

科学数据共享工程技术标准

SDS/T XXX—2006

科学数据共享核心元数据标准

(试行稿)

(本稿完成日期：2006年8月)

××××-××-××发布

××××-××-××实施

中华人民共和国科学技术部 发布

目 次

前 言.....	I
科学数据共享核心元数据.....	2
1 范围.....	2
2 规范性引用文件.....	2
3 术语和定义.....	2
4 符号与约定.....	3
5 科学数据共享核心元数据.....	6
附 录 A 规范性附录 科学数据共享核心元数据字典.....	13
附 录 B 规范性附录 科学数据共享核心元数据 XML SCHEMA.....	17
附 录 C 资料性附录 科学数据共享核心元数据举例.....	27
参考文献.....	28

前 言

本标准为第一次制定。

本标准的附录A和B为规范性附录，附录C为资料性附录。

本标准由中华人民共和国科学技术部基础研究司提出。

本标准由中华人民共和国科学技术部基础研究司归口。

本标准起草单位：国家信息中心。

本标准主要起草人：徐枫、宦茂盛、石雯雯、武晋平。

科学数据共享核心元数据

1 范围

本标准给出了科学数据共享工程核心元数据的基本内容,定义了科学数据共享核心元数据内容的元数据实体和元数据元素,适用于科学数据共享工程门户、中心(网)的元数据汇交。

2 规范性引用文件

下列规范性引用文件通过本部分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(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。

GB/T 7408-2005 数据元和交换格式 信息交换 日期和时间表示法

SDS/T 2111—2004 元数据标准化原则与方法

SDS/T 2122—2004 科学数据共享工程数据分类编码

RFC 1738 统一资源定位器(RFC1738——Uniform Resource Locators (URL))

3 术语和定义

本标准采用下列术语和定义。

3.1

数据集 **dataset**

可以标识的数据集合。

注: 数据集在物理上可以是更大数据集的较小的数据组。从理论上讲,数据集可以小到更大数据集内的单个要素或要素属性。一张硬拷贝地图或图表可以被认为是一个数据集。本标准所指的数据集是指不可再细分的数据集,即能够用一个数据字典唯一描述的数据集合。

3.2

元数据 **metadata**

关于数据的数据。

3.3

元数据元素 **metadata element**

元数据的基本单元。

注: 与UML术语中的属性同义。

3.4

元数据实体 **metadata entity**

一组说明数据相同特性的元数据元素。

注1：与UML术语中的类同义。

注2：可以包含一个或一个以上元数据实体。

3.5

元数据子集 **metadata section**

元数据的子集合，由相关的元数据实体和元素组成。

注：与UML术语中的包同义。

3.6

资源 **resource**

能满足一种需求的资产或手段。

例如：数据集、服务、文档、人力或机构。

3.7

核心元数据 **core metadata**

在国家科学数据共享工程中，描述科学数据集最基本属性、领域在制定其元数据内容标准时必须选择的元数据实体和元数据元素。

3.8

公共元数据 **common metadata**

在国家科学数据共享工程中，大多数领域具有共性的元数据内容。公共元数据包含核心元数据的全部内容，是领域在制定其元数据内容标准时的工作起点。

3.9

领域 **domain**

在科学数据共享工程中，按照学科及研究范围划分的域，包括资源环境领域、农业领域、人口与健康领域、工程与技术科学领域等。

4 符号与约定

本标准采用两种方式定义和描述元数据：摘要表示和字典描述。

4.1 摘要表示

在以摘要表示的方式定义和描述元数据元素时，采用以下属性，包括中文名称、定义、英文名称、数据类型、值域、短名、注解、子元素和扩展巴氏范式。

4.1.1 中文名称

本标准第5章5给出了科学数据共享核心元数据的摘要表示。其中，各元数据元素或元数据实体的中文名称用第5章中各小节的标题来表达。

例如：

5.2.1 数据集名称

4.1.2 定义

描述元数据的基本内容和属性。

4.1.3 英文名称

元数据的英文名称，一般用英文全称。

4.1.4 数据类型

元数据的有效值域和允许对该值域内的值进行有效操作的规定。

例如整型、实型、布尔型、字符串、日期等。

4.1.5 值域

说明元数据元素取值范围。

4.1.6 短名

元数据的英文缩写名称，具体缩写规则如下：

- 1) 短名在本标准范围内必须唯一。
- 2) 采用与国际标准类似的英文名称作为短名。

4.1.7 注解

对元数据的含义的进一步解释，包括该元数据的约束/条件（必选、可选或条件必选）和最大出现次数。当该元数据为条件必选时，应注明其约束条件。

4.1.7.1 约束/条件

说明元数据实体或元数据元素是否必须选取的属性。包括必选（M）和可选（O）。

必选 M

表明该元数据实体或元数据元素必须选择。

可选 O

根据实际应用可以选择也可以不选的元数据实体或元数据元素。已经定义的可选元数据实体和可选元数据元素，可指导领域元数据内容标准制定人员充分说明其数据。

可选元数据实体可以有必选元素；但这些元素只当可选实体被选用时才成为必选的。如果一个可选元数据实体未被选用，则该实体所包含的元素（包括必选元素）也不选用。

4.1.7.2 最大出现次数

说明元数据实体或元数据元素可以具有的最大实例数目。只出现一次的用“1”表示，重复出现的用“N”表示。不为1的固定出现次数用相应的数字表示，如“2”、“3”、“4”等。

4.1.8 子元素

子元素是通过一定的表示规则以确定一个元数据子集或元数据实体中包含的下一级的元数据实体或元数据元素。表示规则为：“标识符 = 表达式”。表达式中各符号的含义如下：

表 1 表达式的符号含义

符 号	含 义
=	由……替换、生成，由……组成
+	与
	或（选择）——在由“ ”分开的两项之中选择其一

$0\{a\}1$	表示{}中的元数据元素 a 为可选项/条件必选项, 且最大出现次数为 1; 若为条件必选项, 约束/条件具体参见其注解
$0\{a\}n$	表示{}中的元数据元素 a 为可选项/条件必选项, 且最大出现次数为 N; 若为条件必选项, 约束/条件具体参见其注解
a	表示元数据元素 a 为必选项, 且最大出现次数为 1
$1\{a\}n$	表示{}中的元数据元素 a 为必选项, 且最大出现次数为 N

在子元素表示中, {}中均使用元数据元素或实体的中文名称。

例如:

子元素: 数据集分类 =

类别名称+

类别编码+

分类标准

4.1.9 扩展巴氏范式

扩展巴氏范式可以更加规范化的表示一个元数据子集或元数据实体与其下一级的元数据实体或元数据元素之间的关系, 便于系统实现。与子元素的表示法不同的是, 扩展巴氏范式用“;”代替子元素中的“+”表示“与”关系, {}中均使用该元数据元素的短名, 并以“;”作为表达式的结尾。

例如:

扩展巴氏范式: `tpCat = catename, catecode, catestd;`

4.2 字典描述

本标准的附录 A 给出了科学数据共享核心元数据的数据字典。数据字典以表格的形式描述元数据的特征属性。字典中加灰的行定义的是元数据实体。数据字典通过以下七个属性定义元数据实体和元数据元素。

4.2.1 名称/角色名称

名称/角色名称是赋给元数据实体或元数据元素的一个标记。

元数据实体名称以一个大写字母开头。元数据实体名称中没有空格, 而是多个单词连写, 其中每一个新的单词开头为大写字母 (如: XnnnYmmm)。元数据实体名称在本标准的数据字典中是唯一的。

元数据元素名称在元数据实体中是唯一的, 但在本标准的数据字典中并不是唯一的。通过元数据实体和元数据元素名称的组合, 可使元数据元素名称在一个应用中唯一 (如: 元数据.元数据字符集)。

角色名称用以标识元数据抽象模型关联, 并由“角色名称:”开头, 将其与其它元数据元素相区分。

4.2.2 短名和域代码

短名的定义及命名规则参见 4.1.6 节。

对于代码表和枚举构造型，本标准为每一个可能的选择均提供了一个代码。这些域代码由三位数字表示，并在该代码表中是唯一的。每个代码表或枚举的第一行包含一个英文短名，是该代码表或枚举的英文名称缩写。

4.2.3 定义

参见 4.1.2 节。

4.2.4 约束/条件

参见 4.1.7.1 节。

4.2.5 最大出现次数

参见 4.1.7.2 节。

4.2.6 数据类型

参见 4.1.4 节。

4.2.7 域

对元数据实体而言，域说明其包含的行数。

对元数据元素而言，域说明其有效值或使用自由文本。“自由文本”表明对字段的内容没有限制。应使用基于整型的代码表示包含代码表的域值。

5 科学数据共享核心元数据

科学数据共享核心元数据元素为元数据子集和实体中必选的元数据元素，可用于数据集编目、数据交换网站活动和对数据集的描述。

5.1 内容组成

科学数据共享核心元数据包括 9 个必选的元数据实体和元数据元素，分别描述了数据集的标识、分发、质量等方面的信息以及元数据的标识信息。

5.2 科学数据共享核心元数据定义

5.2.1 数据集名称

定 义：用以识别某一数据集的专门称呼

英文名称：title

数据类型：字符串

值 域：自由文本

短 名：resTitle

注 解：必选项；最大出现次数为 1

5.2.2 数据集出版日期

定 义：发布数据集的日期

英文名称：date of publication

数据类型：日期型

值 域：按 GB/T 7408-2005 执行，格式为 CCYY-MM-DD

短 名：pubDate

注 解：必选项；最大出现次数为 1

5.2.3 数据集摘要

定 义：数据集内容的简单说明

英文名称：abstract

数据类型：字符串

值 域：自由文本

短 名：abstract

注 解：必选项；最大出现次数为 1

5.2.4 数据集负责方

定 义：对数据集的完整性、正确性、真实性等负有责任的个人或单位的名称和地址信息

英文名称：PointOfContact

数据类型：复合型

短 名：IdPoC

注 解：必选项；最大出现次数为 N

子 元 素：数据集负责方 =

0{数据集负责人姓名}1 + 数据集负责单位 +
负责方联系信息

扩展巴氏范式：IdPoC = 0{rpIndName}1, rpOrgName}1, rpCntInfo ;

5.2.4.1 数据集负责人姓名

定 义：数据集负责人姓、名、头衔，用分隔符隔开

英文名称：individualName

数据类型：字符串

值 域：自由文本

短 名：rpIndName

注 解：可选项；最大出现次数为 1

5.2.4.2 数据集负责单位

定 义：数据集负责单位名

英文名称：organisationName

数据类型：字符串

值 域：自由文本

短 名：rpOrgName

注 解：必选项；最大出现次数为 1

5.2.4.3 数据集负责方联系信息

定 义：与数据集负责人和/或负责单位联系所需的信息

英文名称：Contact

数据类型：复合型

短 名：Contact

子 元 素：数据集负责方联系信息 =

数据集负责方电话信息+

0{数据集负责方地址}1

扩展巴氏范式：Contact = cntPhone , 0{ cntAdd }1;

5.2.4.3.1 数据集负责方电话信息

定 义：与数据集负责人或负责单位通话的信息

英文名称：phone

数据类型：复合型

短 名：cntPhone

注 解：必选项；最大出现次数为 1

子 元 素：数据集负责人电话信息=

1{数据集负责人语音电话}n +

0{数据集负责人传真}n

扩展巴氏范式：cntPhone = 1{ voiceNum }n , 0{ faxNum }n;

5.2.4.3.1.1 数据集负责方电话

定 义：与数据集负责人或负责单位通话的语音电话号码

英文名称：voice

数据类型：字符串

值 域：自由文本

短 名：voiceNum

注 解：必选项；最大出现次数为 N

5.2.4.3.1.2 数据集负责方传真

定 义：数据集负责人或负责单位的传真号码

英文名称：facsimile

数据类型：字符串

值 域：自由文本

短 名：faxNum

注 解：可选项；最大出现次数为 N

5.2.4.3.2 数据集负责方地址

定 义：与数据集负责人或负责单位联系的物理地址和电子信箱

英文名称：address

数据类型：复合型

短 名：cntAdd

注 解：可选项；最大出现次数为 1

子 元 素：数据集负责方地址 =

1{数据集负责方详细地址}n +

0{数据集负责方邮政编码}1

扩展巴氏范式：cntAdd = 1{ delPoint }n , 0{ postCode }1;

5.2.4.3.2.1 数据集负责方详细地址

定 义：数据集负责方所在位置的详细地址

英文名称：delilveryPoint

数据类型：字符串

值 域：自由文本

短 名：delPoint

注 解：必选项；最大出现次数为 N

5.2.4.3.2.2 数据集负责方邮政编码

定 义：数据集负责方邮政编码

英文名称：postalCode

数据类型：字符串

值 域：自由文本

短 名：postCode

注 解：可选项；最大出现次数为 1

5.2.5 关键词

定 义：用于描述数据集主题的通用词、形式化词或短语

英文名称：keyword

数据类型：字符串

值 域：自由文本

短 名：keyword

注 解：必选项；最大出现次数为 N

5.2.6 数据集分类

定 义：数据集的分类信息

英文名称：topicCategory

数据类型：复合型

短 名：TpCat

注 解：必选项；最大出现次数为 N

子 元 素：数据集分类 =

类别名称+

类别编码+

分类标准

扩展巴氏范式：TpCat = cateName, cateCode, cateStd;

5.2.6.1 类别名称

定 义：用于描述数据集类别的通用词、形式化词或短语

英文名称：categoryName

数据类型：字符串

值 域：自由文本，参见 SDS/T 2122—2004 《科学数据共享工程数据分类编码》和各领域数据分类编码标准中各种分类的取值规定

短 名：cateName

注 解：必选项；最大出现次数为 1

5.2.6.2 类别编码

定 义：类别名称对应的编码

英文名称：categoryCode

数据类型：字符串

值 域：自由文本，参见 SDS/T 2122—2004 《科学数据共享工程数据分类编码》和各领域数据分类编码标准中各种分类的取值规定

短 名：cateCode

注 解：必选项；最大出现次数为 1

5.2.6.3 分类标准

定 义：分类标准名称

英文名称：categroyStandard

数据类型：字符串

值 域：A.2.1 数据集分类标准代码<<代码表>>

短 名：cateStd

注 解：必选项；最大出现次数为 1

5.2.7 数据志说明

定 义：数据集生产者对数据源和处理步骤的一般说明

英文名称：statement

数据类型：字符串

值 域：自由文本

短 名：statement

注 解：必选项；最大出现次数为 1

5.2.8 数据量

定 义：数据集的数据量大小

英文名称：dataQuantity

数据类型：字符串

值 域：自由文本

短 名：dtqty

注 解：可选项；最大出现次数为 1

5.2.9 数据集在线资源链接地址

定 义：可以获取数据集的下载链接地址，或可以浏览数据集的链接地址

英文名称：online

数据类型：复合型

短 名：onLineSrc

注 解：必选项；最大出现次数为 N

子 元 素：数据集在线资源链接地址=

1{数据集下载地址}n +

1{数据集浏览地址}n

扩展巴氏范式：onLineSrc = 1 { dtdllinkage } n, 1 { dtbrlinkage } n;

5.2.9.1 数据集下载地址

定 义：可获取到数据集的链接地址，一般为 ftp 或 http 下载地址

英文名称：dataDownloadlinkage

数据类型：字符串

值 域：URL (IETF RFC 1738)

短 名：dtdllinkage

注 解：必选项；最大出现次数为 N

5.2.9.2 数据集浏览地址

定 义：可浏览数据集的链接地址，一般指向该数据集的应用链接地址

英文名称：dataBrowselinkage

数据类型：字符串

值 域：URL (IETF RFC 1738)

短 名：dtbrlinkage

注 解：必选项；最大出现次数为 N

5.2.10 元数据标识符

定 义：元数据的唯一标识

英文名称：metadata Identifier

数据类型：字符串

值 域：自由文本

短 名：mdId

注 解：必选项；最大出现次数为 1；必须是第一个著录项目、标识符须唯一；由前段码和后段码组成，前段码和后段码之间用下划线（_）相连；前段码用以标识该元数据所属领域，使用领域名称的汉语拼音首字母大写表示，如医学元数据标识符的前段码为“YX”；后段码由字母（含下划线（_）短划线（-）点（.）斜线（/）逗号（,）和空格（ ）或数字组成，可由系统自动随机产生；如医学领域某元数据的标识符为“YX_metadata001”

附录 A 规范性附录
科学数据共享核心元数据字典

A.1 科学数据共享核心元数据字典

	名称/角色名称 (中文)	名称/角色名称 (英文)	短名	定义	约束/条件	最大出现次数	数据类型	域
0	元数据	Metadata	Metadata	科学数据共享核心元数据的根实体	M	1	复合型	第 1-24 行
1	数据集名称	title	resTitle	用以识别某一数据集的专门称呼	M	1	字符串	自由文本
2	数据集出版日期	date of publication	pubDate	发布数据集的日期	M	1	日期型	按 GB/T 7408-2005 执行，格式为 CCYY-MM-DD
3	数据集摘要	abstract	abstract	数据集内容的简单说明	M	1	字符串	自由文本
4	数据集负责方	pointOfContact	IdPoC	对数据集的完整性、正确性、真实性等负有责任的个人或单位的名称和地址信息	M	N	复合型	第 5-13 行
5	数据集负责人姓名	individualName	rpIndName	数据集负责人姓、名、头衔，用分隔符隔开	O	1	字符串	自由文本

6	数据集负责单位	organisationName	rpOrgName	数据集负责单位名称	M	1	字符串	自由文本
7	负责方联系信息	Contact	Contact	与数据集负责人和/或负责单位联系所需的信息	M	1	复合型	第 8-13 行
8	数据集负责方电话信息	phone	cntPhone	与数据集负责人或负责单位通话的信息	M	1	复合型	第 9-10 行
9	数据集负责方电话	voice	voiceNum	与数据集负责人或负责单位通话的语音电话号码	M	N	字符串	自由文本
10	数据集负责方传真	facsimile	faxNum	数据集负责人或负责单位的传真号码	O	N	字符串	自由文本
11	数据集负责方地址	address	cntAdd	与数据集负责人或负责单位联系的物理地址和电子信箱	O	1	复合型	第 12-13 行
12	数据集负责方详细地址	delilveryPoint	delPoint	与数据集负责人或负责单位联系的物理地址和电子信箱	M	1	字符串	自由文本
13	数据集负责方邮政编码	postalCode	postCode	数据集负责方邮政编码	O	1	字符串	自由文本
14	关键词	keyword	keyword	用于描述数据集主题的通用词、形式化词或短语	M	N	字符集	自由文本
15	数据集分类	topicCategory	TpCat	数据集的分类信息	M	N	关联	第 16-18 行

16	类别名称	categoryName	cateName	用于描述数据集类别的通用词、 形式化词或短语	M	1	字符串	自由文本, 参见 SDS/T 2122—2004《科学数据 共享工程数据分类编 码》和《XX 领域数据 分类编码》各种分类的 取值规定
17	类别编码	categoryCode	cateCode	类别名称对应的编码	M	1	字符串	自由文本, 参见 SDS/T 2122—2004《科学数据 共享工程数据分类编 码》和《XX 领域数据 分类编码》各种分类的 取值规定
18	分类标准	categoryStandard	cateStd	分类标准名称	M	1	字符串	数据集分类代码<<代 码表>> (A.2.1 数据集 分类标准代码表)
19	数据志说明	statement	statement	数据集生产者对数据源和处理 步骤的一般说明	M	1	字符串	自由文本
20	数据量	dataQuantity	dtqty	数据集的数据量大小	O	1	字符串	自由文本
21	数据集在线 资源链接地 址	online	onLineSrc	可以获取数据集的下载链接地 址, 或可以浏览数据集的链接地 址	M	N	复合型	第 22-23 行
22	数据集下载 地址	dataDownloadLinkage	dtdllinkage	可获取到数据集的链接地址, 一 般为 ftp 或 http 下载地址	M	N	字符串	URL (IETF RFC 1738)

23	数据集浏览地址		dtbrlinkage	可浏览数据集的链接地址，一般指向该数据集的应用链接地址	M	N	字符串	URL (IETF RFC 1738)
24	元数据标识符	metadata Identifier	mdId	元数据的唯一标识	M	1	字符串	自由文本

A.2 代码表

A.2.1 数据集分类标准代码表<<代码表>>

	名称（中文）	名称（英文）	域代码	定义
	数据集分类标准代码	categoryStandard Code	catestd Cd	数据集分类标准名称
1	科学数据共享工程数据分类编码		001	科学数据共享工程数据分类编码标准
2	XX 领域科学数据分类编码		002	XX 领域科学数据分类编码标准

附录 B 规范性附录
科学数据共享核心元数据 XML SCHEMA

```

<?xml version="1.0" encoding="GB2312"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified"
attributeFormDefault="unqualified">
  <xs:annotation>
    <xs:documentation>标准版本: v1.6 2006-08-25</xs:documentation>
    <xs:documentation>标准名称: 科学数据共享核心元数据标准</xs:documentation>
    <xs:documentation>制定机构: 国家信息中心</xs:documentation>
  </xs:annotation>
  <xs:element name="metadata">
    <xs:annotation>
      <xs:documentation>中文名称: 科学数据共享核心元数据根元素</xs:documentation>
      <xs:documentation>英文名称: Metadata</xs:documentation>
      <xs:documentation>定义: 科学数据共享核心元数据根元素</xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:sequence>
        <xs:element name="resTitle" type="xs:string">
          <xs:annotation>
            <xs:documentation>中文名称: 数据集名称</xs:documentation>
            <xs:documentation>英文名称: title</xs:documentation>
            <xs:documentation>定义: 用以识别某一数据集的专门称呼</xs:documentation>
          </xs:annotation>
        </xs:element>
        <xs:element name="pubDate" type="xs:date">
          <xs:annotation>
            <xs:documentation>中文名称: 数据集出版日期</xs:documentation>
            <xs:documentation>英文名称: date of publication</xs:documentation>
            <xs:documentation>定义: 发布数据集的日期</xs:documentation>
          </xs:annotation>
        </xs:element>
        <xs:element name="abstract" type="xs:string">
          <xs:annotation>
            <xs:documentation>中文名称: 数据集摘要</xs:documentation>
            <xs:documentation>英文名称: abstract</xs:documentation>
            <xs:documentation>定义: 数据集内容的简单说明</xs:documentation>
          </xs:annotation>
        </xs:element>
        <xs:element name="IdPoC" maxOccurs="unbounded">
          <xs:annotation>
            <xs:documentation>中文名称: 数据集负责方</xs:documentation>
            <xs:documentation>英文名称: PointOfContact</xs:documentation>
            <xs:documentation>定义: 对数据集的完整性、正确性、真实性等负有责任的个人或单位

```

的名称和地址信息</xs:documentation>

</xs:annotation>

<xs:complexType>

<xs:sequence>

<xs:element name="rpIndName" type="xs:string" minOccurs="0">

<xs:annotation>

<xs:documentation>中文名称：数据集负责人姓名</xs:documentation>

<xs:documentation>英文名称：individualName</xs:documentation>

<xs:documentation>定义：数据集负责人姓、名、头衔，用分隔符隔开

</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="rpOrgName" type="xs:string">

<xs:annotation>

<xs:documentation>中文名称：数据集负责单位</xs:documentation>

<xs:documentation>英文名称：rpOrgName</xs:documentation>

<xs:documentation>定义：数据集负责单位名</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="Contact">

<xs:annotation>

<xs:documentation>中文名称：负责方联系信息</xs:documentation>

<xs:documentation>英文名称：Contact</xs:documentation>

<xs:documentation>定义：与数据集负责人和/或负责单位联系所需的信息

息</xs:documentation>

</xs:annotation>

<xs:complexType>

<xs:sequence>

<xs:element name="cntPhone">

<xs:annotation>

<xs:documentation>中文名称：数据集负责方电话信息

</xs:documentation>

<xs:documentation>英文名称：

cntPhone</xs:documentation>

<xs:documentation>定义：与数据集负责人或负责单位通

话的信息</xs:documentation>

</xs:annotation>

<xs:complexType>

<xs:sequence>

<xs:element name="voiceNum" type="xs:string"

maxOccurs="unbounded">

<xs:annotation>

<xs:documentation>中文名称：数据集负

责方电话</xs:documentation>

voice</xs:documentation>	<xs:documentation>英文名称:
人或负责单位通话的语音电话号码</xs:documentation>	<xs:documentation>定义: 与数据集负责人
minOccurs="0" maxOccurs="unbounded">	</xs:annotation>
责方传真</xs:documentation>	</xs:element>
facsimile</xs:documentation>	<xs:element name="faxNum" type="xs:string"
或负责单位的传真号码</xs:documentation>	<xs:annotation>
</xs:documentation>	<xs:documentation>中文名称: 数据集负
address</xs:documentation>	<xs:documentation>英文名称:
系的物理地址和电子信箱</xs:documentation>	<xs:documentation>定义: 数据集负责人
责方详细地址</xs:documentation>	</xs:annotation>
delilveryPoint</xs:documentation>	</xs:element>
所在位置的详细地址</xs:documentation>	</xs:sequence>
minOccurs="0">	</xs:complexType>
</xs:documentation>	</xs:element>
</xs:documentation>	<xs:element name="cntAdd" minOccurs="0">
</xs:documentation>	<xs:annotation>
</xs:documentation>	<xs:documentation>中文名称: 数据集负责方地址
</xs:documentation>	<xs:documentation>英文名称:
</xs:documentation>	<xs:documentation>定义: 与数据集负责人或负责单位联
</xs:documentation>	</xs:annotation>
</xs:documentation>	<xs:complexType>
</xs:documentation>	<xs:sequence>
</xs:documentation>	<xs:element name="delPoint" type="xs:string">
</xs:documentation>	<xs:annotation>
</xs:documentation>	<xs:documentation>中文名称: 数据集负
</xs:documentation>	<xs:documentation>英文名称:
</xs:documentation>	<xs:documentation>定义: 数据集负责方
</xs:documentation>	</xs:annotation>
</xs:documentation>	</xs:element>
</xs:documentation>	<xs:element name="postCode" type="xs:string"
</xs:documentation>	<xs:annotation>

责方邮政编码</xs:documentation> <xs:documentation>中文名称：数据集负责方
 postalCode</xs:documentation> <xs:documentation>英文名称：
 邮政编码</xs:documentation> <xs:documentation>定义：数据集负责方

```

        </xs:annotation>
    </xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="keyword" type="xs:string" maxOccurs="unbounded">
    <xs:annotation>
        <xs:documentation>中文名称：关键词</xs:documentation>
        <xs:documentation>英文名称：keyword</xs:documentation>
        <xs:documentation>定义：用于描述数据集主题的通用词、形式化词或短语
    </xs:documentation>
    </xs:annotation>
</xs:element>
<xs:element name="TpCat" maxOccurs="unbounded">
    <xs:annotation>
        <xs:documentation>中文名称：数据集分类</xs:documentation>
        <xs:documentation>英文名称：topicCategory</xs:documentation>
        <xs:documentation>定义：数据集的分类信息</xs:documentation>
    </xs:annotation>
    <xs:complexType>
        <xs:sequence>
            <xs:element name="catename" type="categoryNameStandardCode">
                <xs:annotation>
                    <xs:documentation>中文名称：类别名称</xs:documentation>
                    <xs:documentation>英文名称：categoryName</xs:documentation>
                    <xs:documentation>定义：用于描述数据集类别的通用词、形式化词或短
    
```

语。 \r\n

例如： 生物技术 & 生物信息科学数据（代码"B"）； 地球系统
 科学数据（代码"G"）； \r\n
 天文与空间科学数据（代码"C"）； 信息技术
 科学数据（代码"I"）； \r\n
 材料科学数据（代码"M"）； 先进制造

科学数据（代码"P"）； \r\n	气象科学数据（代码"W"）；	水文水资
源科学数据（代码"H"）； \r\n	海洋科学数据（代码"S"）；	地矿与土地
资源科学数据（代码"L"）； \r\n	地震科学数据（代码"D"）；	环境科学
数据（代码"E"）； \r\n	基础地理科学数据（代码"G"）；	生态监测
科学数据（代码"Z"）； \r\n	对地观测科学数据（代码"R"）；	农业科学数
据（代码"S"）； \r\n	农村科技数据（代码"T"）；	林业科学数
据（代码"F"）； \r\n	交通运输科学数据（代码"J"）；	建筑工程科
学数据（代码"A"）； \r\n	能源工程科学数据（代码"E"）；	化工科学数
据（代码"C"）； \r\n	公共安全科学数据（代码"S"）；	人口与计
划生育科学（代码"P"）； \r\n	医药卫生科学数据（代码"M"）；	可持续发
展信息（代码"S"）； \r\n	高等院校科学数据（代码"A"）；	中国科学
院科学数据（代码"C"）； \r\n	国家自然科学基金科学数据（代码"N"）；	国家科技
信息管理科学数据（代码"M"）； \r\n	青藏高原科学数据（代码"Q"）	

```

</xs:documentation>
</xs:annotation>
</xs:element>
<xs:element name="catecode" type="categoryCodeStandardCode">
  <xs:annotation>
    <xs:documentation>中文名称：类别编码</xs:documentation>
    <xs:documentation>英文名称： categoryCode</xs:documentation>
    <xs:documentation>定义： </xs:documentation>
    <xs:documentation>定义：类别名称对应的编码。类别名称和类别编码的
  
```

对应关系如下： \r\n

科学数据（代码"G"）； \r\n	生物技术与生物信息科学数据（代码"B"）；	地球系统
科学数据（代码"I"）； \r\n	天文与空间科学数据（代码"C"）；	信息技术
科学数据（代码"P"）； \r\n	材料科学数据（代码"M"）；	先进制造
源科学数据（代码"H"）； \r\n	气象科学数据（代码"W"）；	水文水资

资源科学数据（代码"L"）； \r\n	海洋科学数据（代码"S"）；	地矿与土地
数据（代码"E"）； \r\n	地震科学数据（代码"D"）；	环境科学
科学数据（代码"Z"）； \r\n	基础地理科学数据（代码"G"）；	生态监测
据（代码"S"）； \r\n	对地观测科学数据（代码"R"）；	农业科学数
据（代码"F"）； \r\n	农村科技数据（代码"T"）；	林业科学数
学数据（代码"A"）； \r\n	交通运输科学数据（代码"J"）；	建筑工程科
据（代码"C"）； \r\n	能源工程科学数据（代码"E"）；	化工科学数
划生育科学（代码"P"）； \r\n	公共安全科学数据（代码"S"）；	人口与计
展信息（代码"S"）； \r\n	医药卫生科学数据（代码"M"）；	可持续发
院科学数据（代码"C"）； \r\n	高等院校科学数据（代码"A"）；	中国科学
信息管理科学数据（代码"M"）； \r\n	国家自然科学基金科学数据（代码"N"）；	国家科技
	青藏高原科学数据（代码"Q"）	

```

</xs:documentation>
</xs:annotation>
</xs:element>
<xs:element name="catestd" type="categoryStandardCode" fixed="科学数据共享工程
数据分类编码">
    <xs:annotation>
        <xs:documentation>中文名称：分类标准</xs:documentation>
        <xs:documentation>英文名称： categoryStandard</xs:documentation>
        <xs:documentation>定义：分类标准名称</xs:documentation>
    </xs:annotation>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="statement" type="xs:string">
    <xs:annotation>
        <xs:documentation>中文名称：数据志说明</xs:documentation>
        <xs:documentation>英文名称： statement</xs:documentation>
        <xs:documentation>定义：数据集生产者对数据源和处理步骤的一般说明
</xs:documentation>
</xs:annotation>

```

```

</xs:element>
<xs:element name="dataQuantity" type="xs:string" minOccurs="0">
  <xs:annotation>
    <xs:documentation>中文名称：数据量</xs:documentation>
    <xs:documentation>英文名称：dataQuantity</xs:documentation>
    <xs:documentation>定义：数据集的数据量大小</xs:documentation>
  </xs:annotation>
</xs:element>
<xs:element name="onLineSrc" maxOccurs="unbounded">
  <xs:annotation>
    <xs:documentation>中文名称：数据集在线资源链接地址</xs:documentation>
    <xs:documentation>英文名称：online</xs:documentation>
    <xs:documentation>定义：可以获取数据集的下载链接地址，或可以浏览数据集的链接地址</xs:documentation>
  </xs:annotation>
  <xs:complexType>
    <xs:sequence>
      <xs:element name="dtdllinkage" maxOccurs="unbounded">
        <xs:annotation>
          <xs:documentation>中文名称：数据集下载地址</xs:documentation>
          <xs:documentation>英文名称：dataDownloadlinkage</xs:documentation>
          <xs:documentation>定义：可获取到数据集的链接地址，一般为ftp或http
          下载地址</xs:documentation>
        </xs:annotation>
      </xs:element>
      <xs:element name="dtbrlinkage" maxOccurs="unbounded">
        <xs:annotation>
          <xs:documentation>中文名称：数据集浏览地址</xs:documentation>
          <xs:documentation>英文名称：dataBrowselinkage</xs:documentation>
          <xs:documentation>定义：可浏览数据集的链接地址，一般指向该数据集
          的应用链接地址</xs:documentation>
        </xs:annotation>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="mdId" type="xs:string">
  <xs:annotation>
    <xs:documentation>中文名称：元数据标识符</xs:documentation>
    <xs:documentation>英文名称：metadata Identifier</xs:documentation>
    <xs:documentation>定义：元数据的唯一标识</xs:documentation>
    <xs:documentation>注解：由前段码和后段码组成，前段码和后段码之间用下划线（_）
    相连；前段码用以标识该元数据所属领域，使用领域名称的汉语拼音首字母大写表示，如医学元数据标识符的前段码为
    “YX”；后段码由字母（含下划线（_）短划线（-）点（.）斜线（/）逗号（，）和空格（ ））或数字组成，可由系统
  
```

自动随机产生；如医学领域某元数据的标识符为“YX_metadata001” </xs:documentation>

```

    </xs:annotation>
  </xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:simpleType name="categroyStandardCode">
  <xs:restriction base="xs:string">
    <xs:enumeration value="科学数据共享工程数据分类编码"/>
    <xs:enumeration value="气象科学领域科学数据分类编码"/>
    <xs:enumeration value="海洋科学领域科学数据分类编码"/>
    <xs:enumeration value="水文水资源科学领域科学数据分类编码"/>
    <xs:enumeration value="国土资源科学领域科学数据分类编码"/>
    <xs:enumeration value="地震科学领域科学数据分类编码"/>
    <xs:enumeration value="农业科学领域科学数据分类编码"/>
    <xs:enumeration value="林业科学领域科学数据分类编码"/>
    <xs:enumeration value="地球系统科学领域科学数据分类编码"/>
    <xs:enumeration value="医药卫生科学领域科学数据分类编码"/>
    <xs:enumeration value="基础科学领域科学数据分类编码"/>
    <xs:enumeration value="可持续发展科学领域科学数据分类编码"/>
    <xs:enumeration value="材料科学领域科学数据分类编码"/>
    <xs:enumeration value="先进制造与自动化科学领域科学数据分类编码"/>
    <xs:enumeration value="能源科学领域科学数据分类编码"/>
    <xs:enumeration value="交通科学领域科学数据分类编码"/>
    <xs:enumeration value="环境科学领域科学数据分类编码"/>
    <xs:enumeration value="区域综合科技信息领域科学数据分类编码"/>
    <xs:enumeration value="测绘科学领域科学数据分类编码"/>
  </xs:restriction>
</xs:simpleType>
<xs:simpleType name="categoryNameStandardCode">
  <xs:restriction base="xs:string">
    <xs:enumeration value="生物技术与生物信息科学数据"/>
    <xs:enumeration value="地球系统科学数据"/>
    <xs:enumeration value="天文与空间科学数据"/>
    <xs:enumeration value="信息技术科学数据"/>
    <xs:enumeration value="材料科学数据"/>
    <xs:enumeration value="先进制造科学数据"/>
    <xs:enumeration value="气象科学数据"/>
    <xs:enumeration value="水文水资源科学数据"/>
    <xs:enumeration value="海洋科学数据"/>
    <xs:enumeration value="地矿与土地资源科学数据"/>
    <xs:enumeration value="地震科学数据"/>
    <xs:enumeration value="环境科学数据"/>
    <xs:enumeration value="基础地理科学数据"/>
  </xs:restriction>
</xs:simpleType>

```

```

<xs:enumeration value="生态监测科学数据"/>
<xs:enumeration value="对地观测科学数据"/>
<xs:enumeration value="农业科学数据"/>
<xs:enumeration value="农村科技数据"/>
<xs:enumeration value="林业科学数据"/>
<xs:enumeration value="交通运输科学数据"/>
<xs:enumeration value="建筑工程科学数据"/>
<xs:enumeration value="能源工程科学数据"/>
<xs:enumeration value="化工科学数据"/>
<xs:enumeration value="公共安全科学数据"/>
<xs:enumeration value="人口与计划生育科学"/>
<xs:enumeration value="医药卫生科学数据"/>
<xs:enumeration value="可持续发展信息"/>
<xs:enumeration value="高等院校科学数据"/>
<xs:enumeration value="中国科学院科学数据"/>
<xs:enumeration value="国家自然科学基金科学数据"/>
<xs:enumeration value="国家科技信息管理科学数据"/>
<xs:enumeration value="青藏高原科学数据"/>
</xs:restriction>
</xs:simpleType>
<xs:simpleType name="categoryCodeStandardCode">
  <xs:restriction base="xs:string">
    <xs:enumeration value="A"/>
    <xs:enumeration value="B"/>
    <xs:enumeration value="C"/>
    <xs:enumeration value="D"/>
    <xs:enumeration value="E"/>
    <xs:enumeration value="F"/>
    <xs:enumeration value="G"/>
    <xs:enumeration value="H"/>
    <xs:enumeration value="I"/>
    <xs:enumeration value="L"/>
    <xs:enumeration value="M"/>
    <xs:enumeration value="N"/>
    <xs:enumeration value="P"/>
    <xs:enumeration value="Q"/>
    <xs:enumeration value="R"/>
    <xs:enumeration value="S"/>
    <xs:enumeration value="T"/>
    <xs:enumeration value="W"/>
    <xs:enumeration value="Z"/>
  </xs:restriction>
</xs:simpleType>
</xs:schema>

```


附录 C 资料性附录

科学数据共享核心元数据举例

```

<?xml version="1.0" encoding="GB2312"?>
<metadata xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="E:\科学数据共享工程\20060821科学数据共享核心元数据\科学数据共享核心元数据内容_v0.8.xsd">
  <resTitle>中国地面气候资料日值数据 </resTitle>
  <pubDate>2004-02-21</pubDate>
  <abstract>本数据集为中国740个地面气象观测站1951-2000年地面日资料集,包括日平均气压、日平均气温、日平均相对湿度、日平均风速、日降水量、日日照时数、日最小相对湿度、日最高气压、日最低气压、日最高气温、日最低气温、日极大风向风速、日最大风向风速等13个要素的日资料。</abstract>
  <IdPoC>
    <rpOrgName>国家气象信息中心</rpOrgName>
    <Contact>
      <cntPhone>
        <voiceNum>(010)68407499 </voiceNum>
      </cntPhone>
      <cntAdd>
        <delPoint>中国气象局 国家气象信息中心 气象资料室</delPoint>
      </cntAdd>
    </Contact>
  </IdPoC>
  <keyword>地面、日值</keyword>
  <TpCat>
    <catename>气象科学数据</catename>
    <catecode>W</catecode>
    <catestd>科学数据共享工程数据分类编码</catestd>
  </TpCat>
  <statement>气压(日平均、最高、最低)在地面月报信息化资料基础上,制作者根据《地面气象观测规范》规定统计计算,其它要素均取自现气象资料室数据库日值表。</statement>
  <onLineSrc>
    <dtdllinkage>http://cdc.cma.gov.cn/noauth_download.jsp</dtdllinkage>
    <dtbrlinkage>http://cdc.cma.gov.cn/noauth.jsp</dtbrlinkage>
  </onLineSrc>
  <mdId>QX_metadata001</mdId>
</metadata>

```

参考文献

- [1] 都柏林核心元数据集 (Dublin Core Metadata Element Set,Version 1.1: Reference Description.2003-06-02)
- [2] ISO 19115 地理信息——元数据(Geographic information—Metadata)
-