

• 医学信息组织与利用 •

国家人口与健康科学数据共享平台资源建设模式分析*

李赞梅 孙海霞

(中国医学科学院医学信息研究所/国家人口与健康科学数据共享平台工程技术中心 北京 100020)

〔摘要〕 总结国家人口与健康科学数据共享平台资源建设模式特点,分析平台资源建设在管理、标准化、规范化方面存在的主要问题,在此基础上提出思考和建议,指出平台应加快形成完善的资源评价与准入机制,积极推进标准化工作进程,增强科学数据开放共享意识。

〔关键词〕 国家人口与健康科学数据共享平台;资源建设模式;科学数据

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2015.10.016

Analysis of the Resources Construction Mode of the National Scientific Data Sharing Platform for Population and Health LI Zan-mei, SUN Hai-xia, Institute of Medical Information, Chinese Academy of Medical Sciences/ Engineering Center of National Scientific Data Sharing Platform for Population and Health Platform, Beijing 100020, China

〔Abstract〕 The paper summarizes features of the resources construction mode of the national scientific data sharing platform for population and health and analyzes main problems existing in the management, standardization and normalization of platform resources construction. On this basis, it proposes thinking and suggestion, and points out that a complete mechanism for resources assessment and access should be formed as soon as possible for the platform, the work progress of standardization should be promoted positively and the awareness of opening and sharing scientific data should be strengthened.

〔Keywords〕 National scientific data sharing platform for population and health; Resources construction model; Scientific data

1 引言

国家人口与健康科学数据共享工程是科技部国家科技基础条件平台科学数据共享工程重要试点项目之一,由国家卫生和计划生育委员会、解放军总

后卫生部、国家中医药管理局和国家食品药品监督管理总局 4 个主管部门联合推荐,中国医学科学院、中国疾病预防控制中心、解放军总医院、中国中医科学院、国家食品药品监督管理总局信息中心和军事医学科学院 6 家国家级科研院所联合申报创建^[1]。项目总体目标是按照统一标准规范、统一资源规划和统一技术架构,建立一个“逻辑上高度统一,开放共享;物理上合理分布,分工合作”的国家人口与健康科学数据管理与共享服务系统^[2],服务于科技创新、政府管理决策与医疗卫生事业发

〔修回日期〕 2015-07-06

〔作者简介〕 李赞梅,实习研究员,发表论文数篇。

〔基金项目〕 “国家人口与健康科学数据共享平台标准规范制定”研究成果之一。

展,为创新型医学人才培养和健康产业发展提供科学数据共享服务,从而提高我国医疗卫生服务整体水平和国际竞争力。

国家人口与健康科学数据共享工程承担着国家科技重大专项、科技计划、重大公益专项等人口健康领域科学数据汇总、数据加工、数据存储、数据挖掘和共享服务的任务。自 2004 年正式启动后,经过 10 余年的发展,国家人口与健康科学数据共享工程建设成果突出,已通过国家人口与健康科学数据共享平台(以下简称人口健康平台)^[3]为全社会提供人口健康科学数据共享与支撑服务。科学数据资源建设是人口健康平台资源共享与服务的基础,是平台建设的重要内容。本文对人口健康平台联合共建共享资源建设模式特点做了简要总结,同时指出平台资源建设面临的主要问题,提出几点建议,旨在为国家人口与健康科学数据共享发展提供参考。

2 人口健康平台资源建设模式特点

2.1 概述

人口健康科学数据资源分布在行业内业务、科研、生产、管理等多个环节,数据类型复杂多样、学科差异性大且持续发展变化。任何一家单位都很难满足所有类型的资源建设与共享服务需求,必须通过各个层次、各种形式的分工协作,联合人口健康领域内各科学数据资源优势单位共同参与,才能有效地推动与实现人口健康平台资源建设工作进展与目标。因此,联合多家单位共建优质资源实现科学数据共享是人口健康平台资源建设与共享的必然选择。

2.2 全面覆盖,重点突出

人口健康平台资源建设坚持“预防与治疗结合、基础与临床结合、中西医结合、军队与地方结合”的布局方针^[4],旨在构建全面覆盖基础、临

床、药物、公共卫生、疾病防控、人口、计划生育、生殖健康等各类资源的人口与健康科学数据管理与共享服务体系,将国家或地方科技计划项目所产生的科学数据、国家长期布局的科学数据以及权威机构或科学家个人所拥有的科学数据作为先行建设重点资源。据调查,仅 1978–2009 年间,人口健康领域各类科技成果、科研项目累计已超 7 万余项^[5]。面对如此庞大的资源建设前景与需求,人口健康平台资源建设经验与力量相对有限,必须抓重点、分批次、讲实效,逐步推动平台资源建设。为此,依据国家人口健康行业发展趋势,同时充分考虑用户需求,平台优先确立了 6 大资源建设重点领域:(1) 以健康与疾病为主题的基础医学领域。(2) 以临床流行病学调查、临床科学研究及临床实践活动为主旨的临床医学领域。(3) 以全国传染性疾病预防及防治为主旨的公共卫生领域。(4) 为实现食品药品科学监管及医药行业健康发展提供信息服务的药学领域。(5) 中医学领域。(6) 人口与生殖健康领域。

2.3 协同共建,优势互补

人口健康平台依据 6 大资源建设重点领域设立了相应的科学数据中心,见图 1。各中心均依托领域内权威的职能部门或资源优势单位,在平台统一协调组织下,由各中心负责各重点领域科学数据资源建设与管理任务。这种联合各领域优势力量共建国家人口与健康科学数据共享资源体系的模式,既较好地解决了资源隶属与共享之间的矛盾,拓宽了资源渠道,而且还能有效避免资源重复建设造成的各种浪费。除了重点领域间的协同共建外,重点领域内资源建设也存在着联合共建。如药学数据中心选择了国家食品药品监督管理局信息中心及中国医学科学院药物所两家单位共同承担资源建设任务,以充分发挥两家单位在药物管理、药品研发领域优势,保证药学领域资源建设的完整性。

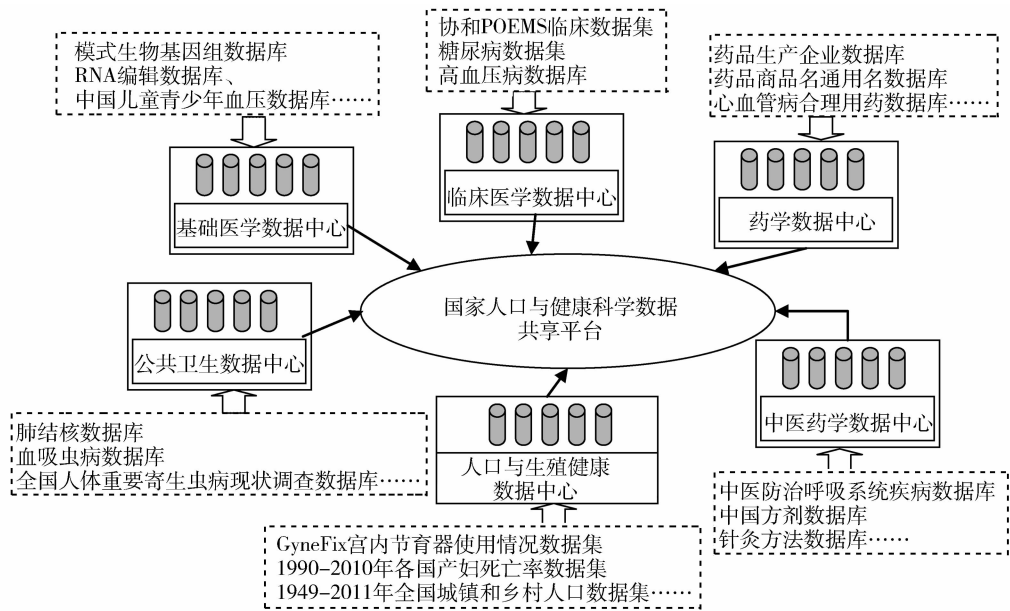


图1 人口健康平台重点领域资源联合共建共享体系

2.4 统一标准，有效共享

联合行业内优势力量共建国家人口与健康科学数据资源共享与保障体系模式的成功建立，既依赖于各资源建设参与方的积极联合，同时还需有效地标准化以保证各类资源建设的规范性，最大限度实现资源共享。平台资源主要来源于国家及地方科研项目产出的科学数据，最初主要是为项目本身服务，标准化程度普遍较低。为此，平台特设专门的标准化工作小组，经过多年的不断探索与长期实践，形成了较为全面、完善的资源建设标准规范体系，制定数据类、管理类、服务类、质量控制、安全类等资源建设有关标准 10 余项。其中，《医药卫生科学数据共享网数据元标准化原则与方法》、《医药卫生科学数据共享网数据集分类与编码原则与方法》、《医药卫生科学数据共享网数据集元数据标准》及《医药卫生科学数据共享网数据模式描述规则与方法》4 项标准已经推广为卫生行业标准^[6]。

3 人口健康平台资源建设存在的问题

3.1 概述

截至 2014 年底，平台 6 个重点领域已建设可共

享数据集 728 个，数据量超 30 TB。资源覆盖领域全面，其中人口统计类资源数据最多，预防医学与卫生学、内科学、中国医学、药学等领域资源数量也颇具规模，各类资源基本覆盖中图法医学学科分类^[7]的所有 2 级类目（表 1），形成了较为完备的国家人口与健康科学数据共享资源体系。但随着工程的持续推进，资源越来越丰富，用户需求日益多样化，平台资源建设存在的问题也显露出来。

表 1 人口健康平台资源分布情况

2 级类目	资源数目 (个)	2 级类目	资源数目 (个)
R1 预防医学、卫生学	155	R73 肿瘤学	3
R2 中国医学	47	R74 神经病学与精神病学	3
R3 基础医学	21	R75 皮肤病学	3
R4 临床医学	9	R78 口腔科学	2
R5 内科学	59	R9 药学	36
R6 外科学	4	R 医药、卫生	5
R71 妇产科学	2	C92 人口学	376
R72 儿科学	3		

3.2 资源建设管理工作有待加强

人口健康平台是基于项目协作的跨部门、跨机构科学数据联合共建共享工程^[8]，涉及面广、参与单位众多。为保证项目的科学管理和有效运行，平台成立理事会及平台管理中心，全面统筹平台发展

战略及运行管理等各项工作,在总平台及分中心设置专门的资源建设小组,负责各项资源建设与管理。平台虽已具备较好的资源建设组织体系及人才队伍,但迄今尚未形成上下配套的资源建设管理机制,以下几点表现尤为突出:(1)资源评价体系缺失。平台现有资源质量层次不齐,急需科学合理的资源评价体系指引资源质量的改进。(2)新资源发现、遴选与准入机制不明晰。平台资源与服务面临的需求随人口健康行业及社会经济发展而不断变化,只有对此做出及时响应,具备准确、及时发现、筛选、建设热点资源的机制与能力,才能不断提高平台的社会影响力。(3)资源更新维护不及时。平台现有资源中,有明确更新维护方案的资源仅 200 余个,尚不足总资源数的 1/3,这可能会对平台资源共享与服务长期可持续发展造成不利影响。

3.3 资源建设标准化、规范化程度有待进一步提高

与我国边实施科学数据共享工程边开展有关政策研制情况相似,人口健康平台各项标准规范也是从无到有、随资源建设而逐步制定实施与完善的。资源建设标准化工作长期以来一直受到平台各级的高度重视。除平台标准化工作小组发布的人口健康领域通用标准规范外,各重点领域也纷纷研制本领域专用标准规范。尽管如此,平台资源建设标准化工作仍存在一些不尽人意之处:(1)部分标准的可用性差。如资源核心元数据标准,长期以来都只能用于描述资源集本身,而不适用于对资源具体数据内容的揭示与描述。(2)相关标准的缺失。平台已为国家及地方众多科研项目提供数据支持,但如何对引用的资源进行标识尚缺乏统一的规范。(3)标准宣传推广力度不足。2014 年底,平台工程技术中心联合平台管理中心对各资源建设单位标准化工作有关情况进行调查,发现部分单位很少甚至从未参加过标准化相关知识的培训^[9]。

4 对人口健康平台资源建设的思考与建议

4.1 加强监管,加快制定切实可行的资源评价与准入机制

权威、主动的组织与管理是人口健康平台各项

活动有序开展的重要保障,是平台资源共建共享健康发展的基础。平台需加强建立健全科学化、规范化、制度化的资源建设监督与管理机制,加强资源形成、成长、成熟、衰亡过程的全链条式管理,形成上下配套的资源管理办法。当前,可以现有资源质量改进及新增资源质量把关两个维度为切入点:一方面加快资源评价与考核机制的制定,以此为基础对现有资源进行全面客观评定,对评价出的优质资源进行适度表彰与推广,对评价结果较差的资源提出有效的整改方案;另一方面要将新资源遴选与准入办法制度化、条例化,从源头上保证资源质量。

4.2 全面动员,积极推动平台资源建设有关标准的研制、修订和贯彻

面对平台标准化工作存在的问题,可以采取针对性措施:一是重点标准修订先行,加紧资源核心元数据、元数据注册与管理、科学数据资源分类与编码等资源内容与管理类标准的修订,可通过选取试点数据集运行修订标准,试运行过程中不断对修订标准进行验证,提高修订标准的适用性,完善后继而推广到平台所有资源中去;二是加快有关标准的制定,如科学数据资源引用增加和数据推广服务需求提出,使得“资源唯一标识符”规则的制定变得更为迫切;三是加强标准的宣传、培训与推广,避免走入“为了标准而做标准”的误区,可从修订标准试运行过程中探索适合平台的标准宣贯路径,确保各项标准的有效落实。但同时也要注意,标准化是一项长期工作,标准化虽能促进平台资源建设有关问题的解决但并非提供全部解决方案,需与管理、技术等措施有效配合,才能充分发挥其在推动平台资源建设与服务中的作用。

4.3 开放思想,增强科学数据资源共享意识

人口健康平台资源建设中存在着重视共建、轻视共享的思想,资源开放共享意识还较为薄弱,导致部分耗资巨大建设的资源只能“孤芳自赏”而不能有效共享,这在一定程度上对平台可持续发展产生不良影响。在大多数人的观念里,开放共享就意味着免费,实则不然。科学数据开放共享可以存在

着公益性运行模式、商业性运行模式亦或二者兼而有之^[10]。平台各参与方应当摒弃在单位内做资源的办法,采取积极措施,主动应对资源开放共享环境,在资源开放共享有关政策、技术、经济、组织等保障下,寻找自身资源共享的最佳运行模式,共同培养可持续发展的国家人口与健康科学数据共享文化。

5 结语

国家人口与健康科学数据共享工程是多部门联合的跨机构、跨学科国家级大型工程建设项目,目的是整合国家人口健康领域优质可共享资源,最终建成覆盖全国、联接世界的人口健康科学数据资源中心与共享服务平台。国家人口与健康科学数据共享平台经过长期实践和不断调整,在资源建设方面已取得了突出的成效。面对资源建设过程中存在的质量层次不齐、新资源发现不及时、资源评价考核机制不明确等问题,平台应从加紧形成完善的资源建设管理机制,积极推进标准化工作进程,加强各方的配合,增强科学数据开放共享意识,使资源建设与共享服务协调发展,将平台资源建设工作提到一个新的高度,充分发挥国家人口与健康科学数据资源在促进社会发展与进步中的价值。

参考文献

- 1 刘德培,尹岭. 医药卫生科学数据共享 [J]. 中国基础科学, 2007, 9 (1): 26-28.
 - 2 赵红,尹岭,王建国,等. 国家医药卫生领域科学数据共享网建设目的与现状 [J]. 中国中医药现代远程教育, 2008, 6 (6): 570-572.
 - 3 国家人口与健康科学数据共享平台 [EB/OL]. [2014-12-16]. <http://www.ncmi.rrrcn/1>.
 - 4 胡凯,王建国,尹岭,等. 浅谈医药卫生科学数据共享网工程建设 [J]. 中国科技资源导刊, 2008, 40 (1): 56-59.
 - 5 国家人口与健康科学数据资源手册 [EB/OL]. [2015-03-15]. <http://www.rrrcnmi.rrrcn/UploadFile/a/c/d1ebfd56307203297a02bb52cb647eca.rrrcpdf>
 - 6 已经推广为卫生行业标准的共享数据类标准 [EB/OL]. [2015-04-06]. <http://www.rrrcnmi.rrrcn/info/108/1046>.
 - 7 林美兰. 中国图书馆分类法专业分类表系列·医学专业分类表 [M]. 北京:北京图书馆出版社, 1999.
 - 8 科学数据共建共享模式探讨 [J]. 安徽农业科学, 2009, 37 (29): 14496-14498, 14501.
 - 9 国家人口健康数据共享平台资源建设与标准规范工作研讨会 [EB/OL]. [2015-04-26]. <http://www.rrrcnmi.rrrcn/info/50/3550>.
 - 10 王晴. 论科学数据开放共享的运行模式、保障机制及优化策略 [J]. 国家图书馆学刊, 2014, (1): 3-9
-
- (上接第 71 页)
- 15 Ke J Y, Kliever K L, Hamad E M, et al. The Flavonoid, Naringenin, Decreases Adipose Tissue Mass and Attenuates Ovariectomy - associated Metabolic disturbances in Mice [J]. Nutrition & Metabolism, 2015, (12): 1.
 - 16 Tu TH, Kim CS, Kang JH, et al. Levels of 4-1BB transcripts and soluble 4-1BB protein are elevated in the adipose tissue of human obese subjects and are associated with inflammatory and metabolic parameters [J]. International Journal of Obesity, 2014, 38 (8): 1075-1082.
 - 17 Ng A C, Wai D C, Tai E S, et al. Visceral adipose tissue, but not waist circumference is A Better Measure of Metabolic Risk in Singaporean Chinese and Indian Men [J]. Nutrition & Diabetes, 2012, (2): e38.
 - 18 Kolehmainen M, Ulven S M, Paananen J, et al. Healthy Nordic Diet Downregulates the Expression of Genes Involved in Inflammation in Subcutaneous Adipose Tissue in Individuals with Features of the Metabolic Syndrome [J]. American Journal of Clinical Nutrition, 2015, 101 (1): 228-239.
 - 19 Lee Y J, Woo S Y, Ahn J H, et al. Health - Related Quality of Life in Adults with Metabolic Syndrome; the Korea National Health and Nutrition Examination Survey, 2007 - 2008 [J]. Annals of Nutrition and Metabolism, 2012, 61 (4): 275-280.