医疗大数据的价值、风险与法律规制
——基于新型冠状病毒肺炎疫情治理的思考

吴何奇

上海财经大学法学院,上海 200433

摘要　在新型冠状病毒肺炎疫情(简称“新冠肺炎”)的应对中,医疗大数据技术体现了多方面的价值:通过大数据实现了对医疗资源的有效调度、基于大数据的AI医疗、通过大数据预测疫情的暴发、疫情趋势的监测等。但在大数据技术成为医疗领域重要技术手段的同时,也产生了数据风险的新问题与新挑战,完善数据规制从而协调数据保护与数据使用之间的冲突成为制度设计的难点。梳理了当前个人医疗数据规制的现状,结合医疗实践分析数据规制存在的不足,通过对国外数据规制的借鉴,提出了完善中国数据规制的建构逻辑,即以社会整体利益为核心,在平衡利益关系的前提下进行制度设计,将数据滥用行为纳入数据规制,引入场景理念修正同意规则并推动数据去识别化标准的形成。

关键词　医疗大数据;疫情治理;数据风险;法律规制

信息数据已成为当下最重要的战略资源之一,其价值堪比石油和黄金[1]。大数据(bigdata)的发展与应用带来了全球范围内的技术变革,也让世界各国高度重视数据资源的战略价值,相继出台国家战略推动大数据的发展。在中国,中国共产党十八届五中全会将大数据上升至国家战略地位,《促进大数据发展行动纲要》与《关于促进和规范健康医疗大数据应用发展的指导意见》两大核心政策文件的相继颁布,为中国健康医疗大数据的快速发展提供了政策上的支持。作为“互联网+健康医疗”命题下的核心构成,随着运行数据的分级管理、分析能力的不断跃升,无论是对医疗服务效率、质量的提高,还是对传统医疗服务模式的转变,抑或是在未来通过智慧医疗模式惠及全民,医疗大数据在医疗卫生事业中的价值与日俱增。

1　重大疫情治理中大数据的价值体现

自2019年12月30日新冠肺炎疫情公告的首次发布之日起,新冠肺炎的确诊人数在1个月的时间里便超过了2003年非典型肺炎(SARS)的确诊人数。对此,在中央全面深化改革委员会第十二次会议上,党中央强调要通过更好地发挥大数据等数字技术以支撑疫情监测分析、病毒溯源、防控救治

收稿日期:2020-04-07;修回日期:2020-09-12

作者简介:吴何奇,博士研究生,研究方向为刑法、国际法、科技法等,电子信箱:122615999@qq.com

引用格式:吴何奇. 医疗大数据的价值、风险与法律规制——基于新型冠状病毒肺炎疫情治理的思考[J]. 科技导报, 2020, 38(23): 31-38;
doi: 10.3981/j.issn.1000-7857.2020.23.004

等。习近平总书记强调:“要运用大数据提升国家治理现代化水平。建立健全大数据辅助科学决策和社会治理的机制,推进政府管理和社会治理模式创新,实现政府决策科学化、社会治理精准化、公共服务高效化。”这次新冠肺炎疫情期间,大数据不仅助力政府在疫情的科学管理和资源优化配置上发挥作用,也让公众及时了解了疫情的发展情况,积极参与到科学防疫的战斗中,大数据在重大疫情治理中的价值主要表现在以下几个维度。

1)通过大数据实现对医疗资源的有效调度。医疗大数据能够实现人力物力成本降低的新型数据应用模式,为健康管理服务提供了新思路。新冠肺炎疫情中,病毒传播的速度之快、范围之广一度导致口罩、防护等医疗物资严重短缺,特别是疫情的集中暴发让湖北省大多数医院的医疗设备、医疗资源应接不暇。通过对医疗大数据的分析有助于政府了解各地区医疗物资的供需现状从而作出协调生产端与需求端的精准决策,通过智能供应链的管理让物资以最短物流路径,最短在途时长从生产线到达疫区,从而保障医疗物资供应、防止疫情扩散。此外,通过对医疗物资下发数据的后续跟进,政府可以及时掌握医疗物资的落实情况,从而确保政策执行的有效性。面对来势汹汹的疫情,最危险的是教条主义和经验主义,比慢作为更可怕的是乱作为和瞎指挥。只有建立起“用数据说话、用数据决策、用数据管理、用数据创新”的决策机制,才能采取超前举措,打破陈规,抢占先机,扭转被动跟随、疲于应付的局面。

2)基于大数据的机器人诊疗,能够降低医护人员的工作风险。无论是AlphaGo的人机对战还是沃森医生的肿瘤诊断,人工智能(AI)本身就是大数据技术的价值体现。AI基于不断积累的医疗大数据、通过不断地训练提升医疗数据分析功能,能够赋予医疗保健范式的转变。在治理类似于新冠肺炎的重大疫情中,AI+大数据不仅可以提前干预人体健康,降低发病概率。还能借助于医疗大数据实现对治疗路径的预测。不仅如此,以大数据为基础的机器人诊疗还能应对“人传人”的病毒传播风险,是保障医务人员与患者生命健康的关键[1]。例如,美国第1例SARS冠状病毒感染者在西雅图确诊并被送往医院后,为了规避病毒在院内的传播,该院的医生一直使用机器人诊治这名患者。

3)通过对数据的挖掘,预测疫情的暴发、防患于未然。在疫情治理过程中,大数据不仅可以分析出人员的流动轨迹,实现对确诊病例、疑似病例的及时追踪,更能通过高位人群的运动情况,借助传播动力学模型、动态感染模型、回归模型等大数据模型和技术,对高危人群提供健康宣教和精准管理,防止疫情的扩散。这体现了在重大疫情的治理中大数据对现状的有效把控。此外,如果能够借助大数据提前预测到疫情的暴发,疫情的形势必然得到有效的改善。疫情预测预警是利用整合、分析健康医疗大数据,通过建立统计分析和数学模型,挖掘其中所传递的传染病发生、发展和流行的规律来实现的。当前,基于大数据对健康风险的预测从而实现精准预防的实践对提高疾病预防和诊治效益、改善人群健康所发挥的作用日益突显[2]。国外医疗实践中,Google公司在2008年便推出了Google Flu,利用用户的搜索查询记录来发现流感的暴发,它甚至比美国卫生部门提前2周发现了2009年H1N1型流感的流行。中国亚型流感病毒检测平台的建立也为中国的流感防控提供了科学依据。实践证明,基于大数据通过先进算法模型能够实现对疫情发展的科学预测与精准研判。

2 医疗大数据运用所存在的数据风险

近年来,在医学和临床研究方面,以高分辨率医学影像技术以及其他高通量分析技术为代表的新一代生物技术的进步,产生了大容量的、结构复杂的组学数据,包括基因组学、影像组学以及蛋白组学等数据。随着时间的积累,多维度、覆盖人类生命周期医疗大数据得以形成。这让医疗技术逐步从基于能力的学科,转变成以大数据为支撑的前沿科学[2]。这也意味着大数据在治理重大疫情中所能发挥的水平高低与医疗数据的多寡与质量息息相关。

伴随着医疗大数据的高速发展,医疗数据的风

险愈发暴露,医疗数据的隐私保护令人担忧。随着医疗信息共享的开展,数据量将不断增加,隐私保护的难度也在逐步提高,这是现阶段发展卫生医疗大数据的主要障碍之一。数据的集中存储会增加数据被泄露、破坏、盗窃的安全风险,网络传播病毒、非授权访问等会造成大数据基础设施的安全威胁,阻碍大数据基础设施的运行,也会引发隐私数据储存的安全问题。近年来,隐私保护和隐私攻击模型同步发展,诸如数据脱敏技术的安全措施无法有效防止大数据在健康医疗过程中的泄露。

在中国,信息泄露的情况不容乐观。根据《中国大数据法治发展报告》中调查的数据显示,个人信息被非法提供和售卖的情况不容乐观。几乎是在公民在向服务和商品提供者提供本人信息的同时,个人信息就遭遇泄露和滥用,并且有半数以上的受访者因此遭受不同程度的直接经济损失[3]。2011年,中国知名信息技术交流平台CSDN上的600万个人用户信息被泄露。在"罗维邓白氏公司非法买卖公民个人信息案"中,2亿条公民个人信息被盗卖。据360互联网安全中心的统计,2011—2014年,11.27亿条中国公民个人信息已证实被泄露,其中包括账号密码、通讯录等私密信息[4]。而在医疗领域,以2016年为例,全国便至少有275例HIV感染者的个人信息被不慎泄露[5]。如果这些数据被不法分子用于不正当的途径,包括隐私权在内的患者的诸多合法权益将会被侵害。

在个人数据已成为大数据时代重要驱动力的同时,以隐私为核心的个人数据保护日益受到关切。一方面,不充分的隐私保护会诱使大数据运用进行监管套利,损害公民个人的合法权益。但另一方面,过于严苛的数据保护又会制约技术的创新,对经济社会的发展产生制约。例如,美国的大型零售百货商塔吉特(Target)在未经用户的同意下,通过收集顾客的浏览记录,用算法得出顾客的个人信息甚至敏感信息。这种数据挖掘是商家精准营销的必要前提,但对于民众隐私的刺探难免引发公众质疑。

尽管不是所有的大数据应用都涉及个人数据,但不可否认的是,大数据应用中的相当一部分都是基于个人粒度数据的展开,目前阶段的大数据应用更是主要集中于对个人数据的分析加工。尽管价值密度低是大数据的特性,并且单一数据的泄露并非都涉及隐私的侵犯,但是多源数据聚合后可能会存在身份识别的风险。大数据时代,公民的行为数据因为精密的电子设备的部署而不断被记录、成为大数据技术运用的原料,但这些记录也引发了日趋严重的个人隐私危机[6]。医疗健康领域所涵盖的海量临床数据以及医学诊断信息是智慧医疗发展的必要前提,而伴随着对海量医疗数据的挖掘和使用,是涉及个人信息的获取、扩散等一系列问题的出现。在新冠肺炎疫情的治理中,就存在工作人员在医疗数据的传输过程中由于疏忽导致他人信息不同程度地被泄露的问题,不仅给被泄露人的生活带来极大不便,甚至会发生地域歧视和人身威胁。但在数据风险不断增加的同时,既能保证数据有效流通又能兼顾个人隐私保护的相关法律或指引尚未形成。与传统数据的分析路径不同,大数据技术的特性决定其在运用中不强调医疗数据与数据处理之间的因果逻辑联系,无法预测被收集的个人医疗数据将以什么样的目的作用于何处[7]。因此,在大数据的运用中,传统数据法律保护机制无法充分发挥作用,导致数据主体的个人医疗信息面临更多更复杂的风险。大数据时代,技术的强大会让个人隐私被侵犯的可能性有所提升,如何加强对数据的保护且不抑制数据的自由流通将成为数据规制应对技术变革带来的新问题的关键。

3　个人医疗数据的合理使用与保护

3.1　法规范视角下的概念厘清

大数据的内涵不仅是数据量的海量,还体现于数据内容、价值的"大"。根据英国《1984年数据保护法》对大数据的界定,判断某项成本可预期的数字技术是否为大数据技术,须至少满足规模(volume)、快速(velocity)、多样(variety)和价值(value)4项标准中的2项。

在世界各国的个人数据保护立法中,既有"个人数据"的称谓,也有"个人信息"的用法,还有"个

人隐私”的表述。自动化领域的“数据”是能够通过设备自动处理、记录的信息。但法律视野下的个人数据意在数据对特定人的识别性而非数据本身。德国的《联邦数据保护法》、英国的《数据保护法》对“个人数据”的定义以及中国《网络安全法》对“个人信息”的定义都采用了类似的概念。虽然个人信息与个人数据的概念存在区别,且个人信息的范围较个人数据要更广,但术语上的差异往往是法律传统、使用习惯以及立法价值取向的不同所导致,术语形式的变化并不影响法律实践的内容[8]。法规范中的“个人信息”与“个人数据”通常只是表述上的差异,往往可以相互替换使用[9],甚至互为解释对象,在法律上区分二者在概念上的差异并未体现出特殊的实践价值。因此,本文讨论的既是数据规制也是信息规制。但从逻辑上看,个人信息并不等同于个人隐私,二者是交叉重合的关系[10]。同理,健康医疗数据中的敏感内容与患者的隐私存在重合,但二者涵盖的范围并不相同。指出二者异同,在于分辨哪些医疗数据属于隐私进而重点保护,哪些医疗数据着眼于共享和利用[11],从而既能实现数据风险的约束,也能保证数据资源的自由流通。

3.2 个人医疗数据法律规制的现状

鉴于大数据时代,现实人格向数据人格的转变,人与人之间社会关系的构建几乎都依赖于数据的展现,因此,敏感性个人数据通常就是个人隐私,这也让个人数据保护重点实则为隐私权的保护。而关于病情与诊疗的个人医疗数据较之于其他个人数据而言更具私密性,由于一般人通常匮乏医学专业知识,个人医疗数据的泄露极易为数据主体带来他人的误解与歧视。诊疗机制的启动表明患者进行医治的环节已由医疗人员共同参与,而诊疗过程难免涉及患者的个人隐私,如何兼顾个人医疗数据的利用与保护成为各国一直以来探索和改进的重点,医疗数据在隐私权的领域内较之于一般个人数据应居于更为重要的地位[12]。这从数据规制的立法沿革可见一斑,尽管联合国在1968年的国际人权会议上才首次提出“数据信息保护”的概念,且直到1970年,世界上第一部关于个人数据保护的立法才诞生于德国黑森州。但早在1974年,《普鲁士一般法》便对医生等特定职业的保密义务予以要求,医生等特定职业的行为主体对患者个人信息的泄露将面临刑罚的苛责。

目前,世界上已有120个国家或地区制定了有关个人数据保护与使用的专门法律,个人医疗数据属于个人数据,因此世界各国针对个人数据的法律规制同样适用于对个人医疗数据的合理使用与保护。医疗领域,新版的《国际卫生条例》在疫情治理的问题上分别从“监测”“核实”“建议”等方面强化了世界卫生组织的活动权力。但上述权力开展的前提是个人医疗数据的公开与流动,换言之,与疫情相关数据的披露对于重大疫情的治理具有无法估量的价值[13]。譬如,疫情的监测与预警机制的实现需要通过增进世卫组织与各国之间的信息沟通与共享来实现疫情的加速通报,进而保证所有疫情信息向世界卫生组织的准确通报。但数据的使用难免触及患者个人信息的私密性问题。尽管世界卫生组织强调要对伦理标准、人权等法律、文化中的敏感性议题予以充分的尊重,但这些规则的提出往往难以约束世界卫生组织成员国在数据使用方面的具体行为。较之于《国际卫生条例》在数据保护方面欠缺拘束力的体现,欧盟堪称个人数据法律规制严苛性的代表,不仅在区域组织层面制定并不断完善指导欧盟整体的通用数据保护条例,还敦促其内部的成员国相继颁布有关个人数据保护的立法,对数据安全、数据主体保护、赔偿责任等方面进行了较为全面、细致的规定。在数据使用方面,《通用数据保护条例》在第6条规定了合法性基础的6种情况,即数据主体的同意、数据主体所参与的合同的履行、履行数据控制者的法定义务、为了保护数据主体或他人的重要利益、实现公共利益的需要以及数据控制者的优先利益。实践中普遍存在的隐私政策、用户协议既是同意原则的落实,也是数据主体与数据处理者之间的合意,即合同履行的体现。通用数据保护条例(GDPR)并未对数据控制者的法定义务划定明确的范围,为条例与国际法、国内法的衔接准备了空间。例如,医疗领域,数据控制者基于疫情防控的需要而承担的强制披露义务理应属于这里的法定义务。数据主体或他人重

要利益的保护以及实现公共利益的需要凸显的是个人数据保护与数据主体的其他合法利益、第三人利益以及数据控制者利益之间冲突的协调。至于何者的利益更具有优先性则需权衡具体个案的实际情况[14]。

在国别层面,发达国家相继建立了较为完善的政策法规体系以加强个人数据的保护。美国于1996年制定的《健康保险可携性与责任法案》,详细规定了医疗信息化中的交换规则、数据安全以及个人隐私等问题。根据该法案,尽管出于科研的目的使用患者的个人医疗数据无须征求患者本人的同意,但仍会委托伦理审查委员会通过审查程序予以保护。美国卫生及公共服务部于2000年依据该法案授权制定《个人可识别健康信息的隐私标准》,标志着美国已为保护患者医疗隐私构建起一个完整且具有可操作性的法律体系。此外,加拿大的《个人信息保护及电子文档法案》对跨省或跨国商业机构使用个人健康信息的行为予以禁止,澳大利亚的《健康档案法案》对隐私保护利益相关者的义务、权限和法律责任等内容做出了严格规定。

围绕数据保护,中国的民事法律、行政法规、刑事法律均有所涉及,一些行业性规范,也对个人敏感信息做了严格的规定。《中华人民共和国民法典》明确宣告了司法对个人隐私、个人信息的保护,第1226条更是针对患者个人隐私、个人信息的保护提出了要求。《政府信息公开条例》第25条确立了个人数据主体知悉"与其自身相关的政府信息"的权利以及要求更正错误信息的权利。严重侵犯他人数据、隐私的行为,还可能构成犯罪,尽管《刑法》并没有直接规定"侵犯隐私罪",也没有直接以隐私权作为独立犯罪客体,但分则中第三、四、五、六章都针对个人信息、数据规定了相关犯罪,即依据个人数据涉及的权益性质提供不同的罪名来予以保护。显然,中国现有法律对个人数据提供了一定的保护,但要在大数据时代兼顾数据保护与数据资源的流通,当前的顶层设计仍值得反思。

3.3　现状下数据规制存在的问题

1)现有的法律针对的只是部分个人数据行为。当前,法律较为关注的是数据的非法收集或泄露,但对数据被收集以后的滥用行为几乎没有涉及[15]。就大部分个人数据而言,侵害往往集中于收集、泄露以后的滥用行为,最为普遍的便是垃圾信息不断、骚扰电话频繁。但现行立法忽视了对个人数据的滥用行为的有效规制。例如,《刑法》在个人信息犯罪的规制逻辑上选择了基于隐私保护模式的仅制裁转移型侵害的思路,仅入罪规制非法提供、获取或泄露个人信息的转移型行为,无法将滥用信息行为纳入规制范围,导致对大数据环境下"合法获取、不当滥用"的典型问题束手无策[16]。

2)数据的处理与使用需要遵守数据规制的合法性前提,中国现有的数据规制是以知情同意原则为基础架构,以数据主体的同意阻止对个人数据收集、利用、提供甚至出售的违法性。而在医疗实践中,知情同意作为一个法律概念,展开于患者自主权的表述过程。随着社会信息技术的繁荣,逐步延伸到个人数据保护的主体框架[17]。中国个人信息的相关立法亦将知情同意作为个人信息合法使用的前提。依据知情同意规范对个人医疗数据的保护意味着医院、疾病预防控制机构在收集数据主体的信息之前,应告知数据主体需要什么样的信息、如何处理和使用什么样的信息,在依法取得患者的明确同意以后,才能对其个人医疗数据进行处理与运用。知情同意原则的建构逻辑是认为信息主体可以理性地自我管理个人信息,从而维护信息主体自己的利益。但在大数据时代,数据二次利用、多环节流转特点将使得知情同意原则发挥的作用越来越有限。对于数据使用者而言,如果每一次信息收集或使用行为均需征得数据主体的同意,将会严重加大数据使用的合法成本。为提供数据收集行为的合法性,数据收集方的隐私政策也更青睐于格式条款的方式,实证研究表明,哪怕隐私政策的文本再通俗易懂,用户也极少真正阅读隐私政策,"知情"的可能性大打折扣。即使用户真正阅读了隐私政策,他们也不见得能够理性地预测自己可能面临的风险[14],立法者建立同意原则的初衷远没有实现。

3)数据脱敏规则的标准有待明确。敏感数据的去识别化有利于对个人隐私的保护和对侵犯个

人隐私的救济。中国传统的匿名模型是对具有个体识别性数据的直接删除,但这种一劳永逸的做法因制约数据的利用价值而饱受诟病。值得注意的是,国外立法对于数据的匿名化程度也未制定具体的标准,美国的《健康保险携带与责任法》也仅仅提及"去身份化"的概念,要求匿名数据不可通过与其他数据的比较实现主体身份的识别[18]。但这样的要求在当前的技术环境下愈发的暴露出缺陷。例如,网络视频公司Netflix曾公布上亿条经过匿名处理的电影评分数据,但德州大学的2位研究人员通过这些匿名数据与公开的IMDB数据的对比便将匿名数据与具体用户对应了起来。Netflix的案例是大数据时代隐私保护和隐私攻击模型同步发展的真实写照,目前国内外还没有形成任何成熟的数据脱敏方案[19],能否通过技术的升级实现数据的完全脱敏尚具有很大的不确定性,即便真的实现是否会导致数据价值的下跌又是值得深思的问题。

4 中国数据规制的理性建构

4.1 中国数据规制建构的基础逻辑

数据规制的基础逻辑是在经济社会的不断进步给传统隐私保护理论带来挑战的过程中产生,不同时代,数据规制的设计都遵循特定的基础逻辑。在大数据时代以前,"数据最小化原则"是数据规制的基础逻辑,强调他人只能收集与数据处理目的直接相关且必要的个人数据,禁止数据收集目的实现以外的数据保留。该时期的数据规制着重于限制个人数据使用者的收集、处理、储存等行为保障公民的人格权益。随着web2.0时代的到来,数据的商业价值刺激了企业收集信息的动机,但为了减少人们数据失控的问题,"数据控制原则"作为对"数据最小化原则"的补充应运而生。该原则在欧盟的立法实践更强调数据收集方的告知义务以及公民对个人数据被使用的同意。进入大数据时代,个人数据第一次被规模化地采集、分析与运用。企业、政府对个人数据的采集和分析,旨在开发一个更受欢迎的产品,抑或是发现潜在的犯罪分子,计算机专家则通过数据预测人类的行为轨迹。数据挖掘把数据分析的范围从"已知"扩展到"未知",从"过去"推向了"将来"[20]。然而,大数据技术缔造了前所未有的数据挖掘和数据利用,也让公民在依赖数据福利的同时陷入数字化的"敞视式监狱"。因此,数据规制的基础逻辑不是单面向地基于数据的利用或者对侵权行为的压制,而是以社会整体利益作为考量,基于利益关系的平衡,进行合目的、合比例、低成本的制度设计。基于此,数据合理使用与侵犯隐私界限绝对不能以静态的视角对其进行一刀切的划分,反之,必须报以动态的场景视角。在此背景下,美国在其数据规制中引入场景为主导的个人数据保护新机制,欧盟在《数据通用保护条例》中也增设了场景导向的新理念,数据规制的基础逻辑逐渐转型为以"兼顾多方利益"为核心的动态平衡,何为个人数据的合理使用,不同的场合应有不同的考虑。医疗领域的场景思维其实有着深厚的国际法渊源。立于人权法的立场,为了疫情治理而对个人权利的限制或克减也是被允许的。例如,《欧洲人权公约》第5条便允许出于预防传染病传播的需要,公共当局有权干预有关的个人隐私。此外,《公民权利和政治权利国际公约》也在第4条第1款赋予政府建立临时医疗保障体系的权力。为了应对危机并恢复正常的社会秩序,将包括隐私权在内的个人权利置于公共利益之后则是这一紧急状况下医疗保障体系的题中之义。

4.2 中国数据规制的发展

1)数据滥用行为应纳入数据规制。如上文所述,当前的民事、行政、刑事法律规制对个人信息保护的设计局限于数据的非法收集与泄露之类的非法数据转移行为,难以约束数据泄露之后的滥用行为,这无疑是数据规制的漏洞,而个人数据的滥用已成为现实中比个人数据非法转移更严重、更普遍的问题[21]。因此,为了弥补当前的规制漏洞,实现对个人数据的全方位的保护,需要将数据滥用行为纳入数据规制。在制度的设计上,应注重民法、行政法、刑法的配合。数据保护的刑法规制主要体现于侵犯公民个人信息罪,该罪是典型的法定犯,因此,为了保证民事责任、行政责任与刑事责任的衔接,同样需要在行政法规中明确数据滥用行为的具

体特征,避免出现“违反国家规定”的解释困境。同时,为了保证数据流通与数据保护的平衡,应以“情节严重”保证刑事责任启动的谦抑。对“情节严重”的划分可参考当前司法解释对非法获取、出售或者提供公民个人信息行为“情节严重”的设置,既有行为对象的数量标准、也有行为的性质标准、后果标准以及再犯标准[22]。

2)引入场景理念修正较为僵硬的同意规则。苛刻的同意规则对于数据合理使用的阻碍不可小觑。虽然倡导数据规制的严格性更有益于公民个人隐私的保护,但数据权利并不只有私权的内涵,数据所具有的共享性以及数据在大数据时代所具有的战略性地位意味着数据保护并不必然产生私权和公共利益、私权与公权的冲突。因此,中国个人数据保护规制的设计应转变传统框架中知情同意的固化思维,认识到个人信息合理使用的判断标准取决于是否符合用户的合理隐私期待,以及是否造成了不合理的隐私风险,而并非僵化审视是否取得了当事人的同意[23]。数据风险源于在具体场景中数据被如何使用以及是否符合数据主体的合理期待。引入场景理念的合理内核,有赖于对个人信息的风险评估,当评估结果为对个人信息造成影响极小时,可不经个人同意采取相应处理,从而弱化数据使用合法性对用户同意的过度依赖。

3)推动数据去识别化标准的形成。中国《网络安全法》及最高人民法院和最高人民检察院司法解释已经认可去识别化的数据不受个人数据规制的约束,为数据流通和使用保留了空间。大数据时代,公民隐私观念的消解已成为必然,在公众平台分享自己的生活与心情近乎成为所有人的日常,个人数据的脱敏日益困难。既然绝对的匿名化并不现实,数据去识别化标准的明确就显得尤为重要。数据去识别化标准的明确有助于将个人数据安全地运用于经济社会,真正意义上地实现数据保护和数据使用的双赢[24]。尽管国外的立法也未规定具体的标准,但数据规制对数据脱敏的基本要求仍值得借鉴。欧盟数据匿名化的法律标准概括为不论是数据接收者还是其他任何人,采用一般合理手段,无法使匿名处理后的数据单独或结合其他数据识别具体个人或推断信息。美国的个人信息去识别化标准不强调数据的不能复原,强调通过风险评估判断数据接受者的再识别能力、动机等因素[25]。相对而言,欧盟的标准严苛于美国,这与不同国家公民的隐私观念的差异息息相关。鉴于中国的国情以及公民的隐私观,结合《个人信息保护法(草案)》的设想,中国的数据去识别化标准应以数据类型的不同而有所区分,对于敏感性的个人数据,对匿名化的要求必须达到不再识别个人的标准,彻底断开数据与具体对象之间的连接。对于个人关联性较弱的数据,匿名化的程度若能达到无法通过与其他数据的对比、分析识别出数据主体的身份即可。

5　结论

大数据技术在医疗领域的价值无可估量,但在其赋能健康医疗的同时,围绕数据使用与保护的议题所呈现的是公共利益和个人利益糅合与冲突的现实图景。中国的个人信息法律保护机制方兴未艾,患者隐私保护立法更是处于初级阶段,如何解决数据使用与个人隐私保护之间的紧张关系进而释放数据的潜在价值,是大数据时代医疗领域不可回避的挑战,也是中国信息立法亟待应对的现实问题。对此,必须建构适宜大数据时代背景的数据法律规制,在理性定位医疗大数据的基础上,弥合现行数据规制的不完备性,引入开放灵活的数据使用规则并推动数据去识别化标准的形成,既要保证数据使用的前景,也要维护公民个人隐私等合法权益的神圣不可侵犯。

参考文献(Reference)

[1] 齐爱民. 大数据时代个人信息保护法国际比较研究[M]. 北京: 法律出版社, 2015.

[2] 谭志明. 健康医疗大数据与人工智能[M]. 广州: 华南理工大学出版社, 2019.

[3] 李爱君. 中国大数据法治发展报告[M]. 北京: 法律出版社, 2018.

[4] 汪东升. 个人信息的刑法保护[M]. 北京: 法律出版社,

2019.
[5] 刘琪, 谷笑颖. 医疗人工智能应用中的伦理困境及对策研究[J]. 医学与哲学, 2019, 40(21): 5-8.
[6] 王融. 大数据时代数据保护与流动规则[M]. 北京: 人民邮电出版社, 2017.
[7] 徐慧丽. 大数据环境中个人医疗信息的法律保护[J]. 图书馆, 2019(11): 38-45.
[8] 吴裘弘. 个人信息的刑法保护研究[M]. 上海: 上海社会科学出版社, 2014.
[9] 叶良芳, 应家赟. 非法获取公民个人信息罪之"公民个人信息"的教义学阐释——以《刑事审判参考》第1009号案例为样本[J]. 浙江社会科学, 2016(4): 71-78, 157-158.
[10] 闫立, 吴何奇. 重大疫情治理中人工智能的价值属性与隐私风险——兼谈隐私保护的刑法路径[J]. 南京师大学报(社会科学版), 2020(2): 32-41.
[11] 于广军, 杨佳泓. 医疗大数据[M]. 上海: 上海科学技术出版社, 2015.
[12] 曾淑瑜. 医疗伦理与法律15讲[M]. 台北: 元照出版公司, 2010.
[13] 龚向前. 传染病控制国际法律问题研究[M]. 北京: 法律出版社, 2011.
[14] 京东法律研究院. 欧盟数据宪章:《一般数据保护条例》GDPR评述及实务指引[M]. 北京: 法律出版社, 2018.
[15] 郭瑜. 个人数据保护法研究[M]. 北京: 北京大学出版社, 2012.
[16] 李川. 个人信息犯罪的规制困境与对策完善——从大数据环境下滥用信息问题切入[J]. 中国刑事法杂志, 2019(5): 34-47.
[17] 田野. 大数据时代知情同意原则的困境与出路——以生物资料库的个人信息保护为例[J]. 法制与社会发展, 2018(6): 111-136.
[18] 刘红. 大数据时代数据保护法律研究[M]. 北京: 中国政法大学出版社, 2018.
[19] 王红凯, 龚小刚, 叶卫, 等. 大数据智能下数据脱敏的思考[J]. 科技导报, 2020, 38(3): 115-122.
[20] 涂子沛. 大数据[M]. 桂林: 广西师范大学出版社, 2015.
[21] 特伦斯·克雷格, 玛丽·E·卢德洛芙. 大数据与隐私: 利益博弈者、监管者和利益相关者[M]. 赵亮, 武青译. 沈阳: 东北大学出版社, 2016.
[22] 石聚航. 侵犯公民个人信息罪"情节严重"的法理重述[J]. 法学研究, 2018, 40(2): 62-75.
[23] 范为. 大数据时代个人信息保护的路径重构[J]. 环球法律评论, 2016, 38(5): 92-115.
[24] 张新宝. 从隐私到个人信息: 利益再衡量的理论与制度安排[J]. 中国法学, 2015(3): 38-59.
[25] 金耀. 个人数据匿名化法律标准明晰——以《网络安全法》第42条为中心[J]. 网络法律评论, 2016(2): 72-87.

Value, data risks, regulation of medical big data thinking: Based on epidemic management

WU Heqi

School of Law, Shanghai University of Finance and Economics, Shanghai 200433, China

Abstract As a national strategy, big data technology aims to benefit all people through the smart medical model. In response to the novel coronavirus pneumonia epidemic, medical big data technology embodies many values: effective scheduling of medical resources, AI medical treatment, prediction of outbreaks, monitoring of outbreaks, trends, etc. However, while big data technology becomes an important technical means in medical field, new problems and challenges in data risks also arise. It has become difficult for regulation design to coordinate conflicts between data protection and data use. This article sorts out the current status of personal medical data regulation, analyzes the shortcomings of data regulation in combination with medical practice, and proposes a construction logic of perfecting China's data regulation by drawing on foreign data regulation, that is, with the overall interests of society being the core, system design is carried out on the premise of balancing interest relationships and data abuse is included in data regulation. Scenario concepts are also introduced to modify consent rules and to promote the formation of data de-identification standards.

Keywords medical big data; public health emergency; data risk; legal regulation

(责任编辑 徐丽娇)