

请输入检索关键词

[首页](#) [权威发布](#) [办公室工作](#) [网络安全](#) [信息化](#) [网络传播](#) [国际交流](#) [地方网信](#) [执法督查](#) [政策法规](#) [互动中心](#) [教育培训](#) [业界动态](#) [工作专报](#)当前位置: [首页](#) > [正文](#)

国家互联网信息办公室关于《生成式人工智能服务管理办法（征求意见稿）》公开征求意见的通知

2023年04月11日 12:51

来源: 中国网信网

[【打印】](#) [【纠错】](#)

为促进生成式人工智能技术健康发展和规范应用，根据《中华人民共和国网络安全法》等法律法规，国家互联网信息办公室起草了《生成式人工智能服务管理办法（征求意见稿）》，现向社会公开征求意见。公众可通过以下途径和方式提出反馈意见：

1. 登录中华人民共和国司法部 中国政府法制信息网（www.moj.gov.cn、www.chinalaw.gov.cn），进入首页主菜单的“立法意见征集”栏目提出意见。

2. 通过电子邮件方式发送至：wajscy@cac.gov.cn。

3. 通过信函方式将意见寄至：北京市海淀区阜成路15号国家互联网信息办公室网络安全协调局，邮编100048，并在信封上注明“生成式人工智能服务管理办法征求意见”。

意见反馈截止时间为2023年5月10日。

附件：生成式人工智能服务管理办法（征求意见稿）

国家互联网信息办公室

2023年4月11日

生成式人工智能服务管理办法

（征求意见稿）

第一条 为促进生成式人工智能健康发展和规范应用，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》等法律、行政法规，制定本办法。

第二条 研发、利用生成式人工智能产品，面向中华人民共和国境内公众提供服务的，适用本办法。

本办法所称生成式人工智能，是指基于算法、模型、规则生成文本、图片、声音、视频、代码等技术。

第三条 国家支持人工智能算法、框架等基础技术的自主创新、推广应用、国际合作，鼓励优先采用安全可信的软件、工具、计算和数据资源。

第四条 提供生成式人工智能产品或服务应当遵守法律法规的要求，尊重社会公德、公序良俗，符合以下要求：

（一）利用生成式人工智能生成的内容应当体现社会主义核心价值观，不得含有颠覆国家政权、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，宣扬民族仇恨、民族歧视，暴力、淫秽色情信息，虚假信息，以及可能扰乱经济秩序和社会秩序的内容。

（二）在算法设计、训练数据选择、模型生成和优化、提供服务等过程中，采取措施防止出现种族、民族、信仰、国别、地域、性别、年龄、职业等歧视。

（三）尊重知识产权、商业道德，不得利用算法、数据、平台等优势实施不公平竞争。

(四) 利用生成式人工智能生成的内容应当真实准确，采取措施防止生成虚假信息。

(五) 尊重他人合法权益，防止伤害他人身心健康，损害肖像权、名誉权和个人隐私，侵犯知识产权。禁止非法获取、披露、利用个人信息和隐私、商业秘密。

第五条 利用生成式人工智能产品提供聊天和文本、图像、声音生成等服务的组织和个人（以下称“提供者”），包括通过提供可编程接口等方式支持他人自行生成文本、图像、声音等，承担该产品生成内容生产者的责任；涉及个人信息的，承担个人信息处理者的法定责任，履行个人信息保护义务。

第六条 利用生成式人工智能产品向公众提供服务前，应当按照《具有舆论属性或社会动员能力的互联网信息服务安全评估规定》向国家网信部门申报安全评估，并按照《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。

第七条 提供者应当对生成式人工智能产品的预训练数据、优化训练数据来源的合法性负责。

用于生成式人工智能产品的预训练、优化训练数据，应满足以下要求：

- (一) 符合《中华人民共和国网络安全法》等法律法规的要求；
- (二) 不含有侵犯知识产权的内容；
- (三) 数据包含个人信息的，应当征得个人信息主体同意或者符合法律、行政法规规定的其他情形；
- (四) 能够保证数据的真实性、准确性、客观性、多样性；
- (五) 国家网信部门关于生成式人工智能服务的其他监管要求。

第八条 生成式人工智能产品研制中采用人工标注时，提供者应当制定符合本办法要求，清晰、具体、可操作的标注规则，对标注人员进行必要培训，抽样核验标注内容的正确性。

第九条 提供生成式人工智能服务应当按照《中华人民共和国网络安全法》规定，要求用户提供真实身份信息。

第十条 提供者应当明确并公开其服务的适用人群、场合、用途，采取适当措施防范用户过分依赖或沉迷生成内容。

第十一条 提供者在提供服务过程中，对用户的输入信息和使用记录承担保护义务。不得非法留存能够推断出用户身份的输入信息，不得根据用户输入信息和使用情况进行画像，不得向他人提供用户输入信息。法律法规另有规定的，从其规定。

第十二条 提供者不得根据用户的种族、国别、性别等进行带有歧视性的内容生成。

第十三条 提供者应当建立用户投诉接收处理机制，及时处置个人关于更正、删除、屏蔽其个人信息的请求；发现、知悉生成的文本、图片、声音、视频等侵害他人肖像权、名誉权、个人隐私、商业秘密，或者不符合本办法要求时，应当采取措施，停止生成，防止危害持续。

第十四条 提供者应当在生命周期内，提供安全、稳健、持续的服务，保障用户正常使用。

第十五条 对于运行中发现、用户举报的不符合本办法要求的生成内容，除采取内容过滤等措施外，应在3个月内通过模型优化训练等方式防止再次生成。

第十六条 提供者应当按照《互联网信息服务深度合成管理规定》对生成的图片、视频等内容进行标识。

第十七条 提供者应当根据国家网信部门和有关主管部门的要求，提供可以影响用户信任、选择的必要信息，包括预训练和优化训练数据的来源、规模、类型、质量等描述，人工标注规则，人工标注数据的规模和类型，基础算法和技术体系等。

第十八条 提供者应当指导用户科学认识和理性使用生成式人工智能生成的内容，不利用生成内容损害他人形象、名誉以及其他合法权益，不进行商业炒作、不正当营销。

用户发现生成内容不符合本办法要求时，有权向网信部门或者有关主管部门举报。

第十九条 提供者发现用户利用生成式人工智能产品过程中违反法律法规，违背商业道德、社会公德行为时，包括从事网络炒作、恶意发帖跟评、制造垃圾邮件、编写恶意软件，实施不正当的商业营销等，应当暂停或者终止服务。

第二十条 提供者违反本办法规定的，由网信部门和有关主管部门按照《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》等法律、行政法规的规定予以处罚。

法律、行政法规没有规定的，由网信部门和有关主管部门依据职责给予警告、通报批评，责令限期改正；拒不改正或者情节严重的，责令暂停或者终止其利用生成式人工智能提供服务，并处一万元以上十万元以下罚款。构成违反治安管理行为的，依法给予治安管理处罚；构成犯罪的，依

法追究刑事责任。

第二十一条 本办法自2023年 月 日起实施。

中共中央网络安全和信息化委员会办公室 中华人民共和国国家互联网信息办公室 © 版权所有 联系我们 站内导航
承办：国家互联网应急中心 技术支持：长安通信科技有限责任公司 京ICP备14042428号

