

《面向人工智能的数据标注合规指南》

团体标准（征求意见稿）

编制说明

一、工作简况

1. 任务来源

本团体标准编制项目由中移互联网有限公司作为牵头起草单位，北京之合网络科技有限公司（智合标准中心）作为组织编制单位，中国电子商会作为归口管理单位，经中国电子商会标准化工作委员会审查批准后正式立项，团体标准项目编号为 CECC 2025-1-070。

2. 起草单位

澳鹏数据科技（上海）有限公司、北京集纳盛广网络科技有限公司、北京热热文化科技有限公司、北京市环球律师事务所、北京云测数据科技有限公司、北京易华录信息技术股份有限公司、北京之合网络科技有限公司、重庆群星引力大数据科技有限公司、德阳旌渝数字科技有限公司、福建中锐电子科技有限公司、公安部第三研究所、广东广悦律师事务所、国网江苏省电力有限公司连云港供电分公司、广西影迅物流有限公司、贵州星华光电产业发展有限公司、杭州数据交易所有限公司、杭州五维数据有限责任公司、金华途样网络科技有限公司、江苏钟吾大数据发展集团有限公司、蚂蚁科技集团股份有限公司、上海之合网络科技有限公司、武汉中韬信息科技有限公司、西藏云通数字科技有限公司、中国电子信息产业发展研究院、中广核智能科技（深圳）有限责任公司、中国民航信息网络股份有限公司、浙江浦源律师事务所、中移互联网有限公司；

3. 主要工作过程

2025年2月：数据标注作为驱动AI模型进化的必备环节，对于AI产业发展至关重要。2025年初，国家发展改革委等部门发布《关于促进数据标注产业高质量发展的实施意见》，为促进产业发展指明方向。然而，行业长期存在的数据来源模糊、标注质量失控、隐私泄露、敏感信息不当处理等合规问题日渐突出。当技术创新与合规风险形成对冲，行业亟需一套低成本、可复制的解决方案。为提升数据标注行业规范化发展，为此，北京之合网络科技有限公司（智合标准中心）提出了制定《面向人工智能的数据标注合规指南》团体标准的建议。

2025年3月：北京之合网络科技有限公司（智合标准中心）发起编制《面向人工智能的数据标注合规指南》团体标准。自标准立项以来，智合标准中心在组建标准起草专家组的同时，面向全国知名高校、科研机构、相关企业、中介机构征集起草单位、起草人。

2025年4月-5月：为提升文本编制的科学性与实用性，标准编制组于4月2日以线上方式，组织召开框架研讨会。会议旨在定稿标准框架并明确编制分工，按计划推进标准文本的编制工作。在会议基础上，编制组历经数月的精心打磨与反复推敲，顺利完成了标准草案稿的编制。

2025年6月：标准面向全体起草单位、起草人启动意见征集工作。6月26日，标准编制组在京组织召开草案稿研讨会，结合产业发展实践和企业实际诉求，共研共议文本待优化内容，会议汇聚50余位企业代表、数商代表、法律专家及行业学者，围绕数据标注全流程合规核心议题展开深度研讨，旨在提升标准的科学性与落地性，推动AI数据标注行业标准化发展，并基于此形成标准修订稿。

2025年7月-8月：经过前述充分研讨论证，标准编制工作组完成《面向人工智能的数据标注合规指南》团体标准（征求意见稿）的编制。结合国内法律法规、现有标准及行业规定，标准编制工作组起草《面向人工智能的数据标注合规指南》团体标准（征求意见稿）编制说明。

二、标准制定的原则与依据及其主要技术内容

1. 标准编制原则

(1) 科学性原则

根据《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等相关法律,《互联网信息服务算法推荐管理规定》《互联网信息服务深度合成管理规定》《生成式人工智能服务管理暂行办法》《人工智能生成合成内容标识方法》等相关部门规章,《GB/T 35273-2020 信息安全技术 个人信息安全规范》《GB/T 41867-2022 信息技术 人工智能 术语》《GB/T 42755-2023 人工智能 面向机器学习的数据标注规程》《GB/T 43697-2024 数据安全技术 数据分类分级规则》和《GB/T 45674-2025 网络安全技术 生成式人工智能数据标注安全规范》等规范性文件,结合我国人工智能产业数据标注的实践发展,制定本文件。

(2) 规范性原则

根据《GB/T 1.1-2020 标准化工作导则第1部分:标准文件的结构和起草规则》的要求进行标准编写。本标准的术语和定义参照 GB/T 35273-2020、GB/T 41867-2022、GB/T 42755-2023、GB/T 43697-2024、GB/T 45674-2025 的要求。

(3) 实用性原则

起草单位广泛覆盖人工智能领先企业、数据标注服务商、法律合规服务机构、技术方案提供商等多个领域,具有丰富实务经验,两次研讨会均聚焦业务实践和实际需求,深入探讨,确保标准内容的实用性和可操作性。

2. 主要技术内容

本标准帮助 AI 企业从源头低成本解决数据标注合规难题,结合“业务场景实操化、人员管理规范化的企业运营稳健化”三大原则,聚焦数据来源、标注内容与过程操作、标注人员管理、数据安全与隐私保护、监督和审计机制五大关键

合规议题展开制定。

在业务场景实操化方面，标准紧扣数据标注业务实践，为企业提供了清晰、可操作的数据来源合规性判断标准和获取指引，从源头杜绝了潜在合规风险。结合具体业务场景，明确了数据标注内容与行为的边界和范围，帮助企业在标注过程中有章可循，避免了因标注内容不当引发的侵权风险。

针对人员管理规范化，标准结合标注人员管理实践，帮助企业建立起科学合理的管理体系，有效提升了一线标注人员的合规意识和操作能力，从而保障标注质量与合规性。

从企业经营稳健化角度出发，标准还探讨了低成本加固数据安全能力的路径，健全了监督和审计机制，帮助企业定期对数据标注工作进行复检，及时发现并纠正存在的合规问题，进一步支撑企业稳健运行。

三、综述报告与预期经济效益

1. 综述报告

数据标注产业是人工智能创新发展的关键，人工智能和机器学习的发展，带来大量数据标注需求，美国、英国、德国、日本、韩国等国家和地区纷纷将数据标注产业作为数字经济发展的基础性、战略性产业。据《全球人工智能产业发展白皮书（2024 年度）》显示，2023 年全球人工智能产业规模达 7078 亿美元，同比增长 19%，全球人工智能标注市场规模超过 1400 亿美元。

美国政府通过国家战略层面的规划和引导，各环节协同配合，形成了完整的产业链，涵盖专业标注公司、基础标注工具开发商、标注众包平台、第三方质量控制机构、标准化组织等主体；欧盟的《人工智能法案》《通用数据保护条例》《数据法案》和《人工智能伦理准则》等对人工智能系统开发、运用中的数据保护作了一系列规定，基本形成了全方面、多层次、重实效的数据安全保护体系，其中涉及对标注数据的要求；英国政府高度重视数据标注产业发展，包括相关立

法、投资计划、人才培养、产业园建设等；德国政府高度重视人工智能和数据领域发展，已出台多项促进政策，包括拨款数十亿欧元投资人工智能、培养人才、建设试点中心等；日本政府出台了《综合创新战略 2024》等政策，将人工智能及相关的标注产业作为重点发展方向；韩国政府则推出“人工智能国家战略”，对与人工智能相关的标注产业也给予高度重视，其中《人工智能发展与信任基础创建法》规定了人工智能技术标准化和人工智能训练数据等相关政策的制定。

随着各类组织对标注产业的日益重视，标注相关标准的研制已成为国内外各标准化组织共同关注的热点。尽管国外目前并没有针对标注的成熟标准与体系，但已有相关国际标准涉及标注。

如下国际标准涉及：**ISO/IEC TR 24027:2021**《信息技术-人工智能-人工智能辅助决策系统偏差》指导标注过程中减少数据偏差，提升模型公平性；**ISO/IEC 23053:2022**《利用机器学习的人工智能系统框架》定义人工智能系统开发流程，要求标注数据的质量和可追溯性；**ISO/IEC 38507:2022**《信息技术-IT 治理-组织使用人工智能的治理影响》要求企业在标注中遵循伦理与合规原则，确保透明性与问责制；**IEEE Std 2801-2022**《医学人工智能数据集质量管理》针对医疗数据规定了相应标注标准，有效提高医疗数据质量。

除此之外，在针对标注的合规指南层面，国际暂无可供参考的标准化解决方案。虽有部分经验可借鉴，但还需有专门的操作指南给予相关机构实践指导。

随着人工智能技术的迅猛发展，对高质量标注数据的需求呈爆发式增长，标注产业引起了市场和政府的高度重视。得益于人工智能技术发展、政府政策支持和各行各业数字化转型的加速，标注市场规模不断扩大，有望成为重要经济增长支撑点。

2023 年 7 月，中央网信办联合多部门发布《生成式人工智能服务管理暂行办法》，标注是其中重要内容，相关内容实施落地亟需配套标准支撑相关工作开展。2023 年工业和信息化部总工程师赵志国，在中国国际服务贸易交易会

数字贸易发展趋势和前沿高峰论坛上表态，要加快制定数据标注等数据要素市场基础制度配套政策。2024 年国家数据局在全国首次数据工作会议上，就已经提出探索建设国家级数据标注基地。同年 5 月，经过国家数据局综合评估，四川省成都市、辽宁省沈阳市、安徽省合肥市、湖南省长沙市、海南省海口市、河北省保定市、山西省大同市共 7 座城市成为首批全国数据标注试点基地；同年 10 月，由国家数据局数字科技和基础设施建设司指导，国家数据局人工智能专班主办的首届“数据标注产业大会暨供需对接会”在北京召开；同时中共中央办公厅 国务院办公厅下发《关于加快公共数据资源开发利用的意见》，提出将数据产业作为鼓励发展类纳入产业结构调整指导目录，支持数据采集标注、分析挖掘、流通使用、数据安全等技术创新应用；同年 12 月，国家发展改革委等部门印发《关于促进数据产业高质量发展的指导意见》，提出强化数据标注、数据合成等核心技术攻关。

同年，各地政府也积极响应国家政策，纷纷出台相关措施，推动数据标注产业在本地区的发展。比如河南省人民政府办公厅发布《河南省加快制造业“六新”突破实施方案》，提出加快建设数据标注行业标准体系，支持商丘、安阳市打造高水平数据标注产业集群；浙江省人民政府办公厅发布《关于加快人工智能产业发展的指导意见》，支持建设数据标注众包服务平台；江西新余市政府发布《新余市关于数据标注与内容审核发展的若干意见（征求意见稿）》，聚焦本地数据标注与内容审核业务，规划发展路径与扶持措施。

2025 年 1 月 13 日，国家发展改革委、国家数据局、财政部、人力资源社会保障部四部门联合发布《关于促进数据标注产业高质量发展的实施意见》，这是国家层面对数据标注产业首次进行系统谋划。《实施意见》提出到 2027 年，数据标注产业专业化、智能化及科技创新能力显著提升，产业规模大幅跃升，年均复合增长率超过 20%，培育一批具有影响力的科技型数据标注企业，打造一批产学研用联动的创新载体，建设一批成效明显、特色鲜明的数据标注基地，形成相对完善的数据标注产业生态，构建创新要素聚集、产业链上下游联动、区域协同发展的新格局。特别提到要支持企事业单位牵头制定数据标注相关标准。

随着“人工智能+”行动的实施，以及对高质量数据集的需求，我国数据标注产业在自动驾驶、医疗影像分析、智能客服、智能制造、金融服务、能源管理等领域的发展空间将更加广阔，也将得到业界更多的关注。

涉及数据标注的国家标准有《人工智能 面向机器学习的数据标注规程》（已发布）、《网络安全技术 生成式人工智能数据标注安全规范》（征求意见稿）等；行业标准有《人工智能医疗器械 质量要求和评价 第3部分：数据标注通用要求》（已发布）、《面向人工智能的数据生产和标注服务能力通用成熟度模型》（报批中）等；地方标准有《人工智能 数据标注通用工作规程》、《人工智能数据标注总体框架规范》、《人工智能 数据标注一般技术要求》等；团体标准有《大数据 数据标注平台技术要求》、《数据标注工程师能力评估规范》、《针对内容安全的人工智能数据标注指南》、《文化资源数据标注系统》和《人工智能数据标注平台技术要求和测试方法》等。

经过调研评估，目前关于数据标注的标准，国际标准暂无专门针对数据标注的标准，更多的是在人工智能技术标准中涉及到部分数据标注要求；国家标准、行业标准、地方标准亦或是团体标准更多侧重技术要求、安全要求或是规范整体业务流程，尚未针对数据标注活动进行专门的合规指导和规范，聚焦数据标注合规层面的标准还处于市场空白。

综上所述，本标准的核心亮点是针对人工智能数据标注合规进行相应规范，给予全流程实践方法，帮助人工智能企业从源头低成本解决数据合规难题，同时促进数据标注行业标准化、规范化发展。

2. 预期经济效益

《面向人工智能的数据标注合规指南》团体标准，以期通过联合行业力量，帮助企业更好地规范和提升人工智能领域数据标注活动的合规与价值创造能力，为企业创新发展提供有力支撑。同时基于产业链上各机构的不同专业视角，共同探索较为公允且一致的数据标注合规的操作指引，促进产业高质量健康发展，为人工智能产业生态建设贡献行业力量。待本标准文本正式发布后，将采取多举措

推进落地应用。一是加强宣贯与推广，不断提高行业内对本标准的认知；二是积极开展本标准应用试点工作。

四、与其他标准的关系

本标准的制定在符合《GB/T 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》的基础上，参考了《GB/T 35273-2020 信息安全技术 个人信息安全规范》《GB/T 41867-2022 信息技术 人工智能 术语》《GB/T 42755-2023 人工智能 面向机器学习的数据标注规程》《GB/T 43697-2024 数据安全 安全技术 数据分类分级规则》和《GB/T 45674-2025 网络安全技术 生成式人工智能数据标注安全规范》等规范性文件，首次制定了《面向人工智能的数据标注合规指南》团体标准，突出了本标准编写应遵循的原则和要求。

本标准与现行的法律法规、国家标准、行业标准和地方标准都不冲突。

五、采用国际标准的程度及水平的简要说明

暂未涉及国际标准的采用。

六、其他应予说明的事项

无。