

# 论人工智能监管<sup>\*</sup>

——国家—市场关系视角下的人工智能技术权力

余南平 栾心蔚

**【内容提要】** 当前世界正面临人工智能发展与应用的转折点，人工智能技术以前所未有的速度和方式迅猛发展，其复杂性、多变性与监管滞后性之间的矛盾愈加尖锐。在大国战略竞争加剧的背景下，人工智能监管呈现独特的权力互动结构和演化轨迹。在国内层面，西方科技公司正凭借自身在技术、经济、社会与政治四个层面的操纵和掌控能力，通过结构性权力、工具性权力与示范性权力的聚合同主权国家围绕监管展开博弈，对国家权力构成一定程度的侵蚀。在国际层面，受技术竞争压力的影响，科技公司作为关键技术的开发者和推动者，其技术权力逐渐被纳入国家战略布局而强化政治性，监管也由此被视为有助于获取决定性优势的战略工具。由此，国家与市场间呈现出更深层次的权力互动和结构性的利益交织，人工智能监管的权力结构也日趋复杂化和多元化，导致全球人工智能治理开始从技术竞争拓展为监管竞争，从而加剧了监管格局的不确定性。有鉴于此，作者提出“人工智能技术权力”的概念，旨在揭示科技公司如何通过结构性、工具性与示范性三类权力的耦合深度嵌入监管过程，从而重塑国家与市场的关系。在此基础上，国际政治经济学应超越传统的利益分配研究视角，对技术主导时代权力结构做出新的理论解释。

**【关键词】** 人工智能监管；国家—市场关系；科技公司；人工智能技术权力；技术政治

**【作者简介】** 余南平，华东师范大学政治与国际关系学院教授；栾心蔚，华东师范大学政治与国际关系学院博士研究生（上海 邮编：200062）。

**【中图分类号】** D815 **【文献标识码】** A **【文章编号】** 1006-9550 (2025) 06-0031-29

<sup>\*</sup> 本文系国家社会科学基金重大项目“颠覆性技术发展对新型国际关系形塑研究”（项目批准号：23&ZD334）的阶段性成果。感谢《世界经济与政治》匿名审稿专家的意见与建议，文中疏漏由笔者负责。

## 一、引言

技术政治学者兰登·温纳（Langdon Winner）在讨论技术自主性时指出，所有技术的发展都最大限度地反映出人类在智力、发明创造力和人文关怀方面的品性。但在超过特定时刻后，这些品性对最终结果产生影响的边际效用将越来越弱，不会再对技术塑造世界的方式产生实质性影响。<sup>①</sup> 目前以 ChatGPT 为代表的新一代通用人工智能不仅仅止于“独立工具”，还以颠覆性通用技术的形式渗透到人类社会的方方面面，在创造新的问题领域的同时引发了系统性风险，尤其体现在技术对政治、经济与社会的再塑造和新型技术权力的构建等方面。

虽然无从得知庞大而充满未知的人工智能系统最终能否发展到“颠倒手段和目的之间的合理关系”<sup>②</sup> 的状态，但现实中科技公司正试图无休止地扩张人工智能赋予它们的利益和权力。与此同时，人工智能的广泛潜在应用及其“军民两用”性质也是当下国家间竞争的关键环节。世界各国能够在多大程度上塑造和参与人工智能价值链将决定它们在未来全球经济秩序中的地位，并影响自身对国际权力结构的塑造能力。当前，在国家与市场多方利益主体不同的利益优先次序下以及合作与竞争博弈并存的互动中，国家与跨国科技公司（市场代言人）正在寻找人工智能监管的新政治经济学逻辑，重新定位国家与市场的关系，进而在监管构建实践中塑造未来的全球人工智能竞争格局。<sup>③</sup> 美国的实践则是其中的典型代表。

## 二、人工智能监管的国际政治经济学再认识

技术的变革影响不仅取决于技术的自身属性特征，更取决于如何接纳、利用和发展技术的制度环境。ChatGPT 的问世标志着生成式人工智能取得了显著进展，将围绕人工智能的监管辩论推向高峰。由于人工智能技术发展受其底层技术、应用场景、社会环境和其他众多因素的影响，不同部门、组织以及国家与地区提出

<sup>①</sup> 兰登·温纳著，杨海燕译：《自主性技术——作为政治思想主题的失控技术》，北京大学出版社 2014 年版，第 269 页。

<sup>②</sup> 兰登·温纳：《自主性技术——作为政治思想主题的失控技术》，第 269 页。

<sup>③</sup> 国际政治经济学研究者提出了“国家与市场”的命题，并以结构性权力为分析框架关注了跨国公司的权力扩张，讨论了政府的两难选择问题。参见苏珊·斯特兰奇著，杨宇光等译：《国家与市场》（第二版），上海人民出版社 2019 年版；约翰·斯托普福德、苏珊·斯特兰奇著，查立友等译：《竞争的国家 竞争的公司》，社会科学文献出版社 2003 年版。

的人工智能技术监管方案也大相径庭，因此本文尝试将对人工智能监管的理解置于更广泛的政治经济学辩论中，旨在从新的技术政治背景中理解当下人工智能的监管动态。

### （一）关于监管的既有研究视角

监管是指通过制定可制裁的规则有目的地干预私人（也包括公共行为者）的行为，也是减轻技术发展所带来的不利影响的重要工具。<sup>①</sup>从这一定义出发，不难发现对监管的既有研究主要集中于国内层面，如第二次世界大战后发展起来的公共利益理论就赋予政府干预和管制的正当性。乔治·施蒂格勒（George J. Stigler）提出了“监管捕获”的概念，认为监管是利益集团相互竞争的结果，将监管的主流解释框架转向以企业理性逐利行为为核心的理性选择理论。<sup>②</sup>一些受古典经济学和经典国际贸易理论影响的学者也将监管俘获理论引入国际政治经济学，<sup>③</sup>将特定行业中占主导地位的公司利益与国际层面的监管设计进行比较，从而证明行业利益与监管结果之间确实存在某种相关性。<sup>④</sup>还有一些研究对监管俘获理论进行了批判性回应，如忽略不同参与者之间的信息不对称、忽视对政策制定者施加影响的团体多样性以及政策制定者本身的多种动机等，<sup>⑤</sup>并在此基础上发展了利益集团政治经济的分析模型。<sup>⑥</sup>不过这些研究仍然只是在理性选择框架中进行局部修正，普遍认为监管最终无法摆脱被俘获的命运。

如果将国内政府与市场之间对抗的这种张弛关系拓展到国际社会，不难发现监管俘获理论基于国家利益和市场利益的二分法在国际政治博弈中是存在偏颇的，原因在于国家利益和市场利益在国际社会中并不总是对立关系。当国家而非监管机构

---

① Robert Baldwin, Martin Cave and Martin Lodge, *Understanding Regulation: Theory, Strategy, and Practice* (2nd edition), Oxford University Press, 2012.

② George J. Stigler, "The Theory of Economic Regulation," *The Bell Journal of Economics and Management Sciences*, Vol. 2, No. 1, 1971, pp. 3-21.

③ David A. Lake, "Open Economy Politics: A Critical Review," *The Review of International Organizations*, Vol. 4, No. 3, 2009, pp. 219-244.

④ Thomas Bernauer and Ladina Caduff, "In Whose Interest? Pressure Group Politics, Economic Competition and Environmental Regulation," *Journal of Public Policy*, Vol. 24, No. 1, 2004, pp. 99-126.

⑤ Daniel Carpenter, "Protection Without Capture: Product Approval by a Politically Responsive, Learning Regulator," *American Political Science Review*, Vol. 98, No. 4, 2004, pp. 613-631; Daniel Carpenter and David A. Moss, eds., *Preventing Regulatory Capture: Special Interest Influence and How to Limit It*, Cambridge University Press, 2014.

⑥ 较具代表性的例子是“委托人—代理人—产业关系”的三层分析模型。参见 Jean-Jacques Laffont and Jean Tirole, "The Politics of Government Decision-Making: A Theory of Regulatory Capture," *The Quarterly Journal of Economics*, Vol. 106, No. 4, 1991, pp. 1089-1127.

本身成为监管行为的主要政治行为体时，监管行为不仅反映传统经济利益，而且深受国际竞争压力、国家安全和技術主导权塑造等因素的影响。罗伯特·基欧汉（Robert O. Keohane）指出，国际政治经济学研究“太少关注利益是如何构建的”。<sup>①</sup>因此回溯监管的政治性并将权力作为监管研究的核心维度，既有助于弥补既往国际政治经济学将监管简化为一种利益平衡的缺陷，还能更有效地剖析科技公司的权力积累及其与监管动态之间的深刻联系，并进一步挖掘国家监管项目在激烈国际竞争中的内生性与复杂的协同机制。

## （二）人工智能监管的国际政治经济学再认识

鉴于传统监管的被动性质，颠覆性创新总会対监管策略提出挑战。<sup>②</sup>相较于其他行业，人工智能的监管显然更加复杂。作为一项面向未来的技术，人工智能正以空前的速度进行自我学习和自我迭代，技术自主性日益增强。由此可见，下一代通用人工智能不仅是一个复杂技术社会系统，而且是以技术链接相互嵌套而构建的“系统的系统”。下一代通用人工智能发展所产生的利益分配、风险产生和权力构建也空前复杂，给监管带来了前所未有的挑战。

当前大多数人工智能监管都采取了基于风险评估的方法。<sup>③</sup>这一思路建立在“绝对安全”不可能是一个合理的监管目标的假设上，<sup>④</sup>因为“绝对安全”需要不成比例地调用公共资源和私人资源，基于风险的监管应该只会干预最高的风险，监管只需要保持企业较低的合规成本。<sup>⑤</sup>在这种监管思路下，适度监管被认为是伦理/安全与创新/经济增长之间的一种平衡。然而这种“平衡式”监管选择也会带来一系列问题，如谁来定义人工智能的伦理？谁能够对“透明”“公平”等规范的定义进行没有歧义的解释？即便当下人工智能监管的应用伦理学不断拓展，并试图针对偏见和缺乏可解释性规则等问题来制定出更精细的法律和技术解决方案，但这些

<sup>①</sup> Robert O. Keohane, “The Old IPE and the New,” *Review of International Political Economy*, Vol. 16, No. 1, 2009, pp. 34-46.

<sup>②</sup> Nathan Cortez, “Regulating Disruptive Innovation,” *Berkeley Technology Law Journal*, Vol. 29, No. 1, 2014, pp. 175-228.

<sup>③</sup> 如欧盟的《人工智能法案》围绕欧盟市场上人工智能系统部署者所需承担的风险和缓解义务的相称性进行了论证。参见 Gabriella Etziani, et al., “A Survey of Artificial Intelligence Risk Assessment Methodologies: The Global State of Play and Leading Practices Identified,” <https://www.trilateralresearch.com/wp-content/uploads/2022/01/A-survey-of-AI-Risk-Assessment-Methodologies-full-report.pdf>.

<sup>④</sup> Giandomenico Majone, “Strategic Issues in Risk Regulation and Risk Management,” *Risk and Regulatory Policy: Improving the Governance of Risk*, OECD, 2010, pp. 93-131.

<sup>⑤</sup> Cass Sunstein, *Laws of Fear: Beyond the Precautionary Principle*, Cambridge University Press, 2005.

规则既难以细化到没有模糊空间,<sup>①</sup>又不可能自我实施。<sup>②</sup>技术理解的匮乏与思维惯性的滞后使风险分析时常沦为一种理想化的技术官僚工具,其实际政策效力相当有限。

如果仅从理性选择的角度理解人工智能监管,进而对相关风险及成本收益进行技术官僚评估,那么整个社会的帕累托最优分配将有效回应“如何监管”这一问题。<sup>③</sup>这种非政治性的监管逻辑认识脱离了实际,理性选择下的市场失灵框架虽然可作为监管分析的起点,但由于预测不确定、解释范围有限及适用性不均等因素影响,难以构成国家监管战略的稳固理论基础。<sup>④</sup>既有研究发现,监管者采用风险评估或其他分析工具通常不是为了解决问题,而是旨在提高其组织合法性或扩大其自身在多层次环境中的监管权力。<sup>⑤</sup>人工智能技术的新颖性及其运作方式具有诸多不确定性,特别是新一代生成式人工智能技术变化的复杂性和自生性已经成为难以掌控的独立变量,通过理性选择把握不同情境下人工智能使用的风险几乎已不可能,任何机构都难以毫无争议地制定出监管的最佳策略。可见,基于风险评估和理性选择的讨论都只是描绘了一个应该如何监管的规范性愿景,而没有解释现实中的监管行为。在此背景下,从权力角度去探讨人工智能监管很有必要。

随着人工智能开启一个颠覆性的智能时代,其显示出的强大技术自主性尤其是未知的涌现点意味着其本身的发展走向和能力边界无法预知,人工智能的技术架构随之成为制定监管框架的决定性因素。在当前人工智能快速革新、技术转化为扎根于物质资源的有形力量的时代,我们亟须直面人工智能发展带来的认知新挑战,对有关监管的国际政治经济学逻辑进行再认识,并将人工智能监管探讨置于更广泛的政治、经济、技术和社会背景中,以期理解和把握国际监管政策制定过程中的权力

① 在实践中,科技公司利用政策措辞中的模糊空间,将其转化为正当化自我监管的合法性工具。参见 Ben Green, “The Contestation of Tech Ethics: A Sociotechnical Approach to Technology Ethics in Practice,” *Arxiv Preprint*, 2022, arXiv: 2106.01784.

② Attila Gyulai and Anna Ujlaki, “The Political AI: A Realist Account of AI Regulation,” *Információs Társadalom*, Vol. 21, No. 2, 2021, pp. 29–42.

③ The Anh Han, et al., “To Regulate or Not: A Social Dynamics Analysis of an Idealised AI Race,” *Journal of Artificial Intelligence Research*, Vol. 69, 2020, pp. 881–921.

④ Christopher Hood, et al., *The Government of Risk: Understanding Risk Regulation Regimes*, Oxford University Press, 2001, pp. 87–89.

⑤ Julia Black, “The Role of Risk in Regulatory Processes,” in Robert Baldwin, et al., eds., *The Oxford Handbook of Regulation*, Oxford University Press, 2010, pp. 302–348; Claudio M. Radaelli, “Rationality, Power, Management and Symbols: Four Images of Regulatory Impact Assessment,” *Scandinavian Political Studies*, Vol. 33, No. 2, 2010, pp. 164–188.

结构转变与权力斗争，特别是要基于新兴主导性权力行为体（科技公司）塑造监管的能力与角色变化，对技术政治予以更深的理解。

### 三、科技公司人工智能技术权力多维度构建

对权力含义的解释是多样化的，权力解释的多维扩散隐含着权力的多维度来源。<sup>①</sup> 本文认为，在人工智能监管政策的场域中，科技公司的人工智能技术权力是一个复杂的权力复合体，包含结构性权力、工具性权力和示范性权力三种相互交织和共同作用的权力形态，它们经由技术、经济、社会和政治四个维度产生复杂的相互影响。

#### （一）人工智能技术结构性权力如何生成

##### 1. 技术维度

从技术维度来看，算力、数据与人才的垄断奠定了科技公司尤其是科技巨头的结构性权力基础。算力是人工智能进步的主要载体和驱动因素，决定了谁能参与技术竞争、如何参与以及竞争的结果。

首先，从技术堆栈的计算底层来看，英伟达的图形处理器是当前人工智能大模型训练的事实标准，其图形处理器架构直接影响了大模型的设计与训练效率，同时与计算统一设备架构平台进行软硬件绑定，几乎垄断了高性能计算领域。这在某种意义上主导了当下大模型的发展方向及进程，并迫使其他参与者不得不在其生态体系内发展。就算力成本而言，大规模人工智能模型使用计算量平均约为普通人工智能模型的100倍。<sup>②</sup> 谷歌的 Gemini Ultra 和 OpenAI 的 GPT-4 的训练成本分别估计为 1.91 亿美元和 7800 万美元。<sup>③</sup> 虽然通过搭建高效的模型架构等方式能够减少计算负荷、降低成本并减弱对顶级图形处理器的依赖，但算力的稀缺性在现阶段仍然是一个不可忽视的结构性约束，这种稀缺性构成人工智能领域结构性权力的核心支柱。

其次，高质量数据是人工智能模型训练和优化的关键。除已知的公开数据集外，

<sup>①</sup> Steven Lukes, *Power: A Radical View (3rd edition)*, Bloomsbury Publishing, 2021; 克里斯蒂安·罗伊-斯米特、邓肯·斯尼达尔主编，方芳等译：《牛津国际关系手册》，译林出版社 2019 年版，第 756—758 页。

<sup>②</sup> Jaime Sevilla, et al., "Compute Trends Across Three Eras of Machine Learning," Paper Prepared for the International Joint Conference on Neural Networks (IJCNN), IEEE, Padua, 2022, pp. 1-8.

<sup>③</sup> "The Impact of Training Costs on AI Models Development," [https://www.gets-usa.com/training-cost-ai.shtml?utm\\_source/](https://www.gets-usa.com/training-cost-ai.shtml?utm_source/).

有研究发现大语言模型在爬取互联网数据生成内容时，关于人类与系统的真实交互的数据将变得越来越有价值。<sup>①</sup> 这使谷歌、Meta 和 X 等传统平台垄断的力量再次突显，庞大的用户基数和生态系统将这种数据转化为非公开资产，科技公司事实上控制了从采集、存储到使用的数据生产链，能够通过掌控数据价值的迁移主导人工智能技术的发展方向与速度。

最后，高技能人才也是决定当下科技公司塑造技术结构能力的关键要素。在人工智能人才供不应求的背景下，科技公司具备更多的财政资源来投资和留住顶尖人才。<sup>②</sup> 科技公司拥有顶尖人才意味着掌握了更多的人工智能创新和知识产权，少数公司由此控制和定义了人工智能研究过程的工具、开发环境、语言和软件，对人工智能发展方向的控制力日益增强。以科技巨头为代表的科技公司对高性能计算资源、独占性用户交互数据以及顶尖技术人才的垄断构成了当前技术结构中明显的资源非对称依赖，确立了自身在技术维度的结构性权力。

## 2. 经济维度

从经济维度来看，科技公司主要通过生态锁定、价值链整合、资本集中与并购等手段塑造了人工智能经济的底层结构，并借由人工智能的通用性将这种结构性塑造拓展至各行各业。

首先，生态锁定使得技术使用者难以摆脱对特定平台的依赖。科技公司围绕自身人工智能产品和服务构建封闭式生态系统，将开发者和用户绑定到自身平台，以闭环商业模式整合了从基础层到应用层的大模型价值链。<sup>③</sup> 而这个过程还会产生间接或直接的网络效应：越来越多的第三方公司基于基座模型开发产品和服务后，其反馈数据将进一步提升基座模型性能，实现收入的乘数效应递增。未来人工智能大模型应用场景的持续增加会刺激领先模型的持有者培育出更大的生态系统，使得某种技术生态成为经济运作的底层支撑。

其次，在生成式人工智能迅猛发展的背景下，科技巨头与初创科技公司之间并不是简单的竞争或垄断与被垄断的关系，而是形成了以科技巨头为算力掌控者、以初创科技公司为创新驱动者的协作型结构性权力关系。在算力限制构成的巨大耗资

<sup>①</sup> Ilia Shumailov, et al., "The Curse of Recursion: Training on Generated Data Makes Models Forget," arXiv Preprint, 2023, arXiv: 2305.17493.

<sup>②</sup> Nur Ahmed, et al., "The Growing Influence of Industry in AI Research," *Science*, Vol. 379, No. 6635, 2023, pp. 884-886.

<sup>③</sup> 如 OpenAI 通过开发 GPT 和 DALL-E 等专有人工智能资产，通过收取 API 调用、模型订阅和人工智能专业服务费用，完成了商业模式闭环。

挑战下，初创科技公司需要科技巨头提供算力与基础设施来进行模型开发和推理，而在模型训练完成后，初创科技公司往往需要科技巨头的平台生态来快速获取用户和实现商业落地、加快市场渗透。<sup>①</sup> 创新范式的变化促使科技巨头通过投资、控股甚至收购行为来吸纳初创科技公司的创新能力。2024年，全球1/3的风险投资流向了初创科技公司，对初创科技公司的风险投资同比增长52%达1315亿美元，这反映出人工智能垂直领域的持续融合。<sup>②</sup> 科技巨头与初创科技公司之间积极协作，形成资源配置、资本注入和平台接口深度耦合的结构性关系，甚至还可能构成一种游离于监管之外的“循环交易”。<sup>③</sup> 而这种“伪竞争”塑造市场格局的背后是经济权力的再集中与结构性异化，企业之间边界的模糊对既有垄断/反垄断的二元视角也造成了新挑战。

最后，科技公司还凭借人工智能技术的通用性向金融、医疗和制造等传统行业全面提供人工智能解决方案。<sup>④</sup> 人工智能技术的多领域渗透催生出跨行业的技术结构性依赖。近年来，各类组织对人工智能的使用显著增加，尤其是生成式人工智能在复杂流程自动化、优化资源配置和提升决策效率等方面展现出广泛价值。<sup>⑤</sup> 然而想要最大限度地发挥生成式人工智能的技术潜力，往往需要深度定制化的解决方案，尤其是基于企业自身及其客户的专有数据进行模型微调。<sup>⑥</sup> 在这一过程中，科技公司又成为人工智能能力的“基础设施提供者”，其他企业的产品和服务不得不在科技公司的技术架构上运行。这种架构性嵌入进一步强化了科技公司对其他行业的主导权，在市场层面形成深层次的路径锁定并产生了持续的经济效应。

## （二）人工智能技术工具性权力如何作用

### 1. 社会维度

从社会维度来看，科技公司在生成式人工智能时代能够借助更有效的技术工具

---

① 如微软成为 OpenAI 最大投资者，享有 OpenAI 技术核心的使用权，目前正通过 OpenAI 开发其 Azure AI 超级计算机并将 OpenAI 的模型集成到其商业和消费产品（GitHub、Office 和 Bing）中，同时 Azure 也成为 OpenAI 算力的独家专供者。

② “AI Startups Grabbed a Third of Global VC Dollars in 2024,” [https://pitchbook.com/news/articles/ai-startups-grabbed-a-third-of-global-vc-dollars-in-2024?utm\\_source/](https://pitchbook.com/news/articles/ai-startups-grabbed-a-third-of-global-vc-dollars-in-2024?utm_source/).

③ “Amazon Spends \$ 2.75 Billion on AI Startup Anthropic in Its Largest Venture Investment Yet,” <https://www.cnbc.com/2024/03/27/amazon-spends-2point7b-on-startup-anthropic-in-largest-venture-investment.html>.

④ Dogan Gursoy and Ruiying Cai, “Artificial Intelligence: An Overview of Research Trends and Future Directions,” *International Journal of Contemporary Hospitality Management*, Vol. 37, No. 1, 2025, pp. 1-17.

⑤ Osman Şahin and Durmuş Karayel, “Generative Artificial Intelligence (GenAI) in Business and Industry: A Systematic Review on the Threshold of Transformation,” *Journal of Smart Systems Research*, Vol. 5, No. 2, 2024, pp. 156-175.

⑥ “The State of AI: How Organizations Are Reweiring to Capture Value,” <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai#/>.

以空前强大的能力塑造用户的认知结构与行为模式，显著扩大了自身对社会结构和价值体系的塑造作用。

首先，从人工智能技术诱发的变化来看，生成式人工智能依靠对数据和算法的技术掌控，深刻影响着人类的认知内容、认知过程和认知能力。数据收集、算法选择、超参数设置和自我学习等“技术黑箱”的集成使科技公司获得了对人类行为敏锐的洞察力，在与用户交互反馈中隐匿地扩展着自身影响力。参与者的自由裁量空间会随着人工智能工具的广泛使用而日益缩减。<sup>①</sup> 嵌入协议设计和选择架构实际上把部分人的价值观以“人类价值”的名义加入了一个原本在统计学意义上客观反映世界的语言模型中，<sup>②</sup> 这是技术性工具权力向意识形态领域延伸的典型反映。例如，用以减轻风险的“护栏设置”本质上是一种高度不透明的认知过滤机制，然而微软和 OpenAI 的科技公司已经签署协议，将隐藏的元提示纳入人工智能工具中。<sup>③</sup> 这种“技术黑箱”的存在虽然在形式上保护了技术知识产权，但同时提供了一种便于规避外部审查与问责的制度性逃逸能力，在政治和伦理上为科技公司施加社会影响制造模糊空间，进一步扩展了其工具性权力的边界。

其次，从生成式人工智能的工具特性来看，生成式人工智能工具通过模板和语言结构实现了准确的响应和指导，在本质上充当着“知识增强器”，能够弥补使用者的认知短板，并将新信息与用户既有认知相联系，在人机交互反馈中建构人类的认知过程。<sup>④</sup> 由此，生成式人工智能正在重新定义解决问题的方式和人类学习的方法，日益成为当代知识生产和认知加工的重要中介。这种双向建构关系不仅改变了人类获取信息与理解世界的方式，并且使科技公司成为价值框架和社会认知的设定者。生成式人工智能大模型在成为其扩展工具性权力的重要支点的同时，也在通过构建知识系统影响社会认知。

最后，从生成式人工智能形成的社会传播效应来看，在当前的互联网数字世界中，搜索引擎虽然提供了丰富的信息选择，但可能隐含某种信息排序上的偏好，用

① Daria Gritsenko and Matthew Wood, "Algorithmic Governance: A Modes of Governance Approach," *Regulation & Governance*, Vol. 16, No. 1, 2022, pp. 45-62.

② Seth Stephens-Davidowitz, *Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are*, Dey Street Books, 2017.

③ Charles Duhigg, "The Inside Story of Microsoft's Partnership with OpenAI," [https://www.newyorker.com/magazine/2023/12/11/the-inside-story-of-microsofts-partnership-with-openai?utm\\_source=](https://www.newyorker.com/magazine/2023/12/11/the-inside-story-of-microsofts-partnership-with-openai?utm_source=)

④ Jingyi Geng, "The Revolution of Learning in the Digital Age: Navigating the Seas of Change with Generative AI," <https://medium.com/@gengjingyi324/the-revolution-of-learning-in-the-digital-age-navigating-the-seas-of-change-with-generative-ai-0ad2d6b3d9e4>.

户需要通过自身的知识认知加以判断、分析和整理。生成式人工智能大模型则与之不同，它与用户的互动方式是对用户问题提供直接“标准”回答且不存在任何间断。无论就解决问题的精准度还是知识传播及时和可达性的广泛性来说，生成式人工智能大模型都强化了知识传播速度与能力，进而放大了科技公司建立在参数选择权基础上的社会影响力。

## 2. 政治维度

从政治维度来看，科技公司通过调动自身强大的资金实力、广泛的游说网络以及专业知识和组织资源，以“超级企业家”身份直接干预政策制定和人工智能的监管进程，使各国出台的政策在优先级和方向上符合其商业利益。

首先，科技公司利用资金优势进行行业游说以致改变政策的力量不容忽视。在2021年欧盟加强数字监管时，科技行业已成为欧盟游说支出的最大行业，其巨额游说经费对欧盟政策制定者和安全特权准入产生了重要影响。<sup>①</sup> 在人工智能领域，截至2023年11月，欧洲议会成员参与的人工智能会议中有66%是与企业利益相关的，比2019—2022年的平均水平高出约10%，欧盟委员会高级官员参与的人工智能会议中有86%都是由工业界举办的。<sup>②</sup> 在美国，2024年有648家公司在联邦层面就人工智能问题展开游说，较2023年增长141%。<sup>③</sup> 其中，OpenAI将其游说支出增加至2023年的7倍，约176万美元。然而，这一金额仍与投入超过2400万美元的Meta具有很大距离。<sup>④</sup> 游说活动不仅仅是一种利益表达，更是一种通过金钱、信息和接触渠道对制度本身运行的“合法”操纵。

其次，科技公司通过稳定的“旋转门”机制实现了对政策网络的深度嵌入。大约3/4的谷歌和Meta的游说者具有在欧盟机构或成员国政府工作的经历，而他们的工作任务往往是游说前同事。<sup>⑤</sup> 在美国则有18%的注册游说人员来自谷歌和微软等大型科技公司，这种高频率与双向的人才流动模糊了公共领域与私人

① “The Lobby Network: Big Tech’s Web of Influence in the EU,” <https://corporateeurope.org/en/2021/08/lobby-network-big-techs-web-influence-eu>.

② “Big Tech Lobbying Is Derailing the AI Act,” <https://corporateeurope.org/en/2023/11/big-tech-lobbying-derailing-ai-act>.

③ “AI Companies Upped Their Federal Lobbying Spend in 2024 Amid Regulatory Uncertainty,” [https://techcrunch.com/2025/01/24/ai-companies-upped-their-federal-lobbying-spend-in-2024-amid-regulatory-uncertainty/?utm\\_source=](https://techcrunch.com/2025/01/24/ai-companies-upped-their-federal-lobbying-spend-in-2024-amid-regulatory-uncertainty/?utm_source=).

④ James O’Donnell, “OpenAI Has Upped Its Lobbying Efforts Nearly Sevenfold,” <https://www.technologyreview.com/2025/01/21/1110260/openai-ups-its-lobbying-efforts-nearly-seven-fold/>.

⑤ “The Revolving Door: From Public Officials to Big Tech Lobbyists,” <https://corporateeurope.org/en/2022/09/revolving-door-public-officials-big-tech-lobbyists>.

领域的界限，使科技公司得以在议程设置、政策咨询和规则谈判中持续保持影响力。

最后，随着政府日益依赖科技公司提供的数字基础设施及其生成式人工智能解决方案，科技公司利用自身技术资源创造了新的依赖关系，并增加了直接参与政策协商的机会。目前生成式人工智能在公共服务、行政管理、教育和政策制定等领域得到广泛应用，日本、新加坡、英国和加拿大等国政府部门都已采纳科技公司的云计算和人工智能算法工具等服务。<sup>①</sup>美国国防部和情报机构也部署了亚马逊云服务提供的联合作战云能力等高度定制化的安全服务体系。<sup>②</sup>生成式人工智能服务不仅为政府提供了数据收集分析和决策支持，更是在技术实现路径和系统运行保障方面具有不可替代性。这种技术嵌入不仅使政府在执行关键职能时必须与科技公司建立长期合作关系，还使科技公司在政策制定与执行中发挥了不可或缺的作用，大大提升了科技公司的政治议价能力。

### （三）人工智能技术示范性权力如何发挥影响

人类生存发展与社会进步本身隐喻了对美好和进步事业的追求，在这个过程中先行者的成功会持续呈现出对其他行为体的吸引力，进而形成示范性权力。

首先，示范性权力并非依靠强制或规范来实现影响，它以模式、范式、模型构建和标杆人物为载体，通过创新吸引他者的模仿与跟随。正如量子物理中的希格斯机制所揭示的，在量子世界中以纠缠方式赋予粒子不同的质量，而某些粒子可以成为典型的“希格斯粒子”，能够带动其他粒子运转，进而使自身具有示范性作用。<sup>③</sup>同样，领先人工智能企业的示范性权力也通过跨领域的渗透性和牵引性影响各类社会行为体的行为选择与认知方向。当某项技术从实验室到企业孵化的过程中中脱颖而出时，领先人工智能企业的技术超越性往往被赋予“人类技术高度”的美誉，引发全社会的关注并被授予各类奖项，其产生的示范性效应不仅会得到业界的广泛认可和追随，也会被赋予同样的社会认可与仰慕，进而产生更大的社会影响力和溢出效应。

其次，OpenAI 推出 ChatGPT 不仅带来了一项在技术路线和商业模式上影响了全

① 参见《新加坡科技局与 AWS 生成式人工智能创新中心合作，扩展生成式人工智能在公共部门的机构中的应用范围》，<https://aws.amazon.com/cn/solutions/case-studies/govtech/>; Geraldine Scott, “How Civil Servants Really Use AI, from Lesson Plans to Recruitment,” <https://www.thetimes.com/uk/politics/article/how-civil-servants-really-use-ai-from-lesson-plans-to-recruitment-k8nzghbn8?region=global>.

② “Cloud Computing for U. S. Defense,” <https://aws.amazon.com/cn/federal/defense/>.

③ 有关对量子世界的“纠缠性”与希格斯机制的详细论述，参见肖恩·卡罗尔著，王文浩译：《寻找希格斯粒子》，湖南科学技术出版社 2018 年版。

球数以亿计的用户和企业的技术轰动产品，更是成为推动下一代通用人工智能这一宏伟愿景走向现实的社会创新标杆。在技术层面，Transformer 架构已成为自然语言处理与多模态 AI 中的主导计算框架。在产业端，技术发展正朝向“垂端大模型”与“人型机器人”两大方向集中演进，其速度之快、影响之深远超政治与社会治理步伐，治理工作难以再主动设定边界而是更多被动地追随；而在知识生产与学习端，大模型已被广泛接受为新的研究辅助范式，改变了科学认知的路径和逻辑。与此同时，中国 DeepSeek 的开源策略又引发了人工智能创新生态的新一轮结构性重组，形成技术扩散的“蒲公英效应”。不少初创科技公司开始将战略重心由模型研发转向具体应用开发，进入后基础模型时代的商业探索路径。

最后，人工智能技术的示范性权力既服务于企业自身的产业主导权与资本扩张，往往也被所在国作为战略性技术资产投入地缘政治竞争，进而驱动国家、企业与个体出于“不掉队”的焦虑而纷纷追随。例如，当下几乎所有国家都确定了自身的人工智能发展战略，以不同路径尝试推动本国的人工智能技术快速发展，甚至采用跨国监管工具对他人工智能发展予以打压。因此，人工智能技术的示范性权力不仅仅是引发模仿，更强化了“科技创新即国家竞争力”的时代共识，而在全社会的技术信仰与创新动员力的强化中，显著扩展了科技公司影响力的边界与深度。

综上，科技公司拥有与以往不同的独特人工智能技术权力，这种权力不仅在内容上是多维的，还以复合型纠缠形式发挥作用，各种权力要素相互集成、相互赋能、自我强化产生了“权力复合体”机制，进而为技术与权力关系的理论拓展解释提供了新视角和新空间。

#### 四、人工智能技术权力的理论解释与再认识

在 2025 年 1 月的离任演说中，拜登类比了艾森豪威尔政府有关军工复合体的讨论，表达了对当前科技工业复合体（tech-industrial complex）对美国公共利益形成潜在威胁的担忧。<sup>①</sup>“科技工业复合体”的概念突显了科技公司在当今社会中的系统性影响。在人工智能迅猛发展的背景下，科技公司在多维度和多领域日益发挥关

---

<sup>①</sup> The White House, “Remarks by President Biden in a Farewell Address to the Nation,” <https://bidenwhitehouse.archives.gov/briefing-room/speeches-remarks/2025/01/15/remarks-by-president-biden-in-a-farewell-address-to-the-nation/>.

键性作用，其权力表现形态已不再是单一、线性且局限于行业与领域的影响力，人工智能技术的复合纠缠性权力作用使得科技公司作为科技工业复合体变得与众不同。由此，对科技公司的人工智能技术复合纠缠性权力也需要在理论层面重新加以解释。

首先，人工智能技术权力具有范式转换性。整体来看，人工智能技术权力的根本基础在于人工智能技术本身，计算资源、算法迭代速度、数据整合能力和模型先进性等技术要素构成了权力生成的技术前提。技术维度不仅为结构性权力、工具性权力与示范性权力三种权力的运作提供了支撑，更决定了其实现机制与辐射边界。这种由技术主导的权力形态在结构逻辑上超越了传统跨国公司基于资本与市场的权力范式。在结构性权力层面，科技公司超越传统物理资源配置，通过API接口和工具依赖等构建起不可替代的系统依赖，实现了嵌入式主导。在工具性权力层面，科技公司不仅运用传统资源展开游说活动，还通过认知工具与知识掌控影响政策议程本身，其权力行使方式更具隐蔽性与内化倾向。在示范性权力层面，科技公司没有止步于产业政策模仿与企业治理仿效，而是通过对技术路径、社会认知方式与未来想象的塑造能力生成了一种叙事性主导力。三种权力并非彼此孤立，而是通过技术支撑形成高度嵌套与循环放大的协同机制：结构性权力为工具性权力提供基础支撑，工具性权力则为示范性权力提供实现路径，示范性权力又为二者创造正当性条件并强化结构性的合理配置与制度嵌入。人工智能技术权力由此具有了结构支配、制度介入与价值塑造的复合性功能，进而实现权力的跃迁。

其次，人工智能技术权力具有高度耦合性与自我强化机制。人工智能不是某一领域的单一技术变量，而是作为战略性基础设施被广泛嵌入生产、金融、安全与知识等多种结构中，形成一种底层技术架构。在各领域的结构性嵌入中，科技公司通过数据采集、模型优化与行为反馈构建起闭环式运作机制。人工智能的反馈性与自我生成性使技术权力相应具备了自我强化特征。人工智能技术权力作用不是一种静态配置，而是呈现为能够自我运作、持续扩张和自我强化的动态机制，并重塑着社会的全部生产和生活方式。科技公司借由高耦合、强反馈的系统动态地构建出难以被替代的深层结构性依赖，在制度、认知与资源层面实现了权力的纵深控制，给现有公共权力结构和治理逻辑带来了空前挑战。

最后，人工智能技术权力具有量子世界典型的纠缠性机制作用。其一，人工智能的设计、部署与监管涉及国际机构、科技公司、标准制定组织和用户等多元行动

主体，不同主体之间又存在不同程度的依赖关系。这种多向度依赖致使任何一方都难以独立掌控权力，也无法对技术风险承担完整责任，责任归属难以明确划分。其二，人工智能技术演进具有极高的不确定性，其快速迭代可能带来短期的权力集中，然而新的突破可能会瞬时改变权力分配和运作方式（如 ChatGPT 和 DeepSeek）。加之人工智能的长期社会影响（如教育和就业的转型）可能在未来数年才会显现，这种时间维度上的错位进一步使得权力路径难以预测和追踪。其三，在全球地缘政治竞争日趋加剧的背景下，主权国家出于安全、产业与大国博弈等多重考量，加深了对科技公司的依赖，将科技公司视为技术竞争的关键节点，使其服务于国家战略，进而使得科技公司与政府日益形成功能互补、权责交错的合作权力结构，科技公司的技术权力也随之在国家加持下获得更加充分的合法性与更大的扩展空间。由此，多维度的纠缠性不仅带来了治理困境，也成为地缘竞争格局中新型政企合作权力场域的象征。

表 1 科技公司的人工智能技术权力分析框架

权力类型	权力维度	权力来源	权力实施结果
结构性权力	技术控制	算力、数据、人才垄断	强化依附关系构成技术高壁垒； 形成路径依赖； 隐性影响决策选择
	市场与资本（经济） 操纵	生态锁定； 竞争控制； 游说	
工具性权力	社会构建	算法与数据操纵； 护栏设置； 知识生产及认知加工	塑造认知和行为； 钝化思维； 显性引导决策
	政治参与	提供基础设施依赖； 政治“旋转门”； 智囊渗透	
示范性权力	技术领先下的综合 维度	创新的魅力； 品牌全球影响力； 科技领袖个人魅力	技术乌托邦主义； 文化与创新依附； 树立全球技术标杆

资料来源：笔者自制。

综上，人工智能结构性权力依赖于科技公司对技术结构和经济结构的深度塑造，由此建立起一种隐蔽但强大的深层次权力格局；人工智能工具性权力依赖于科技公司在社会与政治领域通过技术工具和相关资源的显性应用，旨在实现特定目标或施

加相应影响；人工智能示范性权力则是科技公司通过创新实践和价值观引领，促使社会行为体产生自发性的认同、仰慕与模仿性追随。这三种权力集合塑造了科技公司人工智能技术权力的复合性纠缠，进而产生了特殊的“希格斯场域”效应，推动科技公司加速转化为人工智能技术时代的科技工业复合体，并同时作为权力复合体而存在。权力复合体必然会对政府的政策与监管进程施加深刻影响，从而导致监管场域的复杂性和不确定性。

## 五、科技公司人工智能技术权力对人工智能监管的挑战

权力的存在与行使必然带来相应后果。尽管科技公司在公共场合呼吁加强人工智能领域立法，实际上却强烈抵制对人工智能基础模型进行任何形式的监管。即便是当前对于人工监管影响最为深远的欧盟《人工智能法案》，在其立法过程中也未能抵挡住科技公司的权力侵蚀。<sup>①</sup> 无论是热衷于商业利益还是真正有利于技术创新环境塑造，科技公司的实际行动表明它们正在凭借人工智能技术权力中结构性与工具性权力的复合性纠缠力量，通过技术、经济、社会和政治四个维度来操纵人工智能监管议程，最终形成“合法性垄断”，这给人工智能监管带来前所未有的挑战。

### （一）技术加速主义与监管滞后性的制度张力

从技术维度来看，科技公司凭借对底层技术架构和迭代节奏的垄断掌控，构建起“技术先行—监管滞后”的制度张力。技术加速主义并不是一种单纯的思潮，而是科技公司操控技术发展路径和塑造监管节奏的权力实践逻辑。技术加速主义源于对技术发展的绝对信仰，该观点将技术视为解决所有社会问题的核心力量，主张技术发展应摆脱外部干预，自由加速并充分释放潜力。<sup>②</sup> 在硅谷，技术加速主义的理念已经内化为科技公司推动创新与迭代的实践框架。自 ChatGPT-3 发布以来，全球人工智能模型更新几乎以周为单位。生成式人工智能已从最初的仅为文本处理发展

<sup>①</sup> 在欧洲议会于 2023 年 6 月通过关于《人工智能法案》的最终立场时，大约有 150 名公司高管以公开信的形式呼吁欧盟放宽监管，否则欧洲将“错失重新加入技术前沿的机会”。参见 Javier Espinoza, “European Companies Sound Alarm over Draft AI Law,” <https://www.ft.com/content/9b72a5f4-a6d8-41aa-95b8-c75f0bc92465>.

<sup>②</sup> Andrian Ferry Wijarnarko, “Human’s Relationship with Technology in Nick Land’s Accelerationism,” *Journal Filsafat*, Vol. 34, No. 1, 2024, pp. 1-22.

为“多模态”演化，并给后续控制模型偏差和准确性问题带来严峻挑战。<sup>①</sup>

首先，尽管美国和英国等国已成立人工智能安全机构，但技术发展速度依然远超监管机制的响应能力。仅在 2025 年第一季度，OpenAI 发布 GPT4.5，XAI 发布 Grok3，谷歌发布 Gemini 2.5 Pro；中国国产大模型也层出不穷，除 DeepSeek 外，百度发布文心大模型 4.5X1，阿里发布开源 QwQ-32B。在这种竞速式开发节奏下，尽管一些技术专家持续呼吁暂停开发强大的人工智能系统以避免可能的风险，但人工智能创新仍在继续且不可阻挡地颠覆了原有的产品形态。<sup>②</sup> 因此，在人工智能开发遵循加速主义的客观事实与监管机制固有滞后性之间的缝隙中，科技公司人工智能技术权力不断得到强化，而监管本身日益被视为社会进步的障碍，并在技术加速主义所构建的叙事下更显乏力与形式化。

其次，在技术结构性权力的主导下，监管不仅面临技术认知差距的问题，还在制度执行中遭遇了“滞后合法化”困境。一方面，生成式人工智能的不确定性与涌现特性要求监管具备高度的灵活性与适应性，而科层制政府强调的规范性与流程稳定性难以回应这一变化节奏，二者之间的不匹配容易造成监管延误甚至失效；另一方面，技术官僚对人工智能原理、训练机制与潜在外部性等缺乏系统认知，造成专业能力缺位。这一知识鸿沟使政府在决策时更加依赖科技公司的专业知识和技术支持，从而使得行业标准被转化为政策逻辑，公共决策受科技公司与科技行业权力的嵌入影响，最终加剧了结构性权力对公共治理能力的掣肘。

最后，从人工智能技术作用产生的塑造力和影响力后果来看，技术加速主义在当下不只是一种推动技术创新的理念，更是通过人工智能技术自身的加速构建出一个不断突破制度边界的技术现实并形成强大的技术生态，使监管始终处于“被动响应”的状态。这种制度滞后与认知依赖共同构成了对政府治理能力与独立性的结构性挑战，也突显了科技公司在技术结构性权力作用中对政策导向的实际操控力。

---

① Natalie Hanson, “Feds See ‘Narrow Window’ to Regulate AI as Silicon Valley Innovation Explodes,” <https://www.courthousenews.com/feds-see-narrow-window-to-regulate-ai-as-silicon-valley-innovation-explodes/>.

② OpenAI 正在开发下一个模型 GPT-5 以及一个搜索引擎；谷歌则准备发布 Astra，并在美国以外地区推出人工智能生成的搜索查询；微软正在开发自己的人工智能模型，并聘请了英国企业家穆斯塔法·苏莱曼 (Mustafa Suleyman) 管理一个新设的人工智能部门。数十亿美元的人工智能投资正在涌入各种规模的科技公司。参见 Dan Milmo and Alex Hern, “As the AI World Gathers in Seoul, Can an Accelerating Industry Balance Progress Against Safety?” <https://www.theguardian.com/technology/article/2024/may/18/ai-seoul-global-summit-safety-openai-meta>.

## （二）科技巨头市场垄断对监管的挑战

在经济维度上，科技公司凭借对计算能力、核心数据资源与技术专利的深度整合，构建起高度集中的结构性权力格局，直接制约着监管的有效性与独立性。这一格局带来了三方面挑战。

首先，网络效应和规模经济的累积使得基础模型市场加速集中，只有少数巨头系统性地掌控了算力、数据与部署渠道，具备持续训练和快速迭代模型的优势。虽然 DeepSeek 等初创科技公司通过算法创新与开源路线取得了阶段性突破，拓展了市场可能性，然而整体的结构性壁垒仍然存在，在大规模推理、服务稳定性和多用户并发等场景中，科技公司仍需要依托大型云平台提供的弹性计算与分布式调度能力。此外，初创科技公司在数据整合、用户触达和系统集成等关键环节通常也依赖巨头构建的平台生态，难以在产业生态层面挑战科技巨头的的主导地位。同时，科技巨头依然牢牢控制着从数据到芯片等关键领域的投入与产出，可以通过限制访问来保护自己免于承受竞争压力，并利用网络效应锁定用户，初创科技公司因而面临沦为巨头生态的“结构性附属”的潜在风险。监管制度依然存在科技巨头主导标准和决定资源分配的现实挑战。

其次，科技巨头通过资本性和平台性的绑定将初创科技公司纳入自身体系，掌控了模型发展路径，并在很大程度决定了人工智能创新的方向和速度。目前，美国六家科技巨头通过 90 多个投资和合作伙伴环节产生的互网络构成了人工智能行业的核心。<sup>①</sup> 如微软对美国 OpenAI 和法国 Mistral AI 的投资支持不仅包括模型部署，还涉及 Azure 云平台接入。同样，亚马逊与谷歌对 Anthropic 的投资也将 Claude 模型推理服务限定在自身云平台上。此类交易使少数实力雄厚的公司能够利用其获得的资本特权消除潜在的竞争威胁，已引起美国、欧盟和英国监管机构的担忧。<sup>②</sup> 这种结构性嵌入不仅消解了市场竞争机制，也在事实上造成了技术路径选择的单一化，大大挤压了监管干预的空间。

最后，科技公司可以通过市场的力量反向塑造制度环境。科技巨头更有能力将自身商业利益转化为制度优先事项，限制监管机制的独立空间。2025 年 3 月 5 日，

<sup>①</sup> “UK Has Real Concerns About AI Risks, Says Competition Regulator,” <https://www.theguardian.com/technology/2024/apr/11/uk-has-real-concerns-about-ai-risks-says-competition-regulator>.

<sup>②</sup> “US Launches Inquiry into AI Deals by Microsoft, OpenAI, Google and Amazon,” <https://www.theguardian.com/technology/2024/jan/25/ftc-ai-inquiry-microsoft-alphabet-amazon>; “UK Has Real Concerns About AI Risks, Says Competition Regulator,” <https://www.theguardian.com/technology/2024/apr/11/uk-has-real-concerns-about-ai-risks-says-competition-regulator>.

英国竞争与市场管理局决定终止对微软与 OpenAI 合作关系的反垄断调查，理由是尽管该合作存在“实质性影响力”，但未达到正式审查所要求的“事实控制”标准。<sup>①</sup>这一决定虽然引发了广泛争议，但本质上反映的是现行监管工具难以应对平台化、松散控制与投资网络交织所带来的新型权力形式。监管不得不选择性妥协是其对市场力量无力回应的体现。

### （三）“伪开源”与认知操纵下的监管障碍

在当前的人工智能政策语境中，“开源”被赋予极强的正面话语色彩，常被理解为开放与透明的象征，并被寄望于促进竞争、打破技术垄断。同样，一个系统被定义为“开放”或“开源”也可能对监管方式产生不同影响。在欧盟《人工智能法案》和其他潜在监管方式讨论中，存在一种主张是对开源人工智能给予广泛豁免，以创新更具竞争力的大型人工智能系统。<sup>②</sup>从逻辑上看，如果开放或开源人工智能能够提供透明度，监管当局将可以审核模型性能、识别风险并制定应对措施确保安全，包括更有效地促进创新、提供竞争环境。然而在现实中，开放或开源往往成为科技公司为平抑公众紧张、争取监管豁免的策略性选择。科技公司常借助工具性权力控制技术开源与透明度，这既制造了对技术中立性和开放性的假象，也显著强化了其塑造社会认知的能力，引领了社会对技术意义与方向的集体想象。

首先，“开源”往往由科技公司自己定义，但宣扬的可能是一种“伪开源”。一般来说，开源是指系统具有透明度、可重用性和可扩展性，可以被用来检查、重用和构架。<sup>③</sup>然而，许多标榜开源的人工智能系统实际上是经过包装的封闭模型，它们既没有提供有意义的文档和访问权限，也没有提供附有注释的基于人类反馈的强化学习训练数据（包括公布其研究结果），通过第三方独立评审的情况更是少之又少，只是一种形式上的开源。以 Meta 的 LLaMA-2 为例，虽然模型权重可下载，但其训练数据未公开、部署权限受到限制，用户需要签署特定协议方可使用，并不符

---

① Competition and Markets Authority, “Microsoft Corporation’s Partnership with OpenAI, Inc. —Decision on Relevant Merger Situation,” [https://assets.publishing.service.gov.uk/media/67c841d6d0fba2f1334cf276/1.\\_MS\\_OAI\\_-\\_Summary.pdf#page=4.51](https://assets.publishing.service.gov.uk/media/67c841d6d0fba2f1334cf276/1._MS_OAI_-_Summary.pdf#page=4.51).

② “GitHub CEO Says EU AI Act Shouldn’t Apply to Open Source Devs,” [https://www.theregister.com/2023/02/07/github\\_ceo\\_ai\\_act/?utm\\_source=](https://www.theregister.com/2023/02/07/github_ceo_ai_act/?utm_source=); “BSA Leads Joint Industry Statement on the EU Artificial Intelligence Act and High-Risk Obligations for General Purpose AI,” <https://www.bsa.org/news-events/news/bsa-leads-joint-industry-statement-on-the-eu-artificial-intelligence-act-and-high-risk-obligations-for-general-purpose-ai>.

③ Andreas Liesenfeld, et al., “Opening up ChatGPT: Tracking Openness, Transparency, and Accountability in Instruction-Tuned Text Generators,” Proceedings of the 5th International Conference on Conversational User Interfaces, Eindhoven, July 19–21, 2023, pp. 1–6.

合传统意义上开源系统的关键标准。<sup>①</sup>与此类似，法国的 Mistral AI 尽管最初以开源姿态示人，但其最新发布的旗舰模型 Mistral Large 在接受微软投资后，转为闭源并部署于 Azure 平台，偏离了其一贯立场。<sup>②</sup>因此，“伪开源”的表象混淆了公众与监管者对开放技术的认知判断，实质上构成了一种构建制度合法性的话语操控策略。这不仅削弱了外部社会力量的理解能力，更塑造了一个由科技企业设定的技术解释框架，使“透明”成为被技术包装的意识形态话语工具。

其次，从技术应用的现实出发可以发现，尽管模型本身可以开源，但部署模型所需要的资源依旧昂贵，特别是参数量较高（数十亿级别以上）的开源模型仍需依赖高性能 GPU 集群、分布式训练框架与专业云基础设施，还要求开源模型落地用户自身具备大量的专业人才，进行模型本地化部署所需要的微调。而这一部署的技术门槛却经常在开源话语中被系统性掩盖，导致用户和公众误以为开源等同于人人可用，进而在实际操作中形成认知落差。因而，开源并不必然意味着大规模部署人工智能系统的成本降低，这限制了“可重用性”的定义。<sup>③</sup>“开源”的错觉构成了工具性权力的社会化延伸，使科技企业不仅控制了技术流通的路径，还错误地塑造了人们对技术公平性的认知预期。

最后，即便在理论上能够达成模型代码与开放权重，人工智能系统本身的复杂性与不可预测性也使非专业主体难以真正理解其风险边界。模型输出受架构、超参数和语料分布等多重变量的共同影响，其涌现能力常常超出代码文本所能传达的知识。而这在本质上构成了一种认知封装机制，使科技公司在模型微调中依然主导了公众和监管机构对系统风险的理解框架，削弱了公众对开源的有效理解和掌控能力。应该看到，尽管部分开源模型确实推动了创新扩散，但开源叙事下的权力分配仍具有不对称性，科技公司可以通过自我标榜的开源姿态隐秘实施其工具性权力，使监管在面对一个在形式上开放、实质上封闭的认知空间时容易造成技术误判与治理错位。

#### （四）集体游说与责任转移的监管挑战

在全球首个人工智能监管法案欧盟《人工智能法案》颁布落地之前，欧盟在每

① Steven J. Vaughan-Nichols, "Meta Can Call Llama 2 Open Source as Much as It Likes, But That Doesn't Mean It Is," [https://www.theregister.com/2023/07/21/llama\\_is\\_not\\_open\\_source/](https://www.theregister.com/2023/07/21/llama_is_not_open_source/).

② Daphné Leprince-Ringuet, "Mistral's Deal with Microsoft Is Causing Controversy," [https://sifted.eu/articles/mistral-microsoft-deal-controversy?utm\\_source](https://sifted.eu/articles/mistral-microsoft-deal-controversy?utm_source).

③ David Gray Widder, et al., "Open (for Business): Big Tech, Concentrated Power, and the Political Economy of OpenAI," <http://dx.doi.org/10.2139/ssrn.4543807>.

个政策制定阶段几乎都面临着激烈的企业游说。在欧盟首次提出人工智能监管不到五年的时间内，OpenAI 和谷歌等科技公司就已经成为欧洲人工智能领域的顶级游说者。科技公司在欧盟机构中拥有不成比例且往往为最高层级的政策访问权，这是它们政治维度下工具性权力的典型表达：科技公司不再仅仅是政策对象，而是会通过游说干预和重构监管语义等路径反对甚至暗中破坏监管的严格要求。

首先，在通用人工智能的监管问题及责任转移问题上，下一代通用人工智能以大型语言模型为起点，具有特殊风险且最终用途无法确定，因此对其的监管适用已成为科技公司游说的关键议题。早在 2022 年，谷歌副总裁玛丽安·克罗克（Marian Croak）就公开表示“通用模型不应受欧盟高风险系统规则制约”。<sup>①</sup> 此后，微软和 OpenAI 等公司通过频繁会谈和发布文件意见，力图影响通用模型风险分类的立法逻辑。<sup>②</sup> 在 2023 年的欧洲“三方对话”中，<sup>③</sup> 科技公司工具性权力的运作表现得尤为典型：微软提出“豁免未投放市场的基础模型”来试图规避责任追溯路径，<sup>④</sup> 而 Mistral AI 和 Aleph Alpha 等欧洲本土初创科技公司也以“打造欧洲人工智能冠军”为名，要求降低对模型开发监管的要求。最终，科技公司的集体游说获得了成功，欧盟《人工智能法案》在人工智能价值链的责任公平分担方面，为通用人工智能系统制定了特殊规则，模型提供商只需要简单声明它们的系统只执行准备性或程序性任务即可，这便逃避了《人工智能法案》的核心即应对高风险人工智能系统的要求。<sup>⑤</sup> 此外，用于市场发布前研发的模型也不受法规约束。<sup>⑥</sup>

① 谷歌在提交给欧盟委员会的文件中也再次重申通用人工智能系统本身并不具备很高风险。参见 Camille Schyns, “The Lobbying Ghost in the Machine: Big Tech’s Concerted Defanging of Europe’s AI Act,” <https://corporateeurope.org/sites/default/files/2023-03/The%20Lobbying%20Ghost%20in%20the%20Machine.pdf#page=28.09>。

② 2023 年 3 月，欧洲议会议员提出了通用人工智能相关义务的问题，仅在当月，谷歌、微软和 OpenAI 各自与欧洲议会议员举行了四次会议讨论相关议题。参见 “Byte by Byte: How Big Tech Undermined the AI Act,” <https://corporateeurope.org/en/2023/11/byte-byte>。

③ “三方对话”是指欧盟委员会、欧盟理事会和欧洲议会在《人工智能法案》进入立法程序前的最后阶段进行的三方谈判。

④ 微软的观点参见 “A Balancing Act: Regulating AI to Boost Responsible Innovation in Europe,” <https://blogs.microsoft.com/eupolicy/2021/09/16/a-balancing-act-regulating-ai-to-boost-responsible-innovation-in-europe/>; “EU Artificial Intelligence Act: Microsoft Trilogue Recommendations,” <https://corporateeurope.org/sites/default/files/2023-11/Microsoft%20-%20Key%20trilogue%20recommendations%20-%20Oct%202023%20update.pdf>。

⑤ “Europe’s Approach to AI Regulation: Embracing Big Tech and Security Hardliners,” <https://algorithm-watch.org/en/ai-regulation-europe-big-tech-and-security/>。

⑥ Benjamin Cedric Larsen and Sabrina Küspert, “Regulating General-Purpose AI: Areas of Convergence and Divergence Across the EU and the US,” <https://www.brookings.edu/articles/regulating-general-purpose-ai-areas-of-convergence-and-divergence-across-the-eu-and-the-us/>。

其次，在基础模型的风险自我评估问题上，欧盟《人工智能法案》原本准备要求“基础模型”的提供者对健康、安全、基本权利、环境、民主和法治的合理进行可预见风险评估。<sup>①</sup>然而面对这些要求，科技公司以监管阻碍现有基础模型开发为由，共同反驳了欧洲议会有关提高透明度、披露版权材料以及报告环境影响的三项要求。<sup>②</sup>OpenAI 首席执行官萨姆·奥尔特曼（Sam Altman）以“停止在欧盟运营”相威胁，谷歌宣布了一项关于不可强制执行的自我监管承诺；<sup>③</sup>Meta 甚至在提交的对草案修改意见的文件中单独列出了“Meta 的建议”；<sup>④</sup>微软则表示《人工智能法案》应该推进基础模型，并只需要关注市场上和预期中最强大的模型类别即可。<sup>⑤</sup>与此同时，游说者也加大了同欧盟委员会与欧洲议会高层的接触力度。<sup>⑥</sup>此外，欧盟《人工智能法案》的草案曾经将银行、保险、就业、教育、移民和执法等八类人工智能系统划分在内，然而 Uber 和 Snap 等处应用层的科技公司表达了对其业务领域被划定为高风险的强烈担忧，主张以自我评估代替外部评估。<sup>⑦</sup>最终科技公司的主张再次获得回应，《人工智能法案》条文还是采纳了部分自我评估机制，赋予模型提供商“内部评估与对抗性测试”的主导权，并指出“任何市场监督机构均可出于特殊原因授权将特定高风险人工智能投放市场”。<sup>⑧</sup>这意味着在技术最核心层面，科技公司既是模型的开发者的开发者，也被允许作为模型风险的判定者，从而保留了合规解释空间并模糊了监管边界。这种规则上的可协商性与责任界定上的内嵌性反映出科技公司行使工具性权力的新趋势：通过技术话语与创新压力构建干预的合法性，

① 基础模型是指“在大规模的广泛数据上训练的、旨在提供通用输出并可适应各种不同任务的人工智能模型”。参见 European Parliament, “DRAFR Compromise Amendments on the Draft Report,” [https://www.europarl.europa.eu/meetdocs/2014\\_2019/plmrep/COMMITTEES/CJ40/DV/2023/05\\_11/ConsolidatedCA\\_IMCOLIBE\\_AI\\_ACT\\_EN.pdf](https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05_11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf)。

② “Byte by Byte: How Big Tech Undermined the AI Act,” <https://corporateeurope.org/en/2023/11/byte-by-byte>。

③ Natasha Lomas, “Google to Work with Europe on Stop-Gap ‘AI Pact’,” <https://techcrunch.com/2023/05/24/eu-google-ai-pact/>。

④ “Meta’s AI Act Position—4 Column Document,” <https://corporateeurope.org/sites/default/files/2023-11/Meta%20-%20AI%20Act%20Position%20-%204%20Column%20Doc%20-%20September%202023%20.pdf>。

⑤ “EU Artificial Intelligence Act: Microsoft Trilogue Recommendations,” <https://corporateeurope.org/sites/default/files/2023-11/Microsoft%20-%20Key%20trilogue%20recommendations%20-%20Oct%202023%20update.pdf>。

⑥ 科技公司游说者就《人工智能法案》与欧盟委员会和欧洲议会高级官员的会面情况，参见“Byte by Byte: How Big Tech Undermined the AI Act,” <https://corporateeurope.org/en/2023/11/byte-by-byte>。

⑦ “Meeting with Uber on GDPR, AIA and PWD,” <https://corporateeurope.org/sites/default/files/2023-11/Uber%20-%20Commission%20meeting%2022%20March%202023.pdf>；“Letter to Committee,” <https://corporateeurope.org/sites/default/files/2023-11/Snap%20-%20Letter%20to%20Commission%207%20July%202023.pdf>。

⑧ European Commission, “Artificial Intelligence—Questions and Answers,” [https://ec.europa.eu/commission/presscorner/detail/en/QANDA\\_21\\_1683](https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_1683)。

将合规转化为企业主导的技术程序。这不仅重构了监管的运作方式，也使监管规则面临被结构性预设与功能性稀释的长期风险。

## 六、大国战略竞争下人工智能监管中的国家与市场权力互动

在一个以不稳定和国家间长期竞争为特点的国际体系中，技术和科学发展时常被理解为创造经济和军事不对称的工具。<sup>①</sup> 尽管主权国家并不愿意看到科技公司出于自身利益来抵制监管，然而地缘政治竞争已经将国家安全与技术进步的利益紧密捆绑，技术监管也开始被塑造为国家战略创新和竞争力项目。<sup>②</sup> 在此背景下，主权国家与科技公司之间的权力边界不仅没有随着加强监管的呼声而趋于明晰，反而因战略竞争的现实需要变得更加模糊：科技公司不仅成为国家战略部署的支点，其技术权力也因国家竞争话语被重新赋予合法性，监管实践从原本的对抗性走向了政府与科技公司纠缠结构下的权力互动。这使得人工智能监管由此演化为一种制度场域中的国家—市场互动过程，不仅加剧了传统监管困境，也成为全球科技竞争格局中新型权力结构重组的重要特征。

### （一）欧盟监管的权力平衡与妥协

欧盟《人工智能法案》是其在人工智能时代试图依靠监管力量弥补自身技术差距、加强规范性力量的重要努力。然而，欧盟内部成员国在如何监管问题上的诉求并不一致，在高风险人工智能系统的风险划定与责任分配方面也出现了明显的政策分歧。尤其出乎意料的是，作为欧盟成员国“三巨头”的法国、德国和意大利在关键议题上持反对监管态度，为欧盟的人工智能监管方式的选择带来了实质性压力。

2023年6月以来，法国政府便持续推动放宽《人工智能法案》条款，认为现有方案可能阻碍初创科技公司创新，损害欧洲技术自主前景。以 Mistral AI、LightOn 和 Hugging Face 为代表的法国初创科技公司被法国政府寄予厚望，希望它们与美国 OpenAI、谷歌相竞争，这成为法国政府政策让步的重要推动力量。其中，Mistral AI 的重要游说者、前法国数字事务国务秘书塞德里克·奥（Cédric O）倡导的 150 家欧洲

<sup>①</sup> Robert Reich, "The Rise of Techno-Nationalism," *The Atlantic Monthly*, Vol. 259, No. 5, 1987, pp. 63-69.

<sup>②</sup> Tim Rühlig, "The New Geopolitics of Technical Standardisation: A European Perspective," *Future Europe*, Vol. 3, No. 1, 2023, pp. 102-109.

企业联署公开信声称，欧盟的《人工智能法案》“将危及欧洲的竞争力和技术主权”。<sup>①</sup>出于反对监管基础模型的意图，法国经济部也一直试图影响技术会议以调整相关条款。<sup>②</sup>

德国立场与法国趋同。为支持本土初创科技公司 Aleph Alpha，德国政府与法国、意大利一同在2023年10月会见欧洲科技行业代表，以督促欧盟《人工智能法案》必须以“创新友好型”方式为基础。德国副总理兼经济部部长罗伯特·哈贝克（Robert Habeck）提出，欧洲需要一项有利于创新的人工智能法规，包括如何应对通用人工智能在应用中面临的特定风险。哈贝克的观点呼应了科技公司“反对监管基础模型”以及“将风险转移至模型应用端而非开发端”的主张。<sup>③</sup> Aleph Alpha 也与欧盟和德国政府共同参加了多场有关人工智能监管的听证会，并就技术能力提出了多条建议。<sup>④</sup>

应该看到，欧洲的 Mistral AI 和 Aleph Alpha 两家公司在法德两国政府的支持下，以“打造欧洲人工智能冠军”为名，成功影响了欧洲“三方对话”进程，并推动了欧盟《人工智能法案》对开源基础模型监管做出必要让步。这不仅是安全与创新之间的平衡选择，更体现了人工智能领先国家对企业技术深度依赖下的纠缠结构：科技公司不再只是监管的对象，而是会在国家战略框架中被转化为合作者。2025年2月法国总统埃马纽埃尔·马克龙（Emmanuel Macron）在巴黎人工智能峰会上宣布了一项价值1130亿欧元的人工智能投资计划，并公开呼吁欧盟简化监管规则。<sup>⑤</sup>因此，国家与市场在人工智能领域的纠缠性深化体现为：国家在监管实践中已经无法摆脱对科技公司的战略依赖，这构成了科技公司人工智能技术权力合法化与扩张的通道。

① “Open Letter to the Representatives of the European Commission, the European Council and the European Parliament,” <https://drive.google.com/file/d/1wrtxfvcD9FwfNfWGDL37Q6Nd8wBKXCkn/view?pli=1>.

② Alexandre Piquard, “France Keeps up Pressure on EU’s AI Act, Despite Mounting Criticism,” [https://www.lemonde.fr/en/economy/article/2024/01/27/france-keeps-up-its-pressure-on-the-eu-s-ai-act-despite-mounting-criticism\\_6471038\\_19.html?random=1336260200](https://www.lemonde.fr/en/economy/article/2024/01/27/france-keeps-up-its-pressure-on-the-eu-s-ai-act-despite-mounting-criticism_6471038_19.html?random=1336260200).

③ “Italy, Germany and France Agree on Strengthening Their Cooperation on Artificial Intelligence,” <https://www.mimit.gov.it/en/media-tools/news/italy-germany-and-france-agree-on-strengthening-their-cooperation-on-artificial-intelligence>; “EU Artificial Intelligence Act: Microsoft trilogue recommendations,” <https://corporateeurope.org/sites/default/files/2023-11/Microsoft%20-%20Key%20trilogue%20recommendations%20-%200ct%202023%20update.pdf>.

④ Will Henshall, “EU’s AI Regulation Could Be Softened After Pushback from Biggest Members,” <https://time.com/6338602/eu-ai-regulation-foundation-models/>.

⑤ Scott Nover, “France Puts the AI in Laissez-Faire,” <https://www.gzeromedia.com/gzero-ai/france-puts-the-ai-in-laissez-faire>.

## （二）美国的宽松监管以及对科技巨头权力的工具性转化

与中国进行技术竞赛的压力是美国政府转变人工智能监管态度的关键影响因素。尽管美国曾在“科技抵制潮（techlash）”中推动对科技巨头的反垄断调查，<sup>①</sup>但随着生成式人工智能的快速突破，美国政府逐渐意识到人工智能技术竞争已成为衡量其全球技术领导力和保障国家安全的首要挑战，需要国家层面的努力来保持美国的技术优势。<sup>②</sup>在此背景下，政府对人工智能监管的关注从“抑制巨头”转向“巩固优势”，不再将科技公司的人工智能技术权力视为需要压制的私权力，而是主动将科技公司纳入国家战略体系，使其成为技术竞赛的重要执行主体。

美国政府制定人工智能标准的进程长期落后于其他国家。美国不存在一个通用政策来管理所有人工智能开发者和使用人工智能的组织，联邦和州对人工智能治理采取了明显不同的方法。<sup>③</sup>与欧盟明确的法案形式相比，美国采用了零散的治理框架和宽松的政策方法。虽然2023年10月拜登签署了人工智能总统行政令以确保人工智能安全、可靠和可信发展，但该行政令内容空泛且不具法律效力。除非国会通过广泛立法，否则行政令所能实现的目标有限。<sup>④</sup>2024年，美国国会虽然提出近百项人工智能立法，但始终未有实质性进展。与此同时，美国各州推动超过700项相关法律，但多数流于碎片化，如加利福尼亚州州长加文·纽森（Gavin Newsom）最终宣布否决了备受瞩目的《前沿人工智能模型安全与创新法案》。<sup>⑤</sup>出于对与中国竞争的考虑，美国没有采取过于严格的人工智能监管政策立场。<sup>⑥</sup>监管在美国被概念

① 美国国会指定众议院司法反垄断小组调查科技巨头，亚马逊、Facebook、苹果和谷歌都被进行了反竞争行为调查。参见“48 U.S. States Launch Antitrust Investigation into Google,” <https://www.cbsnews.com/news/google-antitrust-probe-48-u-s-states-launch-antitrust-investigation-of-google-dominance-in-search-ads-and-data/>; Rachel Lerman, “Google Targeted in Antitrust Inquiry by Attorneys General in 48 States,” <https://www.usatoday.com/story/tech/2019/09/09/google-targeted-antitrust-probe-coalition-state-ags/2266639001/>; Adi Robertson, “Congress Just Asked Big Tech Companies for Private Emails About Their Biggest Controversies,” <https://www.theverge.com/2019/9/13/20864135/congress-house-judiciary-committee-antitrust-probe-letters-apple-google-amazon-facebook>。

② “Artificial Intelligence and National Security,” [https://sgp.fas.org/crs/natsec/R45178.pdf?utm\\_source](https://sgp.fas.org/crs/natsec/R45178.pdf?utm_source)。

③ Christopher S. Yoo and Lai Alicia, “Regulation of Algorithmic Tools in the United States,” *Journal of Law and Economic Regulation*, Vol. 13, No. 2, 2021, pp. 7-22.

④ “The Biden Administration’s Executive Order on Artificial Intelligence,” <https://www.csis.org/analysis/biden-administrations-executive-order-artificial-intelligence>。

⑤ Joshua Turner and Nicol Turner Lee, “Misrepresentations of California’s AI Safety Bill,” <https://www.brookings.edu/articles/misrepresentations-of-californias-ai-safety-bill/>。

⑥ Ali A. Guenduez and Tobias Mettler, “Strategically Constructed Narratives on Artificial Intelligence: What Stories Are Told in Governmental Artificial Intelligence Policies?” *Government Information Quarterly*, 2023, DOI: 10.1016/j.qiq.2020.101719.

化为本国在人工智能竞赛中能够获得对其他国家的决定性战略优势的现实手段，这种竞赛类似于人类历史上涉及其他关键政治挑战的技术竞赛（如卫星发射或洲际弹道导弹）。而从这个角度来看，在美国政府监督机构被科技巨头所俘获并不是自然而然或心甘情愿的，行业领袖参与法规制定也构成了美国战略的一个重要方面。

在政治旋转门机制下，人工智能科技巨头与美国政府的关系进入新阶段，进而建立了更为密切的网络关系。作为科技巨头在国会中最坚定的盟友之一，参议员查克·舒默（Chuck Schumer）负责引导拨款法案在参议院获得通过。在2023年他阻止了一系列加强技术反垄断的调查，并鼓励科技公司进行自我监管。<sup>①</sup>同时，国会拨款部门的领导者也同样收到了来自科技公司的大量竞选捐款，他们的预算安排和态度可能决定这些公司的未来。<sup>②</sup>有更多具备科技公司工作经历的专业人士参与了人工智能监管标准制定的讨论。例如，美国联邦贸易委员会定期邀请科技公司的专家参与讨论，国家人工智能咨询委员会成员包括来自谷歌等公司的高级执行官。英特尔等40家科技公司组成了信息技术产业委员会游说集团，并成为美国跨党派智库信息技术和创新基金会的最主要支持者。在2024年美国大选的关键时刻，埃隆·马斯克（Elon Musk）与彼得·蒂尔（Peter Thiel）等硅谷右翼新贵以强大的社会影响力和资本鼎力支持了特朗普当选。而特朗普就任后旋即废除了拜登此前签署的行政令，颁布了新的《消除美国人工智能领导地位障碍》行政令，这也预示着美国人工智能监管未来将走向宽松。

随着地缘政治竞争压力的不断增大，美国政府与科技巨头在人工智能军事应用合作上的迫切需求也成为解读美国监管政策的重要切入点。2024年4月26日，美国国土安全部根据拜登的行政令成立了人工智能安全与保障委员会，该委员会旨在就人工智能在关键基础设施中的作用向政府提供建议，OpenAI、微软、Alphabet和英伟达等公司的首席执行官均作为成员列席。<sup>③</sup>2024年6月，美国国土安全部组建“人工智能军团”，成员包括10位具有科技公司任职经历的人工智能技术专家，这

① David Dayen, "How Chuck Schumer Deep-Sixed the Tech Antitrust Bills," <https://prospect.org/power/2023-01-26-chuck-schumer-tech-antitrust-bills/>.

② "Congressional Appropriators and Big Tech," <https://therevolvingdoorproject.org/new-report-congressional-appropriators-big-tech/>.

③ U.S. Department of Homeland Security, "Over 20 Technology and Critical Infrastructure Executives, Civil Rights Leaders, Academics, and Policymakers Join New DHS Artificial Intelligence Safety and Security Board to Advance AI's Responsible Development and Deployment," <https://www.dhs.gov/archive/news/2024/04/26/over-20-technology-and-critical-infrastructure-executives-civil-rights-leaders>.

是美国政府中规模最大、最具活力的人工智能技术团队。<sup>①</sup> 2024年12月，OpenAI扭转了其不与军方合作的政策，将美国政府部门视为合作伙伴，与美国军方就网络安全工作和其他相关项目达成协议，并宣布将与国防科技公司 Anduril 启动合作开发项目，提供人工智能技术用以帮助击落军事基地周围的无人机。<sup>②</sup>

进一步而言，美国国防部本身就是美国政府中人工智能支出的重点部门，美国国防部与科技公司在人工智能军事应用方面合作的重要性更是不言而喻。<sup>③</sup> 安全与军事部门在人工智能领域的积极活动与主动作为体现了美国为保证其作为全球技术创新领先者的地位并全面推动安全与军事体系人工智能化付出的努力。由此，美国政府与科技公司合作促成了一种国家与市场互动下的监管新逻辑，既能培育美国创新土壤又能提供必要监管，还可以巩固美国在人工智能全球技术领域的全球领先地位，并增强未来竞争力，以战略合力赢得人工智能竞赛的胜利。在这种逻辑下，美国监管制度对科技公司形成了反向依赖，国家与市场之间的边界正在进一步模糊，呈现出一种战略化和以合作为主导的权力互动机制。

### （三）国家与科技公司共同塑造人工智能互操作性以及全球规则的协同输出

在全球人工智能规则逐步成形的过程中，以“互操作性”为名的跨国协调机制正快速发展，其本质上是一种由国家和科技公司共同塑造的规则协同秩序。在国家间地缘政治竞争的显著影响下，科技公司成为国家输出基于价值共识的“互操作性规则”的关键代理，并借助其技术能力和平台优势，对全球治理标准进行主导性输入。

OpenAI、Anthropic 和谷歌等西方科技公司将其模型与它们自身定义的“人类价值观”保持一致，并与美国人工智能安全研究所和英国人工智能安全研究所等组织保持密切合作，不断完善安全协议。政府与科技公司之间的这一权力协同不仅服务于平台治理，还通过技术手段与接口规则在全球范围内形成了事实上的制度扩散机制。如 OpenAI 对中国开发者 API 访问限制已成为人工智能互操作性的典型限制。<sup>④</sup> 尽管中国从未被列入 OpenAI 官方支持的访问国家名单，中国开发者和科技公司此

<sup>①</sup> U. S. Department of Homeland Security, “DHS Hires First 10 Experts in ‘AI Corps’ Recruiting Sprint,” <https://www.dhs.gov/news/2024/06/25/dhs-hires-first-10-experts-ai-corps-recruiting-sprint>.

<sup>②</sup> Kyle Wiggers, “OpenAI Presents Its Preferred Version of AI Regulation in a New ‘Blueprint’,” <https://techcrunch.com/2025/01/13/openai-presents-its-preferred-version-of-ai-regulation-in-a-new-blueprint/>.

<sup>③</sup> Jacob Larson, et al., “The Evolution of Artificial Intelligence (AI) Spending by the U.S. Government,” <https://www.brookings.edu/articles/the-evolution-of-artificial-intelligence-ai-spending-by-the-u-s-government/>.

<sup>④</sup> “How OpenAI’s Decision Not to Operate in China Will Reshape the Chinese AI Scene,” <https://time.com/6992026/openai-china-ai/>.

前仍可通过代理服务器等替代方案绕过限制。但自2024年6月起, OpenAI 陆续加强了对中国开发者访问其 API 的限制。该举措不仅与美国维护国家利益的战略目标高度一致, 也契合美国拟议中的相关立法方向, 即加强对人工智能系统尤其是“可能对国家安全构成威胁的模型”的出口管控。类似地, 微软也在 Azure OpenAI 服务中集成了更严格的内容过滤和数据合规功能, 包括对 API 调用进行实时监控和对敏感数据进行加密存储等, 这使得使用 Azure 的开发者必须遵守微软预设的规则, 从而隐性输出了监管标准。由此可见, 科技公司凭借其在全球部署基础设施、API 控制和数据接入等方面的主导能力已成为重要的监管载体, 使主权国家能够将其价值观和标准内嵌于全球人工智能治理结构、推行符合其地缘政治利益的监管规则。

此外, 科技公司也日益参与到全球治理机制的共建进程中。在七国集团“广岛进程”、布莱切利人工智能安全峰会以及联合国人工智能高级别咨询机构等国际治理平台的讨论中, OpenAI、Google DeepMind、Meta 和 Anthropic 等科技公司代表频繁现身发表主旨演讲, 并对高风险人工智能、责任机制和模型评估方法等关键议题提供技术输入和立场表达。同时, Meta 和 IBM 牵头组建的“人工智能联盟”等行业组织也在推动非约束性准则逐步成为事实上的国际标准。科技公司日益成为国际人工智能治理进程中的“事实参与者”乃至“规则共同制定者”。

值得注意的是, 科技公司在推动规则向全球拓展时始终伴随着技术边界的有形设定以及对技术“他者”的政治化排斥。DeepSeek 在开源大模型上的算法创新和市场突破使中国开源人工智能技术已经被视为非西方技术能力的新变量, 并为美国及其盟友乃至整个西方科技行业带来了挑战。部分国家出于“国家安全”考虑已屏蔽或禁止政府设备访问 DeepSeek 应用, 美国航空航天局和美国海军等一些联邦机构已出于“国家安全”考虑指示员工不要使用 DeepSeek。<sup>①</sup> 同时, OpenAI 在提交特朗普政府“人工智能行动计划”的政策提案中呼吁立法禁止使用 DeepSeek, 并建议政府对本国的人工智能公司提供豁免权和实施三级出口管制。<sup>②</sup> 同样, Inflection 首席执行官兼 DeepMind 联合创始人穆斯塔法·苏莱曼 (Mustafa Suleyman) 公开表示, 美

① Matt Pearl, et al., “Delving into the Dangers of DeepSeek,” <https://www.csis.org/analysis/delving-dangers-deepseek>.

② “A Letter from Christopher Lehane (OpenAI) to Faisal D’Souza (NCO),” <https://cdn.openai.com/global-affairs/ostp-rfi/ec680b75-d539-4653-b297-8bcf6e5f7686/openai-response-ostp-nsf-rfi-notice-request-for-information-on-the-development-of-an-artificial-intelligence-ai-action-plan.pdf>.

国应利用其在半导体领域的领导地位作为“武器”来执行人工智能使用标准。<sup>①</sup> 在这种趋势下，科技公司的计算资源将成为美国实现其目标的最好杠杆。在美国式监管模式下，国家与科技公司间的协作正在推动全球人工智能监管割裂化发展，并试图将人工智能国际治理打造成具备排斥性与地缘色彩的割裂型治理体系。这种以科技公司手中人工智能技术权力为外部工具、以标准作为柔性边界的策略部署恰恰揭示了西方在国家与市场间、在技术竞争与规则设定中的权力协同逻辑。

综上，人工智能监管在大国战略竞争背景下正呈现出不同既往国家与市场权力的新型互动结构和演化轨迹，并向跨国监管领域蔓延。在国内层面，由于国家与市场存在天然的张力，科技公司凭借基于技术、经济、社会和政治力量的人工智能技术权力，与国家机构围绕监管展开权力角逐，并在一定程度上有效对抗并侵蚀了国家的部分权力；在国际层面，受竞争压力影响，科技公司作为关键技术的开发者和应用推动者往往成为国家战略的合作方，其权力也被相应地纳入国家权力，并体现在全球人工智能的竞争上。而在内外双重压力下，科技公司人工智能技术权力的相对权重不仅取决于经济竞争力目标和技术创新目标之间的权衡，也受制于国家自身的战略利益。为争夺人工智能治理话语权，当前存在一种从人工智能竞争到人工智能监管竞争的态势。而当各国制定各自监管规则并重申其主权之时，跨国监管的格局也会相应调整：在竞争压力下，在国际治理格局受制于政治阵营主导而逐渐分裂与碎片化的环境下，国际竞争压力深刻影响了人工智能监管的权力结构，使之日益复杂化与多元化，需要从更细致的角度加以审视。

## 七、结论

正如梅雷迪思·惠特克（Meredith Whittaker）所观察到的那样，“科技公司目前处于一个令人吃惊的有利位置，其人工智能产品在塑造我们的生活和制度的同时，也塑造着我们对人工智能及其相关业务的了解和不了解”。<sup>②</sup> 人工智能的监管问题归根结底就是如何把握人工智能发展轨迹以及谁来控制它的问题。然而现实中，科技公司正在将既有的技术权力转化为新的人工智能技术权力，进而获取接触和引导政策制定者的独特优势。因此，无论是从人工智能监管硬法规还是从治理的软规则制

<sup>①</sup> “US Should Use Chip Leadership to Enforce AI Standards, Says Mustafa Suleyman,” [https://www.ft.com/content/f828fef3-862c-4022-99d0-41efbc73db80?utm\\_source](https://www.ft.com/content/f828fef3-862c-4022-99d0-41efbc73db80?utm_source).

<sup>②</sup> Meredith Whittaker, “The Steep Cost of Capture,” *Interactions*, Vol. 28, No. 6, 2021, pp. 50-55.

定来看,当前西方国家的监管政策正日益呈现迎合领先科技公司利益的特征,其政策制定过程无形中为科技公司的权力扩张提供背书,不仅没有限制反而巩固了科技公司在技术领域的主导地位。而伴随未来人工智能技术复杂性的提升,人工智能监管将持续处于“专业性”被动状态,反而更有助于科技公司人工智能技术权力的扩张乃至形成某种形式的霸权。

人工智能监管历史非常短暂,需要在技术快速发展与创新、在确保安全和在多方利益整合之间实现多维平衡。虽然科技公司常以其人工智能技术权力巩固自身利益,但当这种权力置于国际竞争的格局中时便被赋予了新的战略含义,催生出国家与市场之间基于技术权力“公私协同”的新行为逻辑。这为理解某些国家在人工智能监管中的真实态度提供了启示:在战略竞争需求下,人工智能领先国家与科技公司都倾向于采取相对宽松的监管,政府不仅会默许企业的游说行为,并且会任命更多具有科技公司任职经历的专业人士参与人工智能监管标准制定。同时基于地缘竞争,国家和企业也正协同促进基于“价值共识”的监管互操作性并增加对外的排斥性,进而将技术政治推进至全球治理领域。

国家与市场的关系是国际政治经济学研究的核心议题,也是监管以何种方式存在的宏观背景。尽管科技公司有着基于自身利益的优先事项,然而将其人工智能技术权力置于国际社会之中便被赋予新内涵,促成了技术治理中公私合作的新实践。这也给我们带来不同的启示:与俘获理论不同的是,人工智能竞争中的国家已经是一种积极主动的力量,在技术的政治属性加强的背景下,国家监管不再以单纯的国内认知和公共利益逻辑来迎合商业利益或者政治选票,而是以“国际竞争力和市场化”的名义来扩大实施产业政策干预或是放松监管,进而塑造国家自身经济和技术全球竞争力的有利环境。在西方的政治经济环境下,公司权力和国家权力之间到底是互相侵蚀还是互相辅助?这一问题在未来可能越来越难以得到清晰的答案。但可以预见的是在这种复杂的权力博弈下,人工智能监管框架需要具有足够的动态性和政治性考虑以应对人工智能及其挑战。

(截稿:2025年3月 责任编辑:郭 泉)