

团 体 标 准

T/CHIA 17.1-2020

健康医疗大数据资源目录体系 第 1 部分：总体框架

Health big data information resource catalog system
Part 1: General framework

2020-11-16 发布

2020-12-1 实施

中国卫生信息与健康医疗大数据学会 发布

目 次

前言.....	I
1 范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 概述.....	1
5 目录体系.....	2
5.1 目录体系.....	2
5.2 交换体系.....	2
6 技术总体架构.....	2
6.1 组成.....	2
6.2 资源信息库.....	2
6.3 资源目录中心.....	3
6.4 资源交换中心.....	3
7 流程.....	4
7.1 目录内容服务形成与提供流程.....	4
7.2 共享信息资源的发现定位与交换流程.....	4
8 本标准各部分间关系.....	4

前 言

T/CHIA 17-2020《健康医疗大数据资源目录体系》分为以下五个部分：

- 第1部分：总体框架；
- 第2部分：技术要求；
- 第3部分：基本元数据；
- 第4部分：资源分类；
- 第5部分：资源标识符编码规则。

本部分为T/CHIA 17-2020的第1部分。

本部分按照GB/T 1.1-2020给出的规则起草。

本部分由国家卫生健康委统计信息中心提出并归口。

本部分主要起草单位：国家卫生健康委统计信息中心、华中科技大学同济医学院、中国人民解放军空军军医大学。

本部分主要起草人：胡建平、李岳峰、董方杰、沈丽宁、王霞、夏天、徐勇勇、张学高。

健康医疗大数据资源目录体系

第 1 部分：总体框架

1 范围

本标准规定了健康医疗大数据资源目录体系的总体框架、目录服务形式和流程，描述了总体结构、基本功能、工作流程以及本标准各部分之间的关系。

本标准适用于健康医疗大数据资源目录体系的规划和设计。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件，其随后所有的修改单（不包括勘误的内容）或修订版均不适用于本标准。但是，鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件，其最新版本适用于本标准。

T/CHIA17.2-2020 健康医疗大数据信息资源目录体系 第 2 部分：技术要求

T/CHIA17.3-2020 健康医疗大数据信息资源目录体系 第 3 部分：基本元数据

T/CHIA17.4-2020 健康医疗大数据信息资源目录体系 第 4 部分：信息资源分类

T/CHIA17.5-2020 健康医疗大数据信息资源目录体系 第 5 部分：资源标识符编码规则

3 术语和定义

下列术语和定义适用于本部分。

3.1

健康医疗大数据 **health big data**

健康医疗大数据是指信息网络技术与传统的公共卫生、医疗服务、医疗保障、药品管理、计划生育、综合管理等 6 大业务深度融合，以及互联网+健康医疗、精准医学、人工智能等新兴领域，所产生的与健康相关的数据。

3.2

信息资源目录 **information resource catalog**

信息资源目录是指按照统一的标准规范，对分散在各级各类卫生健康及其相关部门的信息资源进行整合、组织和序化，形成逻辑上集中、物理上分散、可统一管理和服务的目录体系。

4 概述

健康医疗大数据资源分为标准资源、服务资源、数据资源、应用资源。健康医疗大数据资源目录体系由目录体系与交换体系两个重要部分组成，分别提供目录服务和信息交换服务。

5 目录体系

5.1 目录体系

国家健康医疗大数据资源目录体系包括目录结构、基本元数据、扩展元数据、编码模型。目录结构，按照 T/CHIA 17-4 对资源进行分类。基本元数据，按照 T/CHIA 17-3 进行定义。扩展元数据，参照 T/CHIA 17-3 进行扩展定义。编码模型，按照 T/CHIA 17-5 中的编码规则生成资源唯一标识码。

5.2 交换体系

构建基于国家、省、市、县四级卫生健康信息平台的大数据资源目录服务平台，实现健康医疗信息的开放共享、业务协同。其中国家级资源目录服务平台除了实现国家级资源信息的融合外，还提供省际间跨区域资源共享服务，制定省级跨域共享规则，协同调度；省级资源目录服务平台除了实现省内资源信息的融合外，还提供城市间跨区域资源共享服务，制定市级跨域共享规则，协同调度；市级资源目录服务平台除了实现市内资源信息的融合外，还提供城市内跨区（县）域资源共享服务，制定区（县）级跨域共享规则，协同调度；县级资源目录服务平台直接提供各类数据开放服务，实现区（县）域内资源信息的融合。

6 技术总体架构

6.1 组成

国家健康医疗大数据资源目录体系技术总体架构包括资源信息库、资源目录中心和资源交换中心。资源信息库由资源共享数据中心、目录内容信息库、目录内容管理信息库、目录内容服务信息库、元数据库组成。资源目录中心由编目系统、目录内容管理系统、目录内容服务系统组成。资源交换中心由交换管理、交换传输、交换桥接组成。

6.2 资源信息库

6.2.1 资源共享数据中心

资源共享数据中心存储本机构或部门用于机构或部门间共享的信息资源，包括结构化数据、非结构化/半结构化文本、医疗影像数据、流数据、社交媒体数据、系统数据等多种类型的电子化信息资源。

6.2.2 目录内容信息库

目录内容信息库存储由编目系统提取共享信息资源的基本特征编目形成的目录内容。

6.2.3 目录内容管理信息库

目录内容管理信息库存储各卫生健康信息主管部门注册并通过审核到目录服务中心的内容。目录服务中心使用目录内容管理信息库实现对所汇集目录内容的管理。

6.2.4 目录内容服务信息库

目录内容管理信息库存储用于发布的目录内容。目录服务中心使用目录内容服务信息库提供目录查询检索等服务。

6.2.5 元数据库

元数据库存储用于健康医疗信息资源共享活动建立相关目录所需要的基本元数据、扩展元数据内容。

6.3 资源目录中心

6.3.1 编目管理系统

编目系统根据资源共享数据库的内容，提取其基本特征，按照T/CHIA 17-3实现元数据赋值，形成目录内容。应提供资源目录内容编目、资源目录内容注册等功能。

6.3.2 目录内容管理系统

目录内容管理系统实现对目录服务中心的目录内容和目录服务运行的管理，应提供目录内容审核、标识符管理、资源目录分类管理、资源关联性管理、资源共享和交换规则的设置、资源目录内容的发布管理等功能。

6.3.3 目录内容服务系统

目录内容服务系统基于目录内容服务信息库，向用户提供目录内容的资源查询检索、资源下载、资源订阅、资源申请、资源退订、资源评估等服务。目录服务系统应提供两种方式的服务。

6.3.4 元数据管理系统

元数据管理系统实现对元数据的管理，应提供元数据属性管理、元数据实体管理、元数据集管理、元数据版本管理等功能。

6.4 资源交换中心

6.4.1 交换管理子系统

对资源信息订阅共享交换进行运行管理。交换管理子系统提供部署维护、交换配置、运行管理、审计管理等功能。

6.4.2 交换传输子系统

交换传输子系统通过交换服务实现资源信息内容在不同部门或机构的使用者和资源共享数据库之间的传输。交换传输子系统提供消息传输、消息代理、基础服务等功能。

6.4.3 交换桥接子系统

交换桥接子系统提供适配器，用来适应不同类型、不同网段的信息源。

7 流程

7.1 资源目录的形成与发布流程

- 目录内容服务形成与提供流程包括5个环节：
- a) 准备、清洗：首先由健康医疗大数据资源提供者建立资源共享数据库，将各种可共享的健康医疗大数据资源采集至共享数据库；
 - b) 编目：健康医疗大数据资源提供者对共享资源的内容进行特征提取，通过编目管理系统填报资源目录内容；
 - c) 注册：由健康医疗大数据资源提供者通过编目管理系统将目录内容注册至目录服务中心的目录内容信息库；
 - d) 审核：由资源目录管理者通过编目管理系统对注册的资源进行审核，审核通过的传输至目录内容管理信息库；
 - e) 发布：由目录管理者对审核后的信息资源目录内容进行发布。

7.2 资源的发现定位与交换流程

- 资源的发现定位与交换流程包括2个环节：
- a) 目录查询：使用者基于浏览器等客户端查询目录服务中心目录内容，通过目录查询可以发现并定位共享信息。
 - b) 信息获取：使用者根据目录查询得到的定位信息，通过资源交换中心获得具体信息。使用者可以通过网络浏览、查询、下载、订阅推送等各种方式从资源交换中心获取健康医疗大数据资源。

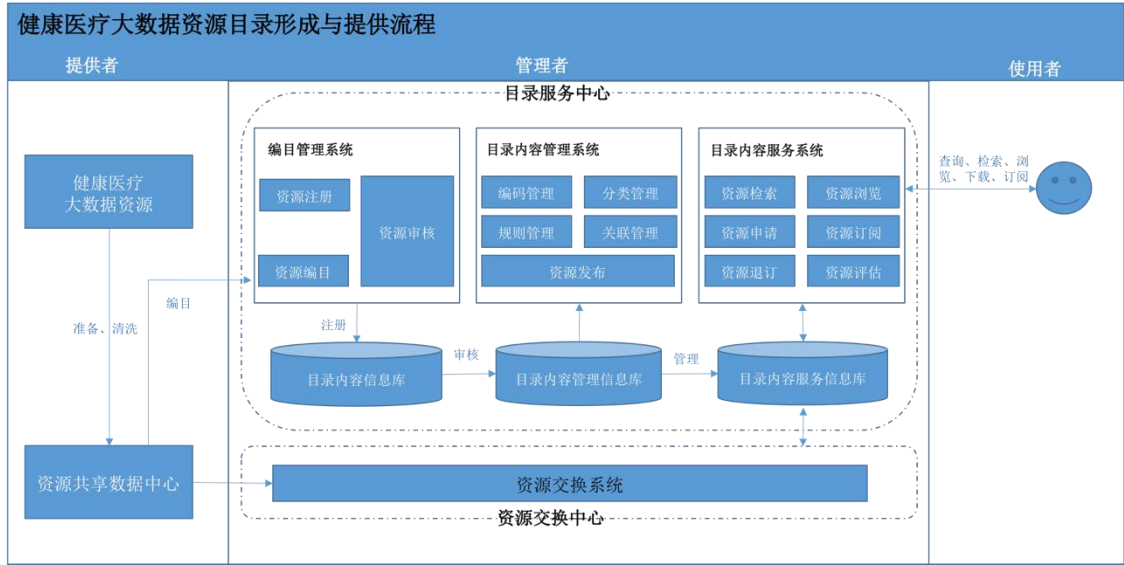


图 1 健康大数据资源目录形成与提供流程

8 本标准各部分间关系

- 第1部分 总体框架：规定了健康医疗大数据资源目录体系的总体框架、目录服务形式和流程，描述了总体结构、基本功能、工作流程以及本标准各部分之间的关系；
- 第2部分 技术要求：规定了健康医疗大数据资源目录体系的基本技术要求；

第3部分 基本元数据：规定了描述健康医疗大数据信息资源特征所需的基本元数据及其表示方式，给出了各基本元数据的定义和著录规则，用以描述健康医疗大数据信息资源的标识、内容、管理等信息；

第4部分 资源分类：规定了健康医疗大数据的资源分类体系和内容；

第5部分 资源标识符编码规则：规定了健康医疗大数据资源标识符的编号结构、基本规则、注册管理架构及流程；

第1部分从总体结构对健康医疗大数据资源目录体系进行规范，第2部分对目录体系的技术要求进行规范，第3部分、第4部分和第5部分对健康医疗大数据资源目录内容进行了规范。
