

大数据在重大流行病疫情防控中的应用及展望

梅小亚¹,赵林畅²

(1.贵州民族大学社会学与公共管理学院,贵州贵阳 550025;2.重庆大学计算机学院,重庆 400044)

摘要:重大流行病疫情都具备范围广、跨区域、全社会、全球性等重要特征,流行病疫情的全要素、全过程管理可以充分发挥现代大数据数字技术功能。以新型冠状病毒肺炎疫情的全生命周期管理为例,论述了大数据在支持临床诊疗持续改进、支持病毒溯源与分析、支撑流行病学调查与诊断、精准监测和研判疫情、有效排查疫情、远程医疗诊疗、准确预警疫情等疫情防控全过程中的应用,提出了大数据助力疫情期形势综合研判、构建适应于国家到社区“六级”联防联控体管理决策系统、疫情期间科学化和精准化社会治理、AI智能创新医疗产品、智慧医疗行业和疫情后期社会经济恢复与发展方面的创新和应用,分析大数据在重大疫情防控应用方面存在的数据准确性、数据割裂、数据公开与隐私保护不足等问题。

关键词:大数据;重大流行病疫情防控;新冠肺炎;经济社会管理

中图分类号:D632.6 **文献标志码:**A **文章编号:**1671-4970(2020)02-0039-09

一、问题的提出

进入21世纪以来,全球自然灾害和人为事故的数量、致死人数上升的趋势明显加速,美国“9·11”恐怖袭击、印度洋海啸、SARS事件、汶川大地震、国际金融危机、甲型H1N1流感等特大灾害事件接踵而至,造成了大量人员伤亡、财产损失和严重的社会失序,灾害及其管理已成为各国政府和民众高度关注的公共议题^[1]。

2019年12月开始,我国武汉暴发了由新型冠状

病毒(COVID-19)引起肺炎的呼吸系统疾病新型冠状病毒肺炎(以下简称新冠肺炎),并迅速传播,形成了“大流行”级别的重大流行病疫情危机。新冠肺炎病原体是一种特殊的新型病毒,感染人数多、传播范围广、潜伏时间长、危险程度大^[2]。这不仅需要对病毒的临床特征、传播特性和病原学特点进行充分的了解和认识,而且出现疫情的防控和危机消除涉及范围广、跨区域、全社会、多行业、多主体,这就需要从重大流行性疾病发生、发展、消亡的全过程分析,并进行多层次、多地区、多主体的合作和协同管理。

收稿日期:2020-02-19

基金项目:国家自然科学基金重点项目(91118005)

作者简介:梅小亚(1979—),女,贵州毕节人,博士研究生,从事社会政策与社会治理研究。

通信作者:赵林畅(1989—),男,贵州铜仁人,博士研究生,从事大数据挖掘分析及人工智能研究。

随着突发事件复杂性的不断提高和自然灾害危机管理的艰巨性不断被人们所认知。大数据分析不仅为风险灾害危机管理提供了全样本、关联性与系统化的研究思维,而且为相应的决策提供了科学、客观的方法与技术支持^[3]。

大数据的价值在于海量与关联。随着移动互联网和物联网的发展,社会的数据量激增,同时,互联网把所有的数据连接在一起,通过数据的采集、转化、挖掘、集成与存储等大数据分析技术的应用,使得有效处理突发事件中不同类型、不同维度以及不同尺度的数据成为可能^[3]。大数据的功能是利用收集的信息和挖掘技术对庞大的数据信息进行专业化分析处理,实现对全部信息的整体把握和核心信息梳理^[4]。在流行病疫情防控中,大数据作用显得尤为重要。大数据不仅能显著提高公共卫生和医务人员对疫情的溯源、追踪、监测、预警和响应能力,使疫情防治“四早”(早发现、早报告、早隔离、早治疗),将防控关口前移,强化社区网格化管理,及时发现和隔离病毒的感染者,降低感染率,提高收治率;能够实现“四集中”(集中患者、集中专家、集中资源、集中救治),加强患者救治,提高治愈率,降低病死率;能够动态监测、了解疫情严重地区及全国的资源供需情况,掌握疫情严重地区的医疗物资和人力资源缺口,加强调度,提供援助。同时,大数据还可以支撑疫情流行期间的社会经济系统运行,为政府管理决策者提供国家和地区的人口、交通、经济、社会、重要生产和生活物资供应保障等信息决策,通过疫情信息公开以及科学工作者对病毒及疫情的解读,为公众提供更完整、连续、准确、及时的数据,保障经济社会秩序健康稳定。

大数据技术在全球的重大公共事件危机管理中如何发挥重要作用,很多专家学者进行了大量研究,过去20年来已在埃博拉、登革热、禽流感等重大流行病疫情防控中大量应用于实践。关于大数据应用于疫情研究方面的主要成果包括2009年谷歌利用搜索关键词大数据成功预测了美国冬季流感的暴发,而且研判非常及时,比美国疾病预防控制中心的数据早一周以上的时间^[5]。Polgreen利用谷歌日志关键词在玻利维亚、巴西、印度、印度尼西亚等地构建登革热传播模型,模型预测值与实际监测数据有良好的相关性^[6]。Bio. Diaspora公司利用大数据,成功预测了下一个可能引爆埃博拉病毒的地区^[7]。

James等基于mHealth策略通过人群流动信号大数据分析进行救济协助、需求评估和疾病监测,十分有利于西非埃博拉疫情的控制^[8]。2003年SARS后,我国已建立了39种法定传染病监测系统、突发公共卫生事件监测系统和预防接种信息系统,完成了从乡镇卫生院至县级及以上医疗机构的大数据监测系统^[9]。加之智能医疗穿戴设备、移动智能设备等提供的大数据实时资讯,为我国监测传染病、及时应对传染病提供了更广阔的空间。虽然我国在医疗方面,大数据已有所应用,但从当前武汉发生的新冠病毒的疫情来看,大数据功能的发挥还显得薄弱,为此,以这次中国的武汉疫情为例,探讨和展望大数据在重大流行病疫情防控中的应用。

二、大数据在新冠肺炎疫情防治中已有的应用

1. 大数据支持临床诊疗持续改进

重大流行性疾病防控的关键之一是对患者进行及时有效的诊治,而诊治的关键是发现流行性疾病发生的症状及其机理,进而对症下药。然而,新的流行性疾病一旦发生,人类并无对其有足够的科学认知,也无对症治疗药物,只能通过以往相似病毒治疗经验以及可能有效的治疗手段和药物,采用循证医疗方法,不断进行治疗手段和发现有效的治疗方案改进,不断提出更完善的诊疗方案。临床病例和药物大数据的利用可以支持临床诊疗不断改进。新冠肺炎疫情暴发,各级政府、医疗卫生系统和医疗专家团队充分利用大数据来辅助疾病临床认知、创新诊疗模式、提升救治效率,全力以赴地救治感染患者。通过回溯相关临床病例研究,从最初的27例不明肺炎患者,表现为呼吸系统常见疾病症状,如:发烧、呼吸困难、双侧肺浸润性病变等^[10];到高福团队基于425例确诊患者的临床资料分析,得出了更为清晰特征的认识和新冠肺炎的流行病学特征(流行倍增时间和基本繁殖数),并声明密切接触者之间发生了人传人现象^[11];钟南山院士团队则基于1099例确诊患者的病例样本大数据分析,得出更全面、更准确的临床特征和新冠肺炎的流行病学特征,并声明基于患者复杂的特征表现,在诊断方面需要多管齐下、精准施策^[12]。说明新冠肺炎疾病的临床表征认知、判断标准制定、治疗药物筛选和试验、诊疗建议方案拟定等的变化过程,均是基于临床病例及其样

本大数据的挖掘分析不断更新、校准和完善的,包括国家卫生健康委员会不停改进的诊断方案标准,目前已经有第7版。

大数据支持临床诊疗的持续改进。由于临床患者轻重不一、基础性疾病不同,临床特征复杂、病症现象不规律、病原学特点不明确,为临床医务工作者科学用药、精准救治带来了困难。另外,全国上市药品约20万种,通用名8000多种,医生难以依靠人脑记忆。而基于大数据及人工智能的合理用药系统能为医生提供用药决策辅助,通过病患治疗用药和效果的大数据分析,不断调整治疗方案,使治疗方案更科学。依据患者病情状况数据,合理地调整用药方案,采用更为经济、有效的用药决策。自火神山医院和雷神山医院共计2500个床位接诊后,一套由杭州逸曜信息技术有限公司提供技术支持、以大数据和人工智能为核心技术的合理用药系统上线^①,该系统通过云计算及云储存、大数据和人工智能技术助力临床医生用药决策,实现精准、科学、高效用药。新冠病毒核酸检测试剂盒的不断升级也是大数据支持临床诊疗持续改进的案例,新冠肺炎疫情蔓延初期,人们对病毒的基因序列、核酸长链(RNA)和蛋白质外壳等结构认识不足,使得传统核酸检测方法耗时、耗力和资源紧张。随着大数据技术对疾病疫情的挖掘分析,多款新型新冠病毒核酸检测试剂盒上市,检测时间从6小时缩短到15分钟左右,缓解了检测试剂盒资源短缺压力,提升了检测准确率,降低了检测人员被感染风险。

2. 大数据支持病毒溯源

病毒溯源对于揭示病毒来源及其进化规律、消除疫情源头、防止疫情扩散是很重要的过程。也是人类彻底通过溯源去发现病毒来源、进化过程及传播规律,并最终战胜病毒的根本出路。自新冠肺炎暴发以来,为发现病毒传染规律、切断传染途径,急需弄清“这个病毒是什么”以及“病毒从哪里来”两个问题,即病原鉴定和病原溯源。对前者,要从成千上万种病毒中锁定病原体何其的难。2003年SARS疫情,全世界科学家经数月努力才确定SARS病原体是冠状病毒。在新冠肺炎发生后,中国科学界充分利用大数据技术,不到一个月就将武汉不明肺炎的病原体锁定为“新型冠状病毒”,为患者快速确诊、治疗、药物疫苗研发及有效防控疫情均起到了关键性作用,但目前对后者尚无明确答案。

实际上,病毒溯源非常困难,世间物种千百万,地域分布广阔,若不运用大数据、人工智能、云计算等数字技术,病毒溯源的希望极其缥缈。为找到新型冠状病毒可能的自然宿主,尤其是其中间宿主,科学家们正积极运用大数据、人工智能等手段进行研究,充分利用开放的国家基因库生命大数据平台,结合已有研究和公开的数据信息,追溯新型冠状病毒的中间宿主。当前,许多研究者已取得了阶段性进展,初步排除了家禽家畜,认为蝙蝠、穿山甲可能是新型冠状病毒的主要来源,水貂、蛇可能是中间宿主。同时,基于大数据、人工智能的疫情溯源系统已上线,由第四范式、南京大学以及苏北人民医院临床专家团队联合研发,他们利用大数据、机器学习技术构建数据驱动的新冠病毒传播数字孪生系统,构建潜在传染的关系网,结合病患信息,在关系网中找到可能的病毒传播源以及潜在的超级传染者^②。

3. 大数据支撑流行病学调查

重大流行性疾病防控的关键是发现病毒感染者和密切接触者,通过收治和物理隔离手段切断传染链,其基础工作在于科学、准确的流行病学调查,掌握流行病病例发病情况、暴露史、接触史等流行病学相关信息,而这一切离不开大数据的支撑。在新冠疫情中,国家卫生健康委组织制定了《新型冠状病毒感染的肺炎流行病学调查方案》,通过流行病学调查去锁定感染者、识别密切接触者、明晰病毒传播路径十分重要。但是,流行病学需要大量的现场调查工作,耗时费力,而许多基层疾病预防控制机构普遍存在人力和能力不足等问题。同时,只有全面精准地采集当事人乘坐交通工具、旅行行程及轨迹、流病史、参加亲朋聚会地点和接触人员等个人信息,才能满足疫情防控的需要。而现实生活中,流行病学调查的科学、严谨要求往往会面临各种主观和客观因素的干扰,从而使调查获得的大量信息真假难辨,无疑会对后期研判和防控产生不利影响,甚至会带来一定的误判风险。在调查中,被调查的当事人可能处于病情危重状态,意识思维不清或表达逻辑

①贾晓芸. 速成火神山医院背后的杭企力量[EB/OL]. [2020-02-17]. http://www.hangzhou.gov.cn/art/2020/2/5/art_812262_41879266.html.

②王心馨. 第四范式人工智能助力疫情防控:精准防控、疫情推演及病毒溯源[EB/OL]. [2020-02-18]. https://www.thepaper.cn/newsDetail_forward_5879766.

混乱等现象,一些主观回忆信息存在谬误也在所难免,甚至会有个别当事人故意隐瞒真实信息或提供虚假信息^[13]。因此,流行病学现场调查取得的大量信息即使经过比较科学的筛选、甄别、清洗和推理,有时也很难完全复原真实的场景,或产生信息遗漏。

在“互联网+”已广泛应用的当下,流行病学现场调查可以广泛运用互联网、大数据等技术,对当事人或知情者提供的有效信息进行甄别和综合梳理分析,不仅可以准确掌握当事人的数据信息,甚至其准确度可能比当事人本人直接提供的还要高。正如李兰娟院士论述:“专家利用大数据技术梳理感染者的生活轨迹,追踪人群接触史,成功锁定感染源及密切接触者人群,为疫情防控提供宝贵信息。”甘肃省利用公安“天眼”系统和大数据平台调取相关数据,根据与基层流行病学调查组的比对,翔实核查出已确诊患者和疑似病例的活动范围及接触人群,缩短了流行病学调查时间,拓展了排查渠道,增强了调查结果的准确性。国家卫生健康委已于2020年2月5日发布了《关于加强信息化支撑新型冠状病毒感染的肺炎疫情防控工作的通知》,要求强化与工信、公安、交通运输等部门的信息联动,形成多源大数据的分析和应用,提供疫情期间当事人交通、住宿、通信、接触人员等信息,为理清传播链条、切断传播路径提供数据支撑。

4. 大数据精准监测和研判疫情

科学、精确地把握疫情演化趋势,是流行病疫情科学防控的关键,而这一切都离不开基于大数据的疫情监测和研判。中国的新冠病毒疫情大暴发始于武汉。当前出现的“第一个”与武汉无接触史的新冠肺炎患者,就是通过大数据发现他实际上和最少3个武汉患者有过接触。武汉人群的移动大数据对监测疫情发展具有重要意义。根据百度地图慧眼大数据的人口迁移分析,截至1月23日,从武汉离开人员流向最多的地方是武汉周边区域,以孝感、黄冈最多,北京、重庆等湖北省外的一二线城市也是武汉流出人口的主要聚集地。根据武汉市政新闻“武汉常住人口将近1100多万人,户籍人口是990多万人,流动人口将近500万人,春节或疫情因素,大概有将近500多万人离开了这座城市”。在武汉外流的人口,65%为省内,35%为省外,即离开武汉在湖北省内的有325万人,流出省外的为175万人。应对疫情,结合交通(铁路、航空、公路等)、宾馆酒

店和通信(中国移动、中国联通、中国电信)大数据全面追踪、排查、监测所有从武汉外出的人员。

截至2020年3月28日24时,国家卫生健康委官方网站显示,全国累计报告确诊病例81439例,累计治愈出院病例75448例,累计死亡病例3300例,现有疑似病例174例。累计追踪到密切接触者701884人,尚在接受医学观察的密切接触者18581人^①。此外,中国疾病预防控制中心新型冠状病毒肺炎应急响应机制流行病学组通过对7万多例的病例研究,探讨了新型冠状病毒疫情的病死率和传染性,并结合疫情确诊病例大数据进行了发布日期回溯性分析,将此次国内疫情的暴发划分为5个阶段,得出2020年1月24—28日为第一个流行病峰值期,2020年2月1日为单日发病高峰,形成了《新型冠状病毒肺炎流行病学特征分析》重要成果^[14]。

5. 大数据有效支持疫情排查

重大流行病疫情往往发生社区聚集性感染,病毒通过感染者与易感人群在公共交通、医院、文化娱乐、旅游、商业等公共场所进行社会接触发生传染。防控的关键环节之一是在社区、单位和公共场合排查和发现感染者以及密切接触者,及时阻断病毒传播链。大数据在新冠疫情时期,防控策略是内防扩散、外防输出,切实做好疫情精准排查工作,做到早发现、早报告、早隔离、早治疗。基于国家卫生健康委、铁路总公司、民用航空局和交通运输部等部门的数据,中国电科云公司研发了疫情防控大数据平台——密切接触者测量仪。在疫情中,该平台被部署在国务院“互联网+监管”平台上,许多地方政府正对接使用。其有两种服务模式:一是公众自查模式,群众登录平台,在输入姓名和身份证号后,即可获知自己是否是密切接触者;二是单位体验模式,即将人员姓名和身份证号信息统一上报后,系统后台进行集中对比,对上报的单位成员进行全面排查。该平台数据权威、模型可信、结果准确。若查询结果是“非密接人员”,则会温馨提示您继续做好个人防护;若是“密接人员”,则会提示您不必过分担心,暂时不要外出,加强隔离防护,如有身体不适,请及时就医。

各地在对重点人员排查中应用“大数据+网格

^① 中华人民共和国国家卫生健康委员会. 截至3月28日24时新型冠状病毒肺炎疫情最新情况[EB/OL]. [2020-03-29]. <http://www.nhc.gov.cn/>.

化”,做好硬核防控。加强多源大数据的分析和应用,坚决做到依法排查、科学排查和精准排查。并结合社区人员实行地毯式排查、网格化追踪,做到“入网入格入户”,将排查工作落实到户、到人。只有做实“人工+大数据+群众监督”摸排调查,才能实现数据取证和技术分析,促进精准排查及时阻断感染源。重庆市政府上线了智能化“疫情排查”信息系统,通过网格化大数据对来渝返渝人员、疑似患者、密切接触者和企业复工复产等进行精准排查。让排查数据“跑起来”、监测数据“动起来”,使排查工作更高效、更科学。浙江省衢州市将人工智能转化为防控疫情新型战斗力,大数据全面赋能防疫一线,全量布防小区道闸、智能停车、城市卡口等关键部位,获得多维感知数据,将重点人员通过交叉匹配关系,经过数据筛查、卡口过滤、网格兜底,巧妙运用大数据“找到人”、大平台“看好门”和大网格“守好门”,将防疫工作真正落细、落地、落实,成为防控阻击战和歼灭战的特种军。江西省瑞昌市运用新开发的“市大数据社会治理信息化平台”和“市疫情防控辅助分析平台”,采取“大数据分析+网格化管理+群众监督”摸排手段,全面排查,精准防控。内蒙古鄂尔多斯市东胜区兴胜街道通过“大数据+网格化”,利用现有人口大数据平台,通过横到边、竖到底的网格化管理模式,将市内流入的、自治区内流入的、内蒙古以外其他省市流入的、有过外省市旅程的本籍人员途经湖北、武汉的人员(未居住)分别排查、分类建立台账,并筛查出重点人员,进行分类管理。大数据助推各项举措“入网、入格、入家庭”,延伸到“最后一微米”,为居民筑起了一道“防护墙”,3天精准摸排21670人,做到不漏一户、不漏一人。

在疫情防控和社会经济发展统筹兼顾阶段,防控疫情常态化手段之一是开发应用了全国疫情防控健康码和“苏康码”“宁归来”等各地使用的健康码系统,通过大数据比对,实现健康人群的识别和通行管控,大大促进了社会运行效率提高。

6. 大数据远程诊疗

重大流行病发生将导致大规模的病患救治和社会医疗咨询需求。有限的医疗资源供给和巨大的病患诊疗需求产生严重的社会矛盾,尤其是高水平感染科、呼吸科、重症医学科专家极为短缺,医院医护人员服务也不能够满足社会需求。大数据助力医疗机构开展网上远程会诊、在线咨询、居家医学观察指

导、诊疗救治培训等服务,形成“线上线下一多空间、多维度、全方位的救治感染患者。贵州省在疾病救治过程中,通过大数据、互联网技术开展了线上线下服务^①。线上通过“朗玛公司”旗下的39互联网医院平台为患者提供网上视频就诊,线下通过国家卫生健康委宣传司与腾讯健康基于大数据联合发布的定点医院和发热门诊导航地图引导患者及时、就近、有序就医。实现资源共享、互联互通、科学救治、精准施策。线上助攻战疫,整合医疗资源。宁夏充分发挥全国“互联网+医疗健康”示范区成果优势,依托远程影像诊断平台,开通了全国新冠肺炎专用通道,充分利用全国的专家资源,对符合疑似新冠肺炎患者的DR/CT片源进行会诊,为抗击疫情工作再添“利器”。该平台利用远程体系进行影像诊断,对疑似患者进行影像学检查,充分发挥CT影像学检查的筛查作用,通过专家会诊对疑似患者进行尽早发现、尽早隔离、尽早治疗。远程影像会诊平台作为这场战“疫”的重要防线,让放射影像成为新冠肺炎鉴别诊断的重要环节,对临床诊断确认提供了重要的参考。

广东省80家互联网医院为居民提供线上问诊服务,居民可通过关注“健康广东”公众号或各家互联网医院微信公众号、手机App进入就诊、咨询平台,在线向医生问诊咨询。首批57家互联网医院上线后,全面开展了针对新冠肺炎的网上义务咨询、居家医学观察指导等服务,2月14日前已累计接到发热门诊问诊近6万人次,新冠肺炎咨询超过2万人次,有效减少了患者线下就诊交叉感染风险。山东省利用“5G远程医疗”助力鲁鄂携手战“疫”。2月9日下午,山东移动7小时打通山东与湖北视频联系通道,实现了湖北黄冈大别山区医疗中心与济南市传染病医院的5G实时视频连线。通过手机、PAD等移动终端,医护人员可直接进行远程情况汇报和会诊,会诊画面清晰、流畅,现场5G网络峰值速率可达1Gbps以上,上传速率超过50Mbps。疫情处置机构、医疗救治定点医院、疾控中心等通过“5G远程医疗”系统连接在一起,实现远程指挥调度、远程诊疗、远程监控、信息共享,既有助于科学指导、精

^①李慧超. 线上线下一合力抗疫——抗击疫情大数据、互联网在行动[EB/OL]. [2020-02-18]. <http://www.gywb.cn/system/2020/01/29/030313231.shtml>.

准施治,提高疫情处置水平与救治效率,又能有效降低医护人员的感染风险。

7. 大数据准确预警疫情

疫情暴发前及演化过程中,防控关键是预警。大数据在疫情预测和警示方面可以发挥至关重要的作用。国家卫生健康委员会利用大数据,结合腾讯地理资讯系统,通过分析全球导航系统、人口迁徙流动和城市公共卫生大数据平台,建立模型,预测下一个可能暴发的疫情区域,为早期预警、优先预防和及早救治抢占先机。天津大学医学工程与转化医学研究院联手中国电子信息产业集团发布了医疗大数据服务云脑平台在线助力疫情防控及远程诊疗^①,利用国家健康医疗大数据中心的平台和技术,开展对疫情数据远程收集、加工处理、科学计算和准确预测等工作。通过大数据、人工智能和科学计算技术,重点实现了疫情预测预警。

大数据在疫情的预测防控上的应用还有很多,中国联通疫情风险预报平台,百度大数据 AI 体温预警以及“摄星智”大数据团队整合海量数据,利用 AI 技术实时预测疫情传播情况^②。“摄星智”团队提供的“摄星人”平台实时更新疫情峰值预测模型可提供近 1 个月的疫情预测数据,帮助医疗专家更快、更精确地了解疫情发展态势、峰值时间和数据等,进而辅助相关部门、医护工作者及广大社会群体对各项事务开展合理计划和安排,助力打赢疫情防控阻击战。

8. 大数据支撑防疫物资调配

重大流行病疫情防控的关键是保障防护物资的及时有效供给。新冠肺炎疫情发生后各地对医用口罩、防护服、呼吸机等医疗物资需求激增,特别是疫情重灾区武汉市,不仅武汉各大医院官网发布了接受急需医疗物资的捐赠公告,学习强国平台、公众号、微博等平台也对外发布了物资短缺、亟须捐赠的消息。消息一经传播,全国各地都主动给政府和防疫一线工作人员捐物捐资。但是,信息发布渠道分散、公信力度低,而且疫情防控机构无法有效监督捐赠信息、捐赠者也无法追踪捐赠结果,慈善机构接受的大量抗疫物资与医院需求物资之间的匹配在一段时间里成为重要社会焦点问题之一。在疫情快速发展阶段,湖北省特别是武汉市防疫物资配发中出现了“肠梗阻”现象^③,通过大数据相关技术作为“润肠剂”可以有效打通阻塞。许多志愿者通过大数据手段,对网上各种公开的需求数据进行爬取、整理、清

洗和分析,构建了湖北医疗物资需求信息平台,该平台能展示物资名称、需求数量、联系方式以及物资运输方式等信息,并及时针对历史物资数据,进行大数据分析,帮助有关部门预测未来物资需求的情况,科学筹划下一阶段资源供应和调配。杭州市锦城街道纪工委利用大数据、云治理技术实现“人防+技防”的精准监督,确保防疫物资来去有踪,并通过大数据、云治理技术,对每一位监察联络员的进出监控记录,做到每一笔物资的流向都有数据可查,这不仅极大地缩短了物资分发时间,减少了人力、物力、财力消耗,还确保了防疫物资“进出有据、去向可查”。

中国工业互联网研究院联合京东集团共同研发了“国家疫情防控物资工业大数据公共服务平台”,该平台充分利用工业互联网的理念、体系和技术,提升产业链上下游数字化、网络化水平,推动生产要素供应链及电商采购物流大数据共享,增强疫情物资智能化调度能力,汇聚了以京东为代表的大型电商采购和物流链企业的成熟资源,建立了覆盖防疫物资需求采集、统筹调度、供给保障、物流跟踪等模块的综合服务体系,实现了防疫物资全要素、全产业链、全价值链的互联互通,推动了防疫物资生产、调用、仓储和供应等流程的协调联动,保障了防疫物资的快速、高效和全面供给。

三、大数据在疫情防控中的新应用展望

重大流行病疫情防控是一项长期的复杂的系统工程,是国家治理和治理能力现代化的重要组成部分。除上述新冠疫情防控、病毒溯源、资源调配等领域的大数据应用技术不断开发、完善和创新之外,仍然有许多可以应用的新领域有待持续开发。

1. 重大流行病疫情形势综合研判和管理决策支持系统

重大流行病疫情防控,需要大数据、互联网、云平台支持下的应急管理和科学决策。因此,开发重

① 刘晓艳. 天津“云脑”上线助力复工疫情预测及远程诊疗 [EB/OL]. [2020-03-27]. <http://news.tju.edu.cn/info/1003/49981.htm>.

② 李季. 以智能技术实时预测疫情传播“摄星人”助力疫情攻坚 [EB/OL]. [2020-03-28]. <http://www.chinanews.com/sh/2020/02-07/9082021.shtml>.

③ 支振锋. 防疫物资调配不妨引入市场力量 [EB/OL]. [2020-03-28]. http://iolaw.ccssn.cn/zxzp/202002/t20200203_5084339.shtml.

大流行病综合管理决策支持系统,形成数字作战地图,解决“数据孤岛”难题,使信息间能够互联互通,构建适应于国家、省、市、县(区)、乡(镇)、村(社区)的“六级”联防联控体系,就十分重要。采用Hadoop、Spark、NoSQL、HDFS、MapReduce、Hbase、Zookeeper、Oozie、Pig、Hive等大数据技术^[15]尽快建立2020新型冠状病毒肺炎综合管理决策支持系统,结合网格化技术、地理信息系统和全球定位系统技术,收集“六级”地区的疫情及防控情况动态,每日动态更新,为中央和各省、市政府提供循证决策支持、满足科学决策需要,为公共管理的高效化、科学化提供保障。提升各级政府新冠肺炎大数据决策支持能力和疫情社会综合治理水平,使基于流行病疫情的大数据决策系统能够提供多元化、个性化、可视化的产品和服务。

为保障决策支持系统的运行,需收集大量的重要数据,包括:各省、市、县(区)、乡(镇)的各类人员(确诊、疑似、发热待查、留观)及四类病人数量;所有医院、各专门医院及床位数、已用床位数、空余床位数、动态预测短缺床位数,医生、护士、管理、服务人员数,主要治疗药品库存、需求数量,主要医疗防护物质及设备数量及需求数量,重要紧缺物资(供医护人员使用的防护服、医用外科口罩、一次性隔离衣、防护面罩等)需求、生产、采购、运输、配置信息等。这些数据是疫情形势综合研判和管理决策的重要依据,即数据科技、数据孪生、数据力量表现出的生命力、竞争力和创造力。依据这些数据构建的决策支持系统能够推动着疫情防控方式的创新,使疫情联防联控与最先进的信息高科技成果相结合,实现疫情阻击目标更加清晰、防控思维更加创新、救治手段更加丰富多彩,推动疫情防控向科学性、系统性、可控性方向发展。积极探索疫情大数据的内在规律,着力释放疫情大数据的内在活力,推进基于疫情大数据决策支持系统治理和综合研判的实现。

2. 大数据支撑疫情期间社会治理

重大流行病疫情期间,需要统筹疫情防控和社会经济正常运行,保障社会治理有效。大数据在疫情防控中发挥了积极的作用,不仅支持医学工作人员更好地做出科学决策,还能借助人工智能、大数据等技术实现疫情数据可视化,助力地方政府精准把握民情民意动态,发现和评估社会治理风险。通过数据挖掘技术和可视化技术,对社会治理问题、治理

程度等实现即时分析和关联性分析,更好更快地查找出疫情期间的社会治理问题源头,为源头治理提供客观依据,监督和协调了各部门间的信息畅通^[16]。通过交通(铁路、航空、公路等)、公安(住宿、道路)和电信大数据,知道疫情严重地区人员在疫情期间的流动去向,把握整体流入流出人群规模和特殊人员(确诊患者)的流动地点。通过交通大数据(铁路、城际公交、城市公共交通、城乡注册出租车)追踪确诊、疑似患者的密切接触者;通过商场、银行、医院、超市等公共大数据可以查询到与确诊患者或疑似患者同一时间段的密切接触者;电信大数据可以用于评估封城期间人员规模的空间分布、交通和日常生活服务需求,为日常生活保障提供基础数据。电商大数据可以了解疫情期间人们的心理动态,物资需求和行为表现。这些大数据信息有利于提高社会治理的科学性、精准性和有效性。

大数据还推动了社会治理从静态化管理走向流动性治理。随着疫情传播速度的加快、联防联控信息共享渠道的扩展和抗疫信息来源的多样化,疫情期间的社会治理变得高度复杂和高度不确定。大数据社会治理可以利用其强大的数据采集和分析能力,即时捕获丰富的抗疫数据信息。通过对海量数据的挖掘整合、统计关联和预测分析,可对社会运行规律、社会偏好(诉求)变化趋势及规律、政府回应机制及效果差异等进行实时性、量化性、可视性观测分析,不断积累疫情期间的社会运行数据特征以应对各类社会风险和提升社会治理效能。加之“大数据+网格化”技术可以快速建立确诊、疑似、发热待查、密切接触者的人数、状态、医疗防护物资、病床位、医护人员等疫情防治数据库系统,梳理清楚人员空间分布、需求、供给等状态,结合GIS、GPS等定位信息技术,快速理清疫情期间的人流、车流和物流分布,为战胜疫情、迅速复工复产、生活保障物流通、交通资源配置等提供数据科学支撑。此外,国外经验表明,大数据以在线社交平台为载体,有助于网络舆情的监测与分析,从而及时地进行危机风险评估与监测对应急管理的全过程和各个环节都发挥着越来越重要的影响^[17]。大数据技术还可用于疫情期间的社交媒体、传统媒体等各种媒体发布信息的采集、统计、监测,对舆情进行识别、分析、评估,对虚假信息予以及时辟谣,对错误信息予以纠正,避免错误信息传播引发恐惧、焦虑情绪和污名化言论的传播。

为疫情期间的社会治理提供“数说有据、依据有声、发声即准”的技术支持,同时也为社会治理模式创新、社会治理精准化、公共服务高效化提供了决策支撑。

3. 大数据助力国民大健康

重大流行病防控,需要利用大数据,支持病患收治、密切接触者隔离、公众疫情健康咨询等应急管理工作。新冠肺炎疫情暴发,催生了大批 AI 智能创新医疗产品,它们既有效助力疫情一线患者救治,还最大限度地保护了医务工作者的健康。上海联影集团针对新冠病毒传染性强的特点,为武汉雷神山医院和火神山医院提供了两类设备:一是搭载“可视化曝光”技术的 X 光设备,医生可以在安全防护范围内实现隔室遥控曝光,既减轻了医生承受辐射剂量,又减少了医患接触;二是具备人工智能技术的天眼 CT,医生在不需进入扫描间的情况下,即可利用设备去智能识别患者身体部位,从而自动完成患者摆位、扫描等全部流程。同时,该集团公司第一时间及时启动并迅速完成了适合武汉方舱医院的“方舱 CT”产品研发、生产和发货,该 CT 具有独立于医院的箱体式设计,具有占地面积小、快速拆装、插电即用等优点。此外,其子公司联影智能及时推出了 uAI 新冠肺炎智能辅助分析系统,该系统是业界第一款针对综合肺炎整体与局部影像特征、肺炎影像精确分诊的 AI 全流程解决方案^①。

2020 年 3 月下旬,百度 AI 多人温度检测系统在北京清河火车站落地,投入新型冠状病毒肺炎疫情防控。该系统基于人脸关键点检测及图像红外温度点阵温度分析算法,可以对一定面积内乘客的额头温度进行检测,即便是佩戴帽子和口罩也能够快速筛查。商汤科技公司的智慧健康团队基于在 AI 基因分析、AI 药理方面的长期科研积累,针对新冠肺炎治疗需要,与多家专业医疗机构、科研单位及高校联合,提出生物大分子治疗方案,以解决病毒对单一药物的耐药性问题,增加对病毒变异的抵抗能力,从而实现抗击病毒目标。思必驰公司的疫情防控机器人成了排查利器,仅用 2 个多小时,就完成了苏州工业园区东韵社区 7 000 余居民的疫情排查,报告详细记录了每个号码的接听情况,并将有武汉接触史和有发热症状的人员进行了特殊标注和登记。在疫情防控期间,该公司的疫情防控机器人高效、快捷,成效显著。截至 2 月 7 日,该公司疫情防控机器人在全部上线地区累计完成互动式访谈通话近 50

万通,通知宣贯类通话近 100 万通。

4. 大数据用于后疫情时代的社会经济恢复

重大流行病疫情最终会被人类控制,疫情从流行到结束,社会经济系统从应急响应到逐步恢复发展常态,大数据是社会经济恢复的关键支撑工具。严防死守、交通管制乃至少数封城的疫情期间,教育行业实行了“停课不停学、宅家不停教”,不耽误每门课程的教育、不落下每一位学生的学习,充分利用大数据、互联网技术推出在线课堂、云上教育,教育部还整合多级优质教学资源,通过国家中小学网络云平台和中国教育电视台空中课堂,免费向广大师生开放,该平台不仅有防疫知识、红色教育、专题教育,还有从小学至普通高中的主要学科课程资源。多省政务服务事项实行 24 小时“不打烊”云服务,利用大数据、云平台系统为群众提供各项政务服务,推行“在家办、网上办、掌上办、预约办、帮代办、邮寄办”等方式,实现行政审批和政务服务“不停摆”、保障疫后社会经济大复苏。

疫情稳定控制之后,大数据助力疫情期间各地企事业单位全面复工复产。据“东经易网”调研结果显示疫后企业复工复产关注度最高,但因各地新冠肺炎疫情的严重程度不一、防控手段和要求也不同,约 80% 厂商经营者对如何申请复工复产存在迷茫。大数据、互联网平台能迅速在复工的二三级厂商之间搭建供需匹配通道,建立直接的业务关联,同时对接网络支付平台和政务服用系统,解决二三级厂商间对复工复产的担忧,助力企业快速复工复产、经济快速复苏。

在疫情基本控制情况下,社会经济发展恢复的关键是如何在疫情防控无重大风险情况下保障人员的有序流动,大数据则帮助管理系统识别出健康人口,保障人员的有序、合理流动。

四、大数据应用的进一步发展建议

大数据在疫情防控方面有着巨大的价值,不仅能够对人口流动、搜索、医疗、社会、经济等数据进行 AI 挖掘、分析、评估、预测、预警,通过趋势分析预见未来,制定对策而防患于未然,还能智能调度人员、物资,甄别新闻信息真伪。人工智能驱动精准治理给社会

^①宋薇萍. 联影集团人工智能解决方案全面助力抗疫[EB/OL]. [2020-03-29]. <https://fin-ance.ifeng.com/c/7u22EGrY6Tw>.

带来巨大福利的同时也潜藏着巨大风险^[18],大数据在疫情防控中应用也存在一些挑战。

1. 大数据结果的准确性需进一步提升

当前大数据在疫情防控中的准确性不够高,除了基准数据集不够完善外,还受大数据挖掘技术的影响。目前收集到的数据有限,基于流行病学数据构建的疾病系统不够完善,为辅助临床医疗、防疫工作和科学研究,需进一步提升数据收集技术和完善系统数据。而且,流行病学的信息源不止一处,数据格式不统一,需实时对数据进行转换、标识、清洗和分类等处理。在原有海量数据的基础上,不断接收新信息、转换新数据,实现关于新冠肺炎疫情大数据的实时动态采集。同时需综合多方面数据信息,注重多学科知识融合和数据与结果间强关联分析,建立更为专业的流行病学研究准则,结合循证医学原则,提升大数据在疫情防控中监测、排查、救治和预测的准确度。

2. 数据开放与隐私保护需进一步加强

大数据技术可快速分析海量数据的相关性、寻找规律性,极大提升疫情防控系统的准确性、实效性和灵敏性,发挥大数据在重大流行性疾病防控中的巨大潜能,但相关大数据的开放和安全隐私保护需进一步加强,从监管和技术角度看,可从以下几方面的改善来逐步实现在不暴露个人隐私前提下实现更多数据的开放共享:①鼓励各机构、医院间协同合作,允许数据开放,联合非限制性数据和非特有数据组成开放公有的“大数据”生态系统;②明确数据访问权限,出台隐私和数据所有权政策,确保所有用户能够正常访问数据并能在网络空间安全前提下实现数据共享;③加强大数据新处理技术的研发,充分利用机器学习、深度学习、人工智能、云计算、云储存、区块链、5G技术、联邦学习以及融合处理等技术,开发新的流行病疫情防控大数据解决方案。

参考文献:

[1] 董星,张海波. 基于中国问题的灾害管理分析框架[J]. 中国社会科学,2010(1):132-146.
[2] ZHOU P, YANG X L, WANG X G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin[J]. Nature, 2020(2):1-20.
[3] 董星,丁翔. 风险灾害危机管理与研究中的大数据分析[J]. 学海,2018(2):28-35.

[4] JIAN L M, ZHAO X, LI J, et al. Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding [J]. The Lancet, 2020(2):391-393.
[5] WILSON N, MASON K, TOBIAS M, et al. Interpreting “Google Flu Trends” data for pandemic H1N1 influenza: the new zealand experience [J]. European Communicable Disease Bulletin, 2009(44):1-3.
[6] BENJAMIN M A, YIH Y N, DEREK A T. Cummings prediction of dengue incidence using search query surveillance [J]. PLoS Neglected Tropical Diseases, 2011(5):1-7.
[7] 辛妍. Bio. Diaspora: 基于大数据的疫情扩散预测 [J]. 新经济导刊, 2014(11):44-49.
[8] JAMES O D, AMALIA B. Controlling ebola through mhealth strategies [J]. The Lancet, 2015(3):22.
[9] 董银峰,刘忠于,王好锋,等. 大数据在疾病预防控制中的重要性 [J]. 实用医药杂志, 2015, 32(7):579-581.
[10] STOCKHOLM ECDC. Pneumonia cases possibly associated with a novel coronavirus in Wuhan, China [J]. European Centre for Disease Prevention and control, 2020(1):1-3.
[11] LI Q, GUA X, WU P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus infected pneumonia [J]. The New England Journal of Medicine, 2020(1):1-3.
[12] GUAN W, NI Z, HU Y, et al. Clinical characteristics of 2019 novel coronavirus infection in China [J]. MedRxiv, 2020(2):1-30.
[13] 汪孔亮. 大数据让现场流行病学调查更精准高效 [N]. 中国人口报, 2020-02-05(003).
[14] 中国疾病预防控制中心新型冠状病毒肺炎应急响应机制流行病学组. 新型冠状病毒肺炎流行病学特征分析 [J]. 中华流行病学杂志, 2020, 41(2):145-151.
[15] BORIS D, JORGE E H, JASON P. Decision support systems V-big data analytics for decision making [M]. Berlin: Springer, 2015.
[16] 洪之旭,陈浩,程亮. 基于大数据的社会治理数据集成及决策分析方法 [J]. 清华大学学报(自然科学版), 2017(3):264-269.
[17] 邵东珂,吴进进,彭宗超. 应急管理领域的大数据研究: 西方研究进展与启示 [J]. 国外社会科学, 2015(6):129-136.
[18] 李利文. 人工智能时代精准治理的隐忧与风险 [J]. 河海大学学报(哲学社会科学版), 2020, 22(1):82-90.

(责任编辑:吴玲)