《基于人工智能技术的未成年人互联网应用建设指南》团体标准编制说明

# 一、工作简况

## （一）标准研制背景及必要性

人工智能作为引领新一代科技革命和产业变革的战略性技术，在提供个性化学习、加强卫生保健和改善社会福利服务等方面的应用正深刻改变社会生活、改变世界，也为在数字世界中实现儿童权利提供了巨大潜力。孩子们已经以许多不同的方式与人工智能技术进行交互，这些技术往往被嵌入玩具、应用程序和视频游戏中，不知不觉中影响着每个孩子的成长。

人工智能系统的使用将从根本上改变未成年人学习和生活的许多领域，例如人工智能与教育教学的系统性融合，促进了教育资源公平化；基于人工智能的青少年健康管理，将实时健康监测成为可能，为疾病预防和健康管理提供保障。但是，由于人工智能系统可以以前所未有的速度和规模来分析大量数据并进行推理，涉及未成年人安全、隐私、公平性与非歧视等问题。加之未成年人具有独特的生理和心理属性，在人工智能系统的应用中需要特别注意，他们的发展和教育将进一步受到人工智能的介入和过滤，并且他们一生中将越来越多地接触人工智能系统。

未成年人互联网保护工作在我国乃至世界范围内都受到了广泛的关注，新修订的《中国人民共和国未成年人保护法》中增设了“网络保护”专章。国家互联网信息办公室发布的《未成年人网络保护条例（征求意见稿）》，在保障未成年人在网络空间的合法权益等方面作出了具体的要求。联合国儿童基金会《人工智能为儿童——政策指南》2.0版包含了以儿童为中心的人工智能的实用建议和原则。

为贯彻落实《国务院关于印发新一代人工智能发展规划的通知》《国家新一代人工智能标准体系建设指南》要求，更好地了解人工智能如何保护、提供和赋予儿童权力，妥善应对人工智能对儿童健康成长可能带来的挑战，助力国家推动人工智能发展和标准化体系建设，为国内和全球的相关政策制定提供有益参考，中网联与联合国儿童基金会（简称“儿基会”）于2021年共同开展人工智能为儿童项目，出版《人工智能为儿童——面向儿童群体的人工智能应用调研报告》，并在2021年联合国互联网治理论坛上面向全球发布。

为更好地利用人工智能技术提高对未成年人的认知和实践，保护未成年人身心健康发展，互联网企业应积极制定用于指导基于人工智能技术的未成年人互联网应用开发等相关标准。

## （二）任务来源

本标准于2021年12月28日由中国网络社会组织联合会批准并列入2022年中网联第一批标准立项目录。

## （三）工作过程

1. 预研与立项申请

2021年标准主要起草单位开展了基于人工智能技术的未成年人互联网应用的预研工作。梳理10余项关于未成年人互联网保护和人工智能技术等相关政策文件、标准等，了解基于人工智能技术的未成年人互联网应用发展现状与存在问题，提出标准立项建议。

2. 起草与征求意见

2021年底，标准正式立项后，牵头起草单位邀请相关企业、科研机构、社会组织等相关方共同组建标准起草工作组。

2022年1月至2月，标准起草工作组集中梳理分析未成年人网络保护相关政策文件要求、人工智能技术相关标准等，走访调研相关企业，逐步完善并形成标准草案。

2022年3月至4月，标准起草工作组多次召开工作会议、研讨会，讨论标准草案技术内容，修改完善并形成征求意见稿。

# 二、主要依据与技术内容

标准起草过程中，主要依据**《人工智能为儿童——政策指南》《关于进一步压实网站平台信息内容管理主体责任的意见》《互联网信息服务算法推荐管理规定》《儿童个人信息网络保护规定》**等相关政策及文件，参考借鉴ISO/IEC TR 24027《信息技术 人工智能 人工智能系统中的偏差与人工智能辅助决策》、 ISO/IEC PDTR 24028《信息技术 人工智能 人工智能可信度概述》、ISO/IEC 23894《信息技术 人工智能 风险管理》、TR《信息技术 人工智能 伦理和社会关注概述》、IEEE P7004《儿童和学生数据治理标准》等标准文件，并结合互联网企业实践经验，规定了**基于人工智能技术的未成年人互联网应用应坚持的基本原则，提出了关键参与角色和基本建议，给出了基于人工智能技术的未成年人互联网应用全生命周期的关键阶段，**力求保证标准的适用性和实用性。

标准在总体原则一章，提出了设计、开发、运维基于AI的未成年人互联网应用时应遵循的基本原则。

标准在关键参与角色一章，给出了与**基于AI的未成年人互联网应用开发、使用密切相关的利益相关方，便于明确各相关方的角色和责任，以便共同协作**。

标准在关键阶段一章，给出了**基于AI的未成年人互联网应用建设过程中以训练模型的建立和训练数据的供给为基础的特点下，AI应用特有的数据获取与准备阶段、模型设计与开发阶段、测试验证与评估阶段、模型部署与使用阶段四个关键阶段**。

标准在技术措施一章，提出了**基于人工智能技术的未成年人互联网应用建设时应注重的风险，及需考虑的关键因素**。

# 三、与现行有关法律、法规和标准的关系

2021年6月1日正式实行的《中国人民共和国未成年人保护法》（2020年修订版）中增设了“网络保护”专章，其中对网络产品和服务提供者提出了“不得向未成年人提供诱导其沉迷的产品和服务”等要求，但没有提及具体的落实措施和建议，本标准与相关法律法规协调一致，不存在冲突和矛盾，是对现行相关法律法规要求的具体细化和落实。当前，尚没有基于人工智能的未成年人互联网应用建设的相关国家标准、行业标准、地方标准，不存在与现行标准冲突矛盾的情况。

# 四、重大分歧意见的处理经过和依据

本标准在制定过程中未出现重大分歧意见。

# 五、贯彻促进会团体标准的要求和措施建议（包括组织措施、技术措施、过渡办法等内容）

建议本标准自批准发布之日起实施，并由中国网络社会组织联合会适时组织标准宣贯工作，推荐互联网企业在**建设基于人工智能技术的未成年人互联网应用时使用**此标准，以保护未成年人合法权益和身心健康，营造良好网络生态。

# 六、其他应予说明的事项

# 无。