团体标准

发 布

中国电机工程学会

20XX—XX—XX实施

20XX—XX—XX发布

能源大数据 数据资源目录

Energy big data resources directory

（征求意见稿）

T/CSEE XXXX—YYYY

代替 T/XXXX

ICS 35.240

CCS L67

目 次

[前 言 1](#_Toc7014)

[1 范围 2](#_Toc31611)

[2 规范性引用文件 2](#_Toc18828)

[3 术语和定义 2](#_Toc26270)

[4 数据资源目录体系及编码规范 2](#_Toc17164)

[4.1 体系构成 2](#_Toc30339)

[4.2 数据分类 3](#_Toc5837)

[4.3 编码规范 3](#_Toc1811)

[5 数据资源目录管理 4](#_Toc32578)

[5.1 管理流程 4](#_Toc18013)

[5.2 主要环节 4](#_Toc519)

[附　录　A （规范性） 能源大数据目录编码表 6](#_Toc31816)

[参 考 文 献 12](#_Toc9523)

前 言

本标准按照《中国电机工程学会标准管理办法（暂行）》的要求，依据GBT 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》的规则起草。

本标准的某些内容可能涉及专利。本标准的发布机构不承担识别这些专利的责任。

本标准由中国电机工程学会提出。

本标准由中国电机工程学会电力信息化专业委员会技术归口并解释。

本标准起草单位：国网河南省电力公司经济技术研究院，国家电网有限公司大数据中心，国网河南省电力公司，国家发展和改革委员会能源研究所，中国能源建设股份有限公司，华为数字能源技术有限公司，河南省发展和改革委员会，国网江苏省电力有限公司，国网山东省电力公司，国网吉林省电力有限公司，西安交通大学，郑州大学，华北电力大学，中国电建集团，河南九域腾龙信息工程有限公司。

本标准主要起草人：白宏坤，王圆圆，周春雷，刘湘莅，王世谦，贾一博，王涵，王思宁，卜飞飞，华远鹏，韩丁，李俊妮，董新微，郝福忠，杨宇方，陈军，邢子涯，邱国卓，魏良，李晓旭，范磊，郑雅楠，刘伯宇，宾军志，童艳丽，张健，牛斌斌，王自强，牛金星，于雪辉，郭正宾，胡吉殿，李干生。

本标准在执行过程中的意见或建议反馈至中国电机工程学会标准执行办公室（地址：北京市西城区白广路二条1号，100761，网址：http://www.csee.org.cn，邮箱：cseebz@csee.org.cn）。

能源大数据 数据资源目录

1. 范围

本文件规定了能源大数据资源目录的定义和构成，确立了数据资源目录的构建原则，明确了数据资源目录包括的各类主题域对象，制定了数据资源目录的编码规则，规范了数据资源目录管理使用流程。

1. 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

DB35/T 1999—2001 自然资源空间数据目录体系

GB/T 21063.1—2007 政务信息资源目录体系 第1部分：总体框架

GBT 21063.2—2007 政务信息资源目录体系 第2部分：技术要求

GB/T 21063.3—2007 政务信息资源目录体系 第3部分：核心元数据

GB/T 20163.5—2007 政务信息资源目录体系 第5部分：政务信息资源标识符编码方案

GBT 21063.2—2007 政务信息资源目录体系 第6部分：技术管理要求

1. 术语和定义

能源大数据资源数据目录是实现能源领域数据资源共享、业务协同和数据应用的基础，可以提供能源大数据领域数据浏览、查找、共享和应用的服务，下列术语和定义适用于本文件。

编码 coding

给事物或概念赋予代码的过程。

1. 数据资源目录体系及编码规范
	1. 体系构成

能源大数据资源目录体系架构包括信息系统、库表、数据目录、目录管理、数据应用。信息系统依托终端设备、数据中台、WEB服务、公开数据作为数据来源，将能源信息数据梳理后配置在各个库表中，N个库表与数据目录形成关联关系，通过目录管理为上层数据应用提供能源信息的访问、管理、获取等服务，如图1所示。



图1 能源大数据资源目录体系构成

* 1. 数据分类

按照主题将能源大数据分为：能源综合、煤炭、石油、天然气、电力、太阳能、风能、水能、生物质能、核能、海洋能、氢能、热能、自来水、新兴用户、碳排放等16个能源主题以及经济社会、政务、气象、环境、其他等5个能源密切相关主题。



图2 能源大数据资源目录体系构成

* 1. 编码规范
1.

数据资源目录编码共二十位，包含行政区划代码、一级主题域、二级主题域、指标名称、指标频度、指标维度。其中行政区划代码6位、一级主题域3位，二级主题域3位，指标名称4位，指标频度2位，指标维度2位，具体结构如图3所示。



图3 能源大数据资源目录编码结构

各部分的说明如下：

1. 行政区划代码：共6位，行政代码是中国国家行政机关的识别符号，第一、二位表示省（自治区、直辖市），第三、四位表示市（地区、自治州），第五、六位表示县（市辖区、县级市）。例如：河南省郑州市中原区410102、河南省洛阳市老城区410302等。
2. 一级主题域：共3位，用于表示数据资源所属的主题域分类，例如：能源综合、煤炭、石油、天然气、电力等。具体的分类代码参考附录A。
3. 二级主题域：共3位，各一级主题域下属的二级主题域划分，例如：煤炭主题域包括煤炭资源、煤炭生产、煤炭加工转换等二级主题域。具体的分类代码参考附录A。
4. 指标名称：共4位，各二级主题域下属的指标划分，例如：煤炭生产二级主题域包含煤炭项目信息、煤炭产能、原煤生产量、煤炭开采量等指标。
5. 指标频度：共2位，标识各指标采集频率，例如秒、分钟、小时、天、周、月、年等。
6. 指标维度：共2位，标识指标的维度，用于表示指标的某种特征，例如性别、地域、时间等。
7. 数据资源目录管理
	1. 管理流程

能源大数据资源目录体系管理流程包括规划、编目、注册、审核、发布、使用6个环节，规划包括数据准备，编目包括资源编目，注册包括目录注册/更新、数据挂接，审核包括目录审核、挂接事项审核；发布包括目录发布，使用包括目录查询、数据获取。具体如图4所示。



图4 能源大数据资源目录管理流程

* 1. 主要环节

能源大数据资源目录管理所涉及的主要环节如下：

* + 1. 规划
1. 将业务需求范围内的数据资源进行数据清洗，检查数据质量，如：数据是否完整，是否有错，发现并纠正数据的缺失值、严重噪声值、不一致值、不完整值等问题。
2. 基于清洗后数据资源规划数据资源目录，规划过程中应保证数据资源目录的完整性、内容规范准确。
	* 1. 编目
3. 通过SQL语句获取能源数据的全量表，涉及信息系统数据库中所有有效用户的表，按照系统配置表规则对全量表进行标注和编排。
4. 按照数据资源目录编码规范，建立数据资源目录与有效数据表的关联关系。
	* 1. 注册
5. 根据表和数据资源的关系，梳理和建立数据目录和表间的挂接关系，将有效数据表挂接至相应的数据目录层级，形成能源大数据资源目录。
6. 将目录内容注册至目录链节点，维护、更新共享库数据及目录内容。
	* 1. 审核
7. 对注册的资源目录内容进行审核，审核通过后待发布上架。发布上架的目录内容变更，需经过审核后重新发布上架。
8. 对数据挂接的技术信息进行审核，审核通过后待发布上架。
	* 1. 发布

a）按照数据资源目录发布管理要求，将通过审核的数据资源目录进行发布上架。

b）发布后将数据资源目录进行实用化推广，并收集数据资源目录优化建议。

* + 1. 使用

a）通过数据应用场景查询数据资源目录内容，提交数据获取申请。

b）数据获取申请审批通过后，可获取实时或定时的数据服务。

* + 1. 迭代更新
1. 当数据资源目录设计有新的业务需求时，结合最新需求和外部条件，对当前数据资源目录进行迭代更新。
2. 基于使用阶段反馈的优化建议，对数据目录进行迭代更新优化。
3. （规范性）
 能源大数据目录编码表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **序号** | **一级主题域** | **类别码** | **二级主题域** | **类别码** | **起始编码** |
|  1 | 能源综合 | ES | 能源生产 | 01 | 410101ES0100010101 |
| 能源供应 | 02 | 410101ES0200010101 |
| 能源效率 | 03 | 410101ES0300010101 |
| 能源消费 | 04 | 410101ES0400010101 |
| 能源贸易 | 05 | 410101ES0500010101 |
| 能源投资 | 06 | 410101ES0600010101 |
|  2 | 煤炭 | CO | 煤炭资源 | 01 | 410101CO0100010101 |
| 煤炭生产 | 02 | 410101CO0200010101 |
| 煤炭加工转换 | 03 | 410101CO0300010101 |
| 煤炭运输存储 | 04 | 410101CO0400010101 |
| 煤炭消费 | 05 | 410101CO0500010101 |
| 煤炭贸易 | 06 | 410101CO0600010101 |
| 煤炭价格 | 07 | 410101CO0700010101 |
| 煤炭经营投资 | 08 | 410101CO0800010101 |
| 3 | 石油 | PE | 石油资源 | 01 | 410101PE0100010101 |
| 石油生产 | 02 | 410101PE0200010101 |
| 石油加工转换 | 03 | 410101PE0300010101 |
| 石油运输存储 | 04 | 410101PE0400010101 |
| 石油贸易 | 05 | 410101PE0500010101 |
| 石油经营投资 | 06 | 410101PE0600010101 |
| 4 | 天然气 | NG | 天然气资源 | 01 | 410101NG0100010101 |
| 天然气生产 | 02 | 410101NG0200010101 |
| 天然气供给 | 03 | 410101NG0300010101 |
| 天然气运输存储 | 04 | 410101NG0400010101 |
| 天然气消费 | 05 | 410101NG0500010101 |
| 天然气贸易 | 06 | 410101NG0600010101 |
| 天然气价格 | 07 | 410101NG0700010101 |
| 天然气经营投资 | 08 | 410101NG0800010101 |
|  5 | 电力 | DL | 电力生产 | 01 | 410101DL0100010101 |
| 电力运输 | 02 | 410101DL0200010101 |
| 电力变换 | 03 | 410101DL0300010101 |
| 电力分配 | 04 | 410101DL0400010101 |
| 电力消费 | 05 | 410101DL0500010101 |
| 电力存储 | 06 | 410101DL0600010101 |
| 电力交易 | 07 | 410101DL0700010101 |
| 电力价格 | 08 | 410101DL0800010101 |
| 电力经营投资 | 09 | 410101DL0900010101 |
| 电力发展 | 10 | 410101DL1000010101 |
| 电力运行 | 11 | 410101DL1100010101 |
| 6 | 太阳能 | SE | 太阳能资源禀赋 | 01 | 410101SE0100010101 |
| 太阳能发电 | 02 | 410101SE0200010101 |
| 太阳能发电消费 | 03 | 410101SE0300010101 |
|  7 | 风能 | WE | 风能资源禀赋 | 01 | 410101WE0100010101 |
| 风能发电 | 02 | 410101WE0200010101 |
| 风能发电消费 | 03 | 410101WE0300010101 |
| 8 | 水能 | RW | 水能资源禀赋 | 01 | 410101RW0100010101 |
| 水能发电 | 02 | 410101RW0200010101 |
| 水能发电消费 | 03 | 410101RW0300010101 |
| 9 | 生物质能 | BE | 生物质资源禀赋 | 01 | 410101BE0100010101 |
| 生物质生产 | 02 | 410101BE0200010101 |
| 生物质贸易 | 03 | 410101BE0300010101 |
| 生物质消费 | 04 | 410101BE0400010101 |
| 10 | 核能 | NE | 核能资源禀赋 | 01 | 410101NE0100010101 |
| 核能发电 | 02 | 410101NE0200010101 |
| 核能消费 | 03 | 410101NE0300010101 |
| 11 | 海洋能 | OE | 海洋能资源禀赋 | 01 | 410101OE0100010101 |
| 海洋能发电 | 02 | 410101OE0200010101 |
| 海洋能消费 | 03 | 410101OE0300010101 |
|  12 | 氢能 | HE | 氢能资源 | 01 | 410101HE0100010101 |
| 氢能生产 | 02 | 410101HE0200010101 |
| 氢能运输存储 | 03 | 410101HE0300010101 |
| 氢能消费 | 04 | 410101HE0400010101 |
| 13 | 热能 | TE | 热能资源 | 01 | 410101TE0100010101 |
| 热能供应 | 02 | 410101TE0200010101 |
| 热能消费 | 03 | 410101TE0300010101 |
| 热力调度 | 04 | 410101TE0400010101 |
| 14 | 自来水 | TW | 水资源 | 01 | 410101TW0100010101 |
| 自来水生产 | 02 | 410101TW0200010101 |
| 自来水运输存储 | 03 | 410101TW0300010101 |
| 自来水消费 | 04 | 410101TW0400010101 |
| 自来水销售 | 05 | 410101TW0500010101 |
| 15 | 新兴用户 | EU | 新能源汽车 | 01 | 410101EU0100010101 |
| 充电设施 | 02 | 410101EU0200010101 |
| 16 | 碳排放 | CE | 能源碳排放 | 01 | 410101CE0100010101 |
| 煤炭碳排放 | 02 | 410101CE0200010101 |
| 石油碳排放 | 03 | 410101CE0300010101 |
| 天然气碳排放 | 04 | 410101CE0400010101 |
| 电力碳排放 | 05 | 410101CE0500010101 |
| 太阳能发电碳减排 | 06 | 410101CE0600010101 |
| 风能发电碳减排 | 07 | 410101CE0700010101 |
| 水能发电碳减排 | 08 | 410101CE0800010101 |
| 生物质碳排放 | 09 | 410101CE1000010101 |
| 核能发电碳减排 | 10 | 410101CE1100010101 |
| 海洋能发电碳减排 | 11 | 410101CE1200010101 |
| 12 | 410101CE1300010101 |
| 氢能碳减排 | 13 | 410101CE1400010101 |
| 其他能源碳排放 | 14 | 410101CE1500010101 |
| 碳交易 | 15 | 410101CE1600010101 |
| 17 | 经济社会 | EC | 国民经济核算 | 01 | 410101EC0100010101 |
| 人口 | 02 | 410101EC0200010101 |
| 固定资产投资 | 03 | 410101EC0300010101 |
| 价格 | 04 | 410101EC0400010101 |
| 人民生活 | 05 | 410101EC0500010101 |
| 规模以上工业 | 06 | 410101EC0600010101 |
| 18 | 政务 | GD | 经济管理 | 01 | 410101GD0100010101 |
| 国土资源 | 02 | 410101GD0200010101 |
| 能源管理 | 03 | 410101GD0300010101 |
| 城乡建设 | 04 | 410101GD0400010101 |
| 档案明细 | 05 | 410101GD0500010101 |
| 19 | 气象 | WD | 气温 | 01 | 410101WD0100010101 |
| 风速 | 02 | 410101WD0200010101 |
| 光照 | 03 | 410101WD0300010101 |
| 降水 | 04 | 410101WD0400010101 |
| 湿度 | 05 | 410101WD0500010101 |
| 20 | 生态环境 |  ED | 生态环境质量 | 01 | 410101ED0100010101 |
| 环境空气质量 | 02 | 410101ED0200010101 |
| 水质量 | 03 | 410101ED0300010101 |
| 声环境质量 | 04 | 410101ED0400010101 |
| 21 | 其他 | ZH |  GIS数据 | 01 | 410101ZH0100010101 |
| 交通数据 | 02 | 410101ZH0200010101 |
|  …… |  |  |

参 考 文 献

[1]T/CSEE XXX—2022 能源大数据 总则

[2]T/CSEE XXX—2022 能源大数据 术语

[3]T/CSEE XXX—2022 能源大数据 数据目录

**━━━━━━━━━━━**