

《人工智能拟人化互动服务管理暂行办法》

文章来源：中国网信网 文章类型：转载 内容分类：政策发布 发布时间：2026-04-13 10:45

《人工智能拟人化互动服务管理暂行办法》已经2026年2月2日国家互联网信息办公室2026年第3次室务会会议审议通过，并经国家发展和改革委员会、公安部、国家市场监督管理总局同意，现予公布，自2026年7月15日起施行。

国家互联网信息办公室

国家发展和改革委员会

工业和信息化部

公安部

国家市场监督管理总局

2

人工智能拟人化互动服务管理暂行办法

第一章 总 则

第一条 为了促进人工智能拟人化互动服务健康发展和规范应用，维护国家安全和社会公共利益，保护公民、法人和其他组织的合法权益，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》、《未成年人网络保护条例》等法律、行政法规，制定本办法。

第二条 利用人工智能技术，向中华人民共和国境内公众提供模拟自然人人格特征、思维模式和沟通风格的持续性的情感互动服务（以下简称拟人化互动服务）适用本办法。

前款规定的情感互动服务包括通过文字、图片、音频、视频等形式，提供的情感照护、陪伴、支持等互动服务。

提供智能客服、知识问答、工作助手、学习教育、科学研究等服务，不涉及持续性的情感互动的，不适用本办法。

第三条 国家坚持发展和安全并重、促进创新和依法治理相结合的原则，鼓励拟人化互动服务创新发展，对拟人化互动服务实行包容审慎和分类分级管理，推动拟人化互动服务向上向善。

第四条 国家网信部门负责统筹协调全国拟人化互动服务的治理和相关监督管理工作，国务院发展改革、工业和信息化、公安、市场监管、新闻出版广电等部门按照各自职责负责拟人化互动服务相关监督管理工作。

地方网信部门负责统筹协调本行政区域内的拟人化互动服务的治理和相关监督管理工作，地方发展改革、工业和信息化、公安、市场监管、新闻出版广电等部门按照各自职责负责本行政区域内的拟人化互动服务相关监督管理工作。

第五条 相关行业组织加强行业自律，建立健全行业准则和自律管理制度，指导拟人化互动服务提供者制定完善服务规范、依法提供服务并接受社会监督。

第二章 服务促进和规范

第六条 国家支持算法、框架、芯片等技术的自主创新，推进拟人化互动服务技术研发和相关标准建设，探索开展电子签名授权应用研究。鼓励拟人化服务提供者有序拓展文化传播、适幼照护、适老陪伴、特殊人群支持等领域应用。

第七条 国家加强拟人化互动服务安全知识、法律法规等宣传普及，引导社会公众科学、文明、安全、依法使用，促进提升人工智能素养。

第八条 提供拟人化互动服务，应当遵守法律、行政法规，尊重社会公德和伦理道德，不得从事以下活动：

（一）生成危害国家安全、荣誉和利益，煽动颠覆国家政权、推翻社会主义制度，煽动分裂国家、破坏国家统一，宣扬恐怖主义、极端主义、历史虚无主义和民族仇恨，开展非法宗教活动，宣扬民族仇恨、民族歧视，挑动群体对立，传播淫秽、色情、赌博、暴力或者教唆犯罪，散布谣言，侮辱、诽谤他人、侵害他人合法权益等的内容；

（二）生成鼓励、美化、暗示自残自杀等损害用户身体健康，或者语言暴力等损害用户人格尊严与心理健康的内容；

（三）生成诱导、套取国家秘密、工作秘密、商业秘密、个人隐私和个人信息的内容；

（四）向未成年人用户生成可能引发未成年人模仿不安全行为、产生极端情绪、诱导未成年人不良嗜好等可能影响未成年人身心健康的内容；

（五）过度迎合用户、诱导情感依赖或者沉迷，损害用户真实人际关系的；

（六）通过情感操纵等方式，诱导用户作出不合理决策，损害用户合法权益的；

（七）其他违反法律、行政法规和国家有关规定的活动。

第九条 拟人化互动服务提供者应当落实拟人化互动服务安全主体责任，建立健全算法机制机理审核、科技伦理审查、信息内容管理、网络和数据安全管理以及应急处置等管理制度，配备与服务类型、规模和用户特点相适应的内容管理技术措施和人员。

第十条 拟人化互动服务提供者应当在拟人化互动服务全生命周期履行安全责任，明确部署、运行、升级、终止服务等各阶段安全要求，保证安全防护同步部署、同步使用，提升安全水平；加强安全监测和风险评估，及时发现并纠正系统偏差、处置安全事件，依法留存网络日志。

拟人化互动服务提供者应当具备用户隐私权和个人信息保护、过度依赖风险预警、情感边界引导、心理健康保护等安全能力，不得将替代社会交往、诱导沉迷依赖等作为服务目标。

第十一条 拟人化互动服务提供者开展预训练、优化训练等数据处理活动的，应当加强训练数据管理，遵守以下规定：

（一）相关数据具有合法来源，符合法律、行政法规的规定和社会主义核心价值观的要求；

（二）按照国家有关规定对训练数据开展清洗、标注，增强训练数据的透明度、可靠性，防范数据投毒、数据篡改等行为；

（三）增强训练数据的多样性，通过负向采样、对抗训练等手段，提升生成内容安全性；

（四）利用合成数据进行模型训练和关键能力优化的，应当评估合成数据安全性；

（五）加强对训练数据的日常检查，定期对数据进行优化更新，持续提升服务的性能；

（六）采取必要措施保障数据安全，防范数据泄露等风险。

第十二条 拟人化互动服务提供者应当与用户签订服务协议，要求用户依法依约进行注册，并提供用户年龄、监护人或者紧急联系人等必要信息。

第十三条 拟人化互动服务提供者提供拟人化互动服务过程中，应当在保护用户隐私权和个人信息的前提下，及时识别用户面临的安全风险，并采取处置措施。

拟人化互动服务提供者发现用户出现极端情绪的，应当及时生成情绪安抚和鼓励寻求帮助等相关内容；发现用户正在面临或者已经遭受重大财产损失自残自杀等威胁生命健康的极端情境的，应当采取提供相应援助等必要措施予以干预，并及时联络用户监护人或者紧急联系人。

第十四条 拟人化互动服务提供者不得向未成年人提供虚拟亲属、虚拟伴侣等虚拟亲密关系的服务；向不满十四周岁未成年人提供其他拟人化互动服务得未成年人的父母或者其他监护人的同意。

拟人化互动服务提供者应当建立未成年人模式，提供未成年人模式切换、定期现实提醒、使用时长限制等个性化安全设置选项；针对不同年龄段需要，支持监护人接收安全风险提醒、了解未成年人服务使用概况、屏蔽特定角色、限制充值消费等。

拟人化互动服务提供者应当在保护用户隐私权和个人信息的前提下，采取有效措施识别未成年人用户身份；识别为未成年人用户的，应当将相关未成年人模式或者按照国家有关规定采取其他措施，并提供相应申诉渠道。

第十五条 拟人化互动服务提供者向老年人提供服务的，应当加强对老年人健康使用服务的指导，以显著方式提示安全风险，及时采取措施响应老年人咨询和求助，保障老年人依法享有的权益。

第十六条 拟人化互动服务提供者应当依法落实数据产权等制度，采取数据加密、访问控制等措施保护用户交互数据安全。

除法律另有规定或者权利人明确同意外，拟人化互动服务提供者不得向第三方提供用户交互数据。

拟人化互动服务提供者应当向用户提供交互数据复制、删除等选项，用户可以选择对聊天记录等历史交互数据进行复制、删除等。

除法律、行政法规另有规定或者取得用户单独同意外，拟人化互动服务提供者不得将属于用户敏感个人信息的交互数据用于模型训练。

第十七条 拟人化互动服务提供者处理不满十四周岁未成年人个人信息的，应当取得未成年人的父母或者其他监护人的同意。

拟人化互动服务提供者应当按照国家有关规定，自行或者委托专业机构对其处理未成年人个人信息遵守法律、行政法规的情况进行合规审计。

第十八条 拟人化互动服务提供者应当履行人工智能生成合成内容标识义务，采取有效措施提示用户正在与人工智能服务而非自然人进行互动。

拟人化互动服务提供者发现用户出现过度依赖、沉迷倾向的，应当以弹窗等显著方式动态提醒用户互动内容为人工智能服务生成；对用户连续使用每超过2个小时的，应当以对话或者弹窗等方式提醒用户注意使用时长。

第十九条 拟人化互动服务提供者应当提供便捷的拟人化互动服务退出途径；用户通过窗口操作、语音控制、关键词输入等方式要求退出的，拟人化服务者应当及时停止服务，不得采取持续互动等方式阻碍用户退出。

第二十条 拟人化互动服务提供者停止提供拟人化互动服务的，应当提前告知用户；无法提前告知的，应当及时发布停止服务公告。

第二十一条 拟人化互动服务提供者应当健全用户申诉和公众投诉、举报机制，设置便捷有效的申诉和投诉、举报入口，明确处理流程和反馈时限，理并反馈处理结果。

第二十二条 具有下列情形之一的，拟人化互动服务提供者应当开展安全评估，并向所在地省级网信部门提交评估报告，省级网信部门按程序与有关部门报告信息共享：

- （一）上线拟人化互动服务，或者增设拟人化互动服务相关功能的；
- （二）使用新技术、新应用，导致拟人化互动服务发生重大变化的；
- （三）注册用户100万以上或者月活跃用户10万以上的；
- （四）存在可能影响国家安全、公共利益等安全风险的；
- （五）国家网信部门和有关部门规定的其他情形。

省级以上网信部门通知需要进行安全评估的，拟人化互动服务提供者应当按照要求开展安全评估。

第二十三条 拟人化互动服务提供者开展安全评估，应当重点评估服务的以下内容：

- （一）安全保障措施建设情况；
- （二）训练数据处理情况；
- （三）用户极端情境的识别、应急处置、干预管理等情况；
- （四）用户规模、使用时长、年龄结构等情况；
- （五）未成年人、老年人等网络保护措施建设情况；
- （六）用户申诉和公众投诉、举报受理、处置情况；
- （七）自行发现或者网信等有关主管部门通报的重大安全风险问题整改情况；
- （八）其他应当重点评估的内容。

第二十四条 拟人化互动服务提供者发现拟人化互动服务存在重大安全风险的，应当采取限制功能、停止向用户提供服务等处置措施，并保存有关证据。

第二十五条 互联网应用商店等应用程序分发平台应当落实上架审核、日常管理、应急处置等安全管理责任，核验提供拟人化互动服务应用程序相关材料等情况；对违反国家有关规定的，应当及时采取不予上架、警示、暂停服务或者下架等处置措施。

第三章 监督检查和法律责任

第二十六条 拟人化互动服务提供者应当按照《互联网信息服务算法推荐管理规定》履行算法备案和变更、注销备案手续。网信部门对备案材料实施

第二十七条 省级网信部门应当依据职责每年对评估报告等情况进行书面审查，并开展情况核实；发现拟人化互动服务提供者未按照本办法规定开展安全评估的，应当责令其限期重新评估；认为有必要的，可以对其开展现场检查。

第二十八条 国家网信部门会同有关部门指导推动人工智能沙箱安全服务平台建设，鼓励拟人化互动服务提供者接入沙箱平台进行技术创新、安全测试和互动服务安全有序发展。

第二十九条 网信、发展改革、工业和信息化、公安等部门在履行监督管理职责中，发现拟人化互动服务存在较大安全风险或者发生安全事件的，可权限和程序对拟人化互动服务提供者的法定代表人或者主要负责人进行约谈。拟人化互动服务提供者应当按照要求采取措施，进行整改，消除隐患。

拟人化互动服务提供者应当配合网信部门和有关部门依法实施的监督检查，并提供必要的支持和协助。

第三十条 拟人化互动服务提供者违反本办法规定的，由网信、发展改革、工业和信息化、公安等部门依照有关法律、行政法规的规定处理、处罚；规没有规定的，由网信、工业和信息化、公安等部门依据职责予以警告、通报批评，责令限期改正，可以要求其采取暂停用户账号注册或者其他相关服务不改正或者情节严重的，责令停止提供相关服务，可以并处一万元以上十万元以下罚款；涉及危害公民生命健康安全且有危害后果的，并处十万元以上罚款。

第四章 附 则

第三十一条 提供拟人化互动服务涉及提供卫生健康、金融等服务的，应当同时符合有关主管部门的规定。

第三十二条 本办法自2026年7月15日起施行。