

DB33

浙 江 省 地 方 标 准

DB33/T XXXXX—XXXX

数字化改革 公共数据目录编制规范

Digital reform—

Specification for compiling public data catalog

(报批稿)

XXXX—XX—XX 发布

XXXX—XX—XX 实施

浙江省市场监管局 发布

目 次

前言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 编制方法	2
5 信息系统普查	2
5.1 普查范围	2
5.2 普查内容	2
6 数据目录梳理	3
6.1 梳理范围	3
6.2 梳理内容	3
7 数据目录审核	4
7.1 部门审核	4
7.2 大数据主管部门审核	4
8 数据目录管理	4
8.1 数据目录分类	4
8.2 数据目录发布和更新	5
附录 A （规范性） 数据目录梳理表	6
A.1 数据目录梳理模板	6
A.2 数据目录梳理模板填写说明	6

前 言

本标准按照GB/T 1.1-2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本标准中的某些内容可能涉及专利。本标准的发布机构不承担识别专利的责任。

本标准由浙江省大数据发展管理局提出并归口。

本标准起草单位：浙江省大数据发展中心、数字浙江技术运营有限公司、北京国脉互联信息顾问有限公司。

本标准主要起草人：徐峰、王路燕、叶剑超、金加和、赵程遥、平正强、施筱玲、甄理、王沁怡、张佳、王静姝、王文达、陈乐燕、朱娇娜。

本标准为首次发布。

数字化改革 公共数据目录编制规范

1 范围

本标准规定了信息系统普查、数据目录梳理、数据目录审核和数据目录管理等阶段的要求。本标准适用于公共数据目录的编制，企事业单位可参照本文件执行。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本标准必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本标准；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本标准。

GB/T 18391. 3—2009 信息技术元数据注册系统(MDR) 第3部分：注册系统元模型与基本属性(ISO / IEC 11179—3: 2003, IDT)

3 术语和定义

下列术语和定义适用于本标准。

3.1

公共数据 public data

国家机关、法律法规规章授权的具有管理公共事务职能的组织（以下统称公共管理和服务机构）在依法履行职责和提供公共服务过程中获取的数据资源，以及法律、法规规定纳入公共数据管理的其他数据资源。

3.2

信息系统 information system

公共管理和服务机构应用信息技术支持依法履行职责的系统。

3.3

元数据 metadata

定义和描述其他数据的数据。

[来源：GB / T 18391. 3—2009, 定义3. 2. 18]

3.4

数据目录 data catalog

以核心元数据为主要描述方式，按照资源分类体系对公共数据进行有序排列的一组信息，用以描述公共数据的特征，以便对公共数据进行检索、定位与获取。

4 编制方法

数据目录编制包括信息系统普查、数据目录梳理、数据目录审核、数据目录管理四个方面，如图1所示：

- a) 信息系统普查。开展本部门非涉密信息系统普查，普查内容包括信息系统的数量、名称、属性等；
- b) 数据目录梳理。开展本部门数据目录梳理，梳理内容包括数据名称、数据项、共享属性、开放属性等；
- c) 数据目录审核。对梳理的数据目录进行整理、汇编、审核，形成数据目录；
- d) 数据目录管理。对形成的数据目录进行分类、发布和更新。

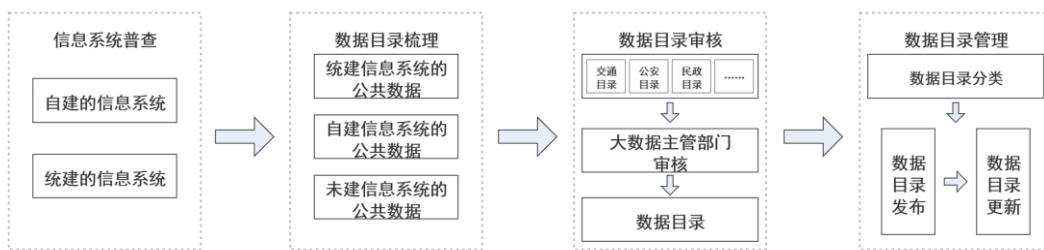


图1 数据目录编制流程图

5 信息系统普查

5.1 普查范围

信息系统普查应覆盖本部门自建的信息系统和行业主管部门统建的信息系统。

5.2 普查内容

信息系统普查应保证普查内容的全面性、及时性和准确性。信息系统普查内容包含以下要素：

- a) 信息系统名称；
- b) 系统所属部门；
- c) 系统功能、服务简介；
- d) 系统所属项目信息；
- e) 系统启用日期；
- f) 系统状态，包括建设中、运行中、停用等；
- g) 系统类型，包括办公运行系统、业务应用系统、门户网站、宣传微博/微信公众号等；
- h) 建设资金来源，包括上级配套资金、财政资金、单位自筹资金等；
- i) 系统归口业务处室情况；
- j) 系统部署情况，包括自建机房、租用机房、政务云和第三方云等；
- k) 网络环境，包括政务内网、政务外网、互联网、业务专网等；
- l) 系统使用范围，包括国家、省、市、县（区）、乡（镇）等；
- m) 系统访问地址；
- n) 建设层级，包括国家、省、市、县（区）、乡（镇）等；
- o) 系统数据情况；

- p) 系统等级保护情况;
- q) 系统厂商情况。

6 数据目录梳理

6.1 梳理范围

数据目录梳理范围包括行业主管部门统建的信息系统、本部门自建的信息系统和未建信息系统的公共数据。

6.2 梳理内容

6.2.1 梳理要素

数据目录梳理应保证数据目录要素完整、内容规范准确。数据目录梳理内容包含以下要素:

- a) 数据资源标识符;
- a) 数据名称;
- b) 数源单位;
- c) 数据摘要;
- d) 数据格式;
- e) 重点领域分类;
- f) 所属信息系统名称;
- g) 更新频率;
- h) 数据项描述:
 - 1) 数据项名称;
 - 2) 数据类型;
 - 3) 数据长度;
 - 4) 字段描述;
 - 5) 共享属性;
 - 6) 共享条件;
 - 7) 开放属性;
 - 8) 是否主键;
 - 9) 是否可为空;
 - 10) 是否字典项;
- i) 创建日期;
- j) 修改日期。

6.2.2 梳理模板

数据目录梳理模板和填写说明见附录A。

6.2.3 梳理要求

涉及行业主管部门统建的信息系统,本部门应将行业主管部门已编制的目录直接纳入;涉及本部门自建的信息系统,应结合信息系统开发过程中产生的规范性文档辅助进行梳理,如《软件需求规格说明书》、《数据库设计说明书》等;涉及未建信息系统的公共数据,应结合业务涉及的材料、表单等进行梳理。

7 数据目录审核

7.1 部门审核

各部门应依据附录A的A.1数据目录梳理模板，汇总、整理、审核后提交至本级大数据主管部门。

7.2 大数据主管部门审核

各级大数据主管部门应依据附录A的A.2各项要素填写说明对各部门提交的数据目录进行审核。

8 数据目录管理

8.1 数据目录分类

8.1.1 应用场景分类

根据数据应用场景分类形成基础库目录、专题库目录和重点领域目录等，具体分类和说明见表1。

表1 应用场景分类

分类	说明
基础库目录	包括人口综合库目录、法人综合库目录、信用信息库目录、可信电子证照库目录等。
专题库目录	包括生态环境库目录、交通出行库目录、教育文化库目录、医疗卫生库目录、社保就业库目录、城建住房库目录、社会救助库目录、法律服务库目录等。
重点领域目录	包括信用服务、医疗卫生、社保就业、公共安全、城建住房、交通运输、教育文化、科技创新、资源能源、生态环境、工业农业、商贸流通、财税金融、安全生产、市场监管、社会救助、法律服务、生活服务、气象服务、地理空间、机构团体、人民防空、应急管理。

8.1.2 目录层级分类

根据目录层级分类形成省级目录和地方目录，具体分类和说明见表2。

表2 目录层级分类

分类	说明
省级目录	包括党委、人大、政府、政协、法院、检察院、企事业单位、民主党派、工商联及群众团体等省级部门数据目录。
地方目录	包括11个市级（含下属各县区）数据目录。

8.1.3 共享属性分类

根据目录共享属性分类形成受限共享类目录、非共享类目录和无条件共享类目录，具体分类和说明见表3。

表3 共享属性分类

分类	说明
受限共享类	列入受限共享类公共数据范围的，应说明理由，并提供有关法律、法规、规章依据。要求使用受限共享类公共数据的，由同级公共数据和电子政务工作机构会同提供公共数据的机构进行审核；审核同意的，开通相应访问权限。受限共享类公共数据可以经脱敏等处理后向公共管理和服务机构提供，法律、法规另有规定的除外。
非共享类	列入非共享类公共数据范围的，应说明理由，并提供有关法律、法规、规章依据。非共享类公共数据可以经脱敏等处理后向公共管理和服务机构提供，法律、法规另有规定的除外。
无条件共享类	除受限共享类和非共享类之外的数据。

8.1.4 开放属性分类

根据目录开放属性分类形成无条件开放类目录、受限开放类目录和禁止开放类目录，具体分类和说明见表 4。

表4 开放属性分类

分类	说明
无条件开放类	无条件开放类数据是指法律、法规明确可以开放，以及除受限开放类和禁止开放类之外的数据。
受限开放类	受限开放类数据指涉及公民、法人和非法人组织，经过脱敏、授权后可以开放的数据，或无条件开放将严重挤占基础设施资源，影响公共数据处理运行效率的数据。
禁止开放类	禁止开放类数据指开放后损害国家安全、商业秘密、个人隐私的数据。

8.2 数据目录发布和更新

8.2.1 数据目录发布

各级大数据主管部门应将本级审核、分类后的数据目录及时发布，形成本级数据目录。

8.2.2 数据目录更新

各级公共数据管理和服务机构，应做好本部门数据目录的更新与维护。

附 录 A
(规范性)
数据目录梳理表

A.1 数据目录梳理模板

数据目录梳理模板见表A.1。

表 A.1 数据目录梳理模板

数据资源标识符	数据名称	数据来源单位	数据摘要	数据格式	重点领域分类	所属系统名称	更新频率	数据项描述									创建日期	修改日期	
								数据项名称	数据类型	数据长度	字段描述	共享属性	共享条件	开放属性	是否主键	是否可为空			是否字典项

A.2 数据目录梳理模板填写说明

数据目录梳理模板填写说明见表A.2。

表 A.2 数据目录梳理模板填写说明

填写要素名称	说明	
1. 数据资源标识符	定义	数据目录的唯一不变的标识字符。
	数据类型	字符型。
	建议	必选项；参照《政务信息资源目录编制指南（试行）》确定的代码结构规则，代码结构由前段码、后段码组成。
2. 数据名称	定义	缩略描述数据目录内容的标题。
	数据类型	字符型。
	建议	必选项；业务类数据目录命名为XXX信息，该类数据含系统、数据库、报表、业务办理记录、业务情况等业务信息；证照类数据一般对应为目录下有两张表，一部分是证照信息表，一部分是证照的版式文件；地理信息类数据命名为XXX地理信息，该类数据为地理位置信息；其他类命名规则如数据名称涉及主表、子表、附表、X月、正本、副本等在数据名称后添加括号表示，同部门同名数据若内容确实为不同数据，且需要加以区分的，在前面加上系统名称简称作为前缀，用“-”连接系统名称与数据名称。

表 A.2 数据目录梳理模板填写说明（续）

填写要素名称		说明	
3. 数据源单位	定义	提供公共数据的部门。	
	数据类型	字符型。	
	建议	必选项；该公共数据部门的规范简称。	
4. 数据摘要	定义	对公共数据内容（或关键字段）的概要描述。	
	数据类型	字符型。	
	建议	必选项；不少于20个字。	
5. 数据格式	定义	公共数据存在方式（可多选）。数据源单位应尽可能提供可机读的电子格式及相关软件版本信息，如只有纸质介质，应尽量提供电子扫描格式。数据格式，可采用但不限于：电子文件的存储格式为 OFD、wps、xml、txt、doc、docx、html、pdf、ppt 等；电子表格的存储格式为 et、xls、xlsx 等；数据库类的存储格式为 Dm、KingbaseES、access、dbf、dbase、sysbase、oracle、sql server、db2 等，同时应明确具体的数据库表结构定义（可细化至“数据项描述”元数据中）；图形图像类的存储格式为 jpg、gif、bmp 等；流媒体类的存储格式为 swf、rm、mpg 等；自描述格式，由提供方提出其特殊行业领域的通用格式，如气象部门采用的“表格驱动码”格式。	
	数据类型	字符型。	
	建议	必选项；按定义的多种格式选择确定。	
6. 重点领域分类	定义	说明公共数据重点领域分类的信息。	
	数据类型	字符型。	
	建议	必选项；按定义的重点领域分类进行选择确定。	
7. 数据项描述	7.1 数据项名称	定义	描述结构化数据中具体数据项的标题。适用于格式为数据库、电子表格类等的公共数据。
		数据类型	字符型。
		建议	必选项；描述电子表格、数据库，以及具有结构化数据内容资源中的各数据项（字段）标题，采用中文表示，一般较为常用的像 ID、IP 这样的英文专有名词可采用英文表达，英文字母应使用大写，其他情况下不应出现英文表达。名称应简洁明了，不应超过 20 个字符长度，数据项中涉及的字典项或特殊说明可在字段描述中进行描述。根据数据属性，关键数据项不应缺失，如涉及人员的信息应提供相对应的证件类型、身份证号码等数据项；涉及法人主体的信息应提供相对应的统一社会信用代码、企业名称等数据项；涉及电子证照数据应按照国家证照目录规范要求提供相应数据项；全省范围的数据，应有地区字段。
	7.2 数据类型	定义	标明该数据项的数据类型，包括：字符型 C、数值型 N、货币型 Y、日期型 D、日期时间型 T、逻辑型 L、备注型 M、通用型 G、双精度型 B、整型 I、浮点型 F 等。
		数据类型	字符型。
		建议	必选项；对应“数据项名称”逐一描述其数据类型。

表 A.2 数据目录梳理模板填写说明（续）

填写要素名称		说明	
7. 数据项描述	7.3 数据长度	定义	标明该数据项在计算机中存储时占用的字节数，适用于结构化数据（数据库类、电子表格类）。属于数据库类的，数据长度即该数据项对应的字段在数据库中的指定长度或默认长度；属于电子表格类的，估算该数据项内容字数的上限，并折算成字节数，该字节数即为数据长度。
		数据类型	字符型。
		建议	可选项；对应“数据项名称”逐一描述其数据长度。
	7.4 字段描述	定义	对重要数据项的含义进行解释、描述。
		数据类型	字符型。
		建议	可选项。
	7.5 共享属性	定义	标明公共数据的共享属性，包括无条件共享类、受限共享类、非共享类三类。
		数据类型	字符型。
		建议	必选项；对每个数据项按定义的共享属性进行选择确定。
	7.6 共享条件	定义	描述不同共享类型公共数据的共享条件。
		数据类型	字符型。
		建议	必选项；无条件共享类和受限共享类的公共数据，应注明使用要求，包括作为行政依据、工作参考，用于数据校核、业务协同等；受限共享类的公共数据，应注明共享条件和共享范围；对于非共享类的公共数据，应注明相关的法律、行政法规或党中央、国务院政策依据。
	7.7 开放属性	定义	标明公共数据的开放属性，包括无条件开放类、受限开放类、禁止开放类三类。
		数据类型	字符型。
		建议	必选项；涉及个人、企业隐私的数据项不应对外开放，如身份证号码、联系电话等。
	7.8 是否主键	定义	标明该数据项是否为数据表的主键，包括“是”和“否”，主键是数据表中的一个或多个字段，用于唯一地标识表中的某一条记录。
		数据类型	字符型。
		建议	可选项；一条数据目录应有一个及以上主键，主键的数据长度不应大于255，已归集公共数据不应更换主键、增加主键和删除主键。
	7.9 是否可为空	定义	标明该数据项是否可以为空，即该字段是否为必填。包括“是”和“否”。
		数据类型	字符型。
建议		可选项；按该类数据的业务理解，关键数据项不应为空，涉及自然人的应有姓名和身份证号，且不应为空；涉及法人的应有名称和统一社会信用代码，且不应为空。	
7.10 是否字典项	定义	标明该数据项是否为字典项，包括“是”和“否”。	
	数据类型	字符型。	
	建议	可选项；若该数据项预设填写内容选项，应填“是”。	
8. 所属系统名称	定义	公共数据所在的信息系统名称。	
	数据类型	字符型。	
	建议	必选项；若该公共数据无信息系统管理产生，应填“无”。	

表 A.2 数据目录梳理模板填写说明（续）

填写要素名称	说明	
9. 更新频率	定义	标明公共数据归集的频率，包括分钟级、小时级、每日、每周、每月、每季度、每半年、每年、不定期等。
	数据类型	字符型。
	建议	必选项。
10. 创建日期	定义	数源单位创建该数据目录的日期。
	数据类型	日期型，格式为CCYY-MM-DD。
	建议	必选项。
11. 修改日期	定义	数源单位最新修改该数据目录的日期。
	数据类型	日期型，格式为CCYY-MM-DD。
	建议	必选项；数据目录首次创建时，修改时间默认等同于创建时间。