

《人工智能 大规模预训练模型总体技术要求及评估方法》

（征求意见稿）编制说明

一、工作简况

《人工智能 大规模预训练模型总体技术要求及评估方法》（计划号：072-T/ISC-23）于2023年8月18日由中国互联网协会下达立项计划。本标准自主制定标准，起草单位为中国移动通信集团有限公司、中国互联网协会人工智能工作委员会、中国信息通信研究院、北京智源人工智能研究院、国能数智科技开发（北京）有限公司、中国石油化工集团有限公司、南方电网人工智能科技有限公司、北京百度网讯科技有限公司、华为技术有限公司、新华三技术有限公司、上海燧原科技股份有限公司、中科寒武纪科技股份有限公司、海光信息技术股份有限公司、摩尔线程智能科技（北京）有限责任公司、中移（苏州）软件技术有限公司、广州趣丸网络科技有限公司、浪潮集团有限公司、中讯邮电咨询设计院有限公司、上海天数智芯半导体有限公司等行业内具有代表性的大规模预训练模型供应方、评测方、需求方等企业和科研机构，本标准提出的技术要求是经行业代表性企业协商一致共同认可的实际做法，具备可行性和通用性。

二、标准编制原则和确定标准主要内容的依据

近年来，随着数据资源的不断丰富以及深度学习计算能力的不断提升，大模型具备了强大的表征能力，在自然语言处理、图像识别、语音识别等领域都展现出了强大的性能和广泛的应用价值，正逐渐成为人工智能领域的研究热点和产业应用焦点。随着产业界的深入参与和合作，大模型的研发和应用从聚焦基础设施构建、数据获取、模型训练及部署等方面技术突破和工程化落地，到关注领域应用、行业应用等方面大模型服务能力评估。然而，针对整个大模型系统，业界还缺乏规范性的技术要求及评估体系，使得不同的研究机构和企业对大规模预训练模型系统难以开展标准化的评估。

本标准按照 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》要求进行编写，主要规定了大规模预训练模型系统参考架构及技术要求、大规模预训练模型系统角色、大规模预训练模型能力评估方法，使得可以通过一个统一、科学的评估标准用于评估不同供应商研发生成的大规模预训

练模型。本标准适用于人工智能大规模预训练模型的设计、研发、评估和应用。

三、工作简要过程

2023年8月18日，该项标准制订计划正式下达，中国移动通信集团有限公司（以下简称中国移动）开始牵头负责标准草案大纲的编写工作。

2023年9月28日，中国移动组织中国互联网协会人工智能工作委员会、中国信息通信研究院、北京智源人工智能研究院、国能数智科技开发（北京）有限公司、华为技术有限公司、上海燧原科技股份有限公司、中科寒武纪科技股份有限公司等10余家首批起草单位召开标准编制启动会，正式成立标准起草组，在会上介绍了该项标准的编制背景和立项情况，就标准草案大纲进行讨论并达成一致。

2023年10月18日，标准起草组开始征集第二批标准参编单位，经过公开征集后增加了8家起草单位。

2023年11月至2024年7月，标准起草组组织召开了多次线上研讨会，围绕各起草单位的标准内容提案进行讨论和审议，共同开展标准草案的编制工作，逐步形成工作组草案。

2024年9月12日，标准起草组在北京组织召开标准编制研讨会，各起草单位共15名代表参会，会上对标准草案内容进行了深入讨论，并提出了相关修改建议。本次会议讨论决定，标准修改完成后进入征求意见阶段。

四、主要试验（或验证）的分析

本标准起草单位中有众多业内领先的人工智能大规模预训练模型研发和生产机构，标准编制过程中，各起草单位结合各自产品在设计、研发、生产、应用过程中的经验和试验验证情况，给出了标准草案相应的技术内容和指标，因此标准草案的相关技术要求和测试方法符合行业共识，经历过实际验证。

五、采用国际标准和国外先进标准情况

本标准暂未采用相关国际标准和国外标准。

六、标准涉及的专利说明

本标准暂未发现相关知识产权问题。

七、预期达到的社会效益、对产业发展的作用

人工智能技术的迅猛发展正推动着从单一模态、小模型向多模态、大模型的

发展模式转变，其中大规模预训练模型在大数据、大算力和强算法的加持下，展现出了卓越的自然语言生成能力和通用性。这一转变不仅极大地提升了生产效率和服务质量，激发了科技创新，改善了用户体验，还在教育、公共服务等领域发挥了重要作用。然而，当前业界对于大规模预训练模型的评估缺乏统一的规范和技术要求，导致不同研究机构和企业难以进行标准化评估。因此，制定一个统一且科学的大模型评估标准显得尤为重要，它能够为不同供应商研发的模型提供一致的功能和性能指标，确保评估结果的客观性、准确性和可靠性，为人工智能大模型的采购方和应用方提供重要的参考依据。这不仅有助于促进我国人工智能产业生态的健康发展和行业的良性竞争，还能加速技术的创新应用，提升产业竞争力，推动社会和经济的全面发展。

八、与现行相关法律、法规、规章及相关标准的协调性

本标准符合我国有关的现行法律、法规和相关强制性标准的相关规定。与本标准相关的标准有 GB/T 41867-2020《信息技术 人工智能 术语》，AIIA/PG 0141-2024《人工智能开发平台通用能力要求 第4部分：大模型技术要求》和 AIIA/PG 0074-2022《大规模预训练模型技术和应用评估方法 第4部分：可信要求》。

已发布的几项相关标准中，大部分标准的标准化对象只聚焦于大规模预训练模型系统的一部分，且不涉及对大规模预训练模型能力的评估方法，本标准不仅结合已有标准，规范了大规模预训练模型系统各个部分的技术要求，包括基础设施层、数据层、模型层、应用层、系统安全层，同时也给出了大规模预训练模型系统的系统角色划分方法，以及大规模预训练模型模型能力评估方法。本标准在填补已有标准空白同时，和已有标准互相配合，共同促进大规模预训练模型技术和产业的发展。

九、重点分歧意见的处理经过和依据

本标准在编制过程中没有重大分歧。

十、标准性质的建议说明

本标准建议作为推荐性团体标准进行发布。

十一、贯彻标准的要求和措施建议

本标准的制定可以为人工智能大规模预训练模型相关产品的设计、研发和生产提供参考和指导，提供统一的标准化的功能与性能指标用于评估不同供应商研发生成的大模型，保证评估结果的客观性、准确性和可靠性，为人工智能大模型的采购方和应用方提供可参考的评估依据，促进我国人工智能产业生态的健康发展和行业的良性竞争。建议在标准发布后尽快组织宣贯、推广和应用实施。

在具体贯彻实施本标准时，一是建议标准归口单位中国互联网协会在会员单位内部进行宣贯，加大标准的宣传力度，扩大影响范围；二是起草单位中各大规模预训练模型生产企业在产品设计、研发过程中，积极开展标准的应用实施工作；三是建议政府、企业等相关机构制定配套实施政策，如在招投标、认证服务等方面引导相关企业提升标准应用积极性；四是积极推动标准的符合性测试和评估工作，为标准的落地实践提供评测服务和保障。

十二、其他应予说明的事项

无。

《人工智能 大规模预训练模型总体技术要求及方法》标准起草组

2024年10月25日