

# 健康医疗大数据信息资源目录体系构建\*

姜 伟 于天贵

苟延农

(山东省精神卫生中心 济南 250014) (山东省卫生健康委员会医疗管理服务中心 济南 250014)

迟蔚蔚 李 磊

韩加亮

(山东省卫生健康委员会规划发展与信息化处 济南 250014) (山东省立第三医院 济南 250002)

〔摘要〕 通过研究山东省健康医疗大数据信息资源目录体系架构、编制流程等方面,提出符合山东省实际的信息资源目录编制方法和架构、共享机制和实际应用。

〔关键词〕 健康医疗大数据;资源目录;共享

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2021.03.004

**Building of Health and Medical Big Data Information Resource Catalog System** JIANG Wei, YU Tianguai, Information Center, Shandong Mental Health Center, Jinan 250014, China; GOU Yannong, Medical Administration Service Center of Health Commission of Shandong Province, Jinan 250014, China; CHI Weiwei, LI Lei, Plan and Information Division, Health Commission of Shandong Province, Jinan 250014, China; HAN Jiali, Shandong Provincial Third Hospital, Jinan 250002, China

〔Abstract〕 Through the study of the architecture, compilation process and other aspects of the health and medical big data information resource catalog system in Shandong province, the paper puts forward the compilation method and architecture, sharing mechanism and practical application of the information resource catalog in line with the actual situation of Shandong province.

〔Keywords〕 health and medical big data; resource catalog; sharing

## 1 引言

〔收稿日期〕 2020-06-11

〔作者简介〕 姜伟,高级工程师,发表论文4篇,参编论著1部;通讯作者:于天贵,教授。

〔基金项目〕 国家卫生健康委规划司委托课题“山东省人口健康信息资源共享目录体系研究”。

国务院办公厅《关于促进和规范健康医疗大数据应用发展的指导意见》(国办发〔2016〕47号)明确要求建立全国健康医疗数据资源目录体系,制定分类、分级、分域健康医疗大数据开放应用政策规范。按照一定分类架构和标准,序化形成的信息资源管理、服务与共享的信息组织形式构成信息资源目录体系,在一定程度上解决“有哪些信息资源”、“信息资源在哪里”、“如何获取信息资源”以及信息资源管理和应用问题<sup>[1-3]</sup>。本文通过研究山东省健康医疗大数据信息资源目录体系,提出符

合实际的信息资源目录编制方法和架构、共享机制和实际应用。

## 2 研究方法

### 2.1 编制思路

按照公共服务驱动、政务管理、应用需求、问题导向并融合数据治理方法论，强化信息标准的原则进行设计，通过共享目录形式发布共享数据。采用信息资源规划、建设主题数据库的整合路径，以信息资源为主线，数据省级大集中方式。全面梳理山东省卫生健康业务和信息资源，基于统一元数据、信息分类等标准规范对信息资源进行描述并编制目录，建立科学的信息资源分类体系。通过构建山东省健康医疗大数据信息资源目录体系，编制信息资源共享目录，对信息资源进行注册登记，规范管理数据采集、处理、存储、公开、共享、服务和应用等环节，掌握数据来源、数量、质量、类型以及分布等具体情况。

### 2.2 体系架构

目录体系架构由信息资源、元数据、资源目录<sup>[4-5]</sup>、共享目录、标准及治理体系、目录编目和交换体系等要素构成，见图1。元数据是描述健康医疗信息内容、管理、表示等属性的数据，核心元数据是健康医疗大数据信息资源元数据的基本子集。通过对卫生健康职能、业务和信息资源进行梳理形成信息资源目录，用元数据进行描述。进一步封装为共享目录，编目对象主要涉及业务协同和动态的数据交换。目录编目、交换体系包括目录管理体系和编目、交换工具。健康医疗信息标准及治理体系包括核心元数据、信息资源分类、标识编码等相关标准。山东省卫生健康信息平台纵向联通国家平台、全省16个市、137个县（市、区），所有三级医院、395家二级医院和1598家基层医疗机构，

全面建成基础资源、全员人口、居民电子健康档案、电子病历4大数据库并实现共享。数据治理以国际数据管理协会（The Global Data Management Community, DAMA）主流<sup>[6]</sup>数据治理方法论、模型为指导，制定数据标准化及治理框架，应用到目录体系研究实践中。数据治理体系包括登记、采集、集中建设、共享与交换、安全与管理、存储、存档与销毁、信息资源再利用等环节，以及相关政策与制度、管理体制等外部条件。

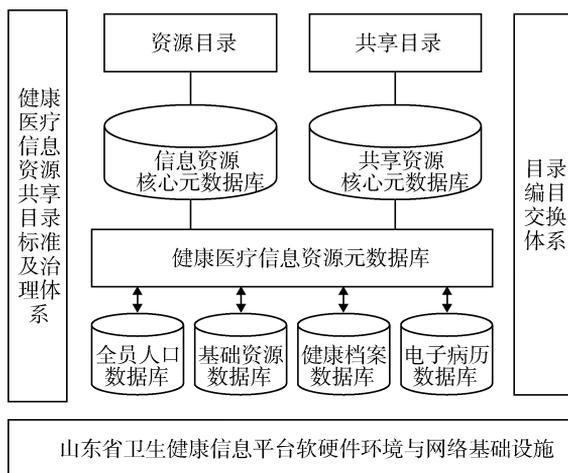


图1 山东省健康医疗大数据信息资源共享目录体系架构

### 2.3 编制流程

包括目录规划、资源调查和目录生成3个阶段<sup>[7]</sup>，见图2。分别从业务梳理、已建信息系统以及两者相结合入手，掌握信息资源现状。汇总、分析、提取业务调查表及信息资源调查表信息，生成山东省健康医疗大数据信息资源目录、共享目录。目录规划阶段制定目录编制工作方案，明确目录编制范围和参与方等。资源调查阶段开展业务及资源情况调查，掌握卫生计生信息资源现状。目录生成阶段对调查内容进行整理、分析、合并，生成业务域政务信息资源目录、省级人口健康信息资源目录和省级人口健康信息资源共享目录。

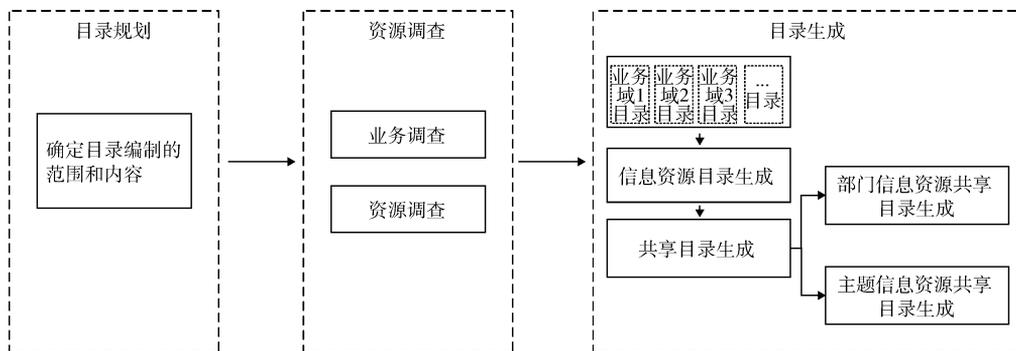


图2 目录编制流程

### 3 建立信息资源目录体系

#### 3.1 开展业务和信息资源调查与梳理

按照山东省卫生健康委“三定方案”职责，从业务梳理入手进行信息资源调查。开展全省信息化专项调查，从已建应用系统入手进行信息资源调查。基于山东省卫生健康信息平台全员人口、基础资源、电子健康档案、电子病历4大资源库，完成健康医疗大数据信息资源梳理。

#### 3.2 建设信息资源目录编目、交换体系

目录编目、交换工具属于健康医疗大数据信息资源范畴，以业务为核心，采用元数据对共享信息资源进行描述，形成统一规范的目录内容并进行管理，以服务方式对外发布，实现横向集成、纵向贯通、全局共享的信息资源目录。信息资源编目是指根据信息资源类型和分类，按照规范的元数据，描述某个具体信息资源特征，形成目录内容的过程。编目过程包括数据源与数据标准维护、资源注册、资源管理、数据监控、资源权限开放等。目录交换主要为服务资源目录，负责服务集中管理展示，方便用户根据分类查找服务。

#### 3.3 构建标准及治理体系

明确目录编码标准和分类模型，以山东省地方标准《人口健康信息化标准体系》(DB37/T 2759 - 2016)为基础建设标准体系，全省统一疾病分类与代码、手术(操作)分类编码、药品编码、检查检

验等项目编码。形成数据治理体系，实现信息资源全生命周期管理。目录分类由类目、亚目两个层次组成，将类目分为4大类：全员人口、基础资源、健康档案和电子病历。亚目是根据各类目领域特点，按照业务内容组成部分或流程先后顺序进行分类，对每个类目增加保留项。亚目通过面向卫生健康业务开展资源调查，对数据进行清洗、拆分、归并后逐一规范化命名形成。

#### 3.4 形成山东省健康医疗大数据信息资源目录

形成全员人口、基础资源、健康档案、电子病历4个类目、220个亚目、累计8 548个数据项，见表1。形成部门信息资源共享目录，结合重点应用实现跨部门、跨地域、跨层级业务协同。

表1 信息资源目录统计(个)

分类	亚目数量	数据项数量
全员人口	4	132
基础资源信息	51	1 499
电子健康档案	91	4 208
电子病历	74	2 709
总计	220	8 548

#### 3.5 形成山东省健康医疗大数据信息资源共享目录

基于山东省健康医疗大数据信息资源目录，形成山东省健康医疗大数据信息资源共享目录。建立面向全省居民全方位、全生命周期的大健康信息服务体系，覆盖医疗资源、公共卫生、计划生育、医疗服务、药品管理、医药人才等信息资源的医疗健

康主题库,实现全省健康医疗大数据信息资源统一汇聚、管理和按需使用,提高卫生健康服务可及性、精准化。

## 4 信息资源共享机制及应用

### 4.1 建立分级分类利用共享机制

建立山东省健康医疗大数据信息资源共享规范,确定实施和责任主体以及信息资源管理和服务、安全管理和监督保障等方面内容。推动健康医疗大数据信息资源科学配置和有效利用,提升卫生健康管理和服务水平,确保健康医疗安全性,保护个人隐私。

### 4.2 实现信息资源“看得见”

通过目录体系建立数据分类、定义和管理规范,理清数据分布,保证信息资源可见。为数据应用提供基础,通过物理模型及逻辑模型梳理,将卫生健康各条线系统数据打上对应业务含义的标签,为健康医疗大数据应用提供支撑。通过梳理业务,现已打通全员人口、基础资源、电子健康档案、电子病历4大资源库与各业务数据库的映射关系,为保障业务流程规范化提供基础保障。

### 4.3 实现信息资源“管得住”

建立目录体系组织架构,制定数据治理管理制度和规范,保证平台信息资源安全可控。通过山东省全民健康信息平台实现全省卫生健康系统内部服务开放,已实现数据开放222项,包括人口基本信息、出生医学证明、住院分娩等。利用山东省政务共享交换平台,共享出生医学证明、中小學生体检等360项信息,实现与其他各部门信息共享。

### 4.4 实现信息资源“用得好”

通过建立和实施目录体系,全面提升数据质量,奠定数据应用基础,促进数据加工应用。在临床科研方面,基于山东省全民健康信息平台的全员

人口、基础资源、电子健康档案和电子病历4大资源库,建立超大规模人群队列,开展山东省全人群全生命周期健康应用研究。在医疗管理方面,实现医疗服务监管部分指标实时监管,建立医院门诊、住院、手术、床位等15类重点医疗服务指标实时监管体系。

## 5 结语

数据作为一种新型生产要素,是数字经济核心生产要素,也是科技创新的突破点。通过建立山东省健康医疗大数据信息资源目录体系,形成卫生健康系统内和部门间相互配合、统一归口的健康医疗大数据共享机制,促进健康医疗大数据应用,为山东省“互联网+医疗健康”示范省和国家健康医疗大数据中心(北方)建设提供数据基础。

## 参考文献

- 1 孟群,胡建平,董方杰,等.我国健康医疗大数据资源目录体系建设研究[J].中国卫生信息管理杂志,2017,14(3):387-391.
- 2 陈敏,黄竹青,秦健.湖北省健康医疗大数据资源目录体系建设研究[J].中国医院管理,2018,38(12):67-68.
- 3 李岳峰,胡建平,张学高.中国健康医疗大数据资源目录体系与技术架构研究[J].中国卫生信息管理杂志,2019,16(3):249-255.
- 4 国家发展改革委,中央网信办.政务信息资源目录编制指南(试行)[EB/OL].[2019-07-13].<http://www.gov.cn/xinwen/2017-07/13/5210203/files/2415d43d2bc1b4dfe9c3f1e1b5c0626c1.pdf>.
- 5 穆勇,彭凯.政务信息资源目录体系建设理论与实践[M].北京:北京大学出版社,2009:43-49.
- 6 DAMA International. The DAMA Guide to the Data Management Body of Knowledge [M]. New York: Technics Publications, 2009: 37.
- 7 山东省质量技术监督局.政务信息资源目录第3部分:编制指南:DB37/T3521.3-2019[S].2019-03-21.