

附件

高质量数据集典型案例名单

（按推荐单位排序，排名不分先后）

序号	案例名称	申报单位	推荐单位
1	芯片陶瓷封装基板表面缺陷二维数据集	东北大学	教育部
2	数字教育应用算法智能诊断公共数据集	北京师范大学、中国信息通信研究院	教育部
3	VenusPod 蛋白质序列数据集	上海交通大学	教育部
4	中国全球海洋融合数据集 1.0	国家海洋信息中心	自然资源部
5	无人机森林防火智能巡护监测系统高质量数据集	中国林业科学研究院资源信息研究所	自然资源部
6	建筑机器人作业场景感知视觉高质量数据集构建与应用实践	中国建筑第四工程局有限公司、住房和城乡建设部科技与产业化发展中心	住房和城乡建设部
7	“交通+能源”高质量数据集	中国交通通信信息中心、中海石油气电集团有限责任公司	交通运输部

序号	案例名称	申报单位	推荐单位
8	长江智慧航运高质量数据集	长江水上交通监测与应急处置中心	交通运输部
9	农业典型作业场景多模态数据集	中国农业科学院农业信息研究所	农业农村部
10	应急管理“久安”AI大模型高质量数据集	应急管理部大数据中心	应急管理部
11	中国地震局“震典”地震监测人工智能数据集建设及其应用	中国地震台网中心	应急管理部
12	中国电信网络大模型高质量数据集	中国电信集团有限公司	国务院国资委
13	肺结核影像精标注高质量数据集	中国联合网络通信有限公司	国务院国资委
14	石油化工领域高质量数据集	中国石油化工集团有限公司	国务院国资委
15	中国移动人时空三元组高质量数据集	中国移动通信集团有限公司、北京航空航天大学、交通运输部规划研究院、中国移动通信集团山东有限公司	国务院国资委
16	跨领域合成图文高质量数据集	中电数据产业集团有限公司	国务院国资委
17	天然气管网运行高质量数据集	国家石油天然气管网集团有限公司油气调控中心	国务院国资委

序号	案例名称	申报单位	推荐单位
18	建筑行业供应链高质量数据集	中国建筑集团有限公司、中建电子商务有限责任公司	国务院国资委
19	电网调度负荷预测高质量数据集	中国南方电网有限责任公司	国务院国资委
20	scCompass: 亿级多物种单细胞转录组 AI 数据集	中国科学院计算机网络信息中心、中国科学院动物研究所	中国科学院
21	面向人群复杂特征的高质量 DNA 甲基化数据集	中国科学院北京基因组研究所（国家生物信息中心）	中国科学院
22	全球海洋环境变化关键参数高质量数据集	中国科学院海洋研究所、中国科学院大气物理研究所、中国科学院南海海洋研究所	中国科学院
23	地球大数据创新青藏高原高质量数据集	中国科学院青藏高原研究所	中国科学院
24	CMA-RA V1.5:10 公里分辨率全球大气再分析数据集	国家气象信息中心	中国气象局
25	面向灾害天气场景的强对流天气人工智能应用训练数据集	国家气象信息中心	中国气象局
26	高精度风资源数据集	国家气候中心	中国气象局
27	高质量电子疾病档案（EDR）数据集	浙江省疾病预防控制中心	国家疾控局

序号	案例名称	申报单位	推荐单位
28	广东省预防接种高质量数据集	广东省疾病预防控制中心	国家疾控局
29	主流价值语料库	人民网股份有限公司	人民日报社
30	中文互联网语料库系列数据集（CCI）	北京智源人工智能研究院	北京市政务服务和数据管理局
31	亿海蓝水运物流高质量数据集	亿海蓝（北京）数据技术股份公司	北京市政务服务和数据管理局
32	全球多口音英语高质量语音数据集	北京海天瑞声科技股份有限公司	北京市政务服务和数据管理局
33	基于“警银网通”AI反诈数据标签高质量数据集	恒安嘉新（北京）科技股份公司	北京市政务服务和数据管理局
34	数据堂教育高质量数据集构建及应用	数据堂（北京）科技股份有限公司	北京市政务服务和数据管理局
35	百度智能云自动驾驶高质量数据集	百度智能云（山西）科技有限公司	山西省数据局
36	农作物芽期、幼苗期高质量数据集构建及应用	内蒙古八爪智能科技有限公司	内蒙古自治区政务服务与数据管理局
37	省域政务热线高质量数据集	辽宁省民心网、辽宁民心出版传媒有限公司、联通（辽宁）产业互联网有限公司	辽宁省数据局

序号	案例名称	申报单位	推荐单位
38	空天地多源融合农业种植全要素高质量数据集	吉林省空天产业发展投资有限公司、北京观微科技有限公司	吉林省政务服务和数字化建设管理局
39	多语种大模型数据集建设	吉林科讯信息科技有限公司、科大讯飞股份有限公司	吉林省政务服务和数字化建设管理局
40	具身智能领域百万真机高质量数据集	智元创新（上海）科技有限公司	上海市数据局
41	中国移动研发大模型高质量数据集	中移（苏州）软件技术有限公司	江苏省数据局
42	无人化智慧农场多模态数据集建设与应用	安徽中科智能感知科技股份有限公司	安徽省数据资源管理局
43	莆田市全域多模态城市治理数据集	莆田市数字集团有限公司	福建省数据管理局
44	宫颈细胞 AI 辅助诊断高质量数据集	武汉兰丁云医学检验实验室有限公司、武汉兰丁智能医学股份有限公司	湖北省数据局
45	抑郁障碍人群体征高质量数据集	武汉东湖大数据科技股份有限公司	湖北省数据局
46	亿级多器官超声影像人工智能高质量数据集	华中科技大学同济医学院附属协和医院	湖北省数据局
47	北斗“通导遥一体化”多模态应急高质量数据集	移动广播与信息服务产业创新研究院（武汉）有限公司	湖北省数据局

序号	案例名称	申报单位	推荐单位
48	融合高精实测、仿真与开放道路场景的多源多模态智能驾驶数据集	湖南湘江智芯云途科技有限公司、湖南湘江智能科技创新中心有限公司、北京大学长沙计算与数字经济研究院、湖南工业职业技术学院汽车工程学院	湖南省数据局
49	惠农数谷智慧农业高质量数据集建设与应用	湖南惠农科技有限公司	湖南省数据局
50	自然资源遥感智能解译综合样本数据集	广东省国土资源测绘院、广东省国土资源技术中心	广东省政务服务和数据管理局
51	电力作业安全管控图像高质量数据集	广东电网有限责任公司信息中心、广东电网有限责任公司江门供电局、南方电网电力科技股份有限公司、广东顺畅科技有限公司	广东省政务服务和数据管理局
52	海南省海洋灾害多维立体监测与智能预报预警高质量数据集	海南欧特海洋科技有限公司、国家海洋信息中心、国家海洋技术中心、海南省海洋监测预报中心	海南省数据局
53	地震监测预报预警多模态联合数据集	四川省地震局	四川省数据局
54	云南动物资源多模态高质量数据集构建及应用	中国科学院昆明动物研究所	云南省数据局

序号	案例名称	申报单位	推荐单位
55	端到端语音大模型高质量数据集	标贝（青岛）科技有限公司	青岛市大数据发展管理局
56	青岛港人工智能多模态高质量数据集	青岛港国际股份有限公司	青岛市大数据发展管理局
57	高质量宫颈细胞数据集	宁波江丰生物信息技术有限公司	宁波市数据局
58	核电 SPV 设备健康诊断、运行异常及故障预测高质量数据集	中国核工业集团有限公司、三门核电有限公司、核电运行研究院（上海）有限公司、杭州安脉盛智能技术有限公司	中国核工业集团有限公司
59	船舶领域科技创新管理高质量数据集	中国船舶集团有限公司综合技术经济研究院	中国船舶集团有限公司
60	面向绿色船舶螺旋桨多性能数据集	中国船舶科学研究中心	中国船舶集团有限公司
61	基于多源数据融合的风电项目高质量数据集	中船科技股份有限公司	中国船舶集团有限公司
62	人形机器人具身操作数据集	人形机器人（上海）有限公司、第伍要素（上海）数据科技有限公司	中国兵器工业集团有限公司
63	司法领域法律专业高质量数据集	中国司法大数据研究院有限公司	中国电子科技集团有限公司

序号	案例名称	申报单位	推荐单位
64	公共信用档案高质量数据集	南京莱斯信息技术股份有限公司	中国电子科技集团有限公司
65	电磁空间环境高质量数据集	中国电子科技集团公司第二十二研究所	中国电子科技集团有限公司
66	平台经济智能治理与服务高质量数据集	中国电子科技集团有限公司电子科学研究院	中国电子科技集团有限公司
67	地球物理测井高质量数据集构建与应用	中国石油集团测井有限公司	中国石油天然气集团有限公司
68	中国石油天然气集团炼化时序数据集构建及应用	中石油（北京）数智研究院有限公司、 中石油兰州石化榆林化工有限公司	中国石油天然气集团有限公司
69	气电集团 LNG 产业协同与态势感知高质量数据集	中海石油气电集团有限责任公司、 中国移动信息技术有限公司	中国海洋石油集团有限公司
70	中国海油海洋环境高质量数据集	中海油能源发展股份有限公司	中国海洋石油集团有限公司
71	透平机组智能运维高质量数据集	中海石油（中国）有限公司	中国海洋石油集团有限公司

序号	案例名称	申报单位	推荐单位
72	油气管道保护高质量数据集	国家石油天然气管网集团有限公司华南分公司、国家石油天然气管网集团有限公司云南分公司、国家石油天然气管网集团有限公司山东分公司、国家管网集团浙江省天然气管网有限公司、国家管网集团西南管道有限责任公司	国家石油天然气管网集团有限公司
73	“双碳”背景下天然气市场预测高质量数据集建设与应用	国家管网集团工程技术创新有限公司、国家管网集团福建公司	国家石油天然气管网集团有限公司
74	天然气管道压缩机监测诊断高质量数据集	国家管网集团储运技术发展有限公司	国家石油天然气管网集团有限公司
75	能源电力行业供应链高质量数据集	国网物资有限公司、国网数字科技控股有限公司	国家电网有限公司
76	配电网智能规划多模态数据集	南方电网人工智能科技有限公司	中国南方电网有限责任公司
77	特高压直流现场作业多模态数据集构建及应用	中国南方电网有限责任公司超高压输电公司、南方电网人工智能科技有限公司	中国南方电网有限责任公司

序号	案例名称	申报单位	推荐单位
78	中国华电发电智能巡检与安全管控高质量数据集	南京南自信息技术有限公司	中国华电集团有限公司
79	国家级光伏储能产品及系统实证实验高质量数据集	国家电投集团青海光伏产业创新中心有限公司、青海黄河上游水电开发有限责任公司、大庆黄和光储实证研究有限公司、国家电投集团四川电力有限公司光伏分公司	国家电力投资集团有限公司
80	光储充一体化综合能源系统高质量数据集	国家电投集团综合智慧能源科技有限公司、中国科学院深圳先进研究院	国家电力投资集团有限公司
81	光伏逆变器健康诊断数据集建设与应用	国核电力规划设计研究院有限公司	国家电力投资集团有限公司
82	龙源电力风电设备诊断和安防防控数据集	龙源电力集团股份有限公司	国家能源投资集团有限责任公司
83	流域水灾害监测及防控高质量数据集建设	国能大渡河流域水电开发有限公司、国能大渡河大数据服务有限公司	国家能源投资集团有限责任公司
84	国家能源集团高质量数据集全生命周期实践	国能信息技术有限公司	国家能源投资集团有限责任公司

序号	案例名称	申报单位	推荐单位
85	工业纺织缺陷检测高质量数据集	中电信人工智能科技（北京）有限公司	中国电信集团有限公司
86	中国联通信息通信领域高质量数据集	中国联合网络通信有限公司	中国联合网络通信集团有限公司
87	车内外人像动作及自动驾驶高质量数据集	联通（广东）产业互联网有限公司、广州祺宸科技有限公司	中国联合网络通信集团有限公司
88	中国移动依托数据飞轮体系打造行业高质量数据集	中国移动通信集团有限公司、中海石油气电集团有限责任公司、复旦大学、数据堂（北京）科技股份有限公司	中国移动通信集团有限公司
89	咪咕公司聚焦“技术+场景”打造文体文旅高质量数据集	咪咕文化科技有限公司、中国移动通信集团湖北有限公司、中旅数智科技（深圳）有限公司、北京航空航天大学	中国移动通信集团有限公司
90	通导融合的高质量人车物时空行为与状态数据集	中移（上海）信息通信科技有限公司、中移动信息技术有限公司、中海石油气电集团有限责任公司、中国移动通信集团安徽有限公司合肥分公司、中移智行网络科技有限公司、北京海天瑞声科技股份有限公司	中国移动通信集团有限公司

序号	案例名称	申报单位	推荐单位
91	基于视觉大模型的低空巡检高质量数据集	中国移动（浙江）创新研究院有限公司、中国移动通信集团安徽有限公司池州分公司、安徽美图信息科技有限公司	中国移动通信集团有限公司
92	医疗卫生领域专科专病高质量数据集	中电四川数据服务有限公司	中国电子信息产业集团有限公司
93	红旗乘用车产品智能领域高质量数据集	中国第一汽车集团有限公司	中国第一汽车集团有限公司
94	铝合金材料金相组织图片数据集	中铝材料应用研究院有限公司	中国铝业集团有限公司
95	农业主粮作物时空耦合高质量数据集	中化现代农业有限公司、中电数据产业集团有限公司	中国中化控股有限责任公司
96	基于行业大模型的钢铁冶金全流程机器视觉数据集建设及应用	中冶赛迪集团有限公司	中国五矿集团有限公司
97	招商局集团交通物流行业高质量数据集	招商局集团有限公司、中电数据产业集团有限公司	招商局集团有限公司
98	轨道交通装备制造行业高质量数据集	中车工业研究院有限公司、中车大连机车研究所有限公司、中车永济电机有限公司	中国中车集团有限公司

序号	案例名称	申报单位	推荐单位
99	中交集团交通基建行业通识类数据集建设与应用	中国交通信息科技有限公司	中国交通建设集团有限公司
100	海洋交通基础设施长期性能野外观测数据集	中交第四航务工程局有限公司	中国交通建设集团有限公司
101	中国交建交通基础设施多维感知安全监测数据集	中国公路工程咨询集团有限公司、中交一公局集团有限公司、中交基础设施养护集团有限公司、中交资产管理有限公司、中国交通信息科技有限公司	中国交通建设集团有限公司
102	节能光伏一体化综合能源高质量数据集建设	中节能太阳能科技（镇江）有限公司	中国节能环保集团有限公司
103	面向千禧大模型的民航高质量数据集	中航信移动科技股份有限公司	中国民航信息集团有限公司
104	视联网多行业空间治理数据集	中国铁塔股份有限公司平台创新中心、铁塔智联技术有限公司	中国铁塔股份有限公司