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 数据应用

Energy big data data applications

（征求意见稿）

T/CSEE XXXX—2022

代替 T/XXXX

ICS 35.240

CCS L 67

目 次

[前言 II](#_Toc106294703)

[1 范围 1](#_Toc106294704)

[2 规范性引用文件 1](#_Toc106294705)

[3 术语和定义 1](#_Toc106294706)

[4 能源大数据应用体系 1](#_Toc106294707)

[4.1 概述 1](#_Toc106294708)

[4.2 服务政府 2](#_Toc106294709)

[4.3 服务企业 6](#_Toc106294710)

[4.4 服务公众 8](#_Toc106294711)

[5 数据应用工作流程 9](#_Toc106294712)

[5.1 概述 9](#_Toc106294713)

[5.2 应用需求 9](#_Toc106294714)

[5.3 数据需求 9](#_Toc106294715)

[5.4 数据准备 10](#_Toc106294716)

[5.5 模型构建 10](#_Toc106294717)

[5.6 成果应用及评估 11](#_Toc106294718)

[5.7 迭代更新 11](#_Toc106294719)

[5.8 应用退役 11](#_Toc106294720)

[参考文献 12](#_Toc106294721)

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。

本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本文件由中国电机工程学会提出。

本文件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。

本文件起草单位：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。

本文件主要起草人：白宏坤，王圆圆，周春雷，刘湘莅，王世谦，王涵，贾一博，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。

本文件首次发布。

本文件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）。](mailto:cseebz@csee.org.cn）。)

能源大数据 数据应用

1. 范围

本文件规定了能源大数据领域的数据应用体系规范，包括数据应用的体系框架以及相关的工作流程。

本文件适用于能源大数据应用的规划、建设、应用及管理工作。

1. 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文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；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。

GB/T 35295-2017 信息技术 大数据 术语

1. 术语和定义

T/CSEE XXX—2022界定的术语和定义适合于本文件。

2. 能源大数据应用体系
   1. 概述

在汇集能源数据的基础上，结合“辅助政府科学决策、支撑企业精益管理、服务公众智慧用能”需求，构建形成多元化数据应用体系，为政府、企业、公众等各主体提供数据应用服务。能源大数据应用体系图如图1所示：



图1 能源大数据应用体系图

* 1. 服务政府
     1. 能源经济

通过汇聚能源、煤炭、石油、天然气、电力、新能源、经济及环境等领域数据，开展分区域、分品类的能源经济发展形态对比，支撑政府进行科学管理和决策。主要内容宜满足如下要求：

1. 分区域经济发展对比分析，包括园区、县、市、省（自治区）、国家等维度的GDP、三次产业、人口、城镇化率及人均GDP等年度发展数据；
2. 分区域、分品类能源发展对比分析，包括园区、县、市、省（自治区）、国家等维度的能源、能源、煤炭、石油、天然气、电力、新能源等年度发展数据；
3. 对重点区域的能源经济发展进行对比分析。
   * 1. 能源规划

辅助支撑煤炭、石油、天然气、电力、新能源等多品类能源系统规划。主要内容宜满足如下要求：

1. 分区域、分品类能源资源量分布分析，包括煤炭矿区分布、油田、气田分布、地质层分布等；
2. 分区域、分品类能源项目分布及网络接线定位，包括已有项目、在建项目、规划项目分布及之间相互关系；
3. 分区域、分品类能源项目管理和统计分析，包括项目名称、规模、位置、核准情况、建设进度、投产时间、项目单位等基本属性信息等；
4. 分区域、分品类能源政策和行业发展动态汇集。
   * 1. 能源监测

通过汇聚煤炭、石油、天然气、电力、新能源等多种能源数据，实现分区域、分品类能源运行可视化监测。主要内容宜满足如下要求：

1. 分区域、分行业发电量、装机结构、利用小时数、用电量等数据分析和可视化展示；
2. 煤炭产量、库存、价格等数据分析和可视化展示；
3. 石油、天然气产量、销量、价格等数据分析和可视化展示。
   * 1. 能流图

通过可视化手段实现能源流向的清晰化、可视化，支撑政府部门精准掌握能源流向趋势。主要内容宜满足如下要求：

1. 基于能源大数据中心的数据，建立区域能源流转分析模型；
2. 依据区域内各类能源的生产与本地消纳规模，对可对外部输出或需从外部输入的能源类型与规模进行分析；
3. 支持将能源从生产、传输直至最终消费的过程数据进行可视化，形成完整的能源流向图。
   * 1. 能源预测

用于辅助支撑煤炭、石油、天然气、电力、新能源等多种能源的趋势预测分析。主要内容宜满足如下要求：

1. 基于能源大数据中心汇聚的近年来电力供需数据，建立电力供需预测模型，对未来发电量、装机结构、负荷等数据进行趋势预测分析；
2. 基于能源大数据中心汇聚的近年来能源供需数据，建立能源供需预测模型，对未来能源的需求量和需求结构进行趋势预测分析。
   * 1. 能源预警

结合区域能源发展规划和运行情况，量化评估能源预警等级。主要内容包括：

1. 综合分析区域内的电力供需形势，通过数据量化区域内发电量、用电量、负荷及有序用电等数据；
2. 综合分析区域内的煤炭、石油、天然气等能源供需形势，通过数据量化区域内能源供应量、需求量、缺口量等数据；
3. 根据电力供需形势和能源供需形势量化评估区域能源预警等级。
   * 1. 碳监测

用于辅助支撑分区域、分行业碳排放的监测分析。主要内容包括：

1. 支持对监测区域内整体的能源消费总量、能源消费结构、碳排放总量、碳排放强度等核心指标进行监测分析；
2. 支持对世界、全国、各省、各地市等维度的碳排放指标进行横向对比分析；
3. 支持对各区域、主要领域、重点行业、重点企业的碳排放监测分析。
   * 1. 碳预测

用于辅助支撑区域碳排放的趋势预测分析。主要内容包括：

1. 基于能源供需预测模型和电力供需预测模型，开展区域碳达峰预测模型的构建；
2. 支持对重点行业、重点领域的碳排放数据进行趋势预测分析，对分行业30·60碳达峰碳中和目标进行预测分析并给出建议；
3. 支持利用区域实际能源消费数据进行预测模型运算，对分区域30·60碳达峰碳中和目标进行预测分析并给出建议。
   * 1. 碳预警

结合区域双碳相关数据和规划情况，量化评估双碳预警等级。主要内容包括：

1. 支持根据分行业的双碳目标预测结果，动态监测距离目标值的偏差程度，量化评估行业双碳预警等级；
2. 支持根据分区域的双碳目标预测结果，动态监测距离目标值的偏差程度，量化评估区域双碳预警等级；
3. 支持当偏差较大时给出预警等级，并对根据分析结果给出相关政策建议。
   * 1. 碳流图

通过可视化手段实现碳排放流向的清晰化、可视化，支撑政府部门精准掌握碳排放流向趋势。主要内容包括：

1. 支持基于能源大数据中心的数据，建立区域碳排放流转分析模型；
2. 支持将碳排放从生产、传输直至最终消费的过程数据进行可视化，形成完整的区域碳排放流向图；
3. 支持对区域碳流图进行多维度的数据对比分析。
   * 1. 新能源规划

结合区域新能源产业现状和发展规划，引导新能源产业有序发展和场站的合理布局。主要内容包括：

1. 支持区域新能源项目信息化管理，支持点击查询项目名称、装机容量、核准情况、建设进度、投产时间、项目单位等基本属性信息，支持分区规模统计分析；
2. 实现区域新能源相关政策和行业发展动态搜集整理，近、中、远期新能源装机容量、并网规模、发电量预测；
3. 支撑提供区域新能源项目并网规模、并网方式、并网时间的推荐方案。
   * 1. 新能源监测

通过接入新能源场站发电等实时数据以及新能源相关项目属性信息，实现区域新能源运行可视化监测。主要内容包括：

1. 支持不同区域新能源装机容量、发电量和利用小时数变化趋势的可视化展示；
2. 支持不同区域新能源资源禀赋、场站分布、出力系数的对比分析；
3. 支持新能源场站环境气象、出力曲线、停检修情况的实时监测。
   * 1. 新能源预警

结合区域新能源发展规划和运行情况，量化评估新能源消纳等级。主要内容包括：

1. 支持区域新能源消纳预测，包括规划年份的可再生能源消纳权重、弃风弃光情况等；
2. 支持区域新能源规划执行进度监测，包括新能源规划并网完成进度、可再生能源配额完成进度、新能源调峰困难弃电比例、新能源网架约束弃电比例等。
   * 1. 行业用电监测

监测各区域、各大类行业、各细分行业、重点企业的用电情况，开展电力看经济大数据分析，为政府部门经济决策提供支撑。主要内容包括：

1. 支持分区域、分行业、分产业、分企业等多维度日用电量监测，支撑运行态势的精细化监测；
2. 支持GDP和用电量耦合关系分析、重点行业发展速度对比分析等功能，支撑基于电力视角的经济运行态势研判；
3. 支持数据查询、报表导出、分析报告自动生成等功能。
   * 1. 规上工业运行监测

通过构建规模以上工业用电监测分析体系，实现规模以上工业的每日运行监测，为政府部门及时决策、精准调度提供数据支撑。主要内容包括：

1. 支持分区域、分行业、分产业、规上企业的日用电量监测分析，从用电同比、电量环比等维度揭示规上工业企业运行态势；
2. 支持规上企业用电活跃度分析，从增产、稳产、减产、停产等维度对规上企业运行态势进行分析；
3. 支持战略新兴、高新技术、传统支柱等相关产业监测分析，支持工业新旧动能产业指数构建，为评估工业新旧动能发展态势。
   * 1. 小微企业运行监测

通过构建小微企业景气指数，从电力视角反映小微企业运行态势和景气程度，为助力小微企业纾困发展提供数据支撑。主要内容包括：

1. 支持分区域、分行业小微企业日用电情况监测分析，从每日用电量、用电报装情况，反映小微企业每日运行态势；
2. 支持小微企业景气指数构建和应用，从用电增长、用电活跃度、用电报装三个方面反映小微企业景气程度；
3. 支持小微企业用电活跃度分析，从增产、稳产、减产、停产等维度，分析分区域、分行业小微企业运行态势和景气程度。
   * 1. 乡村振兴发展监测

从经济社会、民生保障等角度，针对特定区域开展乡村振兴发展进程监测。主要内容包括：

1. 支持区域经济、社会发展、电网建设等数据接入，实现乡村振兴发展进程监测；
2. 支持典型区域（例如：地市、县域、乡镇）乡村振兴发展态势分析；
3. 基于能源大数据中心数据基础，构建乡村振兴电力指数，实现区域乡村振兴态势研判以及区域乡村村振兴情况的量化对比分析。
   * 1. 燃煤机组排放监测

对燃煤机组污染物排放和能源消耗开展实时监测，支撑政府有效监管和电力绿色调度。主要内容包括：

1. 支持采集燃煤电厂机组能耗数据，对区域燃煤发电机组的煤耗进行监测展示和异常数据实时报警；
2. 支持对区域燃煤机组的能源消耗水平及生产情况进行统计分析；
3. 支持构建燃煤机组能耗指标体系，对燃煤发电机组的煤耗进行排序，为实施绿色环保调度提供依据。
   * 1. 环保管控分析

监测重点企业的生产情况，向政府环保部门推送异常用电数据，支撑重点企业精细化环保管控。主要内容包括：

1. 支持全量超限企业、超限电量、超限率动态监测分析，实时推送相关数据；
2. 支持管控企业生产分析预警，向政府环保部门推送异常用电企业清单，支撑远程开展环保管控；
3. 支持报告自动生成，定期对企业超限情况进行分析总结，提报相关管理部门。
   * 1. 电力需求响应管理

针对电力用户开展基于需求响应的负荷管理，实现需求申请、响应签约、响应执行、效果评估等全环节、全过程线上运作。主要内容包括：

a) 支持收集展示区域用电负荷曲线、地区峰谷差率、需求响应能力等基础数据；

b) 根据实际需要，开展特定区域的需求响应执行，并根据实际执行效果进行评估服务。

* + 1. 充电设施管理

针对电动汽车充电设施开展数据统一接入管理，服务政府充电设施监管。主要内容包括：

1. 支持运营商充电设施运行数据的接入，支持对充电设施运营商规范化档案信息查询管理和充电设施运行状态实时监测；
2. 支持对运营商接入的充电站、充电桩、充电枪档案自动生成统一编码，构建充电设施标准数据体系；
3. 支持对充电设施数据信息进行分析整合，实现对充电设施建设、运营、维护、评价等监管环节数据的动态、实时监控及分析；
4. 支持建设运营补贴的申领、审核和发放。
   * 1. 安全监控服务

实现对有安全生产要求的企业进行生产用电监测，分析企业用电规律，实施监测告警模型，实现对在产、关停企业违规生产行为精准监管，助力安全生产治理能力现代化。主要内容包括：

1. 支持实时监测关闭、停产、正常生产的企业用电情况，通过模型算法发现异常用电行为后，形成告警信息；
2. 支持建立安全生产企业画像，整合政务数据、电力客户档案、电力用电数据等多方数据，形成企业画像可视化展示企业安全生产的各项指标；
3. 支持提供安全告警服务，对于明显存在违规生产或需要进一步核查的告警信息，推送至相关管理部门。
   * 1. 双替代清洁供暖

建立区域“双替代”用户级运行情况监测体系，为政府双替代用户电费补贴核算工作提供支撑。主要内容包括：

1. 支持构建区域“双替代”供暖“确村确户”大数据台账，实现对双替代用户基础信息档案进行录入、管理、审核和统计；
2. 支持对区域内“双替代”工程实施情况的监测；
3. 支持对“双替代”用电（气）量、设备类型、均价及降价、配电运行情况进行监测分析。
   1. 服务企业
      1. 用能监测

为企业提供用能监测分析，支撑企业动态掌握用能情况。主要内容包括：

1. 支持对企业用能数据进行在线监测，对企业各类能源的供应、存储、消耗全过程进行实时跟踪；
2. 提供企业用能数据的可视化展示。
   * 1. 能效服务

基于企业用能监测分析，结合企业实际生产和能耗数据，建立能源利用监督评价体系，为企业节能、提高企业能源利用率提供依据。主要内容包括：

1. 支持对企业的生产工艺能效、产品产值能效和设备能效等进行综合分析；
2. 支持为企业制定节能策略、优化能源消费方案，提供综合解决方案。
   * 1. 电力需求响应

针对电力用户开展基于需求响应的负荷管理，实现需求申请、响应签约、响应执行、效果评估等全环节、全过程线上运作。主要内容包括：

a)支持收集展示用户用电负荷曲线、需求响应执行情况等基础数据；

b) 为用户提供参与区域电力需求响应的线上申请、签约、需求响应补贴发放等功能。

* + 1. 企业碳监测

基于企业能耗数据，为企业提供碳排放监测相关服务。主要内容包括：

1. 支持企业能耗信息维护，结合各类能源对应碳排放因子计算碳排放量，实现对企业碳排放的监测；
2. 建立碳排放双控关键指标，支持双控指标分解、执行与监控，实现碳排放双控全过程在线跟踪；
3. 基于企业的历史数据并结合碳排放目标，对企业的碳排放进行预测预警。
   * 1. 企业碳减排

从清洁能源发电、电能替代、碳减排资产分析等多个维度，为企业提供碳减排相关服务。主要内容包括：

1. 支持融合发电侧、用电侧相关电碳数据，对清洁能源发电全链条监测分析，为提升清洁能源消纳、优化区域能源结构提供科学决策支撑；
2. 支持企业的碳减排资产管理，为后续配合做好第三方核查，向政府主管部门申请碳减排资产项目备案、核证减排量等提供数据支撑；
3. 支持构建碳减排测算模型，从区域、行业等方面搭建多维分析场景，分析对比各类电能替代的减排效果。
   * 1. 碳交易服务

提升企业的碳资产管理能力，为企业参与碳交易提供服务。主要内容包括：

1. 提供重点控排企业参与碳市场的辅助服务，包含企业的碳资产信息查询、管理与分析，企业碳配额和碳减排量账户管理；
2. 提供对企业碳交易与履约管理的服务，并为企业推荐履约策略。
   * 1. 碳效码

基于企业碳排放量、产值等数据，根据计算结果得出企业碳效等级。主要内容包括：

1. 支持对各行业的企业的碳排放、产值等数据进行统一归集和管理；
2. 支持对区域内各企业的碳效码进行计算，并进行赋码，量化评价企业单位产值碳排放水平。
   * 1. 新能源增值服务

通过汇集的新能源、新型储能设备级实时运行数据，提供出力功率预测、设备代运维等服务。主要内容包括：

1. 支持新能源场站运营代维服务，包含硬件设备、软件系统运维、运维策略分析等；
2. 支持新能源设备健康管理，包含新能源数据管理、工单管理、检修管理、报表管理等；
3. 支持新能源功率预测，包含未来7日的气象数据、出力曲线预测等；
4. 支持对新能源场站的并网申请服务，促进区域新能源的发展；
5. 支持对新能源场站的补贴申请和发放服务。
   * 1. 电力金融

基于企业用电大数据，构建系列电力金融指数模型，输出得分结果，为金融机构放贷全流程风险防控提供支撑，服务小微企业快速便捷贷款。主要内容包括：

1. 支持分区域、分行业、分县域企业用户分布情况和用电情况可观可测，打造电力金融征信综合服务平台；
2. 支持空壳企业识别、贷前反欺诈、贷中授信辅助、贷后风险预警等模型构建和应用，基于用电档案信息、缴费欠费信息、用电量波动等相关数据，向金融机构输出模型结果，提供电力金融服务产品。
   * 1. 碳金融

为企业在低碳节能改造和发展绿色产业项目等方面提供低成本资金支持。主要内容包括：

1. 支持需求方自主发布金融需求，寻找优质的金融机构为其提供金融服务；
2. 支持金融机构自主发布金融服务产品研发，例如碳金融产品、碳金融衍生品、融资产品、保险产品等；
3. 支持对供需双方对交易账户、交易订单进行管理。
   1. 服务公众
      1. 一网通办

提供能源业务受理、业务过程管理、业务进度查询等一站式服务，实现和政务“一网通办”深度融合，为公众提供更智能、高效、便捷的业务办理。主要内容包括：

1. 支持在政务服务平台完成电力业务申报、办理、查询、实施、办结等，对业务过程实现一站式闭环管理；
2. 支持办电业务过程进度追踪，服务电力客户实时掌握业务办理的进展情况；
3. 搜集用户反馈，形成业务服务质量整改清单并常态化跟踪管理。
   * 1. 用能管家

为公众提供经济、安全、绿色的智慧用能服务。主要内容包括：

1. 支持用能监测服务，支持对居民生活领域的用能情况、设备运行情况、能源账单等进行监测；
2. 支持用能分析服务，支持多维度对客户的用能情况进行分析，包含但不限于用能对标分析、用能结构分析、量价费分析、用能趋势分析、用能安全分析、能耗相关性分析等；
3. 依据统计周期内客户的能源消耗量和能效分析的结果，进行能效诊断并提出相应的专业化建议，节约用能成本。
   * 1. 充电智能服务

研发充电智能服务手机客户端，服务电动汽车车主找桩充电。主要内容包括：

1. 支持地图找桩、导航找桩，充电站点多维度（距离、价格、评价）排序，为用户提供多层次充电选择；
2. 支持扫码充电、充电进度查看、手动、自动停止充电等功能，建立便捷高效充电服务流程；
3. 支持主流支付渠道，建立新型支付模式，实现运营商计费、算费到结算全流程管理，为用户提供多种便捷的支付方式；
4. 支持订单信息查询、电子发票信息提交、发票开具及下载等功能，为电动汽车用户提供在线便捷开票服务。
   * 1. 碳积分

基于用户的绿色低碳行为实现碳减排效果的量化核证，激励居民绿色低碳生活。主要内容包括：

1. 支持建立个人碳账户，记录用户的低碳行为；
2. 支持量化低碳行为的碳排放研究，根据相关碳排放因子测算碳减排数据；
3. 支持构建展示形式与奖励模式，建立个人碳积分体系，对低碳贡献的用户使用积分或者其它形式的奖励，增强用户的粘性。
4. 数据应用工作流程
   1. 概述

能源大数据应用工作基于业务场景设计按照一定的标准步骤开展。能源大数据应用的工作流程包括7个工作步骤：应用需求、数据需求、数据准备、模型构建、成果应用及评估、迭代更新、应用退役。数据应用工作流程图如图2所示：



图2 数据应用工作流程图

* 1. 应用需求

应用需求：梳理业务，理解业务目标并分析需求，使业务设计与能源数据应用需求一致，并形成完整的需求定义。应用需求分为业务梳理、需求分析和业务设计和计划制定四个环节。

1. 业务梳理

明确能源大数据应用需求的业务目标，对所涉及的业务场景、业务数据特征、数据质量、存储方式、管理方法、业务逻辑等方面的业务现状进行调查，并对调查结果进行分析研究。

1. 需求分析

在调研基础上，准确评估形式了分析数据应用需求的具体要求，将业务需求转化为完整的需求定义。

1. 业务设计

以数据应用需求分析结果为依据，对数据应用需求实现的具体业务内容进行确定，明确数据应用的目标。

1. 计划制定

分析拥有的资源、条件和限制，进行风险、成本和效益评估。为整个数据应用制定项目计划，初步确认技术和工具。

* 1. 数据需求

数据需求阶段始于数据需求梳理与溯源，然后是原始数据收集并熟悉数据，标明数据质量问题，探索并初步理解数据，发掘有意义的子集以形成对隐藏信息的假设。具体步骤包括：数据需求梳理溯源、数据采集与提取、数据理解与探索三个环节。

1. 数据需求梳理与溯源

对数据应用的具体数据需求进行梳理和并确定数据来源。从数据应用超市、API市场、外部数据管理平台、算法模型库、内部应用成果库等查询可复用的数据应用组件，评估现有应用成果组件的匹配程度和可复用程度。

1. 数据采集与提取

根据数据应用的设计要求，从内、外部业务系统或线下采集数据，并提取出需要的数据。

1. 数据理解与探索

对数据进行描述和简单的数据统计分析，例如记录数、属性数、关键属性的分布等并给出相应的报告。

* 1. 数据准备

数据准备包括数据选择与清洗、数据构建与变换、数据集成与格式化、数据归约、数据脱敏、数据审计和数据标注等七个环节。

1. 数据选择与清洗

确定包含/剔除数据的基本原则，并选择恰当的属性/字段和记录。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，形成数据清洗报告

1. 数据构建与变换

构造派生字段并生成记录，使用平滑处理、特征构造、聚集、标准化或离散化等方法对数据进行处理，使数据符合目标算法的要求。

1. 数据集成与格式化

将互相关联的分布式异构数据源在逻辑上或物理上有机的集成并进行数据格式化处理，形成所需的数据集。

1. 数据归约

在不影响数据的完整性和数据分析结果正确性的前提下，通过减少数据规模的方式达到减少数据量、提升数据分析效果和效率的目的。

1. 数据脱敏

应根据项目要求制定脱敏规则，并根据规则进行数据脱敏。

1. 数据审计

对数据内容及其元数据的安全性与合规性进行审查，发现各类潜在风险和隐患。对初步集成的数据和API，从数据架构、数据标准、数据质量、数据安全等方面进行审核。

1. 数据标注

为了使数据便于使用，对数据添加词性、颜色、纹理、形状、关键字或语义信息等标签类元数据。

* 1. 模型构建

数据应用模型的构建，包括选择建模技术、模型训练、模型优化、模型应用四个环节。

1. 选择建模技术

基于业务内容、业务规则及数据需求，从业务数据模型或算法模型中选择恰当的建模技术，对业务需求和算法进行抽象化和数据化表达。

1. 模型训练

通过已有数据，对算法模型中未知参数进行确定，算法模型的未知参数确定后即可得到完整的算法模型，再使用已有数据对完整的算法模型的效果进行评价。

1. 模型优化

通过改变模型参数或模型结构提升模型的精度、执行速度等指标。

1. 模型应用

使用构建的算法模型对数据进行挖掘或分析，是对算法模型的实际应用。

* 1. 成果应用及评估

数据应用成果的应用及评估，包括成果交付、成果验收、成果推广和数据应用成果复用四个环节。

1. 成果交付

当数据应用成果开发完成后，开展数据应用成果交付工作。

1. 成果验收

针对交付的数据应用成果开展测试和验证工作，确保成果符合应用需求。

1. 成果推广

对数据应用成果进行实用化推广，并收集数据应用优化建议。

1. 成果复用

复用经审核后的数据应用成果，调用相关数据应用组件嵌入到新的数据应用产品中。

* 1. 迭代更新

数据应用的迭代更新，包括两种情形：

1. 当运行的数据应用又有新的业务需求时，结合最新需求和外部条件，对该数据应用进行功能迭代更新。
2. 基于成果评估阶段提出的优化建议，对应用进行迭代更新优化。
   1. 应用退役

根据外部条件及应用情况，当数据应用不再适用时提出退役需求，经审核后数据应用正式退役。

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

**━━━━━━━━━━━**