

团 体 标 准

T/CHIA 17.5-2020

健康医疗大数据资源目录体系 第 5 部分：资源标识符编码规则

Health big data information resource catalog system
Part 5: Resource identifier coding rules

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 标识符的结构与分配.....	1
4.1 标识符的构成.....	1
4.2 省市代码.....	2
4.3 类目码.....	2
4.4 亚目码.....	2
4.5 细目码.....	2
4.6 信息资源码.....	3
5 注册申请和管理.....	3
5.1 管理架构.....	3
5.2 注册申请.....	3
5.3 健康医疗大数据资源标识符的生成.....	4
附录 A（规范性附录） 省、自治区、直辖市区域代码.....	5
附录 B（规范性附录） 类目代码表.....	6
附录 C（规范性附录） 亚目代码表.....	7
附录 D（资料性附录） 健康医疗大数据资源标识符生成示例.....	9
参考文献.....	10

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5部分。

本标准根据 GB/T 1.1-2020 给出的规则起草。

本部分由华中科技大学提出并归口。

本标准主要起草单位：华中科技大学、国家卫生健康委统计信息中心、湖北省卫生计生信息中心、空军军医大学、华中科技大学同济医学院附属同济医学、武汉市第一医院。

本标准主要起草人：沈丽宁、胡建平、李岳峰、肖兴政、王霞、马敬东、董方杰、张黎黎、张晓祥、王巍、庾兵兵、毕齐擎、陈文强、代伟。

健康医疗大数据资源目录体系

第 5 部分：资源标识符编码规则

1 范围

本标准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。

本标准适用于健康医疗大数据资源的编目、注册、发布、查询、维护和管理，为健康医疗大数据资源标识符提供编码依据。

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/CHIA17.1-2020 健康医疗大数据资源目录体系 第 1 部分：总体框架

T/CHIA17.2-2020 健康医疗大数据资源目录体系 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 健康医疗大数据资源目录体系 第 4 部分：信息资源分类

3 术语和定义

T/CHIA17.1-2020中确立的以及下列术语和定义适用于本部分。

3.1

标识符 identifier

用于标识或命名一个数据项目并可能指出该数据确定性质的一个或一组字符。

[GB/T 14915-1994中定义]

3.2

健康医疗大数据资源标识符 big data for health care resource identifier

用于唯一标识健康医疗大数据资源的一组字符。

3.3

资源分类 resource classification

根据资源的特性及其相互关系而分别排列出的不同次序、类别和体系。

4 标识符的结构与分配

4.1 标识符的构成

基于卫生信息OID标识体系对健康医疗大数据资源进行标识，采用树状分层结构，每个节点码之间用字符“.”隔开，依据我国《卫生信息标识体系对象标识符编号结构与基本规则》和《卫生信息标识体系对象标识符管理注册管理规程》，我国健康医疗大数据资源目录OID根标识为2.16.156.10011.2.100，根节点下一级为省市代码（4位），健康医疗大数

据资源分类由类目、亚目和细目等3个层次组成，每个层次对应一个节点，分别对类目（2 位码）、亚目（2 位码）、细目（4 位码）、机构码（不定长）和信息资源（不定长）进行编码。其形式如图1所示。

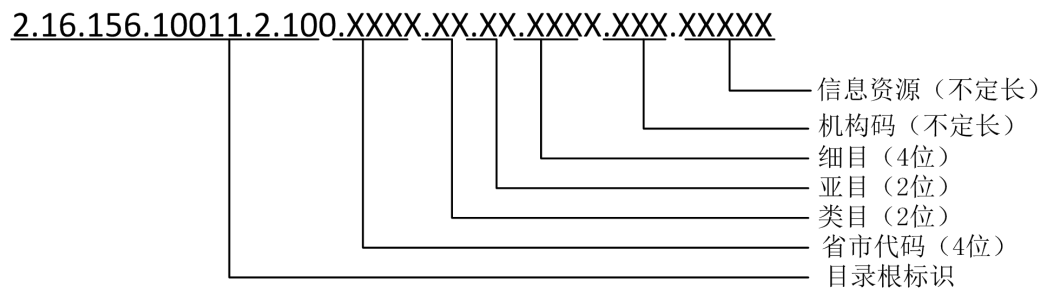


图1 健康医疗大数据资源标识符编码结构图

4.2 省市代码

省市代码是四位数字码，前两位为省级代码，后两位为市级代码。省级代码用于区别委直属事业单位、省、自治区、直辖市和特别行政区，委直属事业单位是指国家卫生计生委直接管辖非盈利性社会服务组织，归属于其他类，编码赋值见附录A。市级编码是市、地区、自治州、盟、直辖市所辖市辖区/县汇总码、省（自治区）直辖县级行政区划汇总码，依据GB/T 2260-2007的规定，选取中国行政区划代码的第三位和第四位作为市级代码，其他地区、不确定地区或省直管机构赋值00。例如北京市东城区的省市代码为“1101”。

4.3 类目码

类目码是省级代码下一级节点编码，用来区分不同的业务应用或者数据业态。类目共划分为 8个大类：依据国家健康医疗信息化“46312-2”框架中对6大业务应用和3大数据库的总结，分为公共卫生、计划生育、医疗服务、医疗保障、药品管理和综合管理等6个基本业务类，将全员人口信息数据库、电子病历数据库和健康档案数据库归为基础信息类，考虑到移动通讯、云计算、物联网和人工智能等新兴技术在健康医疗领域的广泛应用，增加新兴业态类。8个业务类的代码参见附录B。

4.4 亚目码

亚目码是类目码对应节点的下一级节点编码，亚目是根据各类目领域的特点，按照业务内容的组成部分或业务流程的先后顺序进行分类，并对每个类目都增加“其他”项作为已有分类中未包含资源类型的分类。亚目分类明细及代码参见附录C。

4.5 细目码

细目码是亚目码后节点编码，细目编制是一个逐步完善的过程，通过健康医疗大数据资源目录管理系统实现大范围注册后，才可能逐步编目完成一个较为全面、科学的资源目录细目。

细目码共4位，由 GB 18030 中规定的10个阿拉伯数字（0~9）组成。

细目码原则上采用顺序号，即1000~9999，“其他”类的代码规定为9999，即实际可使用的代码范围为1000~9998，由手动编目生成。编制细目码的过程中应考虑代码的有限性、扩展性、规范性、稳定性和简洁性。

4.6 信息资源码

信息资源码是健康医疗大数据资源标识符中最后节点的编码，由省级健康医疗大数据资源目录中心自行分配，信息资源码分配规则或方案由其自主确定，报国家健康医疗大数据资源目录中心备案。在编目元数据时生成信息资源码，以实现健康医疗大数据资源的唯一标识。必须保证在管辖范围内的唯一性。

信息资源码为不定长，由 GB 18030 中规定的阿拉伯数字（0~9）组成，每个节点编码数值应为大于0小于16000000的正整数，赋值方式可以采用顺序号，由程序自动生成，也可以是一个内部为其所管理的信息资源分配的流水号。编制时应考虑代码的稳定性和简洁性。

5 注册申请和管理

5.1 管理架构

健康医疗大数据资源标识符的管理体系原则上采用两级架构，分为国家健康医疗大数据资源目录中心和省级健康医疗大数据资源目录中心。国家健康医疗大数据资源目录中心负责健康医疗大数据资源OID根标识的管理，以及省级健康医疗大数据资源目录中心注册权限的授予，负责健康医疗大数据资源目录中心数据库的管理。省级健康医疗大数据资源目录中心负责各省行政区域内健康医疗大数据资源目录的注册申请和管理，其注册的信息资源需及时向国家健康医疗大数据资源目录中心注册，保持国家数据资源目录的完整性。健康医疗大数据发展较好，有条件开展健康医疗大数据资源注册管理的地区，可建立市级健康医疗大数据资源目录中心，实行国家—省—市三级管理体制，但健康医疗大数据资源标识符的编码方式保持不变。

管理架构如图2所示：

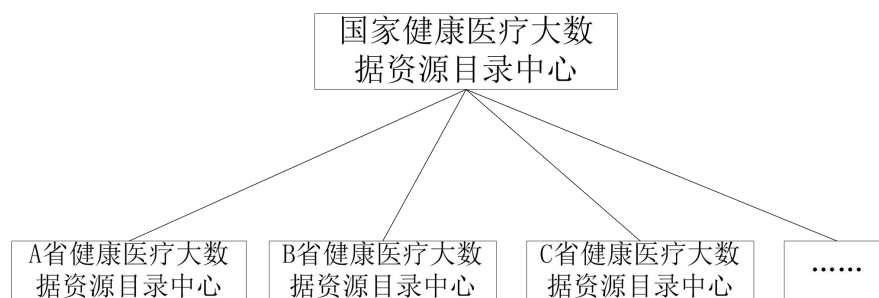


图2 注册管理架构

5.2 注册申请

注册申请流程如下：

a) 申请者对健康医疗大数据资源进行机器或手工编目时，确定该资源所属类目、亚目、细目的类型，由此确定类目码、亚目码、细目码，信息资源码的确定有两种方式：一是申请者根据省级资源目录中心的编码规则，自主赋值；二是省级资源目录中心在申请者提出注册申请后，由其对信息资源码进行赋值。根据自主可控的原则，具体采用哪种方式由各省健康医疗大数据资源目录中心根据自身具体条件决定。

b) 申请者根据省级健康医疗大数据资源目录中心的要求填写注册申请信息，并向省级健康医疗大数据资源目录中心提出健康医疗大数据资源目录的注册申请，省级健康医疗大

数据资源目录中心对申请者的申请信息以及资源标识符编码是否符合规范、标识符的唯一性进行审核。

c) 省级健康医疗大数据资源目录中心审核无误，标识符符合编码规则，则注册申请成功，纳入健康医疗大数据资源目录管理系统。若不符合规范，则注册失败，申请者重新提交注册申请信息。

5.3 健康医疗大数据资源标识符的生成

健康医疗大数据资源标识符的OID根标识、省市代码、类目码、亚目码、细目码以及健康医疗大数据资源目录中心为其分配的信息资源码，即构成了一个完整的、全世界范围内唯一的健康医疗大数据资源标识符。该标识符在健康医疗大数据资源的生命周期内保持不变。生成健康医疗大数据资源标识符的流程参见附录D中的具体示例。

附录 A
(规范性附录)
省、自治区、直辖市区域代码

所在省、自治区、直辖市、特别行政区代码，基于标准GB/T 2260-2007，取其规定的前两位数字作为省（区）级代码，以99作为其他的代码。见表A.1。

表 A.1 省（区）级代码

代码	名称	代码	名称
11	北京市	44	广东省
12	天津市	45	广西壮族自治区
13	河北省	46	海南省
14	山西省	50	重庆市
15	内蒙古自治区	51	四川省
21	辽宁省	52	贵州省
22	吉林省	53	云南省
23	黑龙江省	54	西藏自治区
31	上海市	61	陕西省
32	江苏省	62	甘肃省
33	浙江省	63	青海省
34	安徽省	64	宁夏回族自治区
35	福建省	65	新疆维吾尔自治区
36	江西省	71	台湾省
37	山东省	72	香港特别行政区
41	河南省	73	澳门特别行政区
42	湖北省	99	其他
43	湖南省		

附录 B
(规范性附录)
类目代码表

类目代码共两位，由 GB 18030 中规定的 10 个阿拉伯数字（0~9）组成。见表B.1。

表 B.1 类目代码

代码	名称	说明
11	基础信息	包含全员人口信息数据库、电子病历数据库和健康档案数据库。
12	公共卫生	
13	计划生育	
14	医疗服务	
15	医疗保障	
16	药品管理	
17	综合管理	
18	新兴业态	包含移动通讯、云计算、物联网和人工智能等新兴技术在健康医疗领域的广泛应用。

附录 C
(规范性附录)
亚目代码表

亚目码共两位，由 GB 18030 中规定的 10 个阿拉伯数字（0~9）组成。见表C.1。

表 C.1 亚目代码

类目	亚目	代码
基础信息	全员人口信息数据库	11
	电子病历数据	12
	健康档案数据库	13
	其他	99
公共卫生	疾病控制	11
	监督执法	12
	卫生应急	13
	预防干预	14
	妇幼保健	15
	其他	99
计划生育	人口管理	11
	计生服务	12
	其他	99
医疗服务	临床服务	11
	检查检验	12
	药事管理	13
	患者服务	14
	其他	99
医疗保障	新型农村合作医疗	11
	城镇职工医保	12
	城市居民医保	13
	商业医疗保险	14
	疾病应急救助	15
	其他	99
药品管理	新药研发	11
	生产流通	12
	招标采购	13
	药品配送	14
	药品监管	15
	其他	99
综合管理	统计调查	11
	资源管理	12
	行政管理	13

	科教管理	14
	其他	99
新兴业态	互联网健康医疗	11
	远程医学	12
	精准医学	13
	健康物联网	14
	健康人工智能	15
	其他	99

附录 D
(资料性附录)
健康医疗大数据资源标识符生成示例

假设某一位于A省的单位B准备向A省健康医疗大数据资源目录中心提交健康医疗大数据信息资源的元数据，步骤如下：

- 1) B对健康医疗大数据资源进行机器或手工编目。
- 2) B确定该健康医疗大数据资源所属类目类型，假设其属于公共卫生的范畴，那么类目节点赋值为“12”。
- 3) 确定该健康医疗大数据资源所属亚目类型，假设其属于公共卫生范畴中的疾病控制领域，则亚目节点赋值为“21”。
- 4) 确定该健康医疗大数据资源所属细目类型，假设其属于公共卫生范畴中疾病控制领域中的其他类，则细目节点赋值为“9999”。
- 5) 假设A省为湖北省武汉市，故对应省市代码为“4201”。
- 6) 由于信息资源码分配由省级健康医疗大数据资源目录中心管理，假设根据湖北省健康医疗大数据资源目录中心制定的分配规则，该健康医疗大数据资源信息资源码的字符为“XXX”。
- 7) 将该健康医疗大数据资源标识符的OID根标识、省市代码、类目码、亚目码、细目码，加上省级健康医疗大数据资源目录中心为其分配的信息资源码，即形成一个完整的健康医疗大数据资源标识符，完整的标识符为“2.16.156.10011.2.100.4201.12.21.9999.XXX”。
- 8) B向 A 提交元数据。
- 9) A 验证B所提交的信息资源的标识符元素，检查该标识符的编码是否符合编码规则，其次检查信息资源码是否重复。如果没有异常，则向A给出注册成功的响应。否则提示错误。

参 考 文 献

- [1] 国家OID注册中心. OID国家白皮书（翻译稿）[Z].2012,12.
 - [2] 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).
 - [3] 穆勇, 彭凯. 政务信息资源目录体系建设理论与实践[M]. 北京大学出版社, 2009.
-