

浅谈医药卫生科学数据 共享网工程建设

胡 凯¹ 王建国¹ 尹 岭² 赵 红¹

(1. 中国医学科学院, 北京 100730;

2. 中国人民解放军总医院, 北京 100853)

摘 要: 医药卫生科学数据共享网工程是国家科技基础条件平台建设、国家科学数据共享工程的重大项目。该项目于2004年4月正式启动, 由我国医药卫生领域多家权威科研机构联合承担建设。该项目的目标是建立一个“在物理上合理分布、层次分明, 在逻辑上高度统一、充分共享”的医药卫生科学数据管理与共享服务体系。在统一的管理协调、资源规划、标准制订和平台建设的基础上, 该共享网通过各专业数据分中心对科学数据资源进行收集与整合。该项目已经开展了共享服务并获得了一定的成效与领域影响。本文重点介绍了医药卫生科学数据共享网工程建设的情况。

关键词: 数据共享; 医学信息; 信息资源

中图分类号: G203 文献标识码: A DOI: 10.3772/j.issn.1674-1544.2008.01.011

Brief Introduction to the Project Construction of Medical Science Data Sharing Network

Hu Kai¹, Wang Jianguo¹, Yin Ling², Zhao Hong¹

(1. Chinese Academy of Medical Science, Beijing 100730;

2. The General Hospital of PLA, Beijing 100853)

Abstract: Medical Scientific Data Sharing Network is an important project of National Scientific Data Sharing Project, which belongs to National Science & Technology Infrastructure. The project is jointly constructed by several authoritative medical and health research academies of China. The goal of the project is to build an administrative and service sharing system which is physically well-distributed, clearly layered, logically consistent and fully shared. Based on the unitive management and cooperation, resource programming, standard compiling and platform construction, the sharing network implements the collection and integration of scientific data resource through specialty data centers. The project has started the sharing service and obtained some effect and influence in medical field.

Keywords: data sharing, medical information, information resource

医药卫生科学数据共享网工程是国家科技基础条件平台建设、国家科学数据共享工程的重大项目。该项目由国家卫生部、解放军总后卫生部、

国家中医药管理局、国家食品药品监督管理局4个主管部门联合推荐, 并由中国医学科学院、中国疾病预防控制中心、中国人民解放军总医院、中国

第一作者简介: 胡凯(1976-), 男, 河南洛阳人, 硕士, 主要研究方向是医学信息。

基金项目: 国家科技基础条件平台建设项目“医药卫生科学数据共享网”(2005DKA32400)。

收稿日期: 2007年10月30日。

中医科学院、国家食品药品监督管理局信息中心、中国人民解放军军事医学科学院六大科研院所联合申报,于2003年在科技部立项,2004年4月正式启动。医药卫生科学数据共享网工程的总体目标是建立一个“在物理上合理分布、层次分明,在逻辑上高度统一、充分共享”的医药卫生科学数据管理与共享服务网,通过共享服务网总中心、各级数据中心和共享节点的建设,实现医药卫生领域跨学科、跨部门的科学数据共享。该项目的特点是“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”,其目的是满足政府卫生决策、医药卫生科研业务工作、社会公众卫生服务以及健康产业发展的信息需求,达到国家科技基础条件平台科学数据共享工程有关科学数据共享服务网的建设标准。

1 建设原则

1.1 统筹规划 统一标准

根据国家科技基础条件平台建设和科学数据共享工程的总体规划,加强领域层面的整体性宏观管理,统一部署,统一规划,统一标准,加快医药卫生领域数据共享的标准化、规范化建设。

1.2 联合建设 资源共享

该项目由4个主管部门6家单位联合共建,是一个跨学科、跨行业、多部门资源整合的综合性

建设工程,建设的关键与优势在于多个部门的联合与优势互补,多专业、多学科的数据资源整合。

1.3 需求主导 保障安全

以国家和社会对医药卫生科学数据资源的迫切需求为导向,以全面提升国家医药卫生领域的专业水平和全民健康水平为目的,在保证数据提供者知识产权和个人隐私的前提下,最大限度地实现医药卫生科学数据共享。

1.4 先易后难 稳步推进

基于不同学科和专业对医药卫生科学数据资源的需求以及发展的不平衡性,优先选择条件成熟或需求紧迫的学科和专业进行试点,边建设边共享,边整合边服务,讲求实效,稳步推进。

2 共享网基本架构

医药卫生科学数据共享网的建设目标是要最终实现一个由“三级网络”(总中心、专业数据中心和共享服务节点)和“两级平台”(总平台、专业数据中心平台)的共享服务体系架构,形成一个集资源整合与资源发布于一体、层次分明的服务体系。整个共享网建设的基本框架由共享网总中心、数据中心(目前包括基础医学、临床医学、公共卫生、中医药学和药理学5个数据中心)和共享网地方服务节点等3个方面的建设构成。整个共享网的工程建设框架见图1。

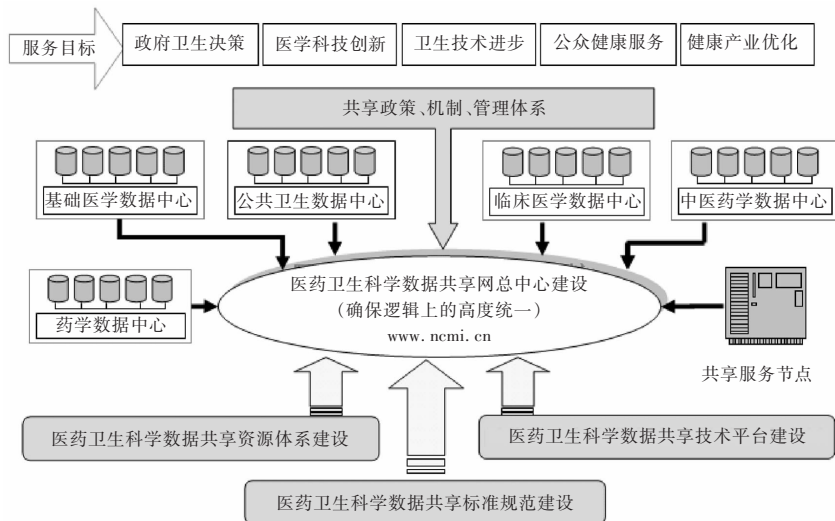


图1 医药卫生科学数据共享网的工程建设框架

2.1 共享网总中心建设

医药卫生科学数据共享网是一个跨学科、跨单位,有多个部委和单位联合参与的工程建设项目。为了保证整个工程建设真正实现“逻辑上的高度统一”,就必须建立起以共享政策与机制研究为指导,包括共享资源体系、标准规范体系和技术平台体系建设的共享网总中心实体。

(1) 共享政策与运行机制的研究制定:负责共享政策与运行机制的制定与实施,建立医药卫生领域的数据共享管理机制,保证整个系统建设的整体协调、规范统一。

(2) 共享资源体系的建设:在了解和分析领域资源情况的基础上建立资源规划与管理体系,研究制定相关规范与机制,对领域层面和各个数据中心的共享资源进行合理规划,保证共享资源建设的规范化与可持续发展。目前,已经完成了资源工作规范、数据集制作规范等文本的征求意见方案,正在进行医药卫生科学数据共享资源手册的编写以及相关工具软件的设计。

(3) 标准规范体系的建设:落实和宣传贯彻科技部标准,在建立领域标准规范体系的基础上进行领域专用标准的研究制订,整合领域现有标准,并进行指导、培训和应用。目前,已经完成了关于元数据、数据元、分类与编码、数据模式等标准文本的征求意见稿,同时还在开展相关工具软件的设计与开发。

(4) 技术平台体系的建设:为整个医药卫生共享网提供技术和软硬件环境支撑的网络平台,通过技术开发与应用,保证数据共享各项服务的进行,并实现对数据中心和共享节点的管理与支持。目前,正在进行总中心网站的进一步设计与开发。

2.2 数据资源建设

医药卫生科学数据共享工程的数据中心建设包括基础医学、临床医学、公共卫生、中医药学、药学、特种医学等专业数据中心或主体数据库的建设:

(1) 基础医学数据中心:由中国医学科学院基础所承担建设任务,信息资源的内容主要包括医

学生物学、人体机能、医学相关学科以及人体解剖等。

(2) 临床医学数据中心:由北京协和医院和解放军总医院联合承担建设任务,信息资源的内容主要包括临床综合信息、临床疾病诊疗、康复医学、临床护理以及医院管理等。

(3) 公共卫生数据中心:由中国疾病预防控制中心承担建设任务,信息资源的内容主要包括疾病预防控制、公共卫生服务、公共卫生管理、公共卫生监督等。

(4) 中医药学数据中心:由中国中医科学院承担建设任务,信息资源的内容主要包括中医药事业、中医信息、中药信息、针灸信息以及中医古籍等。

(5) 药学数据中心:由中国医学科学院药物研究所和国家食品药品监督管理局信息中心联合承担建设任务,信息资源的内容主要包括药物管理、药物资源、药物生产等。

(6) 共享网地方服务节点:已经启动湖南社区、河北农村、北京老年和西部少数民族试点。目的是将医药卫生科学数据资源服务于全国各级医疗机构和边远地区。整合具有区域特征的医学数据资源,包括疾病谱数据集、急诊系列数据集和卫生经济学数据集等。为政府卫生决策、医疗、保健、科研、教学和医疗保险提供数据共享和信息服务。

3 建设思路

3.1 搞好资源规划 明确共享资源域

在医药卫生科学数据资源调查的基础上进行数据资源规划,重点解决医药卫生数据共享资源域的界定与划分问题,明确哪些数据是属于部门的业务数据,哪些数据是可以共享的数据,从而避免因知识产权和保密问题而出现共享系统中的数据不能共享的现象。

3.2 按照需求 分步实施

首先,围绕需求从政府卫生管理、领域业务、科研工作、公众卫生服务以及健康产业等多个需

要角度进行顶层设计,以此作为系统建设的最终目标。然后,根据社会关注、共享需求、数据质量和利用价值等因素确定系统建设的重点与顺序以及每个年度和每个建设阶段的主要任务,从而实现围绕需求,分步实施。

3.3 逻辑上高度统一的一站式服务

医药卫生科学数据共享网的门户网站作为整个工程“逻辑上高度统一”的集中体现,不仅能够实现基于元数据的目录服务,还可以根据用户提出的各种数据服务请求实现对各个物理中心数据资源发现与获取的一站式服务,使用户通过一个门户网站获得整体性和一体化的服务效果。

3.4 政策法规保障体系建设

遵循科技部有关科技基础条件平台建设科学数据共享工程的有关规范和文件,起草中国医药卫生科学数据共享的有关政策法规与运行管理办法,准备在广泛征求意见后修改实施,最终形成规范和保障共享建设和运行的指导性文件。

3.5 建立健全项目管理和运行机制

在医药卫生科学数据共享工程建设中,充分发挥各级组织机构的职能,引入项目管理制度,做到任务分解到人,经费落实到位,考核指标明确。对项目进行定期评估和验收,实现全过程的监督管理。

4 建设成果

整个医药卫生科学数据共享网的物理平台建设主要包括共享网总中心门户和5个专业数据中心的建设。其中,数据中心的门户平台主要进行专业范围内的数据资源整合、汇交与服务工作,总中心门户网站主要作为集中提供一站式服务的数据资源汇交与组织中心,并且针对医药卫生科学数据共享建设发挥新闻发布平台和资源与交流平台等作用。

目前,医药卫生科学数据共享网工程主要通

过总中心门户(www.ncmi.cn)和各个专业数据中心的网站进行共享服务。截止到2007年上半年,已经整合、改造并实现共享的数据集有58个,可共享的数据总量达到2255GB。在线下载数据累计达到1400GB,离线提供数据累计达到6016GB。已经共享的数据相当于盘活了以往3.8亿元的国家投入。目前还引进和改造了83.5GB的国际数据资源。

共享的数据内容支持了领域内的众多科研项目,其中包括29个863项目、21个973项目、53个国家自然科学基金项目、24个国家科技攻关项目、30个国家其他科技计划项目和87个部门和地方重点科技项目。

通过数据共享建设和共享服务的持续提供与发展,参与医药卫生科学数据共享建设的各个部门已经建立了一支面向共享需求的专业性人才队伍。目前,参与医药卫生科学数据共享建设研究的科研与业务单位达到76所,形成了总数达到数百人的共享建设队伍。这些共享专业队伍的知识结构由医学、药学、信息学、计算机科学等构成,一些参与工作的人员在建设过程中已经形成了复合型的知识结构。

参考文献

- [1] 中国生物技术信息网. 刘德培院士谈我国医学数据共享系统建设进展[EB]. [2004-04-27]. <http://www.biotech.org.cn/news/news/show.php?id=14088>.
- [2] 尹爱宁. 中医药学科学数据共享时不我待[N]. 中国医药报,2006-01-12(A7).
- [3] 科学数据共享工程规划课题组. 科学数据共享工程发展规划(2005-2010年)[R]. 北京: 中华人民共和国科技部, 2005.
- [4] 王正兴, 刘闯. 美国国有信息共享的法律基础[J]. 图书情报工作, 2002(6): 60-63.
- [5] 刘闯, 王正兴. 美国全球变化数据共享的经历对我国数据共享决策的启示[J]. 地球科学进展, 2002, 17(1): 151-157.
- [6] 刘闯, 王正兴. 数据开放带来国家振兴-美国国有科学数据“完全与开放”共享国策剖析[N]. 科学时报, 2002-02-07.
- [7] 中华人民共和国科技部. 科学数据共享工程概念与术语第一部分: 概念[EB]. [2004-11-27]. <http://www.sciencedata.cn/pdf/2.pdf>.