

政策建议

敏捷视角下生成式人工智能的治理理论、欧美实践与优化思路

孙茜¹, 贾颖颖², 崔程昱¹, 郭栋^{3*}

摘要 以 DeepSeek、ChatGPT 等为代表的生成式人工智能作为一项新兴数字技术催生了新场景、新业态和新市场,但新兴技术的不确定性风险也对传统治理结构与范式形成冲击,并对治理体系和治理能力提出了新的挑战。面向“善治”的要求,结合敏捷治理理论,提出生成式人工智能敏捷治理的框架,以期为中国未来生成式人工智能治理理念及工具模式选择提供参考。采用理论综述法与案例分析法,从治理理论出发,系统分析了美国和欧盟生成式人工智能治理实践,并从治理目标、治理主体、治理手段和治理关系 4 个方面,探究人工智能国际治理理念的共识性指征和差异性特点,在此基础上,提出中国生成式人工智能治理的优化思路,即遵循柔性的分层治理原则、构建多主体互动的网络治理关系、采取积极主动的治理思路,以适应新技术的动态特征和不确定性风险,积极推动生成式人工智能技术及产业健康有序发展,实现技术向善。

关键词 生成式人工智能;人工智能风险;数字技术治理;敏捷治理

生成式人工智能(generative artificial intelligence, GAI)的发展日新月异。Gartner 提出“2024 年 10 大战略技术趋势”背后的共同主题是人工智能(artificial intelligence, AI),其预测 2026 年超 80% 的企业将使用生成式人工智

能的相关应用程序编程接口或模型,或在生产环境中部署支持 GAI 的应用程序,实现生成式人工智能的普及化(democratized GAI)。国内文心一言、通义千问、讯飞星火等大模型相继推出并且社会应用范围不断扩大,2024 年底,深度求索(DeepSeek)横空出世,更是成为引爆全网的热点。

与此同时,生成式人工智能的“双刃剑”特性也日益突出。一方面,基于深度学习技术与生成式算法,利用现有的数据进行训练,可以发现数据特点与分布规律,进而

生成新的数据^[1],这一过程使其不仅拥有数据分析能力,还演化出数据孪生能力^[2]。另一方面,一系列社会风险也由此衍生。一是法律伦理冲击。世界范围内对 GAI 治理的法律体系尚在建设与完善阶段,传统的治理体系敏捷性不足,导致政策调整滞后于 GAI 的创意化场景风险,如新型隐私泄露、内容诈骗、不良价值观渗透与知识产权保护难题^[3]。二是技术认知冲击。表现为政府治理技术认知落后于技术发展进程,技术认知的信息不对称导致监管者面临平衡

1. 中国科学院大学公共政策与管理学院,北京 100190
2. 中国人民大学公共管理学院,北京 100872
3. 中国传媒大学经济与管理学院,北京 100024

收稿日期:2025-06-04;修回日期:2026-05-06

基金项目:国家自然科学基金青年项目(72404262, 72304262);第十届中国科协青年人才托举工程;中国传媒大学中央高校基本科研业务费专项(CUC26TD15)

作者简介:孙茜,副教授,研究方向为人工智能治理、数字医疗与养老,电子信箱: sunqian@ucas.ac.cn;郭栋(通信作者),副研究员,研究方向为创新管理、科技政策、数字治理,电子信箱: guodong@cuc.edu.cn

引用格式:孙茜,贾颖颖,崔程昱,等.敏捷视角下生成式人工智能的治理理论、欧美实践与优化思路[J].科技导报,2026,44(15):153-160;

doi:10.3981/j.issn.1000-7857.2025.06.00014

政策力度与新兴产业发展的治理困境。三是社会秩序冲击。主要表现在计算机化对低教育与低技术水平的劳动力的替代所引发的失业问题等方面^[4]。数据隐私、内容虚假、产权争议与算法偏见风险,以及透明度、公平性、可解释性等风险敞口,呼吁生成式人工智能治理必须匹配多元协作、动态适应、快速响应和灵活包容的敏捷治理框架。

美国、欧盟等国家和地区积极推进生成式人工智能健康有序发展,并在隐私安全、伦理审查、合作共赢等方面逐步形成初步治理共识。本文基于对治理理论尤其是人工智能治理相关研究的剖析,总结提炼生成式人工智能治理的欧美实践,并从前瞻性的敏捷治理视角提出应对生成式人工智能风险冲击的思路框架,进而探索规范中国生成式人工智能的发展路径,以提升新兴技术的正外部性,并降低技术的经济社会风险。

1 人工智能技术治理的理论分析

1.1 治理理论

20世纪80年代,在经济低迷和福利国家模式难以为继的背景下,以“政治-行政二分法”以及韦伯的科层制为理论基础的传统公共行政模式难以应对科技革命和经济全球化带来的变革,“治理”作为危机的回应方案从企业管理进入到公共管理领域^[5]。西方政治学家主要关注治理的3个议题:一是治理的含义。罗西瑙(Rose-nau)将治理定义为一系列活动领

域里的管理机制,即使未得到正式授权却能有效发挥作用^[6]。罗兹(Rhodes)认为“治理”是“统治”的含义变化,有序统治的条件已经与之前不同^[7]。二是“善治”的治理目标。世界银行将“善治”定义为“好的治理”,意为公共部门管理、问责、法治和信息透明性^[5]。也有学者将“善治”解释为一种新型民主与社会转型相伴而生的治理形态^[8]。三是修正治理失灵。学者指出“元治理”可有效应对去国家化和去政府化产生的治理失灵,是科层治理、网络治理、市场治理3种治理方式的明智组合^[9],意味着政府重新回到治理中心。

治理理论于20世纪末引入中国,俞可平将治理定义为在一定的范围内运用权威以维持秩序,满足公众的需要,并进一步提出“善治”的概念,定义其为“使公共利益最大化的社会管理过程”,将合法性、透明性、责任性、法治、回应、有效概括为善治的基本要素^[10-11]。杨开峰在区分治理与善治时,提出治理研究本土化和国际化的平衡,在不同国家治理的情境特殊性中寻求对话的根基和可能,更好地处理价值的普遍性和特殊性^[12]。

1.2 新兴技术及人工智能治理研究

新兴技术具有完全新颖、成长迅速、影响突出等特征^[13],兼具“经济与社会发展赋能”和“不确定性风险”的技术双重性^[14]。其中,不确定性风险带来的传统治理结构与范式冲击,引发了对技术颠覆性和治理延续性的讨论,掀起了新兴技术治理的研究热潮。21世

纪以来,数字技术发展对公共部门变革产生了深刻影响,传统治理模式过于注重效率而忽视价值与公平的问题开始显现。在此背景下,新兴技术治理理论将为人工智能这类风险与路径未知的新兴技术治理提供理论依据。

在人工智能治理领域,相关研究按照“技术发展-场景赋能-风险治理-理论构建”的研究框架可以分为4类。第1类研究聚焦于人工智能技术本身的规制,从算法模型和训练数据维度防范算法歧视与隐私泄露等问题^[15]。第2类聚焦于人工智能赋能公共部门治理场景,关注人工智能辅助决策,提升公共部门治理效率^[16]。第3类是人工智能应用带来的社会问题与风险治理,通过法律等手段对权利责任、伦理道德进行监管以保障人权。第4类是人工智能治理理论研究,从公共价值与利益相关者等理论视角,形成公共部门、企业、个人共同参与的治理体系,构建多方信任的互动治理关系^[17]。

2 欧美生成式人工智能治理模式探究

美国和欧盟在人工智能技术发展路线方面存在先发优势,在人工智能治理方面各具特色。本文对美国 and 欧盟发布的生成式人工智能治理的主要政策文件进行分析,从治理的目标、主体、手段和关系4个方面进行阐释。美国和欧盟生成式人工智能治理政策法规的相关重点内容梳理见表1和表2。

表 1 美国生成式人工智能治理政策法规重点内容

法律法规	发布年月	制定机构	重点内容
《人工智能权利法案蓝图》	2022年10月	美国白宫	提出人工智能技术应用的五项原则与实践保护措施； 建立安全有效的人工智能系统、建立算法歧视保护措施、保护数据隐私、强调通知和透明度的重要性、开发退出机制
《人工智能风险管理框架》	2023年1月	美国国家标准与技术研究院	提升人工智能可信度,明确人工智能风险； 加强组织文化建设与政策实践,完善风险问责机制； 通过组织内部与组织间合作共担分享,改进决策过程
《联邦法规》37编202部分	2023年3月	美国版权局	对生成式人工智能有关的版权认定和登记指引做出政策规定
《国家人工智能研发战略计划 2023》	2023年5月	美国白宫	提出人工智能研发等9项战略； 提出通过跨学科研究降低人工智能的伦理、法律与社会风险； 加强测试、验证以确保人工智能的安全性
《生成式人工智能技术的开发、部署和使用原则》	2023年6月	美国计算机协会、(ACM)技术政策理事会	提出人工智能8项原则,构建技术安全新标准与指南以规避潜在技术危害； 指出生成式人工智能应遵循部署和使用限制、所有权、个人数据控制、可纠正性原则
《关于安全、可靠和值得信赖地开发和和使用人工智能的行政命令》	2023年10月	美国白宫	根据8项指导原则和优先事项推进人工智能的开发和管理； 制定为人工智能生成的内容加水印的标准
《消除美国在人工智能领域领先地位的障碍》	2025年1月	美国白宫	维持和增强美国的全球人工智能主导地位,以促进人类繁荣、经济竞争力和国家安全； 撤销第 14110 号行政命令

表 2 欧盟生成式人工智能治理政策法规重点内容

法律法规	发布年月	制定机构	重点内容
《通用数据保护条例》	2018年5月	欧洲议会	建立问责机制以保护隐私； 界定数据主体知情权等一系列权利； 限制敏感数据的处理与使用
《可信人工智能伦理指南》	2019年4月	欧盟委员会人工智能高级专家组	确立欧盟人工智能伦理的5项原则与价值观； 从技术方法与非技术方法的视角实现可信人工智能
《数据治理法案》	2022年5月	欧盟理事会	通过协议范本与技术标准防止数据共享时的泄露； 加强企业间数据合同的公平性； 促进欧盟的数据利他主义
《人工智能法案》	2024年6月	欧洲议会	通过罚款的形式禁止销售、使用对人类生计、安全构成威胁的人工智能技术； 对人工智能风险等级和治理策略划分等级,对高风险人工智能严格监管

2.1 治理目标:技术霸权与重视人权

通过梳理美国联邦政府在生成式人工智能治理的行政指令与人工智能战略规划,以及欧盟的政策法案,总结提炼两者的治理目标。(1) 美国: 聚焦技术前沿, 维持先发优势。美国生成式人工智

能治理以减少技术创新与产业应用障碍为主要治理目标, 旨在维持其在生成式人工智能数据采集、大模型、多元生成模态和技术研发生态等技术前沿的先发优势。如白宫发布《赢得竞赛: 美国人工智能行动计划》, 特朗普同时签署 3 项行政命令, 强调促进创新和竞

争并提升美国在海外的领导力, 此外, 《国家人工智能研发战略计划 2023》也将开展可扩展的通用人工智能系统研究列为优先事项。(2) 欧盟: 重视人权, 构建欧洲价值观。欧盟人工智能治理强调尊重个人的权利, 在本土范围内注重人权保护, 并强调隐私保护的全面

性,禁止在执法过程、工作场所和教育领域使用人工智能进行识别。在世界范围内构建“以人为本”的治理观,传播人工智能治理的“欧洲价值观”,打造在人工智能国际治理生态中的规则参照系地位,其生成式人工智能治理也沿袭了这一治理思路。

2.2 治理主体:多元治理与链条监管

美国生成式人工智能治理体系架构主要体现在多元主体参与的治理结构上,欧盟则在多元主体治理架构基础上进一步采取“评估—服务—反馈”的全链条监管。

1) 美国:多元主体参与的治理结构。美国生成式人工智能治理不是由单一机构负责,而是多个联邦机构和政府部门协同工作。例如,卫生与公众服务部(HHS)设立人工智能工作组,负责推动人工智能技术在医疗保健、公共卫生和公众服务领域的安全、负责的使用;基础设施高级研究计划局(ARPA-I)则探索与交通相关的人工智能机遇和挑战。

2) 欧盟:“评估—服务—反馈”全链条监管。欧盟基于多主体的链路监管,将生成式人工智能纳入覆盖事前、事中与事后的全过程监管体系,监管对象包括相关主体在产品或服务提供前、提供中及提供后的各类行为主体。监管与治理范围涵盖欧盟内设机构或位于欧盟内的人工智能系统的提供者、位于欧盟境内的部署者、设立于或位于第三国,但向欧盟输出人工智能系统内容的提供者,并在训练数据集创建阶段、模型训练阶段、使用阶段同步实施域内及

域外管辖相结合的监管机制。

2.3 治理手段:灵活监管与分类治理

从风险源分类视角,美国基于技术的细分领域出发实施分类治理,欧盟则根据4个级别的风险对人工智能系统进行分类治理。“分类—匹配—治理”的敏捷执行机制有助于提升政策工具选择的精准性,显著提升治理效能。

1) 美国:问题导向,细分技术领域。美国生成式人工智能立法以问题为导向,通过对技术领域进行细分,针对风险高发环节实施差异化治理,而不设置全行业规制。如在自动驾驶、人脸识别、音频模仿、算法推荐等不同应用场景确定治理规范,提出限定性要求,提出细分规则。同时,其治理体系在对现实问题及其不良后果的持续修订与完善中趋于完备,呈现回应式治理特征。如针对“缺乏人类创造性贡献”的AI作品不被列入版权法保护范围。

2) 欧盟:分类治理与外部问责。欧盟通过风险源细化分类与分层,构建精细化治理体系。《人工智能法案》首次引入基于风险等级的分类管理模式,依据人工智能的不透明性、复杂性、对数据的依赖性、自主行为性特征,将其划分为不可接受的风险、高风险、有限风险和最小风险,并据此实施差异化监管,同时,对通用型人工智能系统和基础模型要求适用高风险人工智能系统的监管规则。此外,从平台主体义务、监管机构职责等方面入手,《人工智能法案》构建了算法的外部问责机制。

2.4 治理关系:柔性治理与硬性约束

美国与欧盟在人工智能全球化治理中采取了差异化的执行路径。基于创新激励导向的治理目标,美国生成式人工智能治理呈现弱治理的特征,治理架构分散,而欧盟则较为严格,以风险规制为导向,采用强监管的硬性约束手段。

1) 美国:行业自发,柔性治理。美国生成式人工智能发展规则多由行业制定,而非完全依赖政府强制性规制。如美国计算机协会(ACM)计算政策理事会发布《生成式人工智能技术的开发、部署和使用原则》,为行业发展提供了重要指引。

2) 欧盟:全流程控制,硬性约束。欧盟通过《通用数据保护条例》《数据治理法案》《数字服务法案》《数字市场法案》《算法问责及透明度监管框架》《可信人工智能伦理指南》和《人工智能法案》7大政策构建了相对统一的人工智能治理规则体系,以打造区域统一的横向全面治理模式,促进欧盟内部形成合法、安全、可信的人工智能市场,生成式人工智能治理沿用这一体系。

3 生成式人工智能治理理念及工具模式

3.1 欧美生成式人工智能的治理理念比较

美国与欧盟为人工智能治理和促进生成式人工智能健康发展进行了初期路径探索(表3)。总体而言,美国在生成式人工智能监管方面尚缺乏统一且明确的导向,

表 3 欧美生成式人工智能治理经验总结与相关政策法规梳理

治理维度	美国	欧盟
治理目标	聚焦技术前沿,维持技术霸权 《赢得竞赛:美国人工智能行动计划》 《国家人工智能研发战略计划2023》 《消除美国在人工智能领域领先地位的障碍》	尊重人权,构建欧洲价值观 《人工智能法案》 《生成式人工智能与EUDPR》
治理主体	多元参与的治理结构 《关于通过联邦政府进一步促进种族平等和 支持服务欠缺社区的行政命令》	“评估—服务—反馈”全链条监管 《人工智能法案》 《生成式人工智能与EUDPR》
治理手段	问题导向,细分技术领域 《关于安全、可靠和可信的人工智能的行政命令》 《人工智能权利法案蓝图》	分类治理,外部问责 《人工智能法案》
治理关系	行业自发,柔性治理 《关于安全、可靠和可信的人工智能的行政命令》 《生成式人工智能技术的开发、部署和使用原则》	全流程控制,硬性约束《通用数据保护条例》 《数据治理法案》《数字服务法案》《数字市场法案》 《人工智能法案》《算法问责及透明度监管框架》 《可信人工智能伦理指南》

而其政党政治的不确定性也在一定程度上加剧了生成式人工智能技术发展、监管环境和保障条件的波动性。相较之下,欧盟虽然在分类治理、透明性等方面具有优势,但在外部企业的市场准入、技术创新与标准、数据隐私和保护等方面存在诸多限制,也在一定程度上约束了创新活力。

当前,技术创新与风险规制已不再是简单的取舍关系,新兴技术所引发的生产方式和治理方式变革,要求技术与治理实践实现更加深度融合^[13-14]。在此背景下,实现技术发展与治理的平衡,科技前沿治理要与技术应用情境相适应,探

究人工智能国际治理理念的共识性指征和价值对话,可为中国生成式人工智能的前瞻性治理提供经验指引,实现生成式人工智能赋能社会变革与人类进步,把握人工智能国际竞争优势。

3.2 生成式人工智能敏捷治理框架

在 2018 年世界经济论坛提出了敏捷治理(agile governance)的概念,其主要特征包括敏捷性、适应性、持续性、灵活性、自组织性和包容性等^[18]。“敏捷”的先进理念被认为是新兴技术治理范式转变的新方向,回应了新兴技术的经济与社会发展赋能的同时伴生不

确定性风险这一双重属性,可有效解决技术的经济社会风险^[19-21]。敏捷治理内涵与各类治理模式相互联系,并整合了优势(表 4)。

伴随着人工智能技术的发展与演进,“敏捷”的聚焦点也逐步由企业的技术工具外延至技术与治理间的相互作用,在应对生成式人工智能带来的不确定性风险方面展现出更强的适应性。结合四大治理维度,在已有治理理论的基础上,本文依据生成式人工智能的技术特点和风险冲击提出一个前瞻性的生成式人工智能敏捷治理框架(图 1),为中国生成式人工智能治理提供工具模式选择。在治

表 4 敏捷治理相近概念辨析

相近概念	理论核心	与敏捷治理的区别与联系
反应性治理	强调治理模式对突发事件的应对能力与替代方案的设计	敏捷治理意味着治理思路与模式的迅速转变
预见性治理	强调增强技术预见与科技社会共同进步	敏捷治理关注不确定性环境下治理方案与模式设计
适应性治理	建立多中心、嵌套的准自治组织以应对快速变化的不确定环境	敏捷治理包括了治理思路的试探性与弹性
实验性治理	通过学习不断优化迭代政策目标与治理方案	敏捷治理以不确定性环境为起点建立治理体系
探索性治理	通过探索和实验实现政策创新	敏捷治理不过度追求治理模式的创新,而是以敏捷思维为基础应对万变的复杂环境
分布式治理	强调多边治理	敏捷治理包含分布式治理思维
试探性治理	通过循环迭代的方式管理复杂性与不确定性	敏捷治理要求试探性的政策力度与治理手段

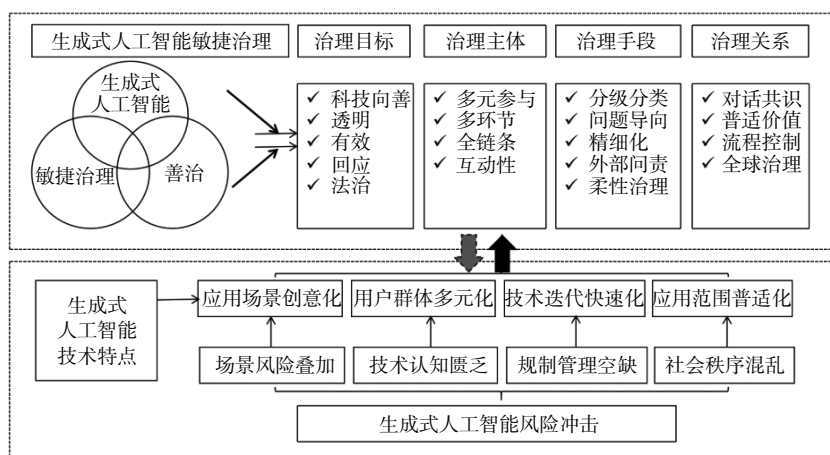


图1 生成式人工智能敏捷治理框架

理目标上,敏捷治理的抽象化治理目标可统筹全局,以适应生成式人工智能应用场景创意化带来的新风险叠加,由此可将善治的合法性、透明性、责任性、法治、回应、有效等要素纳入生成式人工智能治理过程;在治理主体上,敏捷治理下的多元互动的治理主体可前瞻研判人工智能技术发展走向,强调多环节、全链条的多元主体参与;在治理手段上,敏捷治理的柔性治理手段可应对人工智能技术迭代迅速带来的规制管理空缺问题,如吸收分类治理的治理方法,坚持问题导向的精细化监管,采用外部问责的监督模式;在治理关系上,敏捷治理的网络化治理关系呼应了生成式人工智能应用范围扩展所带来的社会秩序扰乱冲击,需坚持全球性对话机制,寻求生成式人工智能治理的普适价值。

4 结论与建议

4.1 结论

本文通过系统性回顾治理理论与新兴技术治理模式,结合人工智能向生成式人工智能迈进的发

展趋势,提出基于敏捷治理理论应对生成式人工智能新特点和不确定性风险冲击的思路。一是系统性回顾了治理理论,比较了新兴技术治理中人工智能治理模式异同,提出敏捷治理模式以应对生成式人工智能的经济社会风险;二是吸收了生成式人工智能治理的欧美经验,对敏捷治理模式应用于生成式人工智能治理进行适用性验证;三是提出了生成式人工智能敏捷治理框架,为中国生成式人工智能治理理念及工具模式提供前瞻性治理参考。

在理论层面,基于敏捷治理的治理模式开展生成式人工智能治理研究,有助于将新兴技术不确定性风险带来的传统治理结构与范式冲击内化于治理模式创新之中,开展算法的前瞻性治理。在实践层面,基于4大治理维度即治理目标、治理主体、治理手段与治理关系,本研究为生成式人工智能治理提供了分析框架,旨在为中国生成式人工智能治理提供研究起点。

4.2 优化中国的生成式人工智能治理的建议

首先,遵循敏捷治理柔性的分

层治理原则,确保治理科学性。第一,瞄准人工智能抽象化治理目标,避免技术无序发展,将技术长期发展目标与社会长期发展目标相统一,依托科技创新增进民众福祉^[22]。第二,借鉴欧美实践经验,利用监管沙盒识别技术前置性风险,确立责任边界和数据治理规范^[23]。确保中国生成式人工智能治理实现刚性与柔性治理协调,把握创新激励和风险防范的平衡,在对企业进行创新激励的同时考虑到技术风险引发的社会失序。此外,在当前阶段数据难以实现高效打通和共享的情境下,可以借助分层治理的思路,构建针对算法黑箱的风险分级响应机制,明确生成式人工智能数据使用者终端责任,并明晰生成内容的权利归属。

其次,构建敏捷治理多主体互动的网络治理关系,提升治理效能。第一,针对人工智能信息和数据交换的实时性、透明性、可共享性、去中心化等特点,治理关系从单向行政命令模式转向去中心化的多中心网络治理模式,提升政策反应的敏捷性。例如,人工智能企业等被监管者可与政府部门等监管者协同,对技术未来发展趋势进行前瞻性研判,从而提升在技术不确定性下的风险应对能力。第二,关注生成式人工智能网络治理关系中权责利的分配和转移,针对生成式人工智能技术和生态中企业、立法机关、政府、公民等利益相关者在认知、功能、资源与能力等方面的差异,应动态调整其在整体网络及各多中心治理单元中的角色定位。如政府可构建全国性人工智能治理监管数据平台、开

发针对公民人工智能侵权自助维权工具、企事业单位开展人工智能赋能计划等。

最后,采取积极主动的治理思路,实现敏捷治理技术善治目标。第一,对于生成式人工智能的治理可借助人工智能技术本身进行治理。传统的治理理论很难适用于新兴产业治理,借助大模型和智能驱动的新型政策工具预测经济与产业发展趋势,运用智能化技术灵活选择治理工具,借助敏捷治理分级、精细、链条式思路不断优化治理路径成为新的治理方向。第二,动态敏捷治理思路的内涵涉及治理边界的模糊化。在生成式人工智能全球治理中,政府的职责不应仅局限于主权边界和法律边界,而应当采取更主动的治理思路,积极参与全球产业发展,寻求人工智能全球化治理中角色地位,在全球治理中实现“善治”。

由于针对生成式人工智能治理的实践和理论研究尚不成熟,未来通用人工智能的技术特性具有极强的不确定性和异质性,仅从针对“弱”人工智能治理的经验中总结,仍存在一定局限性。在研究方向上,首先,现有研究多聚焦于技术特征,未来研究可进一步从企业、产业与国家等不同层面开展跨层级分析,并探索全球跨区域生成式人工智能生态系统的构建路径;其次,可聚焦细分技术模型,探究和比较不同的技术生态系统的治理问题,例如,除Transformer模型外,循环神经网络(RNNs)、词嵌入模型(如Word2Vec、GloVe)、注意力机制(attention mechanism)等也开始有了具体的落

地应用;这些不同的技术对已有的技术系统和生态链进行了重塑,未来研究可针对其技术特点,积极探索技术创新与监管创新的协同共振治理路径。此外,负责任生成式人工智能治理研究是重要议题,深入探讨问责、透明和独立监督条件下的治理模式创新具有重大意义。例如,基于英国研究与创新和物理科学研究委员会(UKRI-EPSRC)提出的AREA(预测—反思—参与—行动)框架研究生成式人工智能在提升社会期望和促进公共福祉的动力作用^[24],进一步探究赋予生成式人工智能人类价值观,发挥学术与科学界在塑造负责任的生成式人工智能系统的核心作用^[25]。

参考文献(References)

- [1] Jovanović M, Campbell M. Generative artificial intelligence: Trends and prospects[J]. *Computer*, 2022, 55(10): 107-112.
- [2] 涂良川. “生成式人工智能”逼近通用智能的哲学叙事: ChatGPT追问智能本质的哲学分析[J]. *东北师大学报(哲学社会科学版)*, 2023(4): 40-47.
- [3] 陈锐, 江奕辉. 生成式AI的治理研究: 以ChatGPT为例[J]. *科学学研究*, 2024, 42(1): 21-30.
- [4] Frey C B, Osborne M A. The future of employment: How susceptible are jobs to computerisation?[J]. *Technological Forecasting and Social Change*, 2017, 114: 254-280.
- [5] Bale M. From crisis to sustainable growth—sub Saharan Africa: A long-term perspective study[EB/OL]. [2025-03-07]. <http://documents.worldbank.org/curated/en/498241468742846138>.
- [6] Rosenau J N. Governance in the twenty-first century[J]. *Global Governance: A Review of Multilateralism and International Organizations*, 1995, 1: 7-40.
- [7] Rhodes R A W. The new governance: Governing without government[J]. *Political Studies*, 1996, 44(4): 652-667.
- [8] Bang H, Esmark A. Good governance in network society: Reconfiguring the political from politics to policy[J]. *Administrative Theory & Praxis*, 2009, 31(1): 7-37.
- [9] Jessop B. Governance and metagovernance: On reflexivity, requisite variety, and requisite irony[C]. Manchester: Manchester University Press, 2003: 142-172.
- [10] 俞可平. 治理和善治引论[J]. *马克思主义与现实*, 1999(5): 37-41.
- [11] 俞可平. 治理和善治: 一种新的政治分析框架[J]. *南京社会科学*, 2001(9): 40-44.
- [12] 杨开峰, 邢小宇, 刘卿斐, 等. 我国治理研究的反思(2007—2018): 概念、理论与方法[J]. *行政论坛*, 2021, 28(1): 119-128.
- [13] Rotolo D, Hicks D, Martin B R. What is an emerging technology?[J]. *Research Policy*, 2015, 44(10): 1827-1843.
- [14] Stilgoe J, Owen R, Macnaghten P. Developing a framework for responsible innovation[J]. *Research Policy*, 2013, 42(9): 1568-1580.
- [15] 张欣. 生成式人工智能的数据风险与治理路径[J]. *法律科学(西北政法大学学报)*, 2023, 41(5): 42-54.
- [16] 曾宇航, 史军. 政府治理中的生成式人工智能: 逻辑理路与风险规制[J]. *中国行政管理*, 2023, 39(9): 90-95.
- [17] O'Shaughnessy M R, Schiff D S, Varshney L R, et al. What governs attitudes toward artificial intelligence adoption and governance?[J]. *Science and Public Policy*, 2023,

- 50(2): 161–176.
- [18] Elmi N, Broekaert K, Larsen A M E, et al. Agile governance: Reimagining policy-making in the fourth industrial revolution[EB/OL]. [2025–03–07]. https://www3.weforum.org/docs/WEF_Agile_Governance_Reimagining_Policy-making_4IR_report.pdf.
- [19] 薛澜, 赵静. 走向敏捷治理: 新兴产业发展与监管模式探究[J]. *中国行政管理*, 2019(8): 28–34.
- [20] 张凌寒, 于琳. 从传统治理到敏捷治理: 生成式人工智能的治理范式革新[J]. *电子政务*, 2023(9): 2–13.
- [21] 薛澜, 贾开, 赵静. 人工智能敏捷治理实践: 分类监管思路与政策工具箱构建[J]. *中国行政管理*, 2024, 40(3): 99–110.
- [22] 张涛, 崔文波, 刘硕, 等. 英国国家数据安全治理: 制度、机构及启示[J]. *信息资源管理学报*, 2022, 12(6): 44–57.
- [23] 吕悦, 陈旭, 彭子璇, 等. 静态到敏捷: 人工智能监管沙盒治理机制研究[J]. *科学学研究*, 2026, 44(5): 910–920.
- [24] Framework for responsible research and innovation[EB/OL]. [2023–11–09]. <https://www.ukri.org/who-we-are/epsrc/our-policies-and-standards/framework-for-responsible-innovation/>.
- [25] Bockting C L, van Dis E A M, van Rooij R, et al. Living guidelines for generative AI: Why scientists must oversee its use[J]. *Nature*, 2023, 622(7984): 693–696.

Governance theory, Euro–American practice and optimization of generative AI from agile perspective

SUN Qian¹, JIA Yingying², CUI Chengyu¹, GUO Dong^{3*}

1. School of Public Policy and Management, University of Chinese Academy of Sciences, Beijing 100190, China

2. School of Public Management, Renmin University of China, Beijing 100872, China

3. School of Economics and Management, Communication University of China, Beijing 100024, China

Abstract As an emerging digital technology, generative artificial intelligence represented by DeepSeek and ChatGPT has given birth to new scenes, new formats and new markets. However, the uncertainty risk of emerging technologies has also brought the impact of traditional governance structure and paradigm, posing new challenges to governance capacity and governance system. Facing the requirements of "good governance" and combining with the theory of agile governance, this paper puts forward the framework of agile governance of generative artificial intelligence, which can provide reference for the future selection of governance concept and tool mode of generative artificial intelligence in China. Based on the theory of governance, this paper systematically analyzes the international governance practices of generative artificial intelligence in the United States and the European Union, and explores the consensus indicators and differences of international governance concepts of artificial intelligence from four aspects: governance objectives, governance subjects, governance means and governance relations. The optimization ideas of the governance of generative artificial intelligence in China were put forward, namely, following the flexible hierarchical governance principle, building a multi-agent interactive network governance relationship, and adopting proactive governance ideas, so as to adapt to the dynamic characteristics and uncertain risks of new technologies, actively promote the healthy and orderly development of generative artificial intelligence technology and industry, and realize the technological goodness.

Keywords generative artificial intelligence; AI risks; digital technology governance; agile governance ●



(责任编辑 王微)