|  |
| --- |
| 附件1重点行业产业数据仓基本信息表 |
| 申报日期： 年 月 日 |
| 申报单位信息 | 单位名称（盖章） |  |
| 统一社会信用代码 |  |
| 单位地址 |  |
| 负责人及联系方式 |  |
| 数据仓简要信息 | 数据仓名称 |  | 数据仓规模（单位：TB） |  |
| 所属工业重点行业 | □农副食品加工业□食品制造业□酒、饮料和精制茶制造业□烟草制品业□纺织业□纺织服装、服饰业□皮革、毛皮、羽毛及其制品和制鞋业□木材加工和木、竹、藤、棕、草制品业□家具制造业□造纸和纸制品业□印刷和记录媒介复制业□文教、工美、体育和娱乐用品制造业□石油、煤炭及其他燃料加工业□化学原料和化学制品制造业□医药制造业□化学纤维制造业□橡胶和塑料制品业□非金属矿物制品业□黑色金属冶炼和压延加工业□有色金属冶炼和压延加工业□金属制品业□通用设备制造业□专用设备制造业□汽车制造业□铁路、船舶、航空航天和其他运输设备制造业□电气机械和器材制造业□计算机、通信和其他电子设备制造业□仪器仪表制造业□其他制造业□废弃资源综合利用业□金属制品、机械和设备修理业□其他 （注：单选1个，每个行业填报一个信息表） |
| 数据模态类型 | □文本 □图形图像 □音频 □视频 □其他  |
| 建设性质 | □已建成 □在建 |
| 数据地域范围 | *数据覆盖的地理区域* |
| 数据时间范围 | *数据的时间跨度* |
| 数据更新机制 | *更新频率和方式* |
| 数据来源 | *数据渠道来源* |
| 数据总条目 |  | 公共数据条目 |  |
| 企业数据条目 |  | 第三方数据条目 |  |
| 本数据仓是否与省产业大脑能力中心产业数据仓进行数据共享 | □是 □否 |
| 本数据仓是否支撑省级行业“产业大脑”建设 | □是 □否 |
| 本数据仓是否符合国家数据安全相关标准和要求。 | □是 □否 |
| 本数据仓是否涉及国家秘密、个人敏感信息。 | □是 □否 |
| 本数据仓中数据的采集、存储、使用等是否符合相关法律法规要求。 | □是 □否 |
| 本数据仓的版权归属和使用授权清晰。 | □是 □否 |
| 针对数据仓名称和单位名称是否同意在省级部门组织的相关供需对接活动发布，并同意媒体宣传。 | □是 □否 |
| 数据仓概况 | （对数据仓进行简要描述，包括项目总投资、建设期限、数据集总量及各模态数据量、主要内容、数据更新机制、目的和用途等，不超过500字）注：数据仓样本数据须作为附件提供，原则上不少于100条，经脱敏后的样本数据电子版直接发送到联系邮箱。 |
| 数据仓质量情况介绍 | （对数据仓的数据清洗情况、数据标注情况、数据完整性、数据一致性、数据准确性、数据时效性、数据唯一性、符合性测试等情况进行简要阐述，不超过500字） |
| 应用场景和人工智能大模型支撑情况介绍 | （对数据仓已落地的或潜在的应用场景及其服务人工智能大模型的情况进行简要阐述，不超过200字） |
| 声明 | 1.我单位对本次申报的数据仓合法性、真实性、完整性和有效性负责，与其他单位或个人无知识产权纠纷。2.经国家、省企业信用信息系统查询，我单位不是严重失信主体。3.该数据仓不包含违反社会主义核心价值观的内容。 |
| 注：如有多页，请加盖骑缝章。 |

|  |
| --- |
| 附件2重点行业知识语料库基本信息表 |
| 申报日期： 年 月 日 |
| 申报单位信息 | 单位名称（盖章） |  |
| 统一社会信用代码 |  |
| 单位地址 |  |
| 负责人及联系方式 |  |
| 语料库简要信息 | 语料库名称 |  | 语料库规模（单位：TB） |  |
| 所属重点行业 | □电子 □机械 □有色 □钢铁 □电力装备 □石化 □轻工 □化工 □建材 □汽车 □基础数学 □金融服务 □其他 （注：单选1个，每个行业填报一个信息表） |
| 语料类型 | □文本 □图形图像 □音频 □视频 □其他  |
| 建设性质 | □已建成 □在建 |
| 语料地域范围 | *语料覆盖的地理区域* |
| 语料时间范围 | *语料的时间跨度* |
| 语料更新机制 | *更新频率和方式* |
| 语料来源 | *语料渠道来源* |
| 本语料库是否符合国家数据安全相关标准和要求。 | □是 □否 |
| 本语料库是否涉及国家秘密、个人敏感信息。 | □是 □否 |
| 本语料库中数据的采集、存储、使用等是否符合相关法律法规要求。 | □是 □否 |
| 本语料库的版权归属和使用授权清晰。 | □是 □否 |
| 针对语料库名称和单位名称是否同意在省级部门组织的相关供需对接活动发布，并同意媒体宣传。 | □是 □否 |
| 语料库概况 | （对语料库进行简要描述，包括项目总投资、建设期限、语料库总量及各模态数据量、主要内容、更新机制、目的和用途等，不超过500字）注：语料库样本数据须作为附件提供，原则上不少于100条，经脱敏后的样本数据电子版直接发送到联系邮箱。 |
| 语料库质量情况介绍 | （对语料库的数据清洗情况、数据标注情况、数据完整性、数据一致性、数据准确性、数据时效性、数据唯一性、符合性测试等情况进行简要阐述，不超过500字）注：本次申请允许使用合成语料，但合成语料的比例不得超过20%，且需明确标注是否为合成语料条目，提供合成语料的来源与合成路径。 |
| 人工智能大模型支撑情况介绍 | （对语料库已落地的或潜在的应用场景及其服务人工智能大模型的情况进行简要阐述，不超过200字） |
| 声明 | 1.我单位对本次申报的语料库合法性、真实性、完整性和有效性负责，与其他单位或个人无知识产权纠纷。2.经国家、省企业信用信息系统查询，我单位不是严重失信主体。3.该语料库不包含违反社会主义核心价值观的内容。 |
| 注：如有多页，请加盖骑缝章。 |

|  |
| --- |
| 附件3重点行业产业数据仓推荐汇总表 |
| 推荐单位：（盖章） 联系人及联系方式： | 2025年 2 月 日 |
| 序号 | 数据仓名称 | 行业领域 | 数据仓模态 | 数据仓数据总量（TB） | 建设性质（已建成/在建） | 主要应用场景 | 申报单位名称 | 申报单位联系人及联系方式 |
| 1 |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |
| …… |  |  |  |  |  |  |  |  |

|  |
| --- |
| 附件4重点行业知识语料库推荐汇总表 |
| 推荐单位：（盖章） 联系人及联系方式： | 2025年 2 月 日 |
| 序号 | 语料库名称 | 行业领域 | 语料库模态 | 语料库数据总量（TB） | 建设性质（已建成/在建） | 拟服务支撑AI大模型的名称 | 申报单位名称 | 申报单位联系人及联系方式 |
| 1 |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |
| …… |  |  |  |  |  |  |  |  |

注：申报单位为高校及科研院所二级学院或机构的，统一由高校或科研院所作为主体进行汇总推荐。