

文章编号:1003-2053(2026)01-0061-14

美国人工智能监管政策的复杂性及临近性研究 ——基于州域层面政策文本的主题模型

杨超, 乔瀚, 李如烟

(中国矿业大学公共管理学院, 江苏徐州 221116)

摘要:捕捉美国人工智能监管动态并探寻中国参与竞争的政策安排是中国参与全球人工智能治理和提升全球科技领导力的关键议题。本文从政策复杂性视角,采用动态主题模型等多种分析技术,深入阐释了美国36个州的397份政策文本,呈现了人工智能监管的州域偏好,并测算了各州与中国人工智能监管的临近度。研究发现:美国人工智能监管政策系统雏形初现,产生了7个政策主题和四大核心政策工具。监管主题分布最丰富的州为加利福尼亚州、伊利诺伊州和纽约州;与中国监管政策最临近的是夏威夷州、俄克拉荷马州、纽约州;加利福尼亚州、马萨诸塞州等科技强州也在临近排名中居前列。最终,面向美国人工智能监管系统复杂性,提出了中国应采取“由近及远”“化整为零”“兼容共生”“群策群力”的政策创新战略。

关键词:人工智能;监管;政策复杂性;政策工具;主题模型

中图分类号:D035;D523

文献标识码:A

人工智能(Artificial Intelligence, AI)引发的新一轮技术变革正在加剧社会转型,但技术的不确定性不仅冲击了现有的法制、伦理乃至社会秩序,而且其颠覆性还超越了国界^[1, 2]。这也激发了全球人工智能治理的热潮。在习近平总书记提出的《全球人工智能治理倡议》、英国等28国发布的《布莱切利宣言》、联合国发布的《为人类治理人工智能》报告等,都强调了监管创新的重要性。美国是人工智能技术强国,也是监管政策数最多的国家^[3],其政策系统有“双轨制”特征。各州政策自主性强,联邦政策对州有指导性影响。尽管新任总统特朗普更重视破除创新的障碍,2025年上任后立即撤销了拜登总统时期《安全、可靠和可信的人工智能开发应用》行政命令,但由于美国三权分立的政治体制,各州的政策已经构成了一个庞杂的政策系统雏形,这种治理模式很难短期发生转变。然而,国内外相关研究多聚焦于国家层面宏观比较,却少有对州域层面的系统性解析。美国州级政策系统雏形呈现出复杂性和碎片化问题^[4],很可能对全球产生重要影响。

推动人工智能“负责任的创新”,亟需提升监管

的敏捷性。诸多监管政策从隐私和数据、透明和可解释性、安全和公平等维度展开设计,重在平衡技术风险和创新之间的关系^[3, 5]。政府的技术监管不仅要防止滥用强权而偏离了国家利益^[6, 7],还要开展政策的“事后评估”,增加多主体参与,提升对新治理场景的适应性^[7, 8]。人工智能监管的复杂性在于技术影响范围大、技术迭代快和应用场景多等方面。传统的“相对静态”的有序单一政策分析很难满足要求,更应选择考虑政府与市场、公众与非营利组织等多种复杂场景的政策分析模式^[9, 10]。政策设计的复杂性会影响决策时长、第三方授权机构的独立性和执行绩效^[11],但也要在监管的同时“不浇灭创新和热情的火焰”^[12, 13]。美国各州自主设计的人工智能监管政策为探讨政策的复杂性设计提供了丰富的样本。中国也需要从国外法规政策、国际规则和本土实践中获取政策创新的思路^[14]。

面对美国人工智能监管政策系统的逐渐形成,如何把握全球人工智能治理规则形塑过程中的政策机遇是当前中国人工智能治理的重要议题。本研究采用隐含狄利克雷分布(Latent Dirichlet Allocation,

收稿日期:2024-11-22;修回日期:2025-02-27

基金项目:国家自然科学基金青年基金项目(72204249);中国矿业大学研究生创新计划项目(2024WLJCRCZL163)

作者简介:杨超(1988-),男,副教授、博士生导师,E-mail:chaoyang@cumt.edu.cn。乔瀚(1999-),男,硕士研究生。李如烟(2002-),女,硕士研究生。

LDA)模型、动态主题模型(Dynamic Topic Modeling, DTM)和内容分析法等技术,结合人工标注、阅读和整理,从系统演进和州域偏好角度展示了美国州级监管政策的复杂性设计,测算了各州与中国人工智能监管的“临近性”,并提出了对中国人工智能监管政策创新的启发性建议,这将为完善中国监管政策系统和企业海外竞争提供参考。

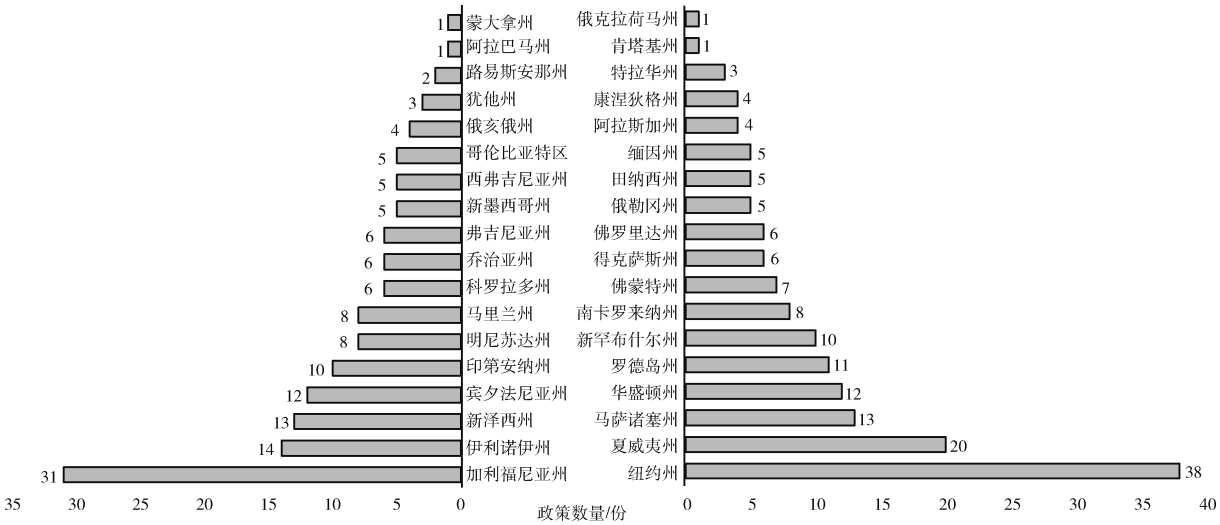
1 庞杂与多变:美国人工智能监管政策系统演进特征

高度复杂的政策环境实际上不利于系统层面的政策实施^[11]。但由于价值、环境和执行差异,初现的美国人工智能监管政策系统呈现出庞杂和多变的特征。

1.1 政策密度区域差异大

本文搜集的 AI 监管政策文本以美国州级层面法规政策数据为基础,政策文本范围主要聚集在 2008 - 2024 年,但政策更多聚集在 2018 年以后。

通过手工整理各州的 AI 监管政策,构建了政策文本数据库,最终从 36 个州搜集到 397 份政策文本。这些政策文本包括已立法实施的政策 140 份和正在提案讨论的政策 257 份,分布情况如图 1 所示。第一,美国各州的 AI 监管政策呈现出从沿海向内陆递减的趋势,产生了纽约州、加利福尼亚州和夏威夷州 3 个政策密集区。纽约州是全球金融中心, AI 在金融领域的应用非常丰富,该州有明显向相邻州的政策扩散效应;加利福尼亚州是全球 AI 技术高地,集聚了大量的 AI 创新资源,是新技术应用的源头;较为有趣的是夏威夷州,该州有丰富的自然资源和发达的旅游业,出台了大量旅游、自然资源监控和灾害预防的 AI 监管政策。第二,截至 2024 年,在美国 50 个行政区中,以加利福尼亚州、纽约州和夏威夷州为代表的 30 个州正式通过了有关人工智能监管的政策;阿拉斯加州、路易斯安那州、阿拉巴马州、南卡罗来纳州、康涅狄格州和肯塔基州等 6 个州的 AI 监管政策则为正在讨论的提案,其余的 15 个行政州尚乏人工智能监管政策。



资料来源:作者根据美国 36 州 397 份政策文本数量分布绘制。

图 1 美国各州 AI 监管政策数量

Figure 1 The number of state - level AI regulatory policies in the United States

1.2 政策主题关联范围广

本文利用了 LDA 主题模型,并结合美国政府对人工智能风险治理的主题标注,绘制了如图 2 所示的 7 个 AI 监管政策的主题词云,依次为:消费者隐私保护;公共舆论和选举;金融科技监管;公私数据治理;咨询委员会建设;自动决策公平性;系统安全有效性。在这 7 个主题中,除了 AI 咨询委员会建设

之外,其余 6 个主题都是 AI 监管政策的热门监管领域。

在此基础上,整理得到如图 3 所示的 AI 监管主题的共性内容。

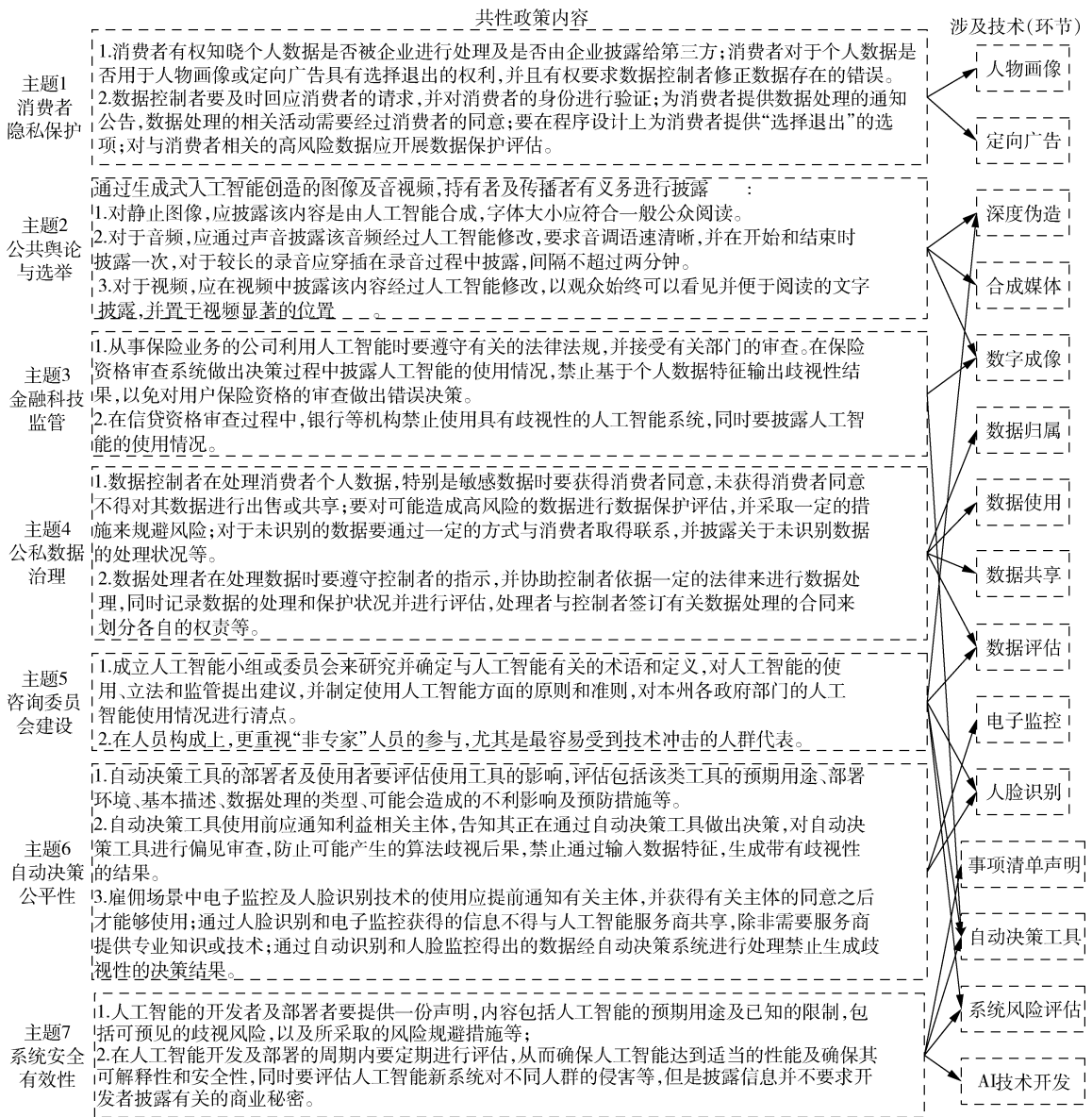
首先,消费者隐私保护是最早开展监管的领域。早在 2008 年,美国就开始对消费者数据隐私保护开展监管,但针对 AI 使用中的消费者隐私数据监管政



资料来源:作者根据美国 36 州的 397 份政策文本绘制。

图 2 美国各州 AI 监管政策主题词云

Figure 2 The word cloud of state – level AI regulatory policies in the United States



资料来源:作者根据美国 36 州 397 份政策文本整理。

图 3 基于 LDA 主题模型的美国 AI 监管政策共性政策内容

Figure 3 Commoncontents of AI regulatory policies in the United States based on LDA Model

策直到 2018 年加利福尼亚州出台后开始在各州扩散。各州重在提升消费者对个人数据披露的知情权、数据控制者与消费者之间的关系界定、数据处理者与数据控制者之间的权责划分等方面。第二,公共舆论和选举是新兴热点监管领域。生成式人工智能不仅可以通过深度伪造视频音频和照片,冒充候选人捏造和发布虚假消息,还会搜集选民投票偏好,产生诱导选民投票等问题。加利福尼亚州、得克萨斯州和犹他州最早于 2019 年出台了监管政策,但这项政策直到 2023 和 2024 年才快速在其他州扩散。第三,金融科技监管集中体现在保险、信贷等业务上,在人工智能监管上更关注自动化决策工具的歧视以及 AI 换脸产生的安全风险问题。该主题政策最早由伊利诺伊州在 2019 年实施。第四,美国对公共和私有数据的治理涉及到数据的归属、使用、共享和评估等多个技术环节,消费者相关的敏感数据应获得个人同意,数据的处理和保护要按法律规范开展评估,以法律法规和合同开展数据治理。加利福尼亚州于 2018 年实施人工智能数据治理的监管,是最早聚焦 AI 领域数据治理的州。第五,各州通过加强建设人工智能咨询委员会,准确地界定人工智能的治理边界,在人员构成中强调“非专家”人员的参与,将被人工智能冲击的人群代表也纳入咨询委员

会。马里兰州和佛蒙特州是最早实施该项政策的州,随后快速扩散到 15 个州。第六,自动决策工具公平性问题也备受重视,使用者应重视自动决策可能产生的不良影响,防止算法歧视。在雇佣场景中,结合电子监控和人脸识别的自动决策要经过同意,并限制共享。纽约州 2021 年开始实施自动决策工具监管,后续扩散到 9 个州。第七,推动人工智能技术的系统安全评估,确保技术使用的可解释性和安全性。该主题较新,20 个州都是在 2023 和 2024 年进入议案讨论或者开始实施。

1.3 政策领域时序变化快

本文进一步采用了动态主题模型刻画出美国各州 AI 监管政策的影响领域、关注焦点和主题时序变化,结果如表 1、表 2 和图 4 所示。首先,金融领域是美国 AI 监管中最早最主要的领域,主要涉及到信贷、保险等方面的技术监管。第二,各州在消费者领域的 AI 监管自 2020 年之后快速增长,已经成为 AI 监管最受关注的领域。第三,尤其值得关注的是,美国各州 2022 - 2024 年最受关注的是选举领域的 AI 监管,选举的词频仅次于消费者居第二位。第四,在健康、就业和教育领域的 AI 监管基本上持平,教育领域的 AI 监管有超越健康和就业的趋势。

表 1 AI 监管政策的影响领域变化 (2008 - 2024)

Table 1 Changes in the impact areas of AI regulatory policies (2008 - 2024)

影响领域	2008	2015	2017	2019	2020	2021	2022	2023	2024
消费者	0	4	0	2	0	32	54	856	954
教育	0	0	0	1	0	2	3	69	115
健康	2	0	0	0	0	4	1	122	99
就业	2	4	0	0	2	25	1	103	105
选举	0	0	0	0	0	0	0	22	194
金融	5	4	1	0	0	3	2	40	61

资料来源:作者根据美国 36 州 397 份政策文本统计。

结合全球人工智能向善峰会 (AI for Good Global Summit)、第 78 届联合国大会人工智能决议、斯坦福人工智能指数报告 (AI Index Report) 等筛选出监管关注焦点词,并绘制了如表 2 所示的美国各州 AI 监管政策中这些词的频率。首先,系统安全 (Security) 是 2008 - 2020 年之间最高的关注焦点词,但在 2022 年以后,隐私 (Privacy) 成为了当前最具关注度的焦点词,伦理安全 (Safety) 2024 年超越系统安全

成为当年排名第二的焦点词。第二,多数政策焦点词都在 2020 年以后才高频出现,排在四至六位的依次是透明性 (Transparency)、深度伪造 (Deepfake) 和歧视 (Discrimination)。第三,较为有趣的是,目前美国各州 AI 监管政策中,当比较透明性和可解释性 (Explainability) 时,多数政策更关注透明性,而可解释性的关注度较低;在比较歧视和公平 (Fairness) 问题时,多数政策更关注歧视,而较少地直接讨论公平。

表 2 监管政策的关注焦点词变化 (2008 – 2024)

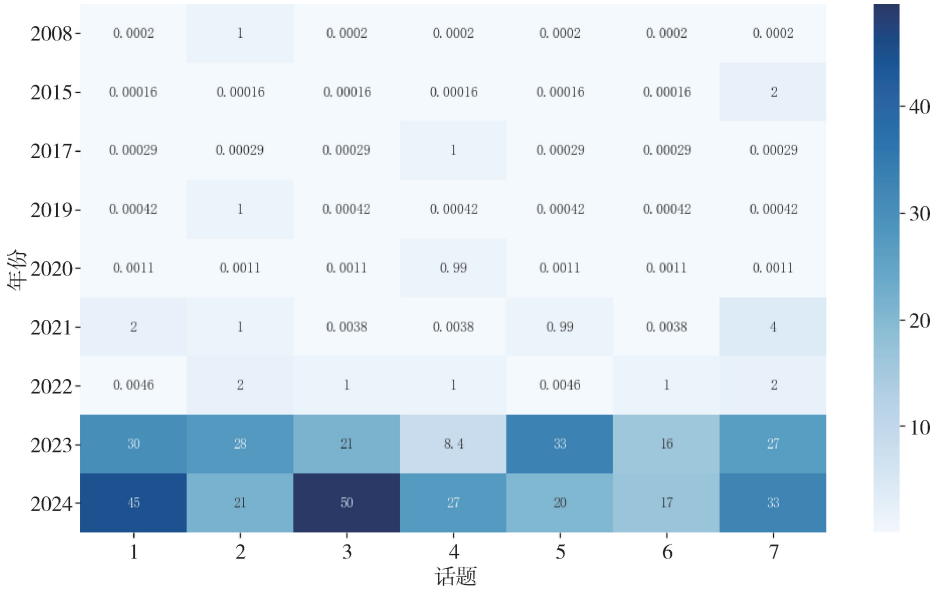
Table 2 Changes of the number of focus words of AI regulatory policy (2008 – 2024)

焦点词	2008	2015	2017	2019	2020	2021	2022	2023	2024
隐私	5	0	1	2	0	4	8	211	190
透明性	0	0	0	0	0	3	2	37	75
可解释性	0	0	0	0	0	0	0	3	5
系统安全	5	2	5	2	0	4	4	148	115
伦理安全	1	2	0	0	0	1	1	64	130
歧视	0	0	0	0	0	8	4	45	50
公平性	0	0	0	0	0	3	0	10	10
深度伪造	0	0	0	0	0	0	0	4	59

资料来源:作者根据美国 36 州 397 份政策文本统计。

图 4 为美国各州的 AI 监管政策主题动态变化结果。首先,2020 年之前的美国 AI 监管基本上是零散出现,聚焦的话题不太稳定,基本上围绕公共舆论、隐私数据、系统安全有效性等相关主题。第二,2021 – 2022 年各州的监管政策开始基本上覆盖 7 个主题领域,并在 2023 – 2024 年实现了完全覆盖。

第三,2023 – 2024 年,各州的 AI 监管政策热度最高的一档是消费者隐私保护、金融科技监管和系统安全有效性;热度居于第二档的主题是 AI 咨询委员会建设、公共舆论和选举;热度居于第三档的主题是公私数据治理和自动决策的公平性。



资料来源:作者根据美国 36 州 397 份政策文本测算。

图 4 AI 监管政策主题的动态变化 (2008 – 2024)

Figure 4 Dynamic changes of AI regulatory policy topics

2 对齐与分散:美国人工智能监管政策州域设计偏好

人工智能监管的目标是实现人工智能对齐 (A-

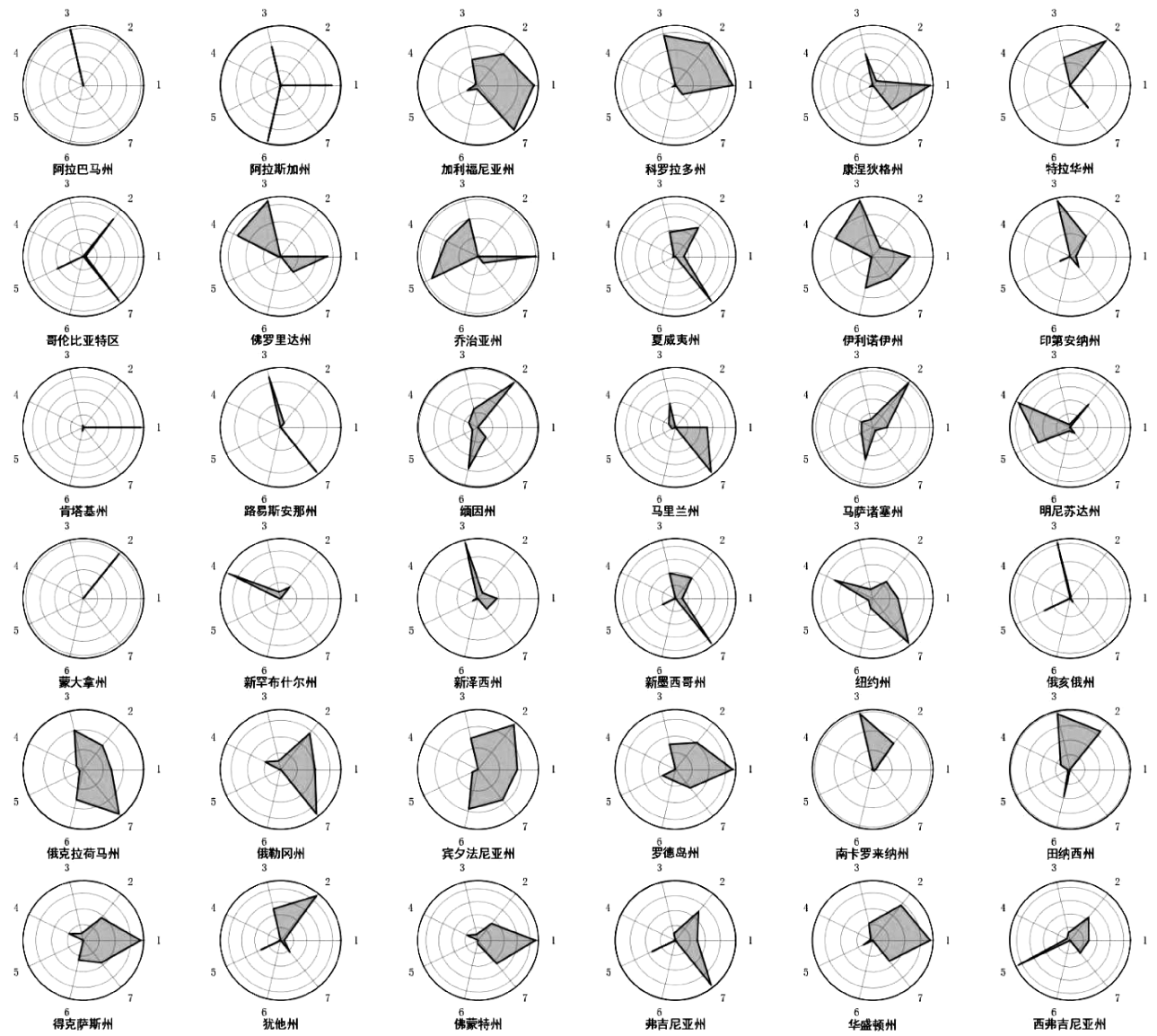
lignment), 技术与人类的价值观和利益相统一, 而现实中的美国各州的 AI 监管政策则非常分散。分散化和差异化的 AI 监管政策使得摸清美国 AI 监管政策信息和推动监管价值对齐的政策学习产生了障碍。这里从各州的监管政策主题、政策工具选择和

监管机构设计上,探讨“分散式”的人工智能监管的对齐思路。

2.1 监管政策主题分散

鉴于政策文本的主题往往存在主题交叉,LDA

主题建模能按照主题比重优先的规则将文本纳入到最佳匹配的主题中,这里将测算出各州在七大主题分布作为主题偏好,如图 5 所示。这能体现各州 AI 监管决策中自发形成的政策偏重。



资料来源:作者基于美国 36 州 397 份政策文本建模后绘制。

图 5 美国各州 AI 监管政策主题分布偏好

Figure 5 Topics distribution of state-level AI regulatory policies in the United States

首先,各州都有明显的监管侧重领域,尚无在七个主题上全部都有得分的州,这体现出美国各州在 AI 监管上更强调在美国联邦政府统筹下的因地制宜决策。例如,在企业监管标准门槛的限定上,即便是俄勒冈州、康涅狄格州和特拉华州都是以企业一年内掌控和处理的个人数据的人数作为标准,但俄勒冈州和康涅狄格州的标准为 10 万人的个人数据,

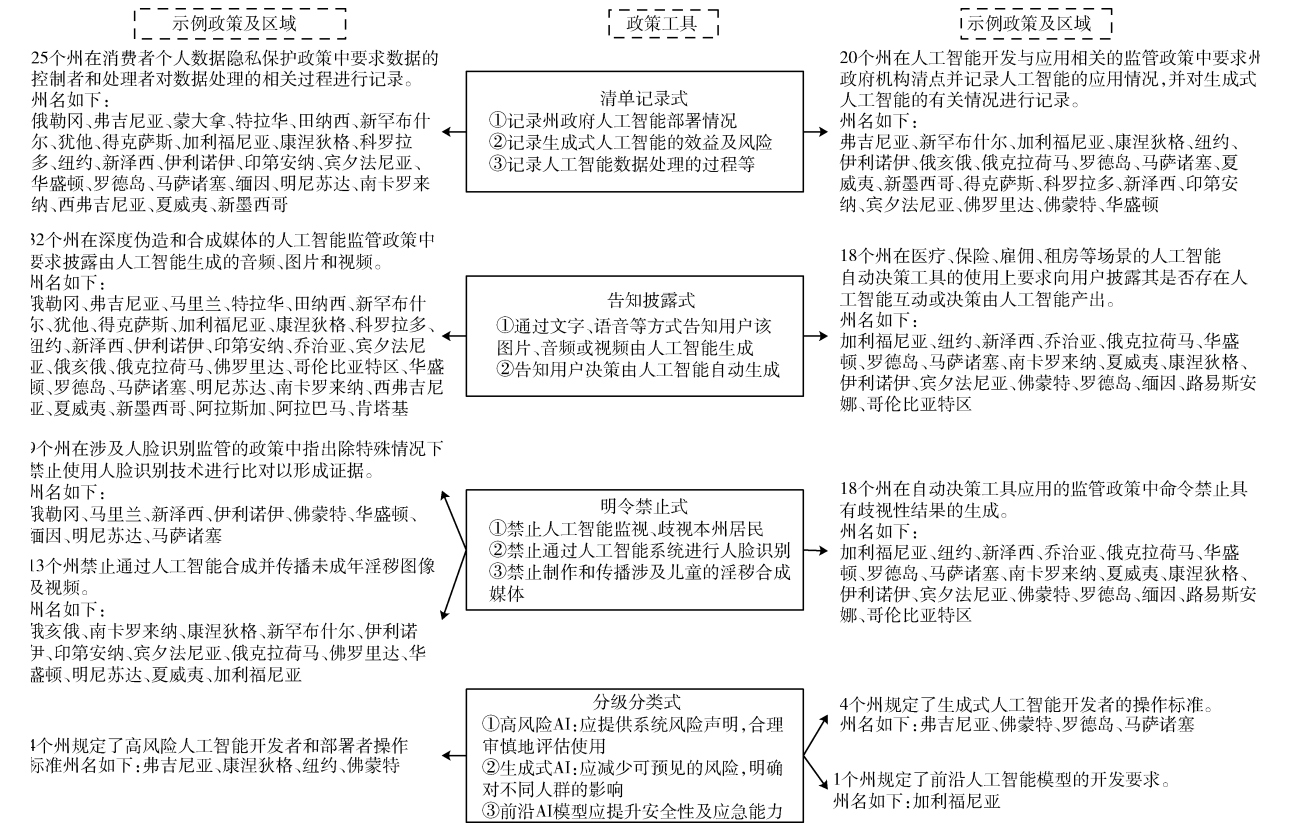
而特拉华州则更为严格地将标准设置为 3.5 万人,且企业还需要达到一定的数据收益标准。第二,从雷达图的面积可知,监管政策主题较为丰富的州包括加利福尼亚州、伊利诺伊州、纽约州、宾夕法尼亚州、俄克拉荷马州、得克萨斯州和华盛顿州,这些再次验证了政策密度分布的数据。在系统安全有效性上,宾夕法尼亚州和路易斯安那州的举措是要求企

业对使用的人工智能模型进行注册和登记,强调对模型系统的安全监管。第三,部分州结合了各州的主要产业开展 AI 监管,仅涉及了少数几个主题。如夏威夷州的 AI 监管政策数量位居全美前三,但都集中在主题 2、3 和 7 上;阿拉巴马州、肯塔基州、阿拉斯加州、哥伦比亚特区、蒙大拿州、俄亥俄州等,这些州涉及的政策主题数也较少,基本都在 3 个主题以内。第四,大多数州在开展 AI 监管中考虑了本地产业的创新发展问题。例如,田纳西州专门针对音乐产业人工智能开展监管,将声音纳入监管范围,规定人工智能合成声音需要获得声音所属者的同意,保

护音乐产业的发展;伊利诺伊州、罗德岛州制定了赌博业的 AI 监管,明令禁止使用自动决策及预测性工具,减少对用户无节制赌博行为的刺激;乔治亚州、宾夕法尼亚州、俄克拉荷马州要求医疗卫生行业和保险业提升行业决策的透明度,降低 AI 决策产生的偏差和歧视。

2.2 监管政策工具对齐

美国各州 AI 监管政策设计主要有 4 种主要的政策工具:清单记录式、告知披露式、明令禁止式和分级分类式。各州 AI 监管政策工具对齐情况如图 6 所示。



资料来源:作者根据美国 36 州 397 份政策内容整理。

图 6 美国各州 AI 政策监管政策工具对齐情况

Figure 6 Alignment of state-level AI regulatory policy tools in the United States

清单记录和告知披露两种软性监管手段是目前美国各州监管政策的主导方式,分级分类和明令禁止式的政策工具也在逐步引入。清单记录式是通过记录和清点人工智能部署、效益、风险及数据处理的过程,从而确保政府对人工智能的潜在风险和责任有充分的信息掌握。告知披露式是通过文字、语音等多种方式要求人工智能企业告知用户潜在的风险和错误。明令禁止式指的是在部分州禁止利用人工

智能监视和歧视,禁止使用人脸识别,禁止制作以未成年人为对象的淫秽视频。分级分类式指的是对人工智能模型的安全风险水平、影响领域进行评估,区别对待高中低风险的 AI 模型实施差异化监管,并针对医疗健康、金融、自动驾驶、舆情与选举等关键领域实施特定的监管。尽管各州在政策工具有各种各样的偏重,但在美国联邦政府的人工智能宏观政策引导下,开始形成了“目标一致、协同合作和有效

沟通”的人工智能监管政策的战略性对齐。

在 4 种政策工具中,分级分类监管是最为热门的政策领域,这也与 2024 年欧盟发布的《人工智能法案》(The EU AI Act)的第三章“高风险的 AI 系统”有一定的相关性。全球常见的治理方式被分为回应式治理、目标式治理和针对性治理,而这三类都

在一定程度上体现了分类的内涵^[15]。美国佛蒙特州、马萨诸塞州、康尼狄格州等地区将 AI 模型分为高风险、生成式人工智能、前沿人工智能模型作分级分类监管,从而在保护人工智能技术创新的同时推动 AI 模型监管的敏捷治理能力。这里整理了这些州的 AI 分级分类标准如表 3 所示。

表 3 分级分类人工智能监管操作标准要求
Table 3 Graded and classified operational standard requirements for AI regulation

义务/责任	高风险人工智能系统		生成式人工智能系统		前沿人工智能
	开发者	部署者	开发者	部署者	开发者
透明义务	应在网站披露系统的风险信息;提供数据处理的声明;披露风险管理和减少歧视的措施	向用户提供一份通俗易懂的有关人工智能系统的声明	提出合适的数 据治理措施	模型编程要具有独特水印以防止伪造	对所有的数据测试进行相关的记录与留存;起草并保留安全协议、测试程序,定期进行审查,在特定网站公示
	提供人工智能系统的评估文件;提供帮助部署者了解系统输出和系统风险的文件	采取措施保护用户免受算法歧视风险;实施并维护风险管理政策及程序;定期审查和更新人工智能系统	制定减少可预见风险的措施;在系统运行周期内定期进行评估以确保有效性	不得基于特征对个人或群体输出歧视性结果;制定安全措施以保护相关数据;定期进行风险评估	防止未经授权访问和获取模型;具有完全关闭模型的能力;采取合理措施防止模型伴随的风险;对所有的数据测试进行相关的记录与留存

资料来源:作者根据美国 36 州 397 份政策内容整理。

2.3 监管机构配置偏好

面对人工智能技术的不确定性,建立一个“专班制”的人工智能监管咨询机构,是各州普遍采用的 AI 监管方式。这类咨询委员会不仅要考虑技术监管的问题,还得分析人工智能监管和地方经济社会之间的关联,尤其是要评估对就业、教育、医疗卫生等重要领域的影响,主要服务于地方人工智能政策改革和人工智能项目的评估。美国各州在人工智能监管咨询委员会的建设上,主要采用多主体、多行业、多中心治理的团队建设模式。在一个州的 AI 监管专家团队中,常见的代表包括以州长、参众议院院长为代表的政府管理者,以教育界、法律界、行业协会为代表的领域精英,在部分团队中还考虑了特殊人群代表。以华盛顿州人工智能咨询小组为例,在常见的官方代表之外,增加了三位易受到 AI 算法伤害的群体代表,比如少数民族代表、特殊性别群体代表、零售业代表和酒店行业代表等。这种设计将落实了将公众纳入人工智能治理决策的国家战略^[16]。通过推动更广泛的监管参与,增强了决策咨询的全面性,能够帮助政策制定者获取人工智能安全风险的真实反馈,从而推动人工智能的向善。

在此基础上,美国各州的人工智能监管委员会产生了两种主要的分类:第一类是综合型委员会。例如,在俄勒冈州、马里兰州、印第安纳州成立的 AI 咨询工作组就是典型的综合型委员会,处理与人工智能相关的医疗、教育、安全等各方面的咨询事项,是一种较为全面的专家咨询团队类型。第二类是专业型委员会。一些州结合特定的领域建立了 AI 咨询委员会。如新墨西哥州成立了自动决策工具的专业型咨询专家组,纽约州成立了人工智能、机器人与自动化的专家组,加利福尼亚州成立了深度伪造问题咨询专家组和算法创新办公室,西弗吉尼亚州成立了人工智能和自然语言处理工作组。这也反映出美国不同州在设计监管机构时会考虑该州的人工智能监管焦点。监管机构专家团队的配置偏好差异化,是各州推动人工智能价值对齐的政策设计策略。

3 差异与距离:美国各州与中国人工智能监管的临近性

本文进一步呈现了中美 AI 监管政策主题的差异,并测算了美国各州与中国 AI 监管政策的临

近性。

3.1 中美 AI 监管政策主题的差异测算

中国的 AI 监管已经形成了较为全面的治理规范和体系,主要针对新一代人工智能可能引发的伦理问题形成了较为丰富的政策内容。中国的 AI 监

管政策是自上而下的,主要由国务院、国家新一代人工智能专业委员会、国家互联网信息办公室发布。首先,选取了 2017 – 2023 年中国发布的 AI 监管政策作为参考文本,如表 4 所示。筛选时重点考虑了 AI 监管政策,并将这些中文文本全文英译后使用。

表 4 中国“AI 治理”主要政策文本
Table 4 Key policy texts on "AI Governance" in China

年份	发布单位	政策标题
2017	国务院	新一代人工智能发展规划
2019	国家新一代人工智能治理专业委员会	新一代人工智能治理原则
2020	国家标准化管理委员会	国家新一代人工智能标准体系建设指南
2021	国家新一代人工智能治理专业委员会	新一代人工智能伦理规范
2022	国家互联网信息办公室	互联网信息服务算法推荐管理规定
2022	国务院	互联网信息服务深度合成管理规定
2023	国务院	生成式人工智能服务管理

资料来源:作者根据中国政府发布的政策整理。

第二,使用信息熵 KL 散度 (Kullback – Leibler divergence) 测算美国州级政策文本的 7 个监管主题与这 7 份典型中国 AI 治理政策论述的信息差异,公式如下。

$$D_{KL}(P||Q) = \sum_{\omega} P(\omega) \log \frac{P(\omega)}{Q(\omega)} \tag{1}$$

其中, $P(\omega)$ 是 ω 词在参考文本的概率分布,即 ω 词在中国的 AI 监管政策文本中的概率分布; $Q(\omega)$ 是 ω 词在评估文本的概率分布,即 ω 词在美国各州政策文本的概率分布。

监管距离是美国各州政策文本的 7 个主题与中国 AI 监管政策的单词概率分布差异的均值。一个主题的监管距离的值越大,则它在参考文本和评估文本概率分布差异越大。结果表明:首先,美国各州产生的 7 个 AI 监管主题与中国 AI 监管的距离从远到近及距离值情况为:咨询委员会建设(主题 5, 3.82)、消费者隐私保护(主题 1, 3.60)、数据治理(主题 4, 3.46)、自动决策公平性(主题 6, 3.43)、公共舆论与选举(主题 2, 3.41)、金融科技监管(主题 3, 3.36)和系统安全有效性(主题 7, 3.21)。第二,美国各州的 AI 咨询委员会建设更加注重多元化的专家配置,参与 AI 监管的人员类别更加丰富,增加了技术监管的视角。相比之下,中国的 AI 监管咨询专家委员会配置更强调专业化,非专家人员的参与度较低。第三,系统安全有效性是中美两国在 AI 监

管上信息分布最相似的主题,这反映出两国在推动 AI 监管存在价值共通性,都强调透明、安全、可解释等价值理念。第四,金融科技监管、数据治理和舆论治理等主题仍然存在较大差异。例如,美国各州对人脸识别技术的监管更为严格,要求在日常应用场景中,人脸识别和电子监控获得的信息不得与人工智能的服务商共享数据;又如,美国要求对深度伪造技术合成的音频在开始和结束时必须披露其使用过人工智能修改;而这些方面,中国的监管要求则相对宽松。

3.2 美国各州与中国 AI 监管的临近性

本文进一步测算美国各州与中国 AI 监管的临近性。监管临近性主要考虑了美国各州与中国 AI 监管政策的主题分布及监管距离。临近性数值越高则说明该州与中国的 AI 监管政策的主题分布和文本相似度越高。由于各州的主题分布和各主题的平均 KL 散度的测度方向相反,因此,这里采用主题分布与平均 KL 散度的商,并引入了放大因子 k (取值为 10),最终得到了各州的主题临近度 $C_{i,t}$,如公式 (2) 所示。

$$C_{i,t} = k \cdot \frac{P_{i,t}}{Q_t} \cdot KL_{i,t} \tag{2}$$

其中, i 为州, t 为各个主题, $P_{i,t}$ 是州 i 在主题 t 上的概率分布; Q_t 是主题 t 的平均分布; $KL_{i,t}$ 是州 i 和参考文本分布 Q_t 的 KL 散度; k 为放大因子。

在计算每个州的整体临近度 C_i 时,对所有主题的临近度加总后,引入了权重 w_i , w_i 表示各州的非零主题数与总主题数之比,从而降低了因该州主题数少却导致临近性过高的问题,提升各州临近度分析的合理性。 C_i 计算公式如下。

$$C_i = \frac{\sum_{t \in T_i} C_{i,t} \cdot w_i}{N_t}$$

(3)

其中, i 为州, t 为各个主题, $\sum_{t \in T_i} C_{i,t}$ 是州 i 所有非零主题的临近度加总, $w_i = \frac{n_i}{N}$ 为权重, n_i 是州 i 的非零主题数, N 为总主题数。

最终得到如表 5 所示的整体临近性测算结果。由结果可知:首先,夏威夷州、俄克拉荷马州、纽约州不仅有较为丰富的政策主题分布,而且是与中国的

监管临近性水平较高的 3 个州。第二,与中国 AI 监管距离最远的 3 个州为蒙大拿州、阿拉巴马州和肯塔基州,它们目前的 AI 监管政策不太完善,政策出台数量少,且比较聚焦于少数几个主题。第三,加利福尼亚州通常被认为是美国的 AI 产业创新最强的州,集聚了谷歌、苹果、特斯拉等巨型 AI 企业和斯坦福大学、加州大学伯克利分校等顶尖高校,也是硅谷的所在州。该州与中国 AI 监管的临近度为 2.927,远高于平均值 2.166,在最接近中国的 AI 监管州排名中位居第六。第四,马萨诸塞州是美国人工智能产业创新发展的科研强州,拥有哈佛大学和麻省理工学院等顶尖高校,该州在人工智能监管政策上与中国 AI 监管的临近性紧随加利福尼亚州,位居第七。

表 5 美国各州与中国 AI 治理政策文本的“监管临近性”

Table 3 The regulatory proximity of AI governance policies between China and states of the U. S.

州名	主题 1	主题 2	主题 3	主题 4	主题 5	主题 6	主题 7	临近度
阿拉巴马州	0.000	0.000	2.975	0.000	0.000	0.000	0.000	0.425
阿拉斯加州	0.954	0.001	0.801	0.001	0.001	1.125	0.002	1.648
加利福尼亚州	0.782	0.598	0.397	0.036	0.159	0.055	0.900	2.927
科罗拉多州	0.900	0.887	0.863	0.000	0.050	0.000	0.201	2.072
康涅狄格州	1.230	0.127	0.744	0.000	0.071	0.000	0.738	2.078
特拉华州	0.000	1.466	0.744	0.000	0.000	0.000	0.777	1.280
哥伦比亚特区	0.050	1.031	0.000	0.000	0.555	0.000	1.301	1.678
佛罗里达州	0.749	0.000	0.981	0.795	0.027	0.000	0.358	2.079
乔治亚州	0.835	0.000	0.603	0.532	0.701	0.000	0.139	2.008
夏威夷州	0.179	0.815	0.561	0.042	0.056	0.004	1.342	3.000
伊利诺伊州	0.492	0.163	0.815	0.568	0.018	0.453	0.419	2.927
印第安纳州	0.141	0.658	1.496	0.000	0.253	0.006	0.383	2.518
肯塔基州	2.519	0.001	0.102	0.001	0.000	0.172	0.001	1.198
路易斯安那州	0.001	0.140	1.346	0.001	0.000	0.000	1.552	1.303
缅因州	0.000	1.141	0.380	0.197	0.109	0.843	0.260	2.512
马里兰州	0.708	0.000	0.590	0.167	0.074	0.000	1.434	2.123
马萨诸塞州	0.278	1.189	0.171	0.251	0.228	0.684	0.088	2.889
明尼苏达州	0.000	0.641	0.062	1.213	0.678	0.079	0.168	2.436
蒙大拿州	0.000	2.931	0.000	0.000	0.000	0.000	0.000	0.419
新罕布什尔州	0.000	0.533	0.250	2.084	0.031	0.000	0.000	1.657
新泽西州	0.503	0.199	1.632	0.031	0.147	0.008	0.415	2.934
新墨西哥州	0.145	0.584	0.595	0.000	0.287	0.000	1.364	2.126
纽约州	0.410	0.373	0.164	0.716	0.071	0.167	1.050	2.952
俄亥俄州	0.000	0.056	1.852	0.000	0.807	0.000	0.156	1.641

续表 5

州名	主题 1	主题 2	主题 3	主题 4	主题 5	主题 6	主题 7	临近度
俄克拉荷马州	0.391	0.452	0.600	0.111	0.052	0.457	0.897	2.960
俄勒冈州	0.579	0.829	0.158	0.296	0.014	0.000	1.083	2.536
宾夕法尼亚州	0.506	0.781	0.446	0.000	0.075	0.553	0.569	2.512
罗德岛州	0.952	0.605	0.461	0.000	0.226	0.196	0.443	2.472
南卡罗来纳州	0.084	1.013	1.766	0.000	0.025	0.000	0.066	2.110
田纳西州	0.001	0.974	1.161	0.201	0.041	0.555	0.004	2.526
得克萨斯州	1.008	0.539	0.136	0.281	0.001	0.373	0.567	2.490
犹他州	0.073	1.297	0.747	0.000	0.434	0.000	0.357	2.077
佛蒙特州	1.227	0.490	0.140	0.265	0.000	0.067	0.715	2.489
弗吉尼亚州	0.408	0.733	0.140	0.000	0.452	0.000	1.191	2.090
华盛顿州	1.001	0.834	0.327	0.040	0.170	0.000	0.519	2.478
西弗吉尼亚州	0.390	0.652	0.176	0.080	1.122	0.000	0.379	2.399

资料来源:作者根据美国和中国政策文本建模测算。

4 结论与建议:美国人工智能监管系统的政策创新启发

人工智能善治,亟需推动理念开放、主体多元、对象分层和工具灵活的治理思路^[17]。中美两国在人工智能技术领域都表现卓越,但美国近年来将中国视为全球科技竞争中关键对手。在兰德公司发布的《2017 – 2024 年美国针对中国的经济政策有效性》报告中向美国政府提出建议:“应增加对中国出口管制和投资提案审查的经费支持,拒绝政府资助建立的研究实验室等在应用型技术领域开展正式交流,应联合盟国对华实施出口管制政策的协调性”。中美两国人工智能竞争正在重塑国际权力结构,但西方主导的治理观念占据了国际话语权优势^[18]。本文从政策的复杂性视角,采用 LDA 和动态主题模型等研究方法,深入解读了美国 36 个州的 397 份政策文本,呈现了美国人工智能监管系统监管州域偏好,并测算了各州与中国人工智能监管的临近度。研究表明:美国人工智能监管政策系统雏形出现,产生了 7 个监管主题:消费者隐私保护、公共舆论和选举、金融科技监管、公私数据治理、咨询委员会建设、自动决策公平性和系统安全有效性,其中系统安全有效性是中美 AI 监管共识性最高的主题;形成了四大核心政策工具:清单记录式、告知披露式、明令禁止式和分级分类式;监管政策主题分布最丰富的州为加利福尼亚州、伊利诺伊州和纽约州;夏威夷州、

俄克拉荷马州、纽约州与中国 AI 监管政策的临近度最高,加利福尼亚州、马萨诸塞州等 AI 科技强州在临近排名上也居前列;各州人工智能监管委员会既有综合型又有专业型,重视“非专家”群体的参与。基于此,进一步提出美国复杂人工智能监管系统对中国政策创新的启发性建议。

4.1 由近及远,丰富监管系统的场景层次

美国的 AI 监管已经初步形成了联邦和州两级的复杂政策系统,而各州的政策差异既有共通性,又有较大差异。中国的 AI 监管系统显然不同,更强调从宏观层面的统一监管。采用“由近及远”的创新策略,丰富中国人工智能监管系统的场景和层次。第一,以场景相近为突破口,因地制宜地提升中国人工智能监管的省域场景契合度和区域敏捷性。例如,夏威夷州在 AI 监管上出台了以旅游、自然资源监管和灾害防治为场景的监管政策。中国的海南省与该州的区域场景相似度很高,应推动在海南建立“AI 旅游资源监管试验区”,尝试探讨中美人工智能监管的“对话协作区”;第二,以区域技术水平相近为切入点,鼓励中国的 AI 科技强省参考美国的科技强州的监管政策经验,建立国际区域间的“监管互动”基础。尽管中美在人工智能领域竞争激烈,美国对华的人工智能出口管制不断升级,但监管政策上仍然存在互动交流的空间。加利福尼亚州作为美国的硅谷,在整体的临近排名中也位居第六,广东依托粤港澳大湾区 AI 科技发展水平上与加州的旧金山湾区有很大的相关度,应推动粤港澳大湾区建立

“AI 全场景监管试验湾”,开展各种监管场景试验,建设大湾区国际监管接口。第三,应推动中国 AI 监管政策系统层次建设,激活省市地方政府的政策创新能力,以探索地方性 AI 监管标准为入口,推动国际认可的 AI 监管标准,提升中国技术的安全可信度,发掘国际技术竞争的政策优势。目前中国在 AI 监管领域的政策文本数量不多,涉及范围还不够广,在很多具体的场景中缺乏政策安排,在对外的系统性对话和宣传上缺乏更详细的政策和案例支撑。中国应加快政策研讨的步伐,发布更多契合国家安全与地方实际的政策,建立具有可比性的政策系统,在确保技术向善的目标下激发技术创新。

4.2 化整为零,完善监管政策的主题细节

美国的 AI 监管政策系统的复杂性体现在区域分布差异大,主题的关联范围广,政策变化的速度快和政策扩散的模式多样。但值得关注的是,美国 AI 监管政策已经形成了 7 个较为明确的主题,各州政策在相近主题上也有一定比例的共性条例。这些主题构成了美国州级政府对 AI 监管的核心关切。首先,以提升系统安全有效性及其国际可信度为目标,中国应加强在美国政府 AI 监管核心议题上的政策建设,充分利用联合国国际电信联盟等国际组织,增强中国参与全球 AI 治理的政策话语权。第二,自动决策公平性和数据治理是人工智能治理的底层基石。中国的 AI 监管决策部门应加强对美国 AI “政策大州”的跟踪关注。伊利诺伊州在这两个方面最为典型。单论公私数据治理,则还可以关注明尼苏达州、佛罗里达州、新罕布什尔州、乔治亚州等。单论自动决策公平性监管,宾夕法尼亚州、缅因州和马萨诸塞州也可作为政策学习重点。第三,消费者隐私保护、公共舆论与选举、金融科技监管主要应对的是人工智能在商业、政治和金融等的伦理和安全问题。个人消费者数据保护是目前差异较大的领域,中国应增强在数据保护领域的合规认证,助力中国企业海外投资的合规审查,降低发达国家借助隐私数据保护审查限制中国 AI 技术发展的几率。在公共舆论等领域的监管上,中国应以共同监管目标为方向,制定 AI 技术的对外宣传战略。金融领域是监管临近度高的专业领域,应优先加强该领域的监管政策合作与交流,应推动中国人民银行、国家金融监督管理总局、中国证监会与美联储等加强在金融科技领域的监管联络。中国的 AI 监管必须化整为零,注重 AI 监管的顶层设计的同时,强化细分领域的政

策积累,完善监管主题和政策交流方案。

4.3 兼容共生,引导政策工具的价值对齐

AI 监管的复杂性的一个重要原因在于“不确定性”,在美国各州的 AI 监管的政策内容中已经凸现了“四大核心政策工具”:清单记录式、告知披露式、明令禁止式、分级分类式。尽管中美在人工智能法规政策的制度系统差异很大,但中国应考虑在已有的政策工具框架下开展政策创新,提升 AI 风险和伦理监管的国际兼容性,释放监管政策的创新动能。首先,中国应学习美国各州的 AI 监管清单,结合中国各省的监管场景,加快本地监管政策的“列清单”工作,形成清晰的政策体系架构,让技术创新敢于安全地释放动能。第二,中国应进一步明确人工智能合规审查和安全评估的披露方式和整改要求,为中国人工智能企业和使用人工智能技术的企业“出海”发展提供政策依据和标准,也为国外企业来华技术投资提供准入的机会。第三,中国应加快完善“告知披露式”和“明令禁止式”的政策工具,促进中国的 AI 监管与不同类型的价值观、道德标准和社会利益兼容共存。美国已有 32 个州明确在深度伪造和合成媒体监管上采用告知披露作为政策工具。这说明告知披露工具已经被广泛使用。第四,中国应加快推进具有系统性和激励性的分级分类监管工作,建立能够与美国、欧盟等相对应的分级分类政策。尽管美国目前在分级分类监管上也处于起步阶段,仅有加利福尼亚、弗吉尼亚、纽约州、马萨诸塞州等少数州正在开展这项工作,但已经有对高风险人工智能、生成式人工智能和前沿人工智能模型的具体分类方式。中国应以“安全向善和激励创新”为目标,建立适应中国人工智能技术的政策工具集,以“一带一路”倡议为抓手,提升政策工具价值的国际认可度。

4.4 群策群力,提升监管决策的敏捷效能

美国在州级人工智能咨询委员会建设上不仅强调多主体参与,提升监管政策的参与度,部分州还建设了人工智能专业细分领域的咨询委员会,增强咨询和监管决策的专业性。中国的人工智能治理应从监管决策参与和专业化角度,提升中国 AI 监管决策的敏捷效能。一方面,中国应建立场景型 AI 咨询机构,增强多元主体参与,构建更便捷高效地建言机制,提升人工智能相关的科技专家、社会专家和各界人士的决策参与能力。在医疗、教育、交通等人工智能安全影响最为显著的领域,中国相关省市应结合

本省实际,参考俄勒冈州、马里兰州、印第安纳州等在建设人工智能咨询团队的做法,筹建省市层面的监管决策小组,服务于人工智能企业的创新试验及影响评估。另一方面,中国应在北京、上海、江苏、浙江等人工智能技术发展领先的省市建立专业技术型人工智能咨询团队,探索建立人工智能细分领域的监管咨询机构,人员配备可以关注美国部分州的做法。如,建立机器人与自动化领域的专家团队,可参考美国纽约州的咨询团队人员安排;建设深度伪造和算法创新领域的专家团队,可参考加利福尼亚州的做法。场景性决策和专业技术型决策咨询团队应保持较强的跨部门协作能力和迭代更新能力,形成“技术应用试点—治理决策反馈—技术应用推广”的多部门协作链,鼓励在科技部门、司法部门、网信办等人工智能紧密相关的部门建立工作小组,负责稳定的团队运营,其他政府部门按需参与,增强中国监管决策的敏捷效能。总体而言,中美人工智能领域的竞争不单是技术竞争,更是政府监管决策敏捷效能的竞争。面对高度复杂和不确定的美国监管政策系统,应取长补短,加强政策学习,并发掘更符合中国技术发展的政策创新。

参考文献:

- [1] Kuziemski M, Misuraca G. AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings[J]. Telecommunications Policy, 2020,44(6): 101976.
- [2] Larsen A G, Følstad A. The impact of chatbots on public service provision: A qualitative interview study with citizens and public service providers[J]. Government Information Quarterly, 2024,41(2): 101927.
- [3] AI Index Report 2024 - Artificial Intelligence Index [EB/OL]. <https://aiindex.stanford.edu/report/>. 2024-08-16.
- [4] 王彦雨,李正风,高芳. 欧美人工智能治理模式比较研究[J]. 科学学研究, 2024,42(3): 460-468. Wang Y Y, Li Z F, Gao F. A comparative study of European and American AI governance models[J]. Studies in Science of Science, 2024,42(3): 460-468.
- [5] 朱旭峰,楼闻佳. 发展还是监管? 人工智能政策的国际比较研究[J]. 学海, 2024(4): 76-85. Zhu X F, Lou W J. Development or regulation? An international comparative study of artificial intelligence policies[J]. Academia Bimestris, 2024(4): 76-85.
- [6] Ulnicane I, Knight W, Leach T, et al. Framing govern-
- ance for a contested emerging technology: Insights from AI policy[J]. Policy and Society, 2021,40(2): 158-177.
- [7] Calandrillo S. Responsible regulation: A sensible cost-benefit, risk vs. risk approach to federal health and safety regulation[J]. Boston University Law Review, 2001,81(5): 957-1032.
- [8] De R, Pal A, Pandey N. Responsible regulation for digital services in India[J]. Journal of Information Technology Case and Application Research, 2024,26(4): 285-295.
- [9] 田昊,李娉. 中国情境下公共政策执行主要问题的转变及其制度分析[J]. 中国行政管理, 2019(5): 77-83. Tian H, Li P. The transition of main issues of public policy implementation in China and its institutional analysis[J]. Chinese Public Administration, 2019(5): 77-83.
- [10] 郑石明,李红霞. 复杂性公共政策: 范式变迁、研究框架与展望[J]. 政治学研究, 2024(2): 180-193. Zheng S M, Li H X. Complex public policy: Paradigm change, research framework and prospect[J]. CASS Journal of Political Science, 2024(2): 180-193.
- [11] Haag M, Hurka S, Kaplaner C. Policy complexity and implementation performance in the European Union[J]. Regulation & Governance, 2025,19(3): 656-674.
- [12] Hurka S, Haag M. Policy complexity and legislative duration in the European Union[J]. European Union Politics, 2020,21(1): 87-108.
- [13] Tsang J. Coping with policy complexity in the globalized world[J]. Public Administration Review, 2016, 76(6): 840-841.
- [14] 张凌寒. 中国需要一部怎样的《人工智能法》? ——中国人工智能立法的基本逻辑与制度架构[J]. 法律科学(西北政法大學学报), 2024,42(3): 3-17. Zhang L H. What kind of artificial intelligence law does China need? Basic logic and institutional architecture of China's artificial intelligence legislation[J]. Science of Law(Journal of Northwest University of Political Science and Law), 2024,42(3): 3-17.
- [15] 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. Xue L, Jia K, Zhao J. Agile governance practices in artificial intelligence: Categorizing regulatory approaches and constructing a policy toolbox[J]. Chinese Public Administration, 2024,40(3): 99-110.
- [16] Wilson C. Public engagement and AI: A values analysis

- of national strategies[J]. *Government Information Quarterly*, 2022, 39(1): 101652.
- [17] 梁正. 新兴技术视域下的人工智能善治——发展现状、风险挑战与治理思路[J]. *人民论坛·学术前沿*, 2024(14): 12–22. Liang Z. Good governance of artificial intelligence from the perspective of emerging technologies: Development status, risks & challenges and governance ideas[J]. *Frontiers*, 2024(14): 12–22.
- [18] 高奇琦. 全球善智与全球合智: 人工智能全球治理的未来[J]. *世界经济与政治*, 2019(7): 24–48. Gao Q Q. Global good intelligence and collective intelligence: The normative goal of artificial intelligence in global governance[J]. *World Economics and Politics*, 2019(7): 24–48.

The complexity of U. S. Artificial Intelligence regulatory policy and Sino – U. S. regulation policy similarity: Based on topic model analysis of state – level policy texts of the U. S.

YANG Chao, QIAO Han, LI Ru – yan

(School of Public Policy and Management, China University of Mining and Technology, Xuzhou 221116, China)

Abstract: To understand the dynamics of artificial intelligence (AI) regulation in the United States and explore policy arrangements in the context of Sino – U. S. AI competition is crucial for China’s participation in global AI governance and its pursuit of global technological leadership. However, existing research has predominantly focused on macro – level comparative analyses of national AI policies, with limited exploration into state – level AI policy comparisons. The U. S. regulatory policy framework, shaped by its federal political system, exhibits significant complexity and fragmentation. This unique structure offers valuable insights for policymakers in other countries navigating the complex and rapidly changing artificial intelligence regulatory policy development settings. Using Latent Dirichlet Allocation (LDA), Dynamic Topic Modeling (DTM), and content analysis, this study conducts an in – depth examination of 397 policy texts from 36 U. S. states, illustrating state – level regulatory preferences, policy impact areas, and shifts in policy focus over time. Along with the aforementioned, the study also compares the regulatory proximity of China’s artificial intelligence policy to that of the United States at the state level. The results show that the U. S. has established a foundational AI regulatory policy system, characterized by seven policy topics and four core policy instruments. The seven policy topics include consumer privacy protection, public opinion and election, fintech regulation, public – private data governance, advisory committee building, automated decision – making fairness and system security and effectiveness. Among these, system security and effectiveness emerge as the most aligned topic between Chinese and U. S. AI regulatory frameworks. Beyond the seven policy topics, four key policy tools were adopted by states in the U. S., including checklist documentation, mandatory disclosure requirements, prohibition and risk classification. Artificial intelligence regulatory committees in U. S. states fall into two categories: comprehensive committees, like those in Oregon, Maryland, and Indiana, addressing broad AI issues; and specialized committees, such as New Mexico’s group on automated decision – making and New York’s panel on AI, robotics, and automation. The study also identifies that states such as California, Illinois, and New York exhibit the most diverse distribution of policy topics. In terms of regulatory proximity to China, Hawaii, Oklahoma, and New York rank highest, with technologically advanced states like California and Massachusetts also featuring prominently. Finally, drawing on the complexity of the artificial intelligence regulatory policy system in the state of the U. S. and the need for artificial intelligence development and regulation, it is recommended that China adopt policy innovation approaches that are tailored to regulatory scenarios and enrich the scenario hierarchy of the regulatory system, deconstruct the entire system into components and improve the details of the subject of regulatory policy, prioritize compatibility and symbiosis to align the value of policy tools, and emphasize collaborative efforts to enhance agile effectiveness in artificial intelligence regulatory decision – making. The study presents the overall landscape of artificial intelligence governance arrangements in the United States and measures the proximity of Sino – US regulation policies. This could serve as a vital reference for Chinese policymakers, providing evidence – based recommendations to guide the innovation and refinement of AI regulatory policies. It is critical for aligning China’s regulatory strategies with global best practices, so increasing its ability to reduce risks, encourage ethical AI development, and boost its leadership in the global AI governance environment.

Key words: artificial intelligence; regulatory; policy complexity; policy instrument; topic model