

第一章

新媒体数据分析概述

引言

自 20 世纪 60 年代末互联网的前身阿帕网（ARPAnet）诞生以来，互联网伴随人类已经五十余年。随着全球互联网用户的不断增加，依托 Internet 形成的世界范围内的新媒体体系正在不断地扩大，由此带来的以亿计的受众数量，时时刻刻都在形成各种类型的新媒体数据，其数据规模是人类历史上前所未有的。对于这些新媒体数据我们要给予重视，从中挖掘价值并善加运用，以促进人类信息传播事业的长足发展。本章在勾勒新媒体数据现状的基础上阐述新媒体数据分析的意义，并就新媒体数据类别与来源进行说明，简要介绍新媒体数据分析基本步骤，对常见的新媒体数据分析工具及国内外新媒体数据机构进行介绍。

根据互联网世界统计（Internet World Stats）网站数据，截至 2022 年 1 月底全球约有网民 52.51 亿，普及率约为 66.2%。而在中国，2022 年 2 月 25 日，中国互联网络信息中心（CNNIC）发布的第 49 次《中国互联网络发展状况统计报告》显示，截至 2021 年 12 月底，我国网民规模达 10.32 亿，手机网民规模达 10.29 亿，互联网普及率达 73%。如今，网络就像水和空气一样，与我们密不可分——截至 2021 年 12 月，在网民中，即时通信、网络视频、短视频用户使用率分别为 97.5%、94.5%和 90.5%，用户规模分别达 10.07 亿、9.75 亿和 9.34 亿。在线办公、在线医疗用户规模分别达 4.69 亿和 2.98 亿，同比增长分别为 35.7%和 38.7%，成为用户规模增长最快的两类；网上外卖、网约车的用户规模增长率紧随其后，同比增长分别为 29.9%和 23.9%，用户规模分别达 5.44 亿和 4.53 亿……庞大的网民与丰富的应用构成了中国蓬勃发展的新媒体市场。无论国外还是国内，巨大的互联网用户基数产生了海量的新媒体数据，这些数据是当前全球新媒体事业赖以存在和发展的基础，互联网机构与用户生产并传播数据，数据繁荣并便利互联网机构与用户，形成了人类发展史上的信息奇观。数据无言，但自有价值的光辉，人类从新媒体大数据的丛林中汲取思想、提炼精髓，将让未来的新媒体事业的发展受益匪浅。

第一节 新媒体发展现状

1969 年，互联网前身——阿帕网诞生。1983 年，Internet 产生，Windows、IE 浏览器、资讯网站等随之出现。1990 年后，计算机、网络技术、通信技术升级，数字化浪潮席卷全球。2000 年后，互联网受众规模、资讯、应用、服务骤增，新媒体深度渗透到社会各个领域。

虽然当前大部分人每天都在使用互联网、应用新媒体，但是尚没有全世界公认的权威机构对全球的新媒体的现状进行统计，更谈不上对新媒体产业、受众、市场、数据进行系统的、有针对性的调查与分析。自从互联网作为新的信息传播形式正式登上人类媒介传播的舞台，人们对新媒体的调查研究就开始了，最早对互联网受众的调查可以被视为最早的新媒体数据的统计，这要从 1996 年美国 Jupiter Media 公司算起。Jupiter Media 是美国最早也是全球最早从事互联网用户访问率（ratings）研究测量的公司，1996 年 1 月，该公司发布了第 1 份美国互联网使用情况报告。进入 21 世纪后，Jupiter Media 不再引领互联网调查机构，而是转身进入其他领域。但人们对互联网发展状况的统计、对新媒体数据的挖掘才刚刚开始，如互联网世界统计（<https://www.internetworldstats.com/>）、国际电信联盟（ITU）、AC 尼尔森、Jupiter Media、NetValue、Alexa、GfK（<http://www.gfk.com/>）、IDC（<https://www.idc.com/>）、玛丽·米克尔（Mary Meeker）的互联网趋势报告、皮尤报告（Pew Research Center）等，足以让我们在同一时段从多个角度对全球新媒体的全貌有所了解。

1987 年，中国发出第一封 E-mail，迈进了互联网大门。2000 年至 2005 年四大门户网站崛起，2005 年至 2012 年博客社交媒体社区视频社交电商矩阵形成，2012 年至 2017 年 BATJ（百度、阿里巴巴、腾讯、京东）崛起，2018 年后互联网进入“下半场”，TMD（今日头条、美团、滴滴）成为新媒体行业冉冉升起的新巨头。2020 年，中国互联网发展已三十余年。国内进行互联网及新媒体受众和产业调查的机构主要有中国互联网络信息中心（CNNIC）、艾瑞咨询集团（iResearch）、易观国际、中文互联网数据研究资讯中心网（199IT）等。中国互联网络信息中心（China Internet Network Information Center，CNNIC）是经国家主管部门批准，于 1997 年 6 月 3 日组建的互联网信息管理和服务机构。从 1998 年起，中国互联网络信息中心于每年 1 月和 7 月左右发布《中国互联网络发展状况统计报告》。统计报告发布后，受到各个方面的重视，被国内外广泛引用。

一、全球新媒体发展现状

由于全球新媒体市场尚没有权威的数据统计与分析，根据学界、业界通常使用的数据来源，可把互联网世界统计网站的数据作为公认的可信数据进行使用。根据互联网世界统计网站数据，截至 2022 年 1 月底全球约有网民 52.51 亿，普及率约为 66.2%，其中亚洲网民约为 27.9 亿、非洲网民约为 6.01 亿，如表 1-1 所示。

表 1-1 截至 2022 年 1 月 31 日全球各大洲网民总数及普及率

世界地区	人口 (2022 年估计)	世界人口 占比/%	互联网用户 2021 年 12 月 31 日	渗透率/%	2000—2022 年 增长/%	互联网世 界占比/%
非洲	1 394 588 547	17.60	601 327 461	43.10	13220	11.50
亚洲	4 350 826 899	54.80	2 790 150 527	64.10	2341	53.10
欧洲	841 319 704	10.60	743 602 636	88.40	608	14.20
拉丁美洲/ 加勒比	663 520 324	8.40	533 171 730	80.40	2851	10.10
北美	372 555 585	4.70	347 916 694	93.40	222	6.60
中东	268 302 801	3.40	205 019 130	76.40	6141	3.90
大洋洲/ 澳大利亚	43 602 955	0.50	30 549 185	70.10	301	0.60
世界总数	7 934 716 815	100.00	5 251 737 363	66.20	1355	100.00

注：（1）互联网使用情况和世界人口统计估计时间为 2022 年 1 月 31 日。（2）人口数字基于联合国人口司的数据。（3）互联网使用信息来自 Nielsen Online、国际电信联盟、GfK、当地 ICT（信息与通信技术）监管机构和其他可靠来源发布的数据。

在互联网新媒体的发源地美国，Facebook、YouTube、Twitter 等成为全球新媒体重要平台，其影响力远超早期的网站类新媒体产品，受到世界各国网民和政府的广泛关注。马克·扎克伯格等在 2004 年 2 月创建 Facebook，截至 2021 年 3 月底 Facebook 用户达 28.03 亿，渗透率约为 35.6%，全球用户每月共享的信息量约为 300 亿条，70% 用户为美国以外的用户。Google 母公司 Alphabet 控股的 YouTube 每月登录用户超 20 亿，覆盖全球 91 个地区，95% 的用户为海外互联网用户，用户每日视频观看时长超 10 亿小时，有 80 种语言版本的视频内容。英国、美国、日本、印度等一些重要的市场都已成为品牌通过新媒体平台触达目标受众的重要渠道。以美国为例，仅 YouTube 移动端，可触达人群比任何美国电视网还多。创建于 2006 年的 Twitter 通过广播式的即时信息流掀起微博客革命，Twitter 在 2010 年全年发送信息多达 250 亿条，用户约为 6 亿，截至 2020 年第一季度，Twitter 的可货币化日活跃用户达 1.66 亿，成为名副其实的全球新媒体。同时，在全球最贫穷的 48 个国家，90% 的人口尚未接入互联网，在全球大约 7100 种语言中，仅 10 种主要语言被互联网广泛使用。

二、中国新媒体发展现状

在中国，互联网新媒体市场数据统计与分析机构较多，其中，CNNIC 统计分析的结果权威性最高，被引用最多。2022 年 2 月 25 日，CNNIC 发布的第 49 次《中国互联网络发展状况统计报告》显示，截至 2021 年 12 月，我国域名总数达 3593 万个，IPv6 地址数量达 63 052 块，移动通信网络 IPv6 流量占比已经达到 35.15%，建成并开通 5G 基站数达 142.5 万个，全年新增 5G 基站数达 65.4 万个。新闻网门户网站，微博、微信、新闻 App，新兴

的移动媒体、自媒体等与庞大的网民结合，构成了中国蓬勃发展的新媒体市场，产生了数以亿计的新媒体数据。

其实，第 47 次《中国互联网络发展状况统计报告》已经显示我国互联网行业在 2020 年抗击新型冠状病毒肺炎（以下简称新冠肺炎）疫情和疫情常态化防控等方面发挥了积极作用，为我国成为全球唯一实现经济正增长的主要经济体、国内生产总值（GDP）首度突破百万亿元、圆满完成脱贫攻坚任务做出了重要贡献。

“健康码”助 9 亿人畅通出行，互联网为抗“疫”赋能赋智。2020 年，面对突如其来的新冠肺炎疫情，互联网显示出强大力量，对打赢疫情防控阻击战起到关键作用。疫情期间，全国一体化政务服务平台推出“防疫健康码”，累计申领近 9 亿人，使用次数超过 400 亿人次，支撑全国绝大部分地区实现“一码通行”，大数据在疫情防控和复工复产中作用凸显。同时，各大在线教育平台面向学生群体推出各类免费直播课程，方便学生居家学习，用户规模迅速增长。受疫情影响，网民对在线医疗的需求量不断增长，进一步推动了我国医疗行业的数字化转型。互联网将在促进经济复苏、保障社会运行、推动国际抗疫合作等方面进一步发挥重要作用。

网络扶贫成效显著。截至 2021 年 12 月，我国农村网民规模已达 2.84 亿，农村地区互联网普及率为 57.6%。近年来，网络扶贫行动向纵深发展取得实质性进展，并带动边远贫困地区非网民加速转化。老年群体加速融入网络社会。得益于互联网应用适老化改造行动持续推进，老年群体连网、上网、用网的需求活力进一步激发。截至 2021 年 12 月，我国 60 岁及以上老年网民规模达 1.19 亿，互联网普及率达 43.2%。老年群体与其他年龄群体共享信息化发展成果，能独立完成出示健康码/行程卡、购买生活用品和查找信息等网络活动的老年网民比例已分别达 69.7%、52.1%和 46.2%。

数字政府建设加速推进。大数据、云计算、移动互联网、物联网、人工智能等新技术不断应用到政府治理过程，全国政府数字化转型工作稳步开展，政务服务平台、监督平台和全国信用体系已经形成。数字政府建设成为实现政务数字化转型、驱动经济社会高质量创新发展、推进国家治理体系和治理能力现代化的关键抓手。浙江“最多跑一次”改革、江苏“不见面审批（服务）”改革、上海“一网通办”改革等发展模式，有效解决了群众和企业办事难、办事慢、办事烦等问题，驱动互联网政务服务由点到面、由浅及深加快发展。

一体化政务服务能力显著提升成为我国现阶段数字政府建设的典型特征。我国各省市积极探索，通过构建普惠均等、便民高效、智能精准的政务服务“一张网”，以“我为群众办实事”为初心，聚焦网上办事的堵点难点和“急难愁盼”问题，着力在提升服务水平和服务效能上下功夫，努力书写便民利企的民生答卷。全国一体化政务服务平台实名用户超过 10 亿人，其中国家政务服务平台注册用户超过 4 亿人，总使用量达 368.2 亿人次，为地方部门提供身份认证核验服务 29 亿余次，群众满意度、获得感不断提升。

在国内新媒体发展中，除了国家党政相关部门及各级机构类官方媒体主管、主办的新闻网站、政务新媒体平台、互联网公共服务平台等之外，以商业机构为主开设的商业门户网站和社交类新媒体平台同样具有广泛的吸引力和社会影响力。当前，在商业机构开设的新媒体平台中，最具影响力的有新浪微博、腾讯微信、今日头条、哔哩哔哩（即 B 站）、快手、小红书等，其无论是受众数量还是市场占有率，抑或是社会影响力，均为当前国内

商业机构的新媒体产品之翘楚。

2009年8月,新浪推出“新浪微博”。2011年10月,中国微博用户总数达到2.498亿,居世界第一。截至2016年第一季度末,新浪微博月活跃用户达到2.61亿,日活跃用户达到1.2亿。截至2017年第一季度末,新浪微博月活跃用户达3.4亿,已超过Twitter,成为全球用户规模最大的独立社交媒体公司。2018年6月,新浪微博月活跃用户为4.31亿;2019年6月,新浪微博月活跃用户数增至4.86亿;2021年第二季度,新浪微博月活跃用户数达到5.66亿;2022年9月,新浪微博月活跃用户数为5.84亿。

腾讯公司于2011年年初推出微信。2012年4月,微信国际化,面向欧美推出4.0英文版WeChat,之后支持多种语言。2013年1月15日,用户数突破3亿,成为全球下载量和用户量最多的通信软件,影响力遍及海峡两岸及港澳、东南亚、欧美、拉非等地区,覆盖200多个国家,支持20多种语言。2014年1月,微信用户数达到5亿左右。2014年7月活跃用户约为4亿。2016年6月,微信活跃用户已达到5.49亿。2017年6月,微信活跃用户达到9.63亿。2018年6月,微信活跃用户达到10.58亿,同期Whats App活跃用户为15亿,Facebook Messenger活跃用户达13亿。2019年3月,微信活跃用户已达到11.12亿;2021年第一季度,微信及WeChat的合并月活跃用户达12.41亿;2022年第3季度,微信及Wechat的合并月活跃用户为13.09亿。

今日头条于2012年3月创建,2012年8月发布第一个版本,它凭借算法推荐从网易、腾讯等传统商业门户网站的资讯战中脱颖而出,3个月注册用户从0增加到1000万。目前,其新媒体产品有面向国内市场的头条号、抖音、火山小视频、西瓜视频等和面向国际市场的Musical.ly、TikTok等。除角力国内新媒体市场外,今日头条不断拓展全球新媒体市场,一边是不断推出头条产品海外版,包括今日头条海外版TopBuzz、TopBuzz Video,火山小视频海外版Hypstar和抖音短视频海外版Tik Tok等;一边是不断收购其他国家的新媒体应用,2016年10月投资印度最大内容聚合平台Dailyhunt,2016年年底控股印尼新闻推荐阅读平台BaBe,2017年2月全资收购美国短视频应用Flipagram,2017年11月收购全球移动新闻服务运营商News Republic、音乐视频分享和互动社交应用Musical.ly,到2017年年底,今日头条在日本、印度、东南亚、巴西和北美都取得了不错的进展,多款产品长时间占据当地App Store/Google Play总榜前三。据QuestMobile公布的TOP榜单,截至2022年6月,头条App月活用户3.44亿,是唯一一款进入TOP20的新闻资讯类App。

哔哩哔哩(即bilibili)是中国年轻世代高度聚集的文化社区和视频网站,该网站于2009年6月26日创建,被网民们亲切地称为“B站”,哔哩哔哩拥有动画、番剧、国创、音乐、舞蹈、游戏、知识、生活、娱乐、鬼畜、时尚、放映厅等15个内容分区,生活、娱乐、游戏、动漫、科技是B站主要的内容品类,并开设直播、游戏中心、周边等业务板块。B站早期是一个ACG(动画、漫画、游戏)内容创作与分享的视频网站。经过十多年的发展,围绕用户、创作者和内容,构建了一个源源不断产生优质内容的生态系统,B站已经涵盖七千多个兴趣圈层的多元文化社区,曾获得QuestMobile研究院评选的“Z世代偏爱App”和“Z世代偏爱泛娱乐App”两项榜单第一名并入选“BrandZ”报告2019最具价值中国品牌100强。截至2020年第二季度,B站月均活跃用户达1.72亿,移动端月活用户达1.53亿;18至35岁用户占比达78%。截至2020年第三季度,B站月均活跃用户已达到1.97亿,

并在8月首次突破了2亿。

快手是北京快手科技有限公司旗下的产品。快手的前身叫“GIF快手”，诞生于2011年3月，最初是一款用来制作、分享GIF图片的手机应用。2012年11月，快手从纯粹的工具应用转型为短视频社区，成为用户记录和分享生产、生活的平台。随着智能手机、平板电脑的普及和移动流量资费的下降，快手在2015年以后迎来较大的发展。根据Quest Mobile发布的《2022中国移动互联网报告》的统计，2022年，快手月活跃用户达3.9亿。快手于2023年2月22日召开的“2023·增量效应”磁力大会上披露，截止到2022年第三季度，快手平均日活跃用户突破3.63亿。

小红书于2013年6月在上海创立。小红的用户可以通过短视频、图文等形式记录生活点滴，分享生活方式，并基于兴趣形成互动。2015年6月，小红书用户达到1500万；2018年10月，用户突破1.5亿。截至2019年7月，小红的用户数已超过3亿。2019年11月，小红书宣布推出创作者123计划，推出品牌合作平台、好物推荐平台和互动直播平台，从创作者中心、活动和产品三方面帮助创作者。2022年12月22日，在小红书与中国社会科学院社会学研究所、DT研究院联合发布“2023年度生活趋势”报告中披露，小红书月活跃用户达到了2亿。

第二节 新媒体数据分析的意义

新媒体数据分析为各级各类机构提供了纷繁多样的信息分析和信息洞察的可能性。新媒体数据分析可使数据产生价值、促进运营、改善和产生新的机会，使新媒体能够在各个领域提供潜在的收益，从用户的个性化方案，到降低风险、内部运营分析，以及营销传播优化，等等，新媒体数据中隐含的价值能使机构在应用新媒体时所产生的价值发挥到最大。因此，新媒体数据分析对于商业网站新媒体、新兴自媒体、新闻网站、政务新媒体等均有重大意义。对于应用最为广泛的商业新媒体与自媒体，以及大型企业的新媒体营销传播部门来说，其意义在于商业营销及品牌传播的视角，主要有以下七个方面。

(1) 了解运营质量。商业新媒体、政务新媒体、新兴自媒体的新媒体数据分析的第一大意义是了解运营质量。新媒体运营的日常工作包括内容更新、微信公众号推广、微博发布、今日头条推送、朋友圈推送、视频推广、直播分享、粉丝维护、社群运营、微店运营、线上线下活动策划与组织等。这些工作是否有价值、是否能够有效实现传播与营销目标，需要通过数据来了解与判断。对于新媒体运营质量数据，商业新媒体、政务新媒体、新兴自媒体等不同的平台关注点不同。目前，大部分新媒体平台关注的运营数据包括网站流量数据、微信公众号粉丝数据、微博阅读数据、今日头条内容数据、活动转发与评论数据等。

(2) 研判运营方向。新媒体数据分析的第二大意义是研判运营方向。现阶段，百度、腾讯、阿里、今日头条等大型互联网公司已经开放了大量的数据，运营者可以直接登录相关网站进行查看。分析用户数据有助于判断新媒体内容、活动、推广是否要和网络热点结合。常见的新媒体相关大数据包括百度指数、新浪微博指数、微信指数、头条指数等。

(3) 控制运营成本。新媒体数据分析的第三大意义是控制运营成本。新媒体传播与营

销一方面需要关注影响力扩展、销售额的增长及品牌价值的提升,另一方面也需要时刻关注运营成本,尤其是传播成本。目前国内网民上亿,如果机构的新媒体信息投放没有精准的方向,则极有可能使信息传播效果被浪费。因此,新媒体团队需要分析用户的分布城市、购买或阅读时间、常用 App、惯用访问终端等数据,每次信息投放前要综合近期的投放情况进行调整与优化,以控制成本。要通过完善的用户画像体系实现千人千面的运营,结合兴趣标签、用户使用路径、用户画像等维度进行分析,使每个用户看到的内容都不相同。用户画像系统的建立不能单纯依靠自身产品的用户分析,可以借助第三方数据服务商的数据能力,打破数据壁垒,充实用户画像的维度。此外,对于大型企业来说,能够抽取特征数据进行自主建模,但对于中小型企业而言,集成第三方数据服务是较多的选择,如个推的用户画像 SDK,可以提供年龄段、性别、爱好等数百种标签,帮助 App 构建全方位立体的用户模型。App 基于对用户画像的理解,可对不同类型、不同场景的用户推荐不同的内容和服务,实现千人千面的运营模式。

(4) 评估营销传播方案。新媒体数据分析的第四大意义是评估营销传播方案。营销传播方案只是新媒体团队根据以往经验制定的工作规划,但在制定一段时间后,需要通过数据进行评估。一方面,分析最终完成数据,可以反推方案中目标的可行性;另一方面,分析过程数据,可以及时发现方案制定后在执行过程中遇到的问题,作为下次营销传播方案制订的参考。评估营销传播方案常用到的数据包括目标达成率、最终销售额、过程异常数据、失误率等。从整体营销目的来看,企业新媒体营销目的可以分为两大类:提升销量和宣传品牌。首先,为了提升销量,新媒体团队需要通过互联网渠道对企业产品进行推广,引导网民在线下单,或者引导网民在线预约后进行线下消费,从而提升企业的销售业绩。其次,为了宣传品牌,新媒体团队需要借助网民的传播力量,让更多人通过互联网接触企业信息、了解企业品牌、对企业产品进行评价。其中,宣传品牌这一营销目的又可以继续细分,包括提升品牌美誉度、提升品牌知名度和提升品牌忠诚度。因此,新媒体营销的目的可以细分为四个类别,包括提升销量、提升品牌美誉度、提升品牌知名度、提升品牌忠诚度。不同的营销目的需要挖掘与分析的数据不同,因此需要根据不同的营销目的进行不同的营销数据组合设计,以便于后续的数据分析与总结。企业新媒体销售通常来自不同的销售平台,包括淘宝店、天猫店、京东店、微店、独立网站等。因此,为了借助数据分析评估销售计划或分析销售结果,必须围绕用户购买或消费的行为进行逐层分析,需要分析的数据包括页面浏览量、用户访问时长、用户浏览页面数、店铺/网站转化率等。

(5) 提升机构品牌美誉度。机构品牌在互联网的美誉度指的是网上有粉丝或顾客对企业进行友好的评价,好评越多或评价内容质量越高,则美誉度越高。因此,为了利用数据分析企业美誉度的提升效果,需要围绕口碑来展开,应该分析的数据包括百度口碑、大众点评星级、网店评价等。由于一部分美誉度数据无法直接统计,需要运营者进行人工分析,以确保数据的准确性。

(6) 提升机构品牌知名度。在网上,机构品牌的知名度与名气相关,关注企业公众号的人越多、阅读相关信息的人越多,企业的知名度就越高。因此,在借助数据分析、评估企业知名度提升效果时,需要进行挖掘与分析的数据包括微博粉丝数、微信用户数、今日头条粉丝数、喜马拉雅订阅数等。

(7) 提升机构品牌忠诚度。每个网民都会关注大量的微信公众号、微博账号、头条号、抖音号等,但未必会对每个账号都忠诚,因此,粉丝数或订阅数只能作为品牌知名度的考量因素,而用户对品牌做出的响应,才能真正体现其对企业品牌的忠诚度。为了评估用户对品牌的忠诚度,可以统计并分析的数据包括二次购买的顾客数、主动转发的粉丝数、老客户访问比例、主动打赏的粉丝数、留言频次高的用户数等。

对新闻网站和政务新媒体的新媒体数据分析的意义主要从新媒体平台的公共传播效果提升视角进行探究。对于新闻网站来说,作为社会主要公共传播平台,其新媒体数据分析的意义在于,借助数据分析了解新闻内容生产及互联网传播的基本效果,了解新闻受众对本新闻网站的阅读和使用行为,了解本网站网络广告的投放效果,并且更进一步从微观角度了解新闻网站作为国家宣传部门的主要执行机构,在主流声音的宣传与传播方面的基本效果。对政务新媒体而言,作为党政机构主要媒体辅助支撑平台,其新媒体数据分析的意义在于了解其在新媒体场域中的内容生产与传播效果,查看其是否达到服务机构职能转变,是否成为电子政务与主流声音传播的有效渠道等,同时也可以通过一段时间的内容生产与传播策略数据分析,进行传播观念与策略的提升,等等。

第三节 新媒体数据类别与来源

在新媒体运营的工作中每个平台都会产生大量的数据,但因为平台不同,数据展示形式也不一样,因此统计方式和分析方式也会有一定的差别,所以要想快速了解新媒体数据分析方法,就必须先了解和掌握常见的新媒体数据类别,这样在日后的工作中才可以有针对性地收集相关类别的数据并进行分析。

新媒体数据一般包括三种展现形式,分别是数值型、图文型、关系流图型。数值型数据主要由数字组成,可以很直观地通过数字进行对应的统计和分析,可以很好地总结并评估运营的效果。常见的数值型数据包括我们常说的粉丝量、阅读量、评论量、转发量、网店的销售数据、网站的浏览数据等,以及各种线上线下活动的参与统计数据。图文型数据是由数字和图片组合而成的。图文型数据一般是指网站栏目分类、账号粉丝分类、消费者反馈、各种平台矩阵分布等。对于图文型数据,我们了解它的目的并非是制定考核指标的量化结果,而是这类数据可以帮助我们找到正确的运营方向。关系流图型一般是移动端新媒体平台的数据类型,平台对数据以 H5 或微视频的形式,采用多维的数据关系及数据流图的形式,进行数据呈现与展示。

目前国内新媒体平台有很多,一般来说有新闻网站与商业网站、微信公众号、微博、今日头条、西瓜视频、火山小视频、大鱼号、搜狐号、网易号、特定主题 App 等。因此,我们要根据不同新媒体平台的特性了解不同的数据来源。

(1) 新闻网站或商业网站数据。这一类数据来源可以通过网站后台导出查阅,也可以自行在网站后台添加代码,借助第三方自行记录数据。该类数据一般是 PV、UV、IP 等。PV (page view) 即页面浏览量,通常是衡量一个网络新闻频道或网站甚至一条网络新闻的主要指标。监测网站 PV 的变化趋势并分析其变化原因是很多站长定期要做的工作。page

view 中的 page 一般是指普通的 html 网页, 也包含 php、jsp 等动态产生的 html 内容。来自浏览器的一次 html 内容请求会被看作一个 PV, 逐渐累积成为 PV 总数。UV (unique visitor) 是指通过互联网访问、浏览这个网页的自然人。通常, 我们可以用两个数值标准来统计访问某网站的访客, 即独立访客数和访问次数。独立访客数和访问次数是两个不同的概念。独立访客数相当于带身份证参观展览会的访问人数, 每一个出示身份证参观展览的人, 无论出入几次, 都只计作一次独立访问。这里所说的“身份证”, 在网络上就是访客的 IP 地址或 cookie。通俗讲就是“带身份证参观展览会的访问人数”。用户每打开一个网站页面就被记录 1 次。用户多次打开同一页面, 浏览量值累加。所以, 访问次数是访客实际浏览过的网页数的总和, 而不是某网站中的网页数的总和。IP 是指独立用户/独立访客, 指访问某个站点或点击某条新闻的不同 IP 地址的人数。独立 IP 只记录第一次进入网站的具有独立 IP 的访问者。独立 IP 访问者提供了一定时间内不同访客数量的统计指标, 而没有反映出网站的全面活动。

(2) 微信朋友圈数据。这一类数据来源多是通过个人统计或清博大数据等工具预估出来的, 该类数据一般是“朋友圈+社群”运营的结果, 具体的数据来源主要有好友增长数(吸粉加好友)、朋友圈发布信息的点赞评论数以及通过朋友圈购买产品的销售数等。

(3) 微信公众号数据。微信公众号数据本身自有统计平台, 运营者可以很直观地了解当天的运营情况, 包括新增关注数、取消关注数、单篇文章阅读量、全部图文阅读量, 甚至还可以选择时间阶段进行统计, 如最近 30 天的数据统计。

(4) 新浪微博数据。无论是个人还是企业都可以在微博后台查看详细的微博运营数据, 微博常见的数据有阅读量、浏览量、视频播放量、粉丝来源以及粉丝新增及取关数等。

(5) 今日头条数据。打开头条指数的新版界面, 从巨量算数 (<https://trendinsight.oceanengine.com/arithmetic-index>) 可以看出分别有算数指数、算数榜单、行业指南、创作指南、洞察报告等。算数指数有关键词和话题两类, 分别从抖音或头条指数、相关或关联分析、人群画像三个方面进行具体呈现。算数榜单呈现不同类型的品牌在抖音生态下的品牌声量, 品牌指数主要由内容、传播、搜索三个部分组成。行业指南展示各个行业在头条体系中的创意趋势、平台热点趋势, 助力内容创作与运营投放。创作指南则更细致地展示各垂直领域的内容的创意、生产、消费的情况。洞察报告则重点展示部分经深度分析的报告。

(6) 百度指数数据。百度指数 (Baidu index) 是以百度海量网民行为数据为基础的数据分析平台, 是当前互联网乃至整个数据时代最重要的统计分析平台之一, 自发布之日便成为众多企业营销决策的重要依据。百度指数能够告诉用户某个关键词在百度的搜索规模有多大、一段时间内的涨跌态势以及相关的新闻舆论变化, 关注这些词的网民是什么样的、分布在哪里、同时还搜了哪些相关的词, 从而帮助用户优化数字营销活动方案。百度指数的主要功能模块有: 基于单个词的趋势研究 (包含整体趋势、PC 趋势和移动趋势)、需求图谱、舆情管家、人群画像; 基于行业的整体趋势、地域分布、人群属性、搜索时间特征。

(7) 西瓜指数数据。可以对微博和微信公众号进行统计分析, 分别从广告价值指数、周期内公众号发文次数、发文篇数、总阅读量、平均阅读量、平均在看数、平均留言数等方面进行数据展示。

随着新媒体机构的指数矩阵形成, 百度、微信、微博、头条、西瓜、清博等新媒体指

数将会组成新媒体数据的主要来源，分别从用户主动搜索量、用户在社交媒体的讨论量、用户点击新闻量三个角度，成为所有新媒体平台传播营销及品牌推广的重要衡量标准。除了以上七类数据来源，还有很多新媒体平台自身具备后台统计功能，运营者可以很直观地看到运营状况，根据运营现状及时调整运营策略。

以上是对某一类新媒体平台借助自身或第三方平台进行的数据索取，一般来说数据比较容易获取，但也有研究机构有对于海外新媒体或基于特定主题的非单一新媒体平台的数据获取的需求，在这种情况下就需要借助网络爬虫（web crawler）工具与 Python 程序语言等进行新媒体数据采集。同时，国外新媒体平台的数据也有和国内类似的情况，有开放数据也有不开放数据，开放数据可自动获取，对于不开放数据同样需要借助上述的程序语言进行采集。

第四节 新媒体数据分析与应用基本流程

新媒体行业每天都有大量数据产生，任何一个新媒体机构所在的平台在一段时间之后也会产生大量的数据，如 PV、粉丝数、阅读量、下载量等。对这些数据进行采集、处理、分析、呈现、形成报告、进行各层面应用等恰当化的数据分析与应用，会最大化地激发数据的价值，促进机构新媒体运营效率及营销传播效率的提高。

一般来说，在进行新媒体平台数据分析应用之前一定要设定目的，即数据分析是为了什么。商业网站新媒体、新兴自媒体、新闻网站、政务新媒体设立的初衷各不相同。商业网站新媒体重视营销自我内容生产传播的广度与停留时间，新兴自媒体重视内容被平台选拔推广与变现，新闻网站重视内容生产的公共传播效果，政务新媒体重视数据对新媒体政务功能的体现，淘宝网店重视浏览量与购买量之间的促进度，个人自媒体平台重视内容生产与粉丝成长的互动关系，直播带货类自媒体重视实时信息传播效果与商品销售额提高的直接对应关系。企业在选择新媒体营销传播平台时更多考虑平台对企业品牌的传播与产品销售的提高。另外，即使是同一新媒体机构或平台，在不同时期对新媒体运营的需求也大不相同，因此，要根据本次新媒体数据分析的目的，结合需求对相应平台的新媒体数据进行分析与应用。在实际中，新媒体数据分析与应用的过程一般包括新媒体数据采集、新媒体数据处理、新媒体数据分析、新媒体数据呈现、新媒体数据报告撰写、新媒体数据分析基础应用与高级应用等环节。

新媒体数据的采集是指在掌握新媒体数据采集的指导方法、新媒体数据采集的基本原理的基础上对不同类型新媒体平台的数据采集方法和途径的实践。新媒体数据采集环节要根据数据分析应用的目的制订采集计划，根据目标新媒体平台的数据开源与否，采用恰当的采集方法，或直接导出文档，或通过网络爬虫抓取，对指定的数据进行符合要求的采集，最后采集的数据要以数字、图文、数据流图等系列表格文档的形式呈现出来。

新媒体数据的处理是指在掌握新媒体数据处理的基本原理、通用的新媒体数据处理方法的基础上，对不同类型新媒体平台的数据进行处理。新媒体数据处理环节要在目的的框架下，在采集到的数据的基础上，对数据进行清理，剔除无效数据，合并相关数据，组合形成新的数据表格，使新媒体数据产生有序化的结果，为数据分析做好基础性工作。

新媒体数据的分析是指在掌握新媒体数据分析的基本方法、熟悉常用的新媒体数据分析工具的基础上,对不同类型新媒体平台的数据分析进行的框架构建及实施,同时避免常见的新媒体数据分析误区。新媒体数据分析环节要在数据处理基础上进行,基于目的的需求设计分析框架,综合运用多种数据分析方法,对所采集处理的数据进行多维度、多层面的分析,使数据的价值初步显现出来。

新媒体数据的呈现是指在掌握新媒体数据呈现的基本原理与常见的新媒体数据呈现形式的基础上,进行不同类型新媒体数据的可视化处理。新媒体数据的呈现具体来说要对数据分析的结果进行视觉化转化,进行各种可视化图表的选择、设计和应用,对不同类型数据进行可视化处理,具体包括数据属性、视觉通道、图表选择、图表设计、可视化方法等的规定及设计工作,为新媒体数据报告的撰写打下基础。

新媒体数据报告的撰写是指进行日常新媒体数据报告的撰写、专项新媒体数据报告的撰写、行业新媒体数据报告的撰写等。在新媒体数据分析与应用工作中,新媒体数据报告的撰写是一个关键的步骤,要将分析过程与可视化结果写成一份通俗易懂的报告,对许多人来说,是新媒体数据分析工作的终点。新媒体数据报告作为调查报告的内容,有和调查报告相似的部分,同时也有其独特之处。要理解新媒体数据报告的分类和撰写要求,做到报告逻辑清晰。首先,要了解报告给谁看,受众是谁决定了报告的深度和广度;其次,要清楚报告要达到的预期目的,报告一定要逻辑清晰、有理有据;最后,要让思路清晰,在最大程度上反映出数据分析的结论及价值。

新媒体数据分析的基础应用是指数据报告的初步应用,主要工作是根据新媒体数据分析结论和报告对内容生产者的传播行为进行进一步分析与优化,对内容热度与传播链条进行分析与优化,对用户画像与群体聚类进行分析与优化。该部分反映出新媒体数据分析的一般层面的实践价值,是新媒体数据应用最为基础的环节。

新媒体数据分析的高级应用是指分析结论与报告在各类媒体传播、各种商业营销中的应用,以及在各级行政机构社会管理中的应用等。该部分反映出新媒体数据分析的深层次的实践价值,某一方面、某一主题或某一时段的新媒体数据对特定主题传播趋向或特定行为动向的综合分析与研判,是新媒体数据分析应用未来最为广阔的领地。

在实际应用中,由于平台属性与目的等诸多不同,新媒体数据分析与应用的具体角度与深度也大不相同。新闻网站、商业网站、自媒体平台、政务新媒体等同属于新媒体,但各自使用新媒体的初衷不同、对新媒体的定位不同、应用新媒体的需求不同,日常运营组成及对运营的效果要求也各不相同。因此,在不同的新媒体平台或新媒体机构,在基本流程基础上可以结合平台的特点及机构的具体需求进行有针对性的分析应用。

第五节 国内外主要新媒体数据分析机构

一、国外新媒体数据分析机构

国外的新媒体数据分析机构较多,根据平台特性可以分为多种类别,有对全球互联网

数据整体进行汇总分析的,有对某一类网站进行数据分析的,有对传统网站的、对移动端 App 的,还有对纯粹的平台内容访问行为的数据分析,以及基于商业新媒体营销或品牌传播需要的多角度耦合分析。总体而言,在全球知名度较高的新媒体数据分析平台有互联网世界统计、StatCounter Global Stats、Alexa、谷歌趋势 (Google Trends)、App Annie 等。

(一) 互联网世界统计

互联网世界统计 (Internet World Stats, IWS) 是一个国际性的互联网数据统计网站,其网址为 <http://www.internetworldstats.com/stats.htm>。互联网世界统计为全世界 233 个国家和地区提供世界互联网使用、人口统计、旅游统计和互联网市场调研等数据。互联网世界统计网站是由 Enrique de Argaez 于 2002 年 3 月 25 日创建的,Enrique de Argaez 是该网站的编辑、网络管理员和所有者。互联网世界统计是国际互联网市场调研、互联网统计资料、世界互联网普及率、世界人口统计、电信信息报告和脸谱统计等有用的信息来源。互联网世界统计网站在各个网页的底部有活动链接,并在网页的左侧栏目里配有内容目录,使得网站的导航系统明晰,实用性也很强。此外,该网站还附有各种有用的链接,如世界旅游和住宿信息、电信国家报告、国家统计源、本地目录、人口数据、市场调研信息等。如果用户想了解网站的内容以及冲浪提示,可以访问网站导航页面。互联网世界统计网站是一个有用的互联网市场调研工具,分为几大不同的版块,每一个版块都提供了具体且详细的市场调研信息。此外,该网站上有弹出式网站搜索目录和网址导航目录,可以帮助用户快速找到自己需要的主题,如亚洲互联网使用情况。

(二) StatCounter Global Stats

StatCounter Global Stats 是由 StatCounter 免费提供的在线访客统计工具,主要提供全球范围内的网站排名,用户可查看全球范围内 2019 年 2 月—2020 年 2 月世界各个国家和地区的浏览器市场份额,各个浏览器份额排名在浏览器份额图标最上方。

(三) Alexa

Alexa 是 Amazon 公司的一个子公司,成立于 1996 年,是一家专门发布网站世界排名的网站。用户可以使用 Alexa 查询网站浏览率统计和世界排名,Alexa 给出多达几十亿的网址链接,并为每一个网站进行了排名。目前,Alexa 是拥有大量网址链接,排名信息发布最详尽的网站。Alexa 的网站世界排名主要分为两种:综合排名和分类排名。综合排名也叫绝对排名,即一个特定的网站在所有网站中的名次,Alexa 每三个月公布一次新的网站综合排名。此排名的依据是用户链接数 (users reach) 和页面浏览数 (page views) 三个月累积的几何平均值。

分类排名,一是按主题分类,如新闻、娱乐、购物等,Alexa 给出某个特定网站在同一类网站中的名次;二是按语言分类,共分 20 种语言,如英文网站、中文网站 [Chinese (simpl) 和 Chinese (trad)] 等,Alexa 会给出特定站点在所有此类语言网站中的名次。另外还有星级排名,Alexa 会根据网民对网站的评论,在综合排名信息中,用“星”给网站评一个等级,最高为“5 星”,其中文网站为 <http://www.alexa.cn/>。

（四）谷歌趋势

谷歌趋势（Google Trends）是谷歌公司的公共网络设施，基于谷歌搜索显示世界各地的一个特定搜索项的搜索量。2012年9月27日，谷歌将谷歌趋势合并进谷歌搜索。作为谷歌旗下一款基于搜索数据推出的分析工具，谷歌趋势通过分析谷歌搜索引擎每天数十亿的搜索数据，告诉用户某一关键词或者话题各个时期在谷歌搜索引擎中展示的频率及相关统计数据。用户可以通过这些搜索数据了解市场情况、受众信息以及未来的营销方向等相关信息。

用户输入搜索字词或主题，按 Enter 键，即可得到相关信息，包括随着时间推移的变化顺序、区域搜索热度（可以根据地区或者城市进行细分）、相关查询（可以通过“搜索量上升”或者“热门”进行排序）等。谷歌趋势可对互联网热点数据做分析报告，每个报告右侧显示3种用户使用数据的方式：以CSV格式下载数据；以图表的形式嵌入数据（提供嵌入和预览桌面、移动屏幕的JavaScript代码）；在Google+、LinkedIn、Facebook、Twitter、Tumblr等社交平台上分享图表。

关键词研究和热门话题研究是Google Trends的两大主要功能。关键词研究功能可以起到以下作用：了解关键词在不同国家、不同城市、不同时间的表现状况；了解关键词、产品的主要市场；了解与关键词相关的主题，进一步了解受众特征；对比不同关键词在相同国家、相同时间的表现状况；利用特定搜索获得更多见解。以glass bottle产品为例，进入Google Trends，输入关键词进行搜索。可以看到全球以及各个国家2004年以来该词的搜索热度表现。同时，还可以看到按照区域显示的搜索热度，43个国家和地区中新西兰、英国、澳大利亚、美国和新加坡搜索热度最高，这也是产品进行营销推广需要重点考虑的市场。

当然，还可以将市场从国家具体到城市，如在新西兰，威灵顿、怀卡托、坎特伯雷、奥克兰等城市对于glass bottle的搜索频率就比较高。接着可以分析人们都在搜索glass bottle的哪些信息，谈论什么样的话题。这里可以看到热门的搜索主题，也可以看到目前正在上升的搜索主题，可以发现人们搜索比较多的词汇有bottle、glass bottles、glass water bottle等，讨论较多的主题有瓶子、玻璃瓶、水、葡萄酒等。由此，用户可以对比自己的网站和产品，看是否与这些搜索和主题相符。通过相关搜索和主题，我们也能看到glass bottle的受众对什么感兴趣。有的时候同一款产品有很多种叫法，在这里也可以对比分析出来，哪一个词人们最喜欢搜索，比如glass bottle和glass container，我们可以看到相对于glass container，人们更喜欢使用glass bottle，同时也可以看到这两个词在不同国家的表现情况以及相关的搜索情况。

特定搜索功能通过界面可以看到，除了默认的谷歌网页搜索，谷歌还提供了几种特定的搜索选项：图片搜索、新闻搜索、谷歌购物、YouTube搜索。不同的搜索方式会产生不同的搜索数据，如果我们需要获取图片、新闻、谷歌搜索以及视频创建、优化等方面的建议，则可直接采用这几种搜索选项。

热门话题是谷歌趋势的另一个重要功能。通过产品热门话题搜索，商家可以了解客户需求，定制相关网站营销内容，确定季节性营销趋势，在正确的时间创建推广内容。

（五）App Annie

App Annie 是全球领先的移动市场数据供应商，受到超过 100 万用户的信赖。在这个 App 改变商业模式的时代，App Annie 是唯一一个支持整个 App 生命周期的数据平台，它可以为商业公司提供值得信赖的 App 数据与具有执行性的见解，帮助使用者在全球移动经济中取得成功。作为业界第一个将应用程序数据与全面的市场数据、数据科学以及深厚的数据基础和引人入胜的数据体验集成在一起的应用程序数据平台，用户可以在 App Annie 上立即访问该机构所有的最新技术创新和数据集，在合适的时间与合适的人员共享合适的数据，准确把握主要机遇，最关键的是制定一个在商业中制胜的战略。

App Annie 是面向移动端的内容分析与应用平台，这个开创性的平台着力于发掘可行的见解和应用程序策略的基础，这些策略同时也在创造着非凡的成就，并在竞争中保持着持久优势。App Annie 订阅主要面向团队和企业的 App 开发与维护工作，而针对独立开发人员，则提供了 App Annie Connect 和 App Annie Intelligence 免费功能以及大量免费资源，包括洞察报告、App Annie Academy、网络研讨会和各种活动。

App Annie Intelligence 全方位把握 App 的整个生命周期，通过数据驱动的关键词策略，以更低的成本推动付费和自然下载量；充分了解竞争对手的应用商店优化策略，通过付费搜索广告提高 App 的可见性，确保用户在应用商店中的优势；通过监控竞争对手的广告活动，改善用户的移动广告策略；根据各广告平台、国家/地区和类别的网络份额和广告主份额，深入研究顶级广告创意；通过识别各个地区的最佳广告平台优化广告支出；根据各个国家/地区和类别的单次安装成本和转化率平均值，做出明智的广告活动决策。

App Annie Connect 让 App 数据化零为整一触即达，可在同一个视图中追踪每一款 App 的表现。借助四十余个可选来源，App Annie Connect 将来自用户的所有 App 和发行商账户（即使有数千个账户也没有问题）的表现数据整合到一个综合平台，而无须整合 API 或 SDK。App Annie Connect 以需要的方式获得数据。用户可以灵活、便捷地访问数据（借助网页、电子邮件、App、CSV 或 API 自动收集并提供数据），让团队随时了解相关动态并且保障数据安全，确保团队能够访问所需的全部数据，而不必共享敏感的登录凭据。

国外的新媒体数据分析机构除了以上主要五个，还有银行或商业机构自主开发的数据，这些数据部分对外开放，可便利索取。例如，世界银行网站 <https://data.worldbank.org.cn/>，其提供了六大主题领域的数据库：贫困与不平等、人、环境、经济、国家和市场，以及全球链接。每个主题页面都简要地介绍或提供了可得数据的类型、专题指标清单以及关于广泛使用的方法和当前数据难题的信息，网站的所有数据都可以免费使用，当然也有一些最低限度的限制，用户可自行查阅。还有 <http://www.gfk.com/>，<http://www.itu.int/>，<http://www.nielsen.com/us/en.html>，<https://www.idc.com/>，以及玛丽·米克尔的互联网趋势报告等均可提供新媒体方面的数据。

二、国内新媒体分析机构

2022 年互联网来到中国三十余年，国内的新媒体产业纷繁复杂，其数据分析机构较多，除 CNNIC 作为官方的互联网新媒体统计分析机构之外，各类互联网企业创建的分析机构也

林林总总，根据平台特性可以分为多种类别，有总体统计分析类、互联网商业营销类、指数类、对移动端新媒体进行监测分析的第三方机构等。总体而言，国内新媒体业界与学界使用较多的新媒体数据分析平台有如下五类十二个：较早的商业营销类的有艾瑞咨询等，当前使用较多的指数类的有百度指数、头条指数、微信指数、阿里指数等，综合分析类的有 199IT 中文互联网数据研究咨询中心、清博大数据等，新兴抖音直播电商数据分析类的有飞瓜数据、头榜数据、知瓜数据等，互联网舆情数据分析类的有新浪微热点、新浪舆情通等。

（一）艾瑞咨询

艾瑞网是艾瑞咨询（iResearch）精心打造的国内首家新经济门户网站，基于多年深入互联网及电信相关领域研究成果，融合众多行业资源，为业内人士提供丰富的产业资讯、数据、报告、专家观点、行业数据库等服务，多方位透析行业发展模式及市场趋势，呈现产业发展的真实路径，进而推动行业高速、稳定发展。艾瑞咨询是一家专注于网络媒体、电子商务、网络游戏、无线增值等新经济领域，深入研究和了解消费者行为，并为网络行业及传统行业客户提供市场调查研究和战略咨询服务的专业市场调研机构。艾瑞咨询秉承专业、严谨、客观的工作作风，努力发展成为网络经济时代中国最优秀的专业市场调研公司。

作为国内最早进行网民行为研究和网络广告监测的市场研究机构，艾瑞咨询集团通过自主开发，建立并拥有一定规模的数据库，通过多种指标研究帮助行业建立评估和衡量的标准。其中，网民行为研究涵盖家庭办公用户、网吧用户以及无线手机用户等各种应用平台；广告投放监测涵盖网络品牌广告、无线品牌广告、搜索关键词广告等多种媒体类型和多种广告形式；此外基于市场需求，艾瑞建立了针对网络媒体和无线媒体流量审计、网络广告和无线广告投放效果认证，推动行业健康持续发展。为帮助各个行业客户更加深入有效地使用各种数据应用服务，更好地利用数据解决企业在营销、运营、产品、竞争环境等诸多层面的决策支持，艾瑞咨询集团在对客户需求和大量网络监测数据产品的深度挖掘的基础上，还提供各种网络评估数据工具和分析报告，不仅包括媒介方案评估工具、营销活动评估工具、品牌效果评估工具等针对网络营销行业的数据应用服务，还包括针对行业用户细分和忠诚度评估的分析工具，以及针对电子商务、网络游戏等领域的分析工具。

（二）百度指数

百度指数（<https://index.baidu.com/>）是以百度海量网民行为数据为基础的数据分享平台，是当前互联网乃至整个数据时代最重要的统计分析平台之一，自发布之日便成为众多企业营销决策的重要依据。百度指数能够告诉用户：某个关键词在百度的搜索规模有多大，一段时间内的涨跌态势以及相关的新闻舆论变化，关注这些词的网民是什么样的、分布在哪里、同时还搜索了哪些相关的词。通过这些数据，百度指数可以帮助用户优化数字营销活动方案。百度指数主要是用来反映关键词在过去 30 天内的网络曝光率及用户关注度。它能形象地反映该关键词每天的变化趋势。指数是网站提供的以网页搜索和新闻搜索为基础的免费海量数据分析服务，用以反映不同关键词在过去一段时间里的“用户关注度”和“媒体关注度”。用户可以通过所搜索关键词的指数来发现、共享和挖掘互联网上最有价值的

信息和资讯，直接、客观地反映社会热点、网民的兴趣和需求。

以关键词“考研”为例，通过在搜索栏输入“考研”二字，将会得到关于“趋势研究”“需求图谱”以及“人群画像”的分析结果。其中，“趋势研究”主要涉及搜索指数（即互联网用户对关键词搜索关注程度及持续变化情况）、资讯指数（即新闻资讯在互联网上对特定关键词的关注和报道程度及持续变化）和媒体指数概览（即关键词所选时间段的总体资讯关注表现）这些趋势图。例如，从2020年前两个月的三项指数结果可以看出，“考研”二字的搜索频率在2020年2月20日，也就是大部分地区公布考研初试成绩当日达到了最高峰。

搜索关键词“考研”后可得到的“需求图谱”结果栏主要涉及需求图谱（即从用户在搜索该词前后的搜索行为的变化中表现出来的相关检索词需求）和相关热度词（即通过用户搜索行为，细分搜索中心词的相关需求中，最热门词及上升最快词）这两项主要内容。用户通过拉动图谱下方的时间轴即可获得相应时间内其他搜索者最需要的信息，以及与之相关的其他关联词。

最后，“人群画像”结果栏会提供“考研”这一关键词搜索者的地区分布图、人群属性（即关注该关键词的用户的性别、年龄分布）和兴趣分布图（即基于百度搜索用户行为数据以及画像库，刻画所选范围中关注该主题词的人群分布情况以及相对全网平均表现的强弱程度）。可见，百度指数能够对搜索这一关键词的用户形成一个较为全面且整体的特征分析，能够为个人乃至相关企业提供一个详细的趋势分析，引导其做出一个科学准确的趋势判断。

（三）头条指数

头条指数是今日头条算数中心推出的一款数据产品。作为内容生产、传播、营销、舆情监控的重要工具，头条指数致力于用数据服务个人和机构，提供丰富及时的数据维度。基于今日头条大数据分析，头条指数能够反映出用户在智能分发下的阅读及互动行为。用户可以通过使用头条指数捕捉用户的兴趣和关注点，监测社会舆情，为精准营销、舆情应对乃至学术研究提供重要的数据参考。基于今日头条海量激活用户的海量行为数据及文章数据，通过数据挖掘和数据分析，捕捉即时热点，预测可能的热点，为个人或企业提供热点事件的数据分析，并产出个人或企业所需的数据、可视化图表、报告等，提供一站式服务和具有参考价值的数据。用户可以通过搜索关键词来定位和观测某个事件。当查询不到所搜索的关键词时，可根据系统提示修改搜索关键词或将其拆分成相近的关键词之后进行组合使用。

热度指数反映的是某个关键词受用户关注的程度。将关键词的阅读量、评论量、转发量、收藏量等加权求和，以小时级计算出各关键词的热度值，绘制成趋势图。

关联分析是用户在阅读含有某关键词的文章时，与该关键词有关联的其他关键词同时被阅读的频次为基础计算出的词与词之间的相关性。头条指数将及时跟进当天最热门的事件，并进行小时级别的数据实时更新。用户可筛选时间区间和地域范围，更有效地洞悉和观察热门事件，辅助个人或企业了解热点的最新动态。头条数据报告为今日头条算数中心根据热点事件特别制作的数据报告，供个人或企业下载及使用。用户能根据自身需要，

按照报告的领域、来源、标签、关键词进行检索，精准找到感兴趣的相关报告，甚至进行人群画像。

（四）微信指数

微信指数是微信于 2017 年 3 月 23 日推出的基于微信大数据的移动端指数。现阶段微信指数作为内嵌于微信当中的小程序形式存在，便于微信用户了解关键词搜索热度，帮助企业更好地掌握实时搜索舆情状况。但由于其入口较深且微信官方没有做太多推广，所以知道它的用户并不多。目前，有以下两种方法查看所搜索关键词的指数信息。

一是使用“微信指数”小程序，在微信的搜索框中搜索“微信指数”，然后在小程序的分类下，选择“微信指数”小程序；二是直接在微信搜索框搜索“关键字 + 微信指数”，或“微信指数+关键词”，了解特定关键词的指数详情。此外，在“微信指数”小程序中，用户还可以添加不同的关键词，对比每个关键词的热度曲线。

微信指数的适用人群为微信这一社交媒体 App 的所有使用用户，在不涉及隐私的前提下，挖掘用户搜索和浏览行为的数据，抓取微信公众账号文章及微信搜索中的热词，发现、反映微信用户的热点、兴趣和需求，并记录每天的变化趋势。长期以来，搜索引擎平台中的关键词指数查询、主流社交平台的热词榜等因为有明确的量化标准，一直是企业管理者、品牌传播人、媒体、公关人，以及行业研究员等人群掌握潮流趋势、了解社会舆论导向等热点的必备参考。而微信指数的推出，可以为这些用户群体带去多一维度的选择。因此，微信指数的功能及意义主要集中在：微信指数整合了微信上的搜索和浏览行为数据，基于对海量数据的分析，可以呈现关键词动态指数变化情况，方便看到某个词语在一段时间内的热度趋势和最新指数动态。微信指数可以提供社会舆情的监测，能实时了解互联网用户当前最为关注的社会问题、热点事件、舆论焦点等，方便政府、企业对舆情进行研究，从而形成有效的舆情应对方案。微信指数提供的关键词的热度变化可以间接反映用户的兴趣点及变化情况，如日常消费、娱乐、出行等，从而对品牌企业的精准营销和投放形成决策依据，也能对品牌投放效果形成有效监测、跟踪和反馈。

（五）阿里指数

阿里指数是了解电子商务平台市场动向的数据分析平台，2012 年 11 月 26 日阿里指数正式上线，其分析结果主要是根据阿里巴巴网站每日运营的基本数据，包括每天网站浏览量、每天浏览的人次、每天新增供求产品数、新增公司数和产品数 5 项指标统计计算得出的。因为阿里业务调整，2020 年 8 月底阿里指数下线，之后阿里指数的店铺和人群数据可以在完成店铺开设以后，通过后台的方式进入生意参谋网站进行查看，其网址为 <http://sycm.taobao.com/portal/>。

阿里指数是为淘宝卖家提供淘宝数据分析的一款重要的工具。其作用就是让卖家可以更加详细地了解市场行情，把握市场机遇，并能及时调整店铺的运营方向，降低店铺的经营风险。总之就是用数据说话，为店铺节约成本，并让店铺能更好地赚钱。

阿里指数的主要功能有三个方面。其一，行业总体分析与趋势判断。行业大盘：主要包括市场行情、热门行业。以某个行业为视角进行分析。市场行情：市场的综合趋势——价

格、采购、供应的趋势。用户可以通过对淘宝采购、阿里采购以及供应指数的对比的变化观察，分析市场的走向并且预测接下来的市场行情，做好相应的准备。热门行业分析：在这个板块中，阿里指数会做出对各种热门细分子行业的分析，并对各个子行业做出排序。此外，还会做出数据解读。其二，产品的属性与采购量分析。热门基础与营销属性：针对商品本身所固有的基本属性做出属性划分和营销排名，用于表明该行业中这一热门产品受欢迎的属性行情。例如女鞋行业，通过属性细分可以发现，圆头女鞋采购指数和供应指数都是最高的，但是通过计算比例，发现采购率最大的还是新款女鞋，所以新款女鞋的需求量最大，同时竞争力也是最大的，因此必须把产品优化到前面，才有机会展示和成交。采购者素描画像：主要供用户查看搜索、购买自己产品的客户群体，从而确定自己的目标人群。同时，还可以用来对比行业客单价，从而分析自己的商品单价情况，如果过低则应该思考怎样提高客单价。阿里上主要以批发为主，如果客单价过低，说明大部分客户只是拿样，并没有成批拿货，这时就要仔细分析客户拿样却不下单的原因，从而进行优化改正，提高店铺销量。其三，阿里排行榜，即涉及商品交易的产品排行，分为交易排行和流量排行，其中，交易排行是指在一定时间内销售的情况，通过分析交易指数可以发现哪款产品更加容易成交，以及店铺内哪款产品更有爆款的潜质，打造爆款通常是一种快速、有效的方法，阿里上不乏一款爆款撑起一个店铺的实例，所以一定不要忽视这种方法。而公司排行和企业官网排行主要考察店铺供货情况和实力，店铺内产品销量上升了，店铺综合实力自然也会上升，所以做好基础是必要的。除阿里指数，还可以从阿里数据（<http://dt.alibaba.com/alizs.htm>）获取最新行业可视化数据。

（六）中文互联网数据研究咨询中心

中文互联网数据研究咨询中心（199IT）作为大数据分享平台，在数据提供方面主要是以门户网站的形式聚集大数据报告，其自身定位为用数据解读 TMT（telecommunication, media, technology）行业。199IT 提供的内容主要是有关 TMT 各个行业的解读，包括电信、科技以及媒体类，覆盖面广泛。199IT 网站属于内容聚集型平台。首先，199IT 没有自己生产与原创的内容，199IT 所提供的报告由其他大数据平台发布，作为数据报告的“搬运工”，搬运的内容不仅有国内专业性的解读，也有国外数据的专业化解读，采用“搬运”的模式一方面是因为“原创”内容属于收费项目，另一方面 199IT 网站的广泛受众也是其原因之一。其次，199IT 平台对其他数据平台的专业解读报告通过贴标签和专题分类的形式将最新发布的互联网数据平台报告收集并整理到网站当中，在报告内容中表明数据与报告来源，同时数据解读内容及时更新。199IT 通过在数据解读报告下方附上微信公众号的二维码，让用户通过扫码关注微信公众号就可以获取最新的发布内容，从而实现了用户操作的便捷性以及获取内容的及时性。

199IT 报告对于新媒体行业意义重大。

（1）其涉及范围广泛。199IT 数据解读内容当中，新媒体行业的解读包括社交行业中的网络交友、即时通信、微博、SNS、图片分享、问答类网，以及新媒体营销中的移动营销、搜索营销、整合营销、品牌网络广告营销等。

（2）专业性强。199IT 数据报告中虽然大多数为转载的数据解读报告，但其以标签化

的形式，在被授权转载的前提下集合需要付费或使用代理才能享用的原创内容，对于新媒体行业具有指导性意义。一方面，经过授权转载后的付费数据内容可以使更多的人免费享有；另一方面，专业性的行业解读对新媒体行业的走向与趋势的预测具有指导作用。

（3）社群交流，资源共享。199IT 除提供专业性的新媒体行业数据解读，还以付费方式加入新媒体行业的社群，从而获取更多的有关新媒体行业的相关资料。在社群内，可以接收最新的新媒体行业数据以及最新的解读报告，同时新媒体从业人员以社群的方式相互连接在一起，从而构成一个内部资源相互共享、行业数据共同解读的共同体。

（七）清博大数据

清博大数据（<https://www.gsdata.cn/>）是基于网络公开数据的产业、融媒体和舆论大数据人工智能服务商，新媒体大数据评价体系和影响力标准的研究制定者，舆论分析报告和软件供应商，融媒体平台解决方案提供商，一站式行业大数据解决方案服务商。清博大数据为众多政务机构、新闻媒体、品牌企业、互联网公司提供大数据服务，客户涵盖万达集团、海尔集团、阿里巴巴、腾讯、百度、新浪、字节跳动（今日头条、抖音、西瓜视频、火山小视频）等。清博大数据提供清博指数、清博舆情、清博管家、清博广告、清博开放平台、行业大数据、可视化大屏七项数据产品服务；公开提供关于微信、微博、头条号、抖音、快手等视频新媒体应用的排行榜单。

清博指数主要提供微博、微信公众号、头条号、抖音、快手、腾讯视频等移动应用或新媒体平台榜单。其推出的相关指数以时间为指数标准，可以通过选择日榜、周榜、月榜来查看当日、当周与当月的有关微信榜单的指数；也可以根据具体的日期选择进行查看。同时，除了以时间为标准进行分类，对于微信公众号的指数的判定也可以根据所发布内容的形式以及类别进行分类，根据内容的形式可以分为文章与视频，所提供的数据内容分别为排名、公众号、发文数、总阅读数、头条阅读、平均阅读、总在看数、WCI（微信传播指数，从整体传播力、篇均传播力、头条传播力、峰值传播力四个维度进行评价）。清博指数推出的相关指数还可以以地域为标准，微信公众号榜单指数标准的划分可以通过选择不同的地区，进行公众号榜单统计发布。

（八）飞瓜数据

飞瓜数据（<https://www.feigua.cn/>）是西瓜文化传播有限公司旗下产品，是一款提供全网抖音、快手潜质爆款素材、各类排行榜、视频流量监控，以及多号矩阵管理等功能，可以让用户在 1 分钟内掌握抖音、快手运营现状，迅速获取抖音、快手运营数据的专业平台。抖音特有的算法会给创作者的账号打上相关的内容标签，可以让算法系统更精准地将视频推荐给感兴趣的人。如果用户想知道自己账号的标签，可以通过飞瓜数据查询，飞瓜数据会根据账号的视频内容，自动识别相关的账号标签和内容标签。

作为一个短视频热门视频、商品及账号的数据分析平台，飞瓜数据着力于利用大数据追踪短视频流量趋势，提供热门视频、音乐、爆款商品及优质账号，助力账号内容定位、粉丝增长、粉丝画像优化及流量变现。目前，飞瓜数据专门为抖音和快手短视频平台的用户设立具有针对性的数据分析服务。飞瓜数据能够为使用者提供各类最新的热门视频、热

门音乐资源，并通过数据搜索找到时下最受欢迎的热点，提醒用户将其融入内容创作，以此为用户提供获取大量流量的机会。可以说，飞瓜数据能够帮助用户快速发现短视频平台最新热点，把握短视频热门趋势，让用户追热点、生产爆款视频快人一步。飞瓜数据还提供了播主查找功能，即通过挖掘各类榜单和感兴趣的抖音/快手数据，帮助用户了解不同领域 KOL（关键意见领袖）的详细信息，从而帮助用户明确自身账号的定位、受众喜好以及内容方向。飞瓜数据的用户数据监测功能是通过提供查询者的抖音号或快手评论号、视频监控、账号和品牌对比来实现账号发布内容前的时间把控、发布中点赞和转发的数据监控，以及发布后的用户、作品和互动数据分析、数据反哺。飞瓜数据的电商分析功能是指提供抖音或快手全网销售的商品、食品相关数据的检索功能和排行榜单，分析商品热度和在售播主的情况，快速实现购物车商品优选和电商合作选择。

（九）知瓜数据

知瓜数据（<https://www.zhigua.cn/>）是由福州西瓜文化传播有限公司开发的，是一款淘宝直播数据分析的实用工具，提供了主播带货转化量分析、粉丝互动分析、粉丝画像分析等实用功能，也提供了播主销量榜、爆款商品榜、MCN（多频道网络）排行榜等各类电商直播的相关榜单，是一款将直播数据可视化的数据分析监测云平台，可精准、可靠、高效地提供淘宝直播平台的数据分析服务。

在淘宝直播成为新的卖货趋势时，更多卖家希望可以赶上这班顺风车，知瓜数据旨在为从事直播业务的 MCN 机构及淘宝卖家提供从数据到价值的服务支撑，让用户可以更高效、简捷地获取信息。通过淘宝直播可视化数据分析监测云平台，精准、可靠、高效地提供淘宝直播平台的数据分析服务，助力直播广告主精准投放。其中，查看播主榜单分析的操作如下。点击播主排行榜，选择想要查看的行业类目，即可查看该行业类目下的播主排行及播主粉丝数、点赞数等基本信息。

（十）新浪微热点

新浪微热点（<https://www.wrd.cn/>）用于免费监测热门事件，可以监测网站、论坛、博客、微博、微信等多方信息，进行数表统计等。微热点可以根据地域选择查看当日热点事件；热点按排行统计分为一小时热点、24 小时热点、热门事件周榜、热门事件月榜。

以 2020 年 3 月 11 日湖北地区为例，可以通过新浪微热点查看湖北地区热点事件名称、一小时热点读书、一小时环比变化、事件类型，并且可以对相应事件创建全网与全微博分析。同时，点击湖北地区可以查看湖北地区热点事件总数、前五名热点事件与前五名变化最大的事件。在湖北地区 3 月 11 日前五名热点事件中，以“湖北将有序推进企业复工复产”事件为例，事件热度为 39.56，报刊报道热度为 341，新闻报道热度为 2395，政务报道热度为 921，同时还可以查看该事件的敏感程度。点击热点事件可以查询详细内容，以“郑州境外输入病例隐瞒行程被立案”为例，可以看到 24 小时内的热搜指数、热度指数均值、一小时热度变化趋势、热度指数峰值。

在重大事件中可以根据时间降序、信息降序、热度均值降序查看微博对于重大事件的关注程度。通过热门微博可以看到微博中的热门内容以及全网关注的重点事件，平台还对

热门事件进行了分类,可以根据社会、笑话、时尚、星座、电影、体育、理财、萌宠、美食、旅游、健康、视频、音乐、教育、政务等进行分类查看。热门人物搜索可以通过人物类型(全部、事件人物、歌手、演员、主持人、体育人物、企业家)进行分类;并且可以查看人物和单一事件在特定时间段的热度走势对比,如通过人物“王一博”、事件“王一博献唱公益歌曲”看到人物热度与事件在 24 小时内的热度走势。在案例库中,可以根据案例分类查看热点事件,以“视觉中国版权风波”为例,可以查看热点事件走势图、热点事件区间占比图,并且可以同类似事件进行对比。根据微博采集的基础数据可以对热搜或热点事件的大数据报告进行分析。以热点事件“川航‘史诗级’备降微博分析报告”为例,在微博分析报告中可以获知事件简介、事件趋势、热点词、意见领袖、核心传播人、热门信息、传播途径、表情分析、博主分析、数据类型与评论分析,可以得到关于此次川航降落事件的总体微博舆论走势与热度趋势。

(十一) 新浪舆情通

新浪舆情通(<http://www.yqt365.cn/>)是新浪微博政务舆情大数据分析平台,共有十一类信息:新闻报刊政务微博公众号博客论坛视频网站客户端;微博信息实时获取;网站、论坛等信息源最快 3 分钟轮询 1 次;预警信息最快 1 分钟下达;精准设置词距和地域;定向监测指定行业、指定信息源;自定义信息监测时间范围;等等。其主要功能如下:竞品分析:对拟进行对比的多个事件、人物、品牌或地域等进行关键词设置,然后系统自动生成一份包含热度指数、正负面词频、地域分布、媒体报道情况等多个维度的竞品分析报告,通过竞品分析报告可以快速查看竞品相关信息在互联网上的品牌传播概况;舆情监测:可以提供信息监测、舆情事件、舆情预警,以及舆情数据化大屏、敏感信息预警、重点账号监测、舆情态势感知、舆情定制报告等;热点事件分析:以热点舆情事件为例生成网络传播分析,可以把分析结果生成数据化大屏,其舆情智库可以通过舆情数据分析来为行业、品牌或党政机构提供建议。

以上十一个新媒体数据分析机构采集、分析的数据几乎涉及了国内所有的新媒体应用,除了面向 PC 端的采集平台,大部分是微信小程序或 App,可以随时随地进行数据采集与分析,为一般需求的新媒体数据分析与应用提供了较多便利。对有复杂需求的新媒体数据采集分析与应用,可以在充分使用以上平台的基础上进行分析框架的自我扩展,可采用多元数据采集处理分析方法达到更高的目标。

国内的新媒体数据分析机构除了以上十一个,还有自主收集数据形成数据报告的机构,这些数据全部或部分对外开放,可便利索取。例如,CNNIC 每半年度发布的报告、中国社会科学院新闻与传播研究所发布的中国新媒体发展年度报告、中国互联网协会从 2002 年起发布的中国互联网发展年度报告、腾讯公司发布的年度新媒体专题报告、易观分析(<https://www.analysys.cn/>)、虎嗅(<http://www.huxiu.com>)、友盟+(<https://www.umeng.com/>),均可提供新媒体方面的数据。

习 题

1. 根据互联网世界统计网站的数据,分析全球各大洲互联网网民总数及普及率差异性的原因。
2. 根据中国互联网络信息中心最新一次《中国互联网络发展状况统计报告》的数据,归纳我国政务新媒体应用的特点。
3. 比较国外的新媒体数据分析机构的异同。
4. 比较国内指数类新媒体数据分析机构的异同。

实 践 任 务

1. 查看自身微信朋友圈或微博、微信公众号等后台自行统计的数据并导出,试进行简单分析。
2. 根据新媒体数据分析应用的一般步骤,结合实际,制订切实可行的方案,试分析所在学校的某个新媒体平台,并提出应用数据的基本思路。

案 例 讨 论

2020年11月15日,由国家互联网信息办公室指导,CNNIC、中国科学院计算机网络信息中心(CNIC)主办的“第二届中国互联网基础资源大会(China Internet Infrastructure Resources Conference, CNIRC 2020)”11月15日在北京落下帷幕。作为互联网基础资源领域的行业盛会,本次大会的主题为“夯实‘根’基,数‘聚’未来”。会上正式发布了“国家互联网基础资源大数据(服务)平台二期”。据悉,“国家互联网基础资源大数据(服务)平台一期”已于2018年9月20日在北京正式上线。经过一期和二期工程的持续研发与建设,国家互联网基础资源大数据(服务)平台已初步形成涵盖数据采集、清洗、汇聚、管理、分析、挖掘、安全保障等环节在内的互联网基础资源领域全链条大数据技术能力,并形成了多项对外服务成果。

更多信息请登录中共中央网络安全和信息化委员会办公室官网,查找“工作专题”栏目的“第二届中国互联网基础资源大会”专题网站,以及“数据服务”栏目相关报告,进而查阅国内外最新的互联网新媒体基本数据信息。

思考:

在新媒体数据的各项工作中,我们如何做好数据查找工作,如何查找到权威可靠的数据?