引用格式： 郑鹰,张欣亮,张雪飞.生成式人工智能国内外标准化发展状况、面临挑战及对策研究[J].标准科学,2025(12):35-40.
ZHENG Ying,ZHANG Xinliang,ZHANG Xuefei. Research on the Standardization Development Status, Challenges, and Countermeasures of Generative Artificial Intelligence at Home and Abroad [J].Standard Science,2025(12):35-40.

生成式人工智能国内外标准化发展状况、面临挑战及对策研究

郑 鹰　张欣亮　张雪飞

（中国标准化研究院）

摘　要：【目的】通过开展生成式人工智能标准化发展状况、挑战及对策研究，为推动生成式人工智能的规范化和高质量发展提供参考。【方法】通过文献查阅法，研究分析国际标准化组织、欧盟、美国、日本及我国在生成式人工智能标准化领域的发展重点，并通过逻辑框架法，探索构建生成式人工智能标准体系框架。【结果】研究发现，生成式人工智能标准化工作尚在起步阶段，仍存在很多风险与挑战，相关技术及应用服务标准研制数量存在明显不足。【结论】应加强和完善人工智能标准化标准体系建设，加快标准的制定和实施，凝聚各方力量和推动国际合作，以期全面提升生成式人工智能标准化水平，促进其健康有序发展。

关键词： 生成式人工智能；标准化；挑战；对策

DOI编码： 10.3969/j.issn.1674-5698.2025.12.005

Research on the Standardization Development Status, Challenges, and Countermeasures of Generative Artificial Intelligence at Home and Abroad

ZHENG Ying　ZHANG Xinliang　ZHANG Xuefei

(China National Institute of Standardization)

Abstract: [Objective] This study aims to provide reference for promoting the standardized and high-quality development of generative artificial intelligence by investigating the development status, challenges, and countermeasures regarding the standardization of generative artificial intelligence. [Methods] Using the literature review method, this study analyzes the development priorities in the field of generative artificial intelligence standardization of international standardization organizations, the European Union, the United States, Japan, and China. Additionally, the logical framework method is adopted to explore and construct a standards system framework for generative artificial intelligence. [Results] The study finds that the standardization work of generative artificial intelligence is still in its initial stage, with many risks and challenges remaining, and there is a significant shortage in the number of developed standards related to technologies, application

基金项目： 本文受中央基本科研业务费重点项目“基于隐私计算的公共数据流通研究与应用”（项目编号：242024Y-11461）；山西省重点研发计划项目“面向基层治理的政务大数据服务的关键技术研究和应用示范”（项目编号：202302020101010）；中国标准化研究院基本科研项目“数字政府标准化发展指数研究”（项目编号：242024Z-11946）资助。

作者简介： 郑鹰，本科，高级工程师，研究方向为数字政府与公共服务标准化。
张欣亮，硕士，助理研究员，研究方向为数字政府及平台经济标准化。
张雪飞，博士，工程师，研究方向为数字政府与大数据标准化。

services. [Conclusion] Efforts should be made to strengthen the construction of the artificial intelligence standardization system, accelerate the development and implementation of standards, gather forces from all parties, and promote international cooperation. These measures are intended to comprehensively enhance the standardization level of generative artificial intelligence and facilitate its healthy and orderly development.

Keywords: generative artificial intelligence; standardization; challenges; countermeasures

0　引言

近年来，生成式人工智能取得了突破性进展，包括以ChatGPT、DeepSeek为代表的大语言模型以及能够生成逼真图像、视频的人工智能工具[1]。作为一种利用人工智能算法自动生成文本、图像、音频、视频等各种形式内容的新型技术，生成式人工智能的技术特点包括以下几个方面：

（1）创造性生成。突破了传统数据处理的局限，具备强大的内容生成创造能力。

（2）多模态融合。支持文本、图像、音频、视频等多种形式内容的生成与交互。

（3）多样性输出。面对用户同一输入，生成式人工智能能够生成多种不同结果，可满足多样化需求。

（4）开源性共建。呈现“去中心化社区共建”模式，模型架构、训练代码、优化工具等核心技术要素通过开源平台向全球开发者开放，形成跨机构、跨地域的协作网络。

（5）快速迭代。生成式人工智能技术发展迅速，模型能够不断迭代优化，以适应新的任务和场景。

生成式人工智能正深刻改变着信息内容的平台生产、传播与交互方式，在电商、医疗、金融、休闲娱乐等诸多领域展现出巨大的应用前景。国家统计局2024年相关统计数据显示，在电商领域，超65%的大型电商平台借助AI生成商品详情；在医疗领域，38%的三甲医院采用AI辅助生成病历、分析医学影像，使医生的诊断效率平均提高38%；在金融领域，60%的金融机构运用AI生成风险报告、优化投资策略，业务处理速度加快42%，成本降低约28%；在休闲娱乐领域，45%的网民接触过AI生成的影视、游戏、音乐等内容，虚拟产业市场规模在2024年增长了70%。

生成式人工智能技术快速发展的同时也引发了一系列风险与问题，如数据隐私泄露、算法偏见、虚假信息传播、知识产权纠纷等[2]，对社会秩序、经济运行和个人权益构成潜在威胁。在此背景下，推动生成式人工智能标准化工作，已成为全球共同应对生成式人工智能技术风险、应用风险，促进其技术规范、健康发展的关键举措。

1　生成式人工智能国际标准化发展状况

1.1 国际组织

当前，国际组织积极投身生成式人工智能国际标准化工作，发挥着重要的引领与协调作用。国际标准化组织（ISO）、国际电工委员会（IEC）和国际电信联盟（ITU）联合开展世界标准合作倡议[3]，致力于通过全球共识加强人工智能规范体系建设，推动生成式人工智能技术安全、负责、有效地跨领域应用[4]。世界数字技术院（WDTA）于2024 年在第27届联合国科技大会期间，正式发布WDTA AI-STR-01《生成式人工智能应用安全测试标准》和WDTA AI-STR-02《大语言模型安全测试方法》2项国际标准，来自OpenAI、谷歌、微软、英伟达、蚂蚁集团、科大讯飞、百度、腾讯等数十家中外科技企业参与了编制[5]，为国际业界提供了统一的测试框架和明确的测试方法，填补了大语言模型和生成式人工智能应用安全测试领域的空白[6]。

1.2 欧盟

欧盟对生成式人工智能的安全监管较为严格。欧盟于2024年8月1日正式实施《人工智能法案》。该法案对生成合成音频、图像、视频或文本

内容的有限风险明确规定了透明度与明确标识义务，要求相关系统在运行过程中向用户清晰告知其使用的人工智能技术，以及生成内容的来源和性质等信息，旨在保障用户知情权，防范虚假信息和潜在风险传播。欧洲标准化委员会（CEN）、欧洲电工标准化委员会（CENELEC）及欧洲电信标准化协会（ETSI）三大区域性标准化机构分工协作，聚焦于通用技术与安全标准。同时，欧盟在数据隐私保护、算法问责等方面的严格要求，也为生成式人工智能系统在数据使用和算法设计方面提供了重要技术规范，促使企业在技术开发过程中充分考虑数据安全与用户权益保护。

1.3 美国

美国对生成式人工智能国际标准研制参与度高。美国高度重视生成式人工智能技术发展与标准化工作，2024年4月，美国国家标准与技术研究院发布《降低合成内容带来的风险》，详细列出检测、验证和标记合成内容的方法，包括数字水印和元数据记录等技术手段，为识别和管理生成式人工智能产生的合成内容提供技术指导。此外，由于美国在人工智能技术标准研发方面处于世界领先地位，众多科技巨头积极参与国际标准化活动，凭借其技术优势和创新能力，在推动生成式人工智能国际标准制定过程中发挥着重要的影响力。

1.4 日本

日本在生成式人工智能领域的场景应用取得实效。日本在生成式人工智能标准化发展中，遵循坚实的科技基础与审慎的管理政策。2023年，政府组建的特别战略小组制定战略，明确生成式人工智能工具严禁触碰敏感信息，并需配套严密的隐私泄露等风险防范措施。在政务领域，一些部门及地区已在探索生成式人工智能的应用潜力。如总务省已在开展试点，期望借助生成式人工智能简化标准化工作流程，提升行政效率；在医疗领域，自2018年起，厚生劳动省和经济产业省已着手制定医疗领域生成式人工智能的应用标准；在教育领域，自2023年发布《初等和中等教育阶段使用生成式人工智能的暂定方针》，引导学生规范应用生成式人工智能信息能力。总之，日本在各行业领域紧扣安全、责任等核心要素，凭借政策监管与标准化试点稳步推进，为生成式人工智能的健康发展筑牢根基。

2 生成式人工智能国内标准化发展状况

我国在国家层面高度重视生成式人工智能标准化工作，出台了一系列政策法规为其发展保驾护航。2023年7月发布并于8月15日正式实施的《生成式人工智能服务管理暂行办法》明确要求，生成式人工智能服务提供者应使用合法来源的数据和基础模型[7]，并致力于提高训练数据的质量。这一办法体现了国家在促进人工智能创新的同时，对国家安全和公众利益的充分考量，为后续的标准化工作提供了坚实的政策依据与发展导向。

在具体标准制定上，2025年4月25日，国家市场监督管理总局与国家标准化管理委员会发布了3项生成式人工智能领域关键性国家标准，分别为GB/T 45654—2025《网络安全技术 生成式人工智能服务安全基本要求》、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》、GB/T 45652—2025《网络安全技术 生成式人工智能预训练和优化训练数据安全规范》[8]，并于2025年11月1日正式实施。这3项国家标准的研制实施对提升我国在全球生成式人工智能领域的竞争力，保障国家网络安全、社会公共利益及用户合法权益提供了坚实的标准化支撑 。

3 生成式人工智能标准化面临的挑战

尽管生成式人工智能标准化工作取得了一定进展，但目前仍处于起步阶段，其标准化推进过程中仍面临着诸多严峻挑战。

3.1 生成式人工智能的开源及快速迭代特性对标准化工作构成挑战

生成式人工智能的去中心化协作导致技术碎

片化，造成不同开源社区的模型架构、接口协议差异显著，加剧了统一标准的制定难度。另外，生成式人工智能快速迭代的技术节奏与标准化工作所需的稳定性形成矛盾。例如多模态生成技术在近期取得重大突破，能够实现文本、图像、音频、视频等多种模态信息的融合生成，极大拓展了应用边界。然而，标准制定过程通常需要经过预研、起草、征求意见、审查和报批等流程，这使得标准更新速度明显滞后于生成式人工智能技术发展，导致生成式人工智能应用场景往往存在技术标准空白，增加了技术滥用和安全隐患的风险。

3.2 国际的协调困难也阻碍着标准化进程

不同国家和地区在人工智能发展战略、价值观、法律体系及技术实力等方面存在显著差异。欧盟强调伦理道德和隐私保护优先，将其作为制定标准的核心考量；美国注重技术创新与产业发展，更倾向于宽松的监管环境以激发企业创新活力；而发展中国家更关注技术的普惠性和能力建设。这种差异使得在全球范围内达成统一、有效的国际标准规范共识难度较大。不同国家和地区由于文化、价值观的差异，对于伦理道德的认知也不尽相同，这使得制定全球统一的伦理道德标准面临重重困难，进而影响生成式人工智能在全球范围内的标准化推广。

3.3 生成式人工智能伦理与法律问题不容忽视

在伦理方面，生成式人工智能面临诸如算法偏见、虚假信息传播、内容侵权等问题，尚没有明确的法条及标准加以规范。例如，一些基于生成式人工智能的内容创作工具可能会生成带有偏见性的内容或图像，对社会公平和价值观造成负面影响；虚假信息及内容的自动化生成与泛滥传播，也给信息真实性和社会稳定带来挑战。从法律角度看，现有的法律框架在应对生成式人工智能带来的新问题时仍存在诸多不适应性。例如，在训练数据合法性、生成内容版权归属、生成结果侵权认定等方面缺乏明确的规定及标准，这导致司法实践中存在大量争议，增加了企业和开发者的法律风险。

3.4 不同应用场景下的安全风险复杂多样

生成式人工智能面临多种安全风险。内容生成的不可预测性使得模型可能输出有害的言论、虚假信息甚至是敏感内容，对用户和社会造成潜在、持续的危害。例如在政治领域制造虚假新闻，扰乱社会秩序。而目前的内容检测技术还难以完全准确、及时地识别深度伪造的敏感内容。在数据安全隐私方面，生成式人工智能模型的训练方式往往依赖大量用户数据，因而存在关键信息泄露风险。模型本身也可能被恶意使用，增加了隐私数据伪造和滥用风险[9]。针对这些复杂多样的安全风险，生成式人工智能安全标准研制工作面临着巨大挑战，需要综合考虑研制技术防护、风险评估、应急处置等多个方面的安全标准，并针对不同应用场景下的安全需求体现标准内容差异。

此外，行业应用的差异化需求也给标准化工作带来挑战。不同行业对生成式人工智能的性能、安全性、可靠性等要求各不相同。

4 对策建议

为应对上述挑战，推动生成式人工智能标准化朝着更加科学、合理、有效的方向发展，宜重点关注以下几个方面。

（1）要强化国际合作与规则协同。在联合国框架下成立开放式专家组，整合各国技术路径与法律监管经验，推动生成式人工智能风险治理国际公约谈判。国际标准化组织可充分发挥协调作用，搭建全球性的人工智能治理与标准制定平台，汇聚各国政府、国际组织、企业、科研机构和民间团体等各方力量，通过多边对话与协商，制定具有广泛共识和普遍约束力的国际标准规范。推动不同国家和地区间生成式人工智能标准的互认与衔接，鼓励企业、行业协会等非政府组织积极参与国际标准制定，充分发挥其在技术创新和行业实践方面的优势，使生成式人工智能国际标准制定更贴合国际市场需求与技术发展趋势。

（2）规范伦理与安全风险综合治理。在技术

研发阶段，将伦理与安全风险审查机制深度嵌入到生成式人工智能的开发流程规范中。要求开发者在模型训练、算法设计等环节充分考虑伦理及安全风险因素，通过算法优化减少伦理偏见，防止不良、违法信息生成及隐私信息泄露等安全风险。另外，应针对生成式人工智能的特点，进一步制定和完善相关法律法规及标准规范，如制定人工智能生成内容的版权归属通则，构建适应算法决策的风险责任认定规则，确定针对数据隐私和个人信息的保护规范，为生成式人工智能的应用提供准确、明晰的标准化指引。

（3）建立共建、快速的标准制修订机制。设立专门的技术监测小组，实时跟踪生成式人工智能技术发展动态，建立开源社区多方共建式标准制修订机制，吸纳平台、企业、开发者代表共同参与，将不同社区的最佳实践转化为标准技术内容，同时保留技术兼容性接口规范，平衡统一标准与技术多样性，及时识别新技术带来的风险与挑战，为标准制修订提供技术依据。同时，对于市场急需的标准可依据《国家标准化指导性技术文件管理规定》开展标准化指导性技术文件的研制工作，制定周期压缩至12个月，以同步跟踪新兴技术发展，提升生成式人工智能标准文件制定与实施的效率和科学性。

（4）提升生成式人工智能标准研制质量。宜针对生成式人工智能在不同行业应用场景的特定需求，制定高质量、可操作的应用场景标准。例如，在医疗领域，制定生成式人工智能辅助医疗诊断的准确性评估标准、数据安全防护标准，确保人工智能技术在医疗场景中的安全、可靠应用；在金融领域，建立生成式人工智能风险评估与防控标准，规范其在金融交易预测、客户信用评估等方面的应用；在政务领域，规范生成式人工智能系统与政务云平台、各部门数据库的数据接口格式，确保跨领域数据共享的安全性与兼容性，构建隐私保护和应急响应标准[10]。同时，宜加强标准实施质量的监督与评估，建立第三方认证机制，对符合标准要求的企业和产品给予认证标识，提高市场对标准的认可度和遵循度。

（5）构建完善的生成式人工智能标准体系。应构建全面、系统、科学的生成式人工智能标准体系。体系框架拟涵盖基础层、技术层、安全层、合规层、内容应用层五维多层结构（见图1）。在基础层面，构建统一的术语定义、参考模型、内容标识和共性技术规范，确保各层级标准的一致性与互操作性；在技术层面，制定数据处理、算法模型、生成内容、平台系统等的具体技术标准，确保生成式人工智能技术的规范性与可控性；在安全层面，构建覆盖数据安全、模型安全、内容安全、系统安全的安全标准，防范生成式人工智能领域各类安全风险；在合规层面，从伦理治理、合规管理、内容评估及其他等角度制定合规标准规范，平衡技术发展与社会合规及伦理价值的协调关系；在内容应用层面，结合不同行业的特性与需求，制定服务业（涵盖金融、医疗、教育、传媒等领域）、公共服务（涵盖政务服务、城市治理等领域）、制造业、医疗健康及其他等领域的生成式人工智能内容应用及服务标准，推动技术与行业产业的深度融合。标准体系的各层级应相互支撑、有机衔接，为生成式人工智能标准化工作的开展提供全面指引。

5 结 语

通过对国内外标准化发展状况及面临挑战的分析研究可以看出，虽然国际和国内在生成式人工智能标准化工作方面都取得了一定进展，但仍面临着标准缺失、国际协调、伦理合规、安全风险等方面诸多挑战。因此，应从加强国际合作、规范伦理与安全风险，优化标准制定程序、提升标准研制质量、完善标准体系等方面入手，提升生成式人工智能标准化工作质量，构建起统一、安全、合规、快捷、普惠的生成式人工智能标准应用生态，支撑人类社会在数字化、智能化浪潮中稳步前行。

图1 生成式人工智能标准体系框架

参考文献

[1] 朱红儒,静静,彭骏涛,等.人工智能治理国内外政策与标准分析[J].中国信息安全,2023(5):48–52.

[2] 林阳荟晨,上官晓丽.欧美人工智能网络安全标准化最新动态[J].信息技术与标准化,2023(12):57–61.

[3] 王刚,赖海龙,李娟婷,等.有关知识产权创造、保护及其利用的推进计划(日本知识产权促进战略)[J].网络法律评论,2004,5(2):305–354.

[4] 梁桐.加快人工智能标准化建设[N].经济日报,2024–11–21(4).

[5] 李铎.AI技术创新“风起云涌”“智慧”生活[J].大众投资指南,2024(24):3–5.

[6] 周武英.人工智能在全球产业变革中地位凸显[N].经济参考报,2024–12–27(4).

[7] 董毅敏,吴素平.我国数字内容产业发展趋势、挑战与建议:基于2019至2023年数据观察[J].中国出版,2024(5):34–40.

[8] 高松.生成式人工智能的数据安全风险与应对措施[J].中国信息安全,2024(6):32–37.

[9] 袁杰.浅析生成式人工智能个人数据风险及法律规制[N].重庆科技报,2024–12–10(4).

[10] 王芬,王可欣,李玉兰,等.提高数字化管理水平促进城市高质量发展[N].江苏经济报,2025–04–03(T5).