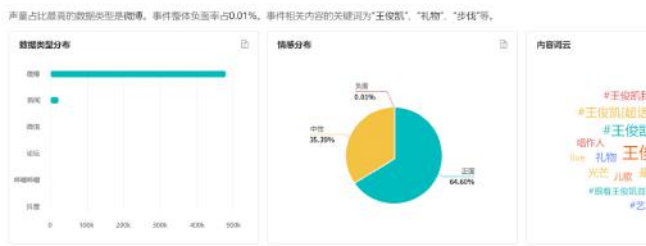
### 受众画像

声量趋势、互动量趋势、情感分布、阵地分布  
关键词词云、微博话题词云、微博话题阅读数排行

事件概况分析，洞察传播情况



渠道KOL和热帖挖掘



### 微博KOL

序号	作者	发帖数	总互动量	平均互动量	转发数	评论数	点赞数
1	TFBOYS-王俊凯	1	3,252,102	3,252,102	1,000,000	1,000,000	1,252,102
2	TFBOYS-王俊凯	28	1,079,832	3,857	1,066,092	4,618	8,121
3	TFBOYS-王俊凯	2	991,104	495,552	991,074	30	30
4	TFBOYS-王俊凯	3	245,943	81,981	245,939	3	3
5	TFBOYS-王俊凯	1	240,619	240,619	240,597	11	9
6	TFBOYS-王俊凯	1	138,014	138,014	98,777	21	17
7	TFBOYS-王俊凯	7	93,991	13,427	180,219	12	3
8	TFBOYS-王俊凯	1	80,238	80,238	95,852	11	3
9	TFBOYS-王俊凯	2	85,867	42,933	86,404	3	12
10	TFBOYS-王俊凯	1	86,409	86,409	86,115	3	3

### 热帖

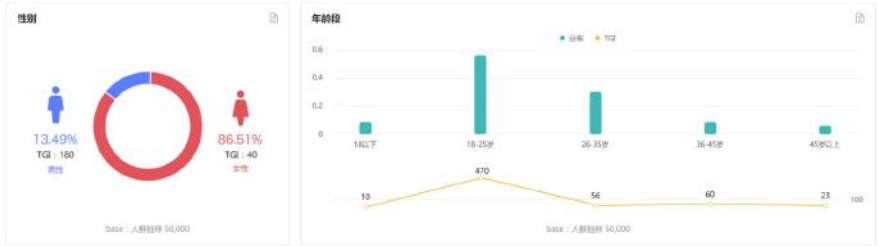
TFBOYS-王俊凯  
儿时听过的每一首歌，都是独特的成长记忆，陪伴我们向未来的方向出发。伊美诺儿童艺术中心活动「父亲节音乐 联合发起艺术行动」「儿童节歌」活动，作为陪伴儿童成长中的人，我们小朋友们准备了一份特别的礼物《闪亮》，希望更多人加入，用音乐点亮未来！

TFBOYS-王俊凯  
儿时听过的每一首歌，都是独特的成长记忆，陪伴我们向未来的方向出发。伊美诺儿童艺术中心活动「父亲节音乐 联合发起艺术行动」「儿童节歌」活动，作为陪伴儿童成长中的人，我们小朋友们准备了一份特别的礼物

### 受众画像

关注该事件的受众中，比例最高的性别是男性，年龄段是18-25岁。

受众画像分析



## 口碑分析

研究目的：基于电商网站数据，运用三元组口碑提取算法，分析产品口碑优劣势，最大化还原消费者的评价和反馈



通过商品URL获取商品评论

适用行业：彩妆、护肤、香水、电子、家电、食品饮料

1

### 购买体验

各个维度总体满意度分析、各个维度总体关注度分析  
每个一级维度中消费者最关注的二级维度，以及该维度下的正面、负面评价词云

2

### 产品体验

每个一级维度的总体满意度和关注度  
分析优先改进维度→关注度高、满意度低的维度  
每个一级维度中消费者最关注的二级维度，以及该维度下的正面、负面评价词云

## 购买体验分析

### 购买体验总览

购买体验总体满意度较低的站点是(京东, 满意度为(53%), 满意度较低的维度是(“卖家”。

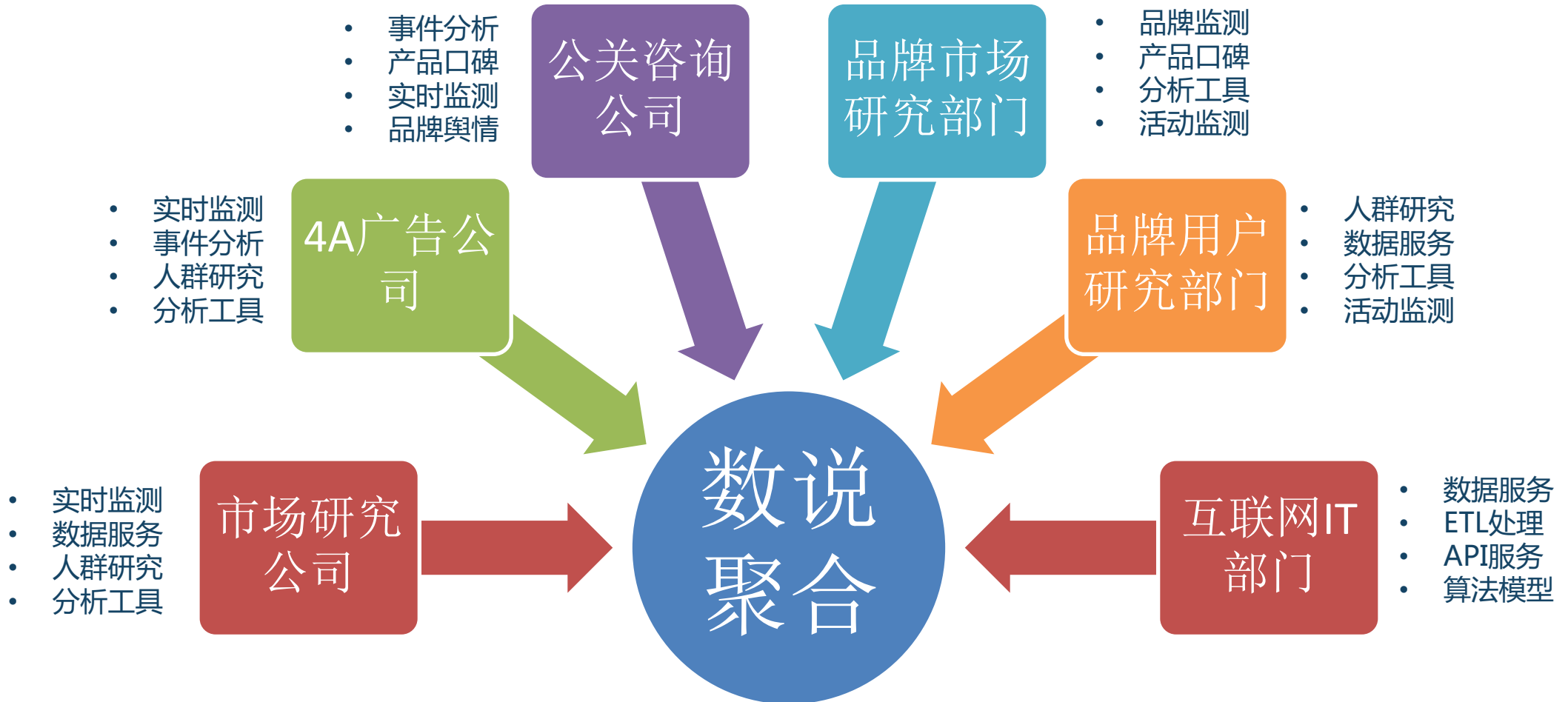


## 产品体验分析

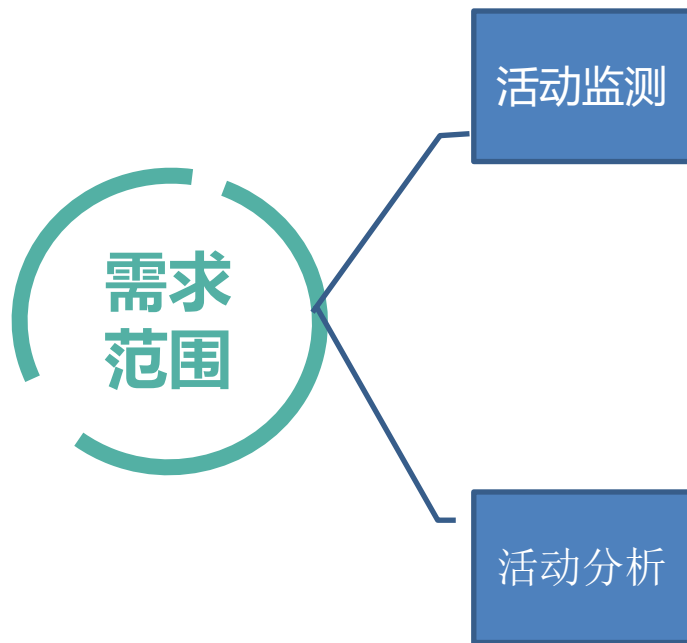
### 产品体验总览

产品体验中, 消费者最关注的是(“硬件”)和(“感官”), 这两个维度的满意度分别为(79%)和(94%)。





## 某互联网公司全年度营销活动效果评估



基于社交媒体数据监测，实时了解营销活动在各阵地  
的声量趋势、声量分布和热点讨论内容，快速把握市场  
表现，帮助品牌营销策略持续优化。

全面分析营销活动的整体传播情况、活动核心讨论内容  
以及细分内容分析，总结提炼活动效果。

通过全网监测，  
及时了解传播  
趋势及情感反  
馈，及时评估  
活动的传播情  
况和热门讨论  
内容；活动结  
束后通过分析  
报告进行深度  
复盘和总结，  
全面评估活动  
效果。

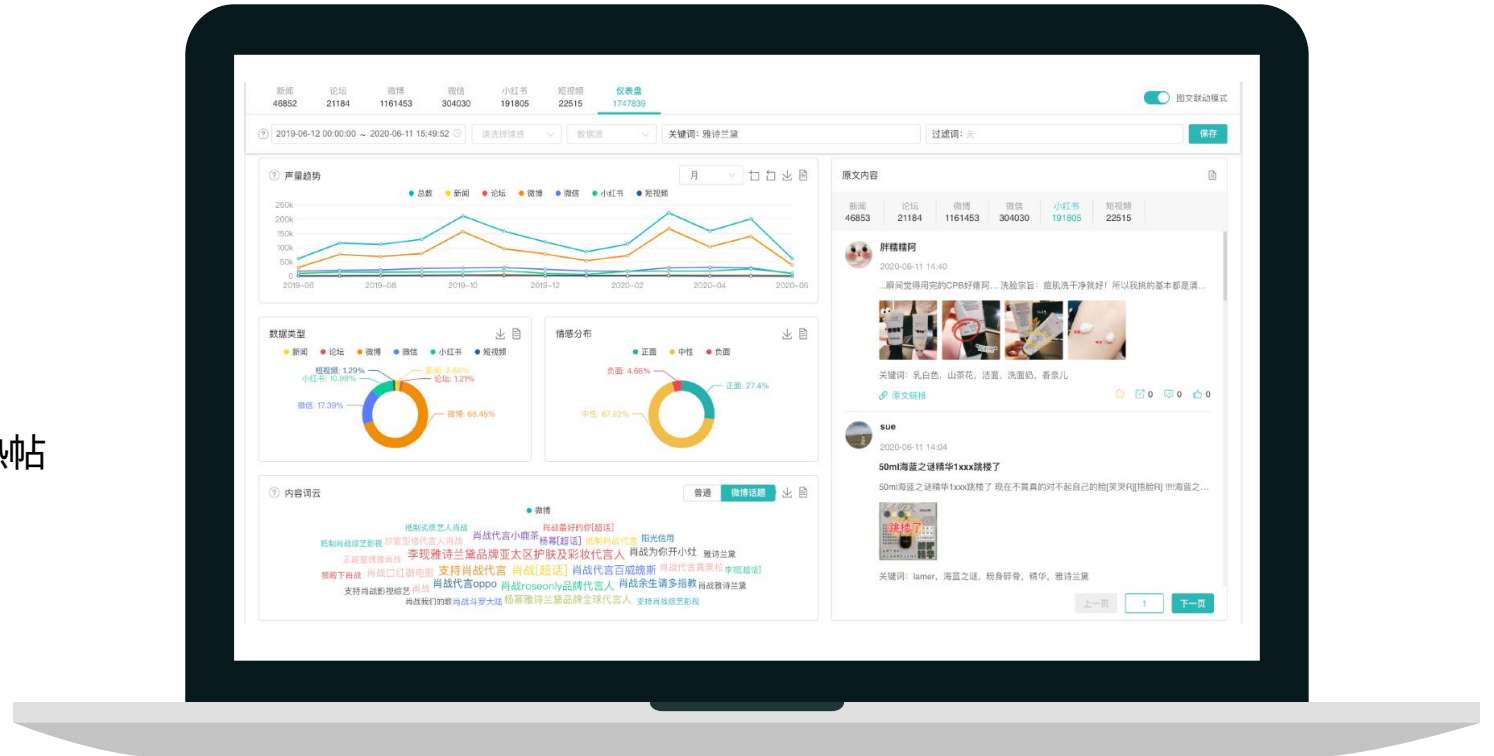
活动过程中使用实时仪表盘及时监测

## 实时监测

- 将活动词添加到采集监控中
- 通过查询功能，获取活动全网传播概况
  - ❑ 活动热度趋势
  - ❑ 活动主要传播阵地
  - ❑ 活动负面信息
  - ❑ 活动热词
  - ❑ 根据互动指标实时查看微博\微信热帖

## 根据数据调整活动策略

- 根据传播效果决定是否调整投放KOL
- 根据用户反馈，即时调整KOL内容投放策略



活动结束后使用分析市场复盘整体效果

## 事件分析

- 配置活动词、阵地、活动时间
- 通过分析报告，全方面评估活动效果
  - ❑ 传播概况，声量趋势、互动量趋势、词云、话题词、话题讨论热度
  - ❑ 渠道分析，获取各个渠道热门传播观点和关键KOL
  - ❑ 受众画像，年龄、性别、地区分布等

## 根据报告结果复盘活动效果

- 对比不同活动的传播效果
- 内容策略总结、KOL投放效果对比等
- 下一次活动策略参考

