



北京大学

本科生毕业论文

题目:

通用人工智能监管决策研究：
以欧盟《人工智能法案》为例

The Dynamics of Regulatory
Decision-Making for General Purpose AI
in the EU Artificial Intelligence Act

姓名: 王子健

学号: 2100015146

院系: 国际关系学院

专业: 国际政治 (国际政治经济学方向)

指导教师: 董昭华

二〇二五年六月

北京大学国际关系学院本科毕业论文导师评阅表

论文题目 (中文)	通用人工智能监管决策研究：以欧盟《人工智能法案》为例				
论文题目 (英文)	The Dynamics of Regulatory Decision-Making for General Purpose AI in the EU Artificial Intelligence Act				
导师姓名	董昭华	职 称	副教授	论文成绩 (百分制)	95
导师评语 (包含对论文的性质、难度、分量、综合训练等是否符合培养目标的等评价)	<p>人工智能技术领域的监管决策是近年来国际政治经济学领域关注的重要议题。论文以欧盟《人工智能法案》的决策过程为案例，采用过程追踪与一致性分析相结合的研究方法，探讨了欧盟立法过程中重大决策转向背后的政治逻辑。论文认为：人工智能监管中的风险界定过程不仅涉及技术判断，也受到政治价值与制度结构的重要影响。欧洲本土中小企业因其与欧盟相近的技术主权等价值观念，使其在立法过程中具有较大型跨国科技企业更强的决策影响力。</p> <p>论文选题具有前沿性和学理性。论文广泛参考欧盟机构和成员国发布的政策文件、官方声明和立法草案、相关企业的立场文件以及相关学术研究和新闻报道，资料翔实，结构清晰，写作规范，语言流畅，相较于既往的研究，提出了颇有新意的观点，是一篇较为优秀的本科毕业论文。</p> <p>论文在资料收集、案例追踪和研究方法上存在一定的难度，是对过去四年本科学习结果的一次有效考察，符合培养目标。</p> <p style="text-align: right;">导师签名：</p>				

内容提要

2022 年以来,以 ChatGPT 为代表的通用人工智能迅速崛起,伴随其机遇一同出现的风险与挑战成为了世界技术监管的新问题。为治理人工智能风险、增强欧洲竞争力,欧盟在 2021 年 4 月提出了人工智能法案(AI Act)的立法提案,试图对人工智能推行严格监管,通用人工智能自兴起后也成为了立法中重点讨论与被监管的对象之一。在 2023 年 6 月前,具有雄心壮志的欧盟立法机构在跨国大型科技企业的游说中仍然不为所动,坚持向通用人工智能分配严格义务。然而,在 2023 年 12 月欧盟三方会谈最终形成的临时协议中,立法机构却在重要的争议领域做出了重大妥协,推翻了诸多对人工智能的严格监管条款。本文使用过程追踪与一致性检验结合的研究方法,利用欧盟机构和成员国发布的政策文件、官方声明和立法草案,有关企业的企业立场文件、游说会议记录以及相关学术研究与新闻报道,细致探索了法案中通用人工智能监管的立法决策过程,试图解释欧盟立法作出上述重大转变的原因。分析发现,本土中小企业因其在游说渠道、信息可信度乃至价值观上的优势是其获得区别于大型跨国科技公司游说效果的重要原因,而决策中科技创新与技术监管间的矛盾实质上是一种风险界定冲突。这一认知对理解欧盟风险界定与监管偏好的形成路径、中小企业在新型技术监管中的独特力量、探索未来国际人工智能领域合作与冲突的边界均具有重要意义。

关键词 人工智能法案 欧盟 技术监管 风险治理 企业游说

Abstract

Since 2022, the rapid rise of general-purpose AI (GPAI), exemplified by ChatGPT, has introduced both opportunities and challenges, presenting new regulatory dilemmas for global technology governance. In response, the European Union proposed the Artificial Intelligence Act (AI Act) in April 2021, aiming to impose strict rules on AI systems to balance innovation with risk management and strengthen Europe's technological sovereignty. Initially, EU lawmakers remained resolute, resisting intense lobbying from major multinational tech firms and insisting on robust obligations for GPAI. However, the provisional agreement reached in the December 2023 trilogue marked a striking shift, significantly relaxing many of the earlier, more restrictive provisions. To unravel the reasons behind this sudden institutional pivot, this study adopts a process tracing and congruence testing approach, drawing on EU policy documents, official statements, legislative drafts, corporate position papers, lobbying records, as well as academic studies and media reports. The analysis reveals that local small and medium-sized enterprises (SMEs) played a pivotal role in this outcome, leveraging their advantages in lobbying access, informational credibility, and alignment with regional values to secure a more favorable regulatory framework. More fundamentally, this compromise reflects a deeper, often overlooked tension in the regulatory process – a conflict over how technological risks should be defined and prioritized, highlighting the broader dilemma between fostering innovation and safeguarding public interests. This perspective not only clarifies the EU's evolving approach to risk governance but also underscores the unique influence of SMEs in shaping the rules for emerging technologies, offering critical insights into the future contours of international AI governance.

Keywords Artificial Intelligence Act, European Union Studies, Digital Regulation, Risk Governance, MNEs Lobbying

目录

绪论	1
第一节 研究背景	1
第二节 研究问题的提出	3
第三节 选题意义与研究价值	4
一、理论意义	4
二、实践意义	4
第四节 研究方法	5
第五节 文献综述	7
一、欧盟人工智能法案研究	7
二、跨国公司与游说相关研究	10
三、技术监管相关研究	14
第一章 人工智能法案与通用人工智能监管	16
第一节 人工智能法案的起源、方法与挑战	16
一、技术风险与规范意愿	16
二、“基于风险的方法”	19
三、通用人工智能的突现性风险与制度反应	21
第二节 通用人工智能监管的主要争议场域	23
一、风险定位与监管方式	23
二、开源与研究型通用人工智能的豁免	24
第二章 决策互动结构与企业决策介入分析	26
第一节 监管决策互动结构与过程	26
一、主要利益攸关者	26
二、立法决策过程分析	27
第二节 企业决策介入分析	29
一、游说频率与游说手段	29
二、游说的竞争力逻辑	31
第三章 监管决策过程追踪	32
第一节 过程追踪节点划分	32

第二节 机制探索与有限妥协：2022年6月-2022年12月	34
第三节 游说高潮与文本严格转向：2022年12月-2023年6月	37
第四节 本土企业介入与最终让步：2023年6月-2023年12月	41
第四章 讨论与解释	48
第一节 本土中小企业与立法的妥协转向	48
一、独家游说渠道	48
二、信息策略与信息内容	50
三、技术信息的可信性	52
第二节 欧洲竞争力与企业—成员意见传导	54
第三节 技术风险界定与偏好表达	57
结语	60
参考文献	63
附录	73
致谢	75

仅供学习交流

绪论

第一节 研究背景

人工智能技术是 21 世纪迄今为止最具颠覆性的技术之一，不仅在语言理解、图像识别、自动驾驶等多个领域展现出快速演进的能力，更正以前所未有的速度重新塑造着全球政治经济。这一技术的迅猛发展不仅带来了巨大的经济机遇，也引发了深刻的监管挑战，在此背景下，欧盟在 2021 年 4 月率先提出了全球第一部试图全面监管人工智能的立法提案——《人工智能法案》（AI Act），确立以“基于风险的方法（risk-based approach）”为核心的监管框架。^①

人工智能法案最初将监管对象聚焦于“具备特定用途（intended purpose）”的高风险人工智能系统，强调技术安全性、数据治理、透明度与人类监督等关键指标，^②这一监管逻辑本质上建立在对人工智能系统功能可预期性的前提之上。但正是在人工智能法案草案提出后不久，以 OpenAI 发布的 ChatGPT 为代表的新一代通用人工智能（General Purpose Artificial Intelligence, GPAI）迅速崛起。就技术而言，通用人工智能的广泛适配性与不可预期性挑战了传统基于“特定用途”的人工智能监管方法，也对“基于风险的方法”提出了复杂的风险界定挑战。而在现实中，以荷兰福利算法丑闻为例的算法滥用也引发了监管者对算法系统性风险的高度警觉。因此，如何为通用人工智能制定恰当的监管措施成为了欧盟人工智能法案的立法决策中最重要的议题之一。^③

在推进人工智能监管的过程中，欧盟立法机构也遭遇了来自科技产业界的许多障碍。随着通用人工智能监管被纳入人工智能法案框架，围绕通用人工智能风险界定与监管手段、开源与研究型人工智能监管豁免等关键条款，各类利益相关者纷纷介入决策过程，在决策议程中构成了强大的外在干预因素。不仅以谷歌、微软为代表的大型跨国科技企业展开高强度的跨国游说，部分本土中小企业如

^① European Commission: “European Approach to artificial intelligence: AI Act”, <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>, 2025-05-12.

^② Martha Cantero Gamito, Christopher T Marsden, "Artificial Intelligence Co-Regulation? The Role of Standards in the EU AI Act", *International Journal of Law and Information Technology*, Vol.32, No.11, pp.1-19.

^③ Melissa Heikkilä: “Dutch scandal serves as a warning for Europe over risks of using algorithms”, March 29, 2022, <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>, 2025-05-12.

Mistral AI、Aleph Alpha 也在 2023 年中期积极介入立法谈判，通过与本国政府形成策略共同体推动政策立场上升为成员国共识，并深刻影响了决策的最终形态。

更具挑战性的是，通用人工智能的发展不仅推动了技术的进步，也引发了治理主权与权力分配的剧烈争夺。在美国，硅谷的大型跨国科技企业（如 OpenAI、谷歌、微软等）主导了大模型开发的绝大多数路径，并形成了基于算法能力、数据垄断与算力积累的治理话语霸权。这些公司普遍主张“轻监管”甚至“自我治理”的模式，强调“先部署、后纠正”的创新优先原则。^①在此语境下，科技公司不再是国家监管的被动对象，而是转化为主动塑造政策议程的制度行动者，通过旋转门机制、赞助研究、游说网络与公众动员等方式，影响人工智能治理的知识基础与政治立场。^②与此同时，美国“技术法西斯主义”逐渐兴起，技术背景官僚入主政府部门，进一步将科技产业利益与政府治理整合；技术巨头试图进一步腐蚀国家权力，以便迅速取而代之。^③在跨国大型科技企业的政治干预下，欧盟同样面对着遭遇监管捕获的风险。

2023 年 6 月前，欧盟立法机构始终维持强硬的监管立场，即便面对跨国科技巨头的密集游说攻势仍不为所动，在核心争议条款上均未作出实质性让步；然而，2023 年 12 月三方会谈（Trilogue）达成的临时协议却出现了戏剧性转折，原争议条款发生重大调整，多项高强度监管义务被废止，通用人工智能监管框架从最初的统一强监管模式转向约束力更弱的分层分类治理体系。这一立法转向不仅是单纯的技术调整，它既折射出新兴技术监管框架构建的探索性特征，也揭示了欧盟本土中小企业通过制度渠道介入决策进而制衡跨国科技巨头的博弈逻辑，还暴露出立法进程中多方主体在风险界定维度的深层冲突。理解这一立法转折的动因与机制，对于解析欧盟人工智能法案决策的内部动态、阐释企业作为国际政

^① Google: “Recommendations of Regulating AI”, May 2023, p.3.

<https://ai.google/static/documents/recommendations-for-regulating-ai.pdf>, 2025-05-12. Microsoft: “Microsoft’s AI Access Principles: Our commitments to promote innovation and competition in the new AI economy”, February 2024, <https://blogs.microsoft.com/on-the-issues/2024/02/26/microsoft-ai-access-principles-responsible-mobile-world-congress/>, 2025-05-12. OpenAI: “Frontier AI regulation: Managing emerging risks to public safety”, July 2023, <https://openai.com/index/frontier-ai-regulation/>, 2025-05-12.

^② Khanal Shaleen, Hongzhou Zhang et al., "Why and How Is the Power of Big Tech Increasing in the Policy Process? The Case of Generative AI", *Policy and Society*, 2024, pp.1-18.

^③ Becca Lewis, “‘Headed for Technofascism’: The Rightwing Roots of Silicon Valley,” *The Guardian*, January 29, 2025, <https://www.theguardian.com/technology/ng-interactive/2025/jan/29/silicon-valley-rightwing-tech-nofascism>, 2025-05-14.

治行为体的力量与行动逻辑以及探索新兴技术监管中的新现象、新方法均具有重要的理论与实践价值。

第二节 研究问题的提出

欧盟《人工智能法案》建立在风险导向框架下，其监管逻辑假设人工智能系统具有明确的预期用途，从而可以根据用途将系统划分为不同等级的风险类别，并据此制定相应的合规义务。然而，通用人工智能并不具备单一的功能目的，其强通用性与不可预期性使立法机构难以将其纳入既有风险框架。与此同时，通用人工智能被视为人工智能系统中最最有能力的一类，在商业、医疗等方面具有巨大的应用潜力，而其所携带的系统性风险与不可预测的影响也使得“是否监管”、“如何监管”成为人工智能法案内部争议最为激烈的问题之一，欧盟的监管决定也被认为将对未来的社会经济产生深远影响。^①

本文通过对人工智能法案自 2022 年 6 月至 2023 年 12 月间人工智能法案总通用人工智能监管决策过程的追踪发现：在早期，欧盟立法机构遵循统一的强监管路径，试图将通用人工智能整体纳入针对高风险人工智能的监管框架，并在大型跨国公司的高频游说下不为所动；而在 2023 年中期以后，随着以 Mistral AI 为代表的欧洲本土中小企业的高频介入，以及法国、德国、意大利等强势成员国的政策立场转向，通用人工智能监管路径逐步朝向分层分类、义务削减的妥协方案转化，最终在 2023 年 12 月达成临时协议，确立了对极少数的“高影响力通用人工智能”施加有限义务、对其余绝大多数通用人工智能则仅鼓励自我规范的弱监管结构。

因此，本文的核心研究问题是：**为什么在通用人工智能监管决策中，欧盟在初期推行全面的严格监管，后期却发生了重大的弱化与妥协？在这一过程中，本土中小企业如何介入政策制定过程，并影响最终的监管制度结果？**这一问题不仅涉及监管路径演化的因果解释，也触及在全球技术治理领域中，资源较弱的本土中小企业是否、以及在复杂的制度环境中施加影响的问题，更关注新兴技术治理中监管方与科技公司间互动引起的政策妥协与制度重构的关键议题，为理解新兴技术监管中围绕风险界定的争议本质提供了窗口。

^① Helberger Natali, Nicholas Diakopoulos, "ChatGPT and the AI Act", *Internet Policy Review*, Vol.12, No.1, 2023, pp.4-6.

第三节 选题意义与研究价值

一、理论意义

从理论上讲，本文首先尝试在既有的欧盟人工智能法案研究上进一步拓展对立法决策内部动态的理解。与主要聚焦法律条款和制度框架的研究不同，研究从国际政治经济学的视角切入，考察了立法决策过程中具有不同利益取向的行为体如何通过互动、博弈与妥协影响最终的立法结果。通过对决策过程的追踪与分析，研究将监管条款的演变与外部科技企业的干预联系起来，为理解这一复杂政策过程提供了新的解释框架，补充了现有研究对人工智能法案目的与影响的讨论。

其次，本文探索了中小企业在立法过程中的政治参与现象。不同于以往主要聚焦大型跨国企业的研究，研究利用人工智能法案这一案例，展示了中小企业如何在特定情境下发挥超越其规模影响力的作用，在一定程度上补充了企业政治参与与游说理论中对中小企业关注的相对欠缺。通过具体案例分析，研究尝试说明，即便在相对资源受限的情况下，中小企业也能够通过利用其在国家竞争力上的重要意义、在价值观接近性与信息可信度上的优势，在关键立法阶段实现对政策结果的实质性影响。

最后，从技术监管的角度来看，本文试图表现人工智能监管中技术创新与基本权利保护之间的复杂权衡过程。不同于传统的产品监管或行业规范，人工智能技术的广泛适用性、不可预测性和潜在的系统性风险使得监管面临更大的不确定性和动态性。基于对过程追踪与决策转向的分析，研究尝试深化对这一复杂平衡的理解，并探索在强价值观监管导向与多种利益主体互动的情况下，监管机构如何在动态环境中进行政策选择。这一视角不仅有助于补充新兴技术治理领域的研究理论，也为理解技术治理中的利益协调与制度选择提供了新的思考路径，同时为揭示通用人工智能监管争端的风险界定本质提供了案例支持。

二、实践意义

从实践上看，人工智能法案是全球首部拟对人工智能建立全面规制的立法，其治理路径无疑将对其他国家和国际组织构建本国（或区域）人工智能治理体系产生重要示范效应。从数字服务法（DSA）到数字市场法（DMA），欧盟已通过“规范先行”的方式在全球数字治理领域形成了制度输出路径，而人工智能法案也是这一战略在人工智能领域的延续。尤其是对通用人工智能的纳入与监管路径的探索，不仅关涉欧盟“数字主权”战略的落地执行，也有望在全球科技制度

竞赛中重塑治理范式。因此，深入考察通用人工智能监管路径从统一强监管向分层弱监管这一转变的内部动态，将为我们理解欧盟数字政策背后的制度逻辑、权力互动与战略取向提供关键视角。

其次，当前以硅谷为中心的大型跨国科技企业正在推动一套“以市场为主导、技术即正义”的治理逻辑，而欧盟在人工智能法案的立法过程中尝试建构一个强监管框架，力图通过以风险为基础的原则将各式各样的人工智能系统纳入其严格监管范围内，以对冲通用人工智能的黑箱特点与巨大的滥用、误用风险。虽然这一雄心在决策过程中反复遭到来自外部科技公司与商业集团的冲击，并最终以一个经历高度妥协的形态呈现，但欧盟对人工智能的全面监管框架仍然走在全球技术治理前沿，在强大的科技公司力图渗透监管决策的背景下，能够为类似监管决策提供方向与策略上的参考。

最后，随着人工智能监管成为地缘政治竞争的新前沿，欧盟在标准设定上依然具有重要话语权。理解欧盟监管的决策演化与内部动态，有助于中国更好判断未来国际合作的空间、冲突的边界。通用人工智能的高度不确定性风险，引发了欧盟对技术安全与风险管控的双重关注。这一安全概念至少包含两重内涵：其一是欧盟在中美人工智能竞争中的技术主权与本土化发展需求，其核心在于通过构建本土人工智能生态体系，摆脱对外部技术的依赖；其二是通过监管框架控制通用人工智能风险的制度需求，而风险控制的实现路径，本质上取决于欧盟对风险的界定逻辑与监管方法论（即监管框架的设计思路）。因此，要理解欧盟对通用人工智能技术的风险界定逻辑与管控路径的形成机制，欧盟内部行为体的利益偏好分布形态，及其偏好通过欧盟制度体系进行加总与表达的具体过程。这一分析有助于中国在人工智能时代把握作为中美竞争大背景下第三股力量的欧盟的风险认知，进而明确未来人工智能时代中国与欧盟间合作与冲突的潜在方向。

第四节 研究方法

在研究方法上，本文采用过程追踪（*process tracing*）与一致性分析（*congruence analysis*）相结合的研究方法，以识别影响欧盟人工智能法案中通用人工智能监管立法决策转向的关键因果机制，并理解本土中小企业在该过程中施加影响的方式与路径。

过程追踪是一种通过对个案进行详细的案例内实证分析来追踪因果机制的方法，它通过系统地检验根据研究问题和假说选择和证据来考察特定案例

中的因果联系，其主要目的是通过重建从假定的事件或因素链或因素来弄清“是什么造成什么”。^①图里亚·法莱蒂（Tulia Falleti）等人在其对政治现象因果机制的研究中认为，过程追踪使人们能够撬开政治现象因果关系的黑匣子，以更精确地发现一个因素或现象与另一个因素或现象如何相关。^②

本文使用过程追踪的方法重建自 2022 年以来欧盟通用人工智能监管立法决策演化的关键节点与因果链条，力图完整呈现人工智能法案的立法争议、外部干预与决策转向的动态过程。研究将关键时间节点的决策演变与行为体干预的互动关系构建为若干反映特定因果机制的阶段性片段，具体涵盖政策倡议的提出、草案内容的修订调整、技术标准争议的显现、制度性妥协的达成等关键环节。通过这些片段的拼接与校验，研究试图还原游说活动如何在政治过程的不同阶段被吸纳、放大或抵制，并最终反映到政策结果之中。

一致性分析则蕴含两层比较分析的内涵。其一，将从理论或解释性框架中推演的预测与特定案例的实证观察进行对照；其二，对不同竞争性解释展开比较，以检验理论预设因素在政治变迁中的实际作用。^③参考约尔希姆·布拉特（Joachim Blatter）等人提出的框架，我们假定多个因素在特定制度环境中相互作用，共同产出政策转向结果。^④

研究过程中，本文追踪并分析将特定游说行动、政策争议和制度反馈与最终结果联系起来的关键事件与机制，识别可能构成导致立法决策转折的必要或充分条件的因素。研究将理论框架下关于不同类型企业影响路径的预测，与实际政策变化过程进行对照。通过比较竞争性解释（独家游说渠道、信息策略与信息内容、信息的可信度等）与实证材料之间的一致性程度，验证各解释路径在特定政策阶段中的解释力，进而评估其对最终妥协方案形成的因果贡献。

本研究在资料使用方面综合运用一手与二手材料，以确保对决策过程的深入把握。一手资料涵盖了欧盟机构、成员国政府及相关专家发布的关键政策文件、磋商意见、官方声明与立法草案，以及介入决策的企业立场文件等。研究基于这

^① Kimberly J. Morgan, "Process Tracing and the Causal Identification Revolution," *New Political Economy*, Vol.21, No.5, 2016, pp.489-492.

^② Tulia G Falleti, Julia F Lynch, "Context and Causal Mechanisms in Political Analysis", *Comparative Political Studies*, Vol.42, No.9, 2009, pp.1143-1166.

^③ Joachim Blatter and Till Blume, "In Search of Co-variance, Causal Mechanisms or Congruence? Towards a Plural Understanding of Case Studies," *Swiss Political Science Review*, Vol.14, No.2, 2008, pp.315-356.

^④ Joachim Blatter, Markus Haverland, *Designing Case Studies*, Palgrave Macmillan UK, 2012, p.22.

些材料构建了清晰的时间序列线索，建立了政策采纳逻辑与争议焦点的核心基础。二手资料则包括权威媒体报道与相关学术研究，用于补充事件背景并进一步揭示各行为体的偏好。为追踪因果路径、识别关键变量，研究将整个政策过程按时间维度划分阶段，系统梳理了各行为体在不同阶段的互动方式、博弈内容及其产出结果，并以此构建政策演化的时间序列图谱，揭示制度转向的关键驱动因素及其背后的机制逻辑。

综上所述，过程追踪法提供了重建因果链条的路径视角，一致性分析则用于理论机制的验证。本文在研究方法上将两者结合，试图在解释通用人工智能监管议程的制度变迁时兼具过程精确性与理论解释性。

第五节 文献综述

一、欧盟人工智能法案研究

当前关于欧盟《人工智能法案》的学术研究大致可分为三个主要方向：监管目的、监管影响与决策内部过程，三者分别对应了该立法在理念目标、实施机制与政治过程三个层面的核心问题。需要说明的是，由于法案的监管技术(regulation techniques)与本文主要讨论的问题并不相关，此部分并未包含对人工智能法案做法律分析的文献。

首先，在监管目标层面，学者们对立法在应然层面上的目的做了多方面的讨论，包括在构建可信赖的人工智能生态、风险规避与基本权利保护以及维护数字主权与全球影响三点。例如，玛莎·卡米托(Martha Gamito)等人通过对人工智能法案立法中技术标准的设定研究指出，这一法案通过提出实现可信人工智能的要素和要求(如遵守法律法规、符合伦理准则、技术和社会稳健性等)，旨在建立一个可信任的人工智能法律框架与信任生态系统，促进欧洲人工智能的创新与可持续发展。^①而这一可信的人工智能生态也将有助于实现风险规避与基本权利保护的目标。伊莎贝尔·库什(Isabel Kusche)基于人工智能法案的草案内容，分析了基于风险的AI监管的潜力和限制，展示了基本权利保护在人工智能法案中相当于法律规则的基础地位，同时表明了基于风险的逻辑在法案中的基石作用。^②殷佳章等人进一步认为，上述对伦理准则的强调实质上是欧盟在人工智能技术

^① Martha Cantero Gamito, Christopher T Marsden, "Artificial Intelligence Co-Regulation? The Role of Standards in the EU AI Act", pp.1-19.

^② Isabel Kusche, "Possible Harms of Artificial Intelligence and the EU AI Act: Fundamental

研发和投资优势消失情况下的替代策略，欧盟试图通过其相对优势的规范实力把握人工智能时代的发展机遇。^①最后，许多学者援引欧盟近年来提出的“数字主权”（digital sovereignty）概念，指出欧盟的人工智能立法实际上是争取人工智能技术上的充分自主权与人工智能监管上的领先地位的表现。丹尼尔·穆格（Daniel Mügge）在对人工智能法案的文本分析中发现，数字主权（或人工智能主权）的概念在话语和政策层面都扮演着关键角色，人工智能法案展现出了欧盟强烈的技术自主的愿望。^②丹尼斯·布罗德斯（Dennis Broeders）等人则从地缘政治角度出发，通过对欧盟多个规范性立法的分析，确定了欧盟技术政策逐渐地缘政治化的趋势，同时描述了制定人工智能政策中的基本权利、自由市场与地缘政治间的互动。^③可以看出，在技术创新和风险控制之间，欧盟更加偏向后者。^④

其次，在监管影响层面，学界主要关注这一法案对技术创新发展以及欧盟地位与全球治理两方面的影响。就技术创新发展而言，虽然仍然有文献讨论人工智能法案中监管沙盒与新设管理机构与对技术创新的潜在促进作用、^⑤法案创造出的公平市场对创新竞争的良性引导效果等正面效用，^⑥但作为世界第一部全面的人工智能监管立法，学者们多关注其对技术创新可能产生的负面影响。崔宏伟对欧盟“数字技术政治化”的研究指出，尽管欧盟在监管方面领先，但未来的严格监管框架可能抑制技术创新和阻碍外国投资。^⑦吴桐、刘宏松基于对人工智能法案立法进程与监管机制的梳理与分析认为，在积极作用上，法案中的制度创新设计不够有力，而在消极影响上，法案中过度的监管会增加数字企业的合规成本和创新压力，提高初创企业进入市场的门槛。^⑧在此基础上，宫云牧结合欧盟数字

Rights and Risk,” *Journal of Risk Research*, 2024, pp.1–14.

^① 殷佳章、房乐宪：《欧盟人工智能战略框架下的伦理准则及其国际含义》，《国际论坛》2020年第2期，第18–21页。

^② Daniel Mügge, "EU AI Sovereignty: For Whom, to What End, and to Whose Benefit?," *Journal of European Public Policy*, Vol.31, No.8, 2024, pp.2200–2225.

^③ Dennis Broeders, Fabio Cristiano and Monica Kaminska, "In Search of Digital Sovereignty and Strategic Autonomy: Normative Power Europe to the Test of Its Geopolitical Ambitions," *JCMS: Journal of Common Market Studies*, Vol.61, No.5, 2023, pp.1261–1280.

^④ 戚凯、崔莹佳、田燕飞，《美欧英人工智能竞逐及其前景》，《现代国际关系》2024年第5期，第118–125页。

^⑤ Celso Cancela-Outeda, "The EU's AI Act: A Framework for Collaborative Governance," *Internet of Things*, Vol.27, 2024, pp.1-9.

^⑥ Phillip Hacker, *AI Regulation in Europe: From the AI Act to Future Regulatory Challenges*

^⑦ 崔宏伟：《“数字技术政治化”与中欧关系未来发展》，《国际关系研究》2020年第5期，第21–40页。

^⑧ 吴桐、刘宏松：《地缘经济转向、数字主权与欧盟人工智能治理》，《国际安全研究》2024年第5期，第81–108页。

治理中的数字主权概念指出，针对大型平台的反垄断监管也可能反过来限制欧洲本土的大型数字企业发展。^①

在欧盟地位与全球治理上，欧盟被期望树立全球人工智能监管的标杆，再通过“布鲁塞尔效应”在全球范围内输出其监管标准，巩固其在全球技术监管竞争中的中心地位。^②李舒（Li Shu）等人在对欧盟通用数据保护条例（GDPR）通过市场机制推动国际监管趋同的机制梳理上，分析了人工智能法案的立法文本中人工智能的定义范围、风险评估方法等内容，指出人工智能法案同样将引发“布鲁塞尔效应”，促使新兴经济体的企业和政府接受欧盟的人工智能监管标准。^③然而，一份对人工智能法案域外影响的分析指出，由于现有国际市场的多极发展局面、国际标准机构的多样存在以及外国政府在规则制定上的努力三点因素，这一法案虽然依旧能够产生全球影响，但不太可能在人工智能法案上获得和通用数据保护条例一样的布鲁塞尔效应。^④从既有文献中可以看出，人工智能法案被预期产生有效的布鲁塞尔效应、巩固欧盟全球领导者地位与促进全球人工智能治理革新，但具体的效果仍有待验证。

最后，部分研究也关注了人工智能法案内部的决策过程。弗朗切斯卡·帕米尔托（Francesca Palmiotto）对人工智能法案的立法过程进行了过程追踪，指出了欧盟不同机构在基本权利保护上的不同偏好与立法过程中基本权利保护程度的起伏，发现人工智能法案的最终版本并未如预期地有效保护欧盟公民的基本权利。^⑤乔纳斯·特尔伯格（Jonas Tallberg）等人则使用了定量方法，利用人工智能法案的公众咨询数据首次分析了非国家行为体内部在立法决策过程中的不同偏好。^⑥类似地，安娜玛丽亚·纳塔萨（Anamaria Nastasa）等人分析了上百名利

^① 宫云牧：《数字时代主权概念的回归与欧盟数字治理》，《欧洲研究》2022年第3期，第21-25页。

^② Charlotte Siegmann, Markus Anderljung, “The Brussels Effect and Artificial Intelligence: How EU regulation will impact the global AI market”, Report by Centre for the Governance of AI, August 2022, p.3.

^③ Shu Li, Béatrice Schütte, Suvi Sankari: “The ongoing AI-regulation debate in the EU and its influence on the emerging economies: a new case for the ‘Brussels Effect’?”, in Mark Findlay, Li Min Ong, Wenxi Zhang, *Elgar Companion to Regulating AI and Big Data in Emerging Economies*, Edward Elgar Publishing, 2023, pp.34-37.

^④ Pagallo Ugo, “Why the AI Act Won’t Trigger a Brussels Effect”, in *AI Approaches to the Complexity of Legal Systems*, Springer, forthcoming, 2024, pp.3-10.

^⑤ Francesca Palmiotto, “The AI Act Roller Coaster: The Evolution of Fundamental Rights Protection in the Legislative Process and the Future of the Regulation,” *European Journal of Risk Regulation*, 2025, pp.1-24.

^⑥ Jonas Tallberg, Magnus Lundgren and Johannes Geith, “AI Regulation in the European Union: Examining Non-State Actor Preferences,” *Business and Politics*, Vol.26, No.2, 2024, pp.218-239.

益相关者对人工智能法案立法提案的反馈，展示了包括透明度、人工智能应用与责任分配等利益相关方的普遍关切，同样揭示了包括企业、非政府组织与研究机构在内的非国家行为体的不同偏好。^①最后，在较为普遍的理论探索中，罗尼特·加斯托-哈那尼（Ronit Justo-Hanani）探讨了以人工智能法案为代表的一系列欧盟监管改革，通过对 2017 年起的欧盟数字监管改革的过程追踪，检验并验证了竞争利益、欧盟制度结构以及本土偏好三个主要因素在塑造欧盟数字监管改革中的共同作用。^②对人工智能法案决策内部动态的探讨在一定程度上展现了欧盟内部国家行为体间与非国家行为体间利益与偏好的复杂分野，以及这些分野对最终决策产生的潜在影响。

在人工智能法案的研究中，研究者对法案的目标和影响做了相对全面的分析，指出了欧盟建立内部统一人工智能市场、控制人工智能风险、输出欧盟规范的主要目的，对这一立法的正面与负面影响也做了相对全面的分析。然而，较少研究深入了人工智能法案的立法过程中观察其内部行为体动态，现有的观察一部分关注行为体偏好，但并未将其置于立法全程的动态过程中，另一部分观察了动态过程，但主要关注立法中的价值观因素，没有将这一过程与立法中的妥协转向并置观察。

二、跨国公司与游说相关研究

在全球化时代，跨国公司成为国家作为唯一权威的衰落的最重要的受益者，他们利用自身的经济与资源优势不断反哺自身的政治实力，并借此介入全球治理与各主权国家内部的政治过程。^③从框架上来看，国际政治经济学对跨国大型跨国科技企业与企业游说的研究已经相对成熟，第二代国际政治经济学学者在跨国公司-国内政治研究中呈现了两个重要转变：一是在分析主体上将母国的利益与跨国公司本身的利益拆分看待，二是在分析对象上从一般关系的分析转向不同领域（例如人工智能技术）的分析。^④后续兴起的从产业或生产要素所有者的角度来理解国家的经济政策选择的开放经济政治学（OEP）范式搭建了个体利益上升转化为政治制度的理论桥梁。^⑤本文建立在这一基于利益的框架上。观察科技公

^① Anamaria Nastasa, Maer Matei Monica Mihaela et al., "Artificial Intelligence: Friend or Foe? Experts' Concerns on European AI Act", *Journal of Economic Computation and Economic Cybernetics Studies and Research*, Vol.57, No.3/2023, 2023, pp.5-22.

^② Ronit Justo-Hanani, "The Politics of Artificial Intelligence Regulation and Governance Reform in the European Union," *Policy Sciences*, Vol.55, No.1, 2022, pp.137-159.

^③ Doris A. Fuchs, *Business Power in Global Governance*, Lynne Rienner Publishers, 2007, pp.43-59.

^④ 王正毅：《国际政治经济学通论》，北京：北京大学出版社 2010 年版，第 244-245 页。

^⑤ David A. Lake, "Open Economy Politics: A Critical Review," *The Review of International*

司关于科技公司的政治参与文献不一而足，下文重点关注了在人工智能这一新兴技术背景下对科技公司政治参与的研究。

在人工智能技术迅速发展的背景下，人工智能技术创新越来越集中于私营部门而非大学等公共研究机构，特别是部分大型跨国科技企业。虽然大型跨国科技企业常常同时与顶尖大学合作以促进创新，但技术巨头实际上掠夺了合作中大部分的技术与智力产出，这一以大型跨国科技企业为中心的技术流的活动本质是为大型跨国科技企业的继续扩张创造有利条件。^①诺尔·阿麦德(Nur Ahmed)等人提出了影响人工智能研究的三个关键要素：计算能力、大型数据集和高技能的研究人员，而这三个方面日益受私营部门主导。^②例如，高级研究人员从学术部门向技术巨头的正流动不断增长，以 Google 为例的技术巨头在国际会议中贡献了越来越多的技术论文，人工智能知识的私有化趋势明显。^③另外，人工智能训练所需的巨额成本和高质量数据集拉大了技术巨头和公共研究机构间的研究能力差距，部分技术巨头也通过收购人工智能初创公司的方式不断巩固其行业和科研地位。^④

大型跨国科技企业的人工智能能力集中(或垄断)如何转化为权力？卡纳尔·沙林(Khanal Shaleen)等人提出了一个基于政策制定流的分析框架，认为政策制定是一个非线性过程，由问题流、政策流和政治流三个独立的流构成。技术巨头利用在技术流中的巨大优势渗透到政策流的三个过程中，提升自身权力。^⑤在问题流中，大型跨国科技企业通过控制社交媒体与推送算法，从而有效地影响什么问题会受到公众和学界的关注，对 700 多篇与大型跨国科技企业和 2019 年至 2020 年反垄断运动相关的文章进行的研究表明，与风险相比，新闻媒体倾向于更多地强调大型跨国科技企业的正面影响。^⑥

Organizations, Vol.4, No.3, 2009, pp.219–244.

^① Cecilia Rikap, “Varieties of Corporate Innovation Systems and Their Interplay with Global and National Systems: Amazon, Facebook, Google and Microsoft’s Strategies to Produce and Appropriate Artificial Intelligence,” *Review of International Political Economy*, Vol.31, No.6, 2024, pp.1735–1763.

^② Nur Ahmed, Muntasir Wahed and Neil C. Thompson, “The Growing Influence of Industry in AI Research,” *Science*, Vol.379, No.6635, 2023, pp.884–886.

^③ Jurowetzki Roman, et al., “The Privatization of AI Research(-Ers): Causes and Potential Consequences”, published in Arxiv, 2021, pp.30-33.

^④ Agora: A Unified Asset Ecosystem Going Beyond Marketplaces and Cloud Services

^⑤ Khanal Shaleen, Hongzhou Zhang et al., “Why and How Is the Power of Big Tech Increasing in the Policy Process? The Case of Generative AI”, *Policy and Society*, 2024, pp.2-13.

^⑥ Dyer-Witheyford Nick, Alessandra Mularoni, “Framing Big Tech: News Media, Digital Capital and the Antitrust Movement”, *The Political Economy of Communication*, Vol.9, No.12, 2021, pp.12-15.

而政策流的分析则指向大型跨国科技企业倡导并合法化特定政策工具的能力。例如，泰马尔·莎伦（Tamar Sharon）等人的研究讨论了公共部门与技术巨头的 AI 产品的合作案例，揭示了大型跨国科技企业如何通过数字化专业知识和优势进入新的社会领域，以及这些行为可能带来的隐私、数据保护、不公平回报和新依赖等风险，同时批评了欧盟的通用数据保护条例未能有效防止这类风险的产生。^①技术巨头通过提供基础设施（分析工具、云等）加强了自身倡导并合法化特定政策工具的能力。

最后，政治流的分析关注大型跨国科技企业与政治本身的互动。游说活动是大型跨国科技企业将科技能力转化为政治权力的重要方式。游说理论为分析这一过程提供了框架：由于决策机构的规模有限，其无法完全自给准确决策所依赖的全面信息，海伦·米尔纳（Helen Milner）对国际合作的研究指出，虽然决策中的信息不对称普遍存在，但行为体具有想方设法获取足够信息的动机，例如，当立法部门缺少关于国际协议谈判的信息时，它将会依赖掌握更多谈判信息的利益集团的信号作出决策。^②这一理论存在两方面的展开：首先，信息是现代决策中的关键要素，而立法决策较为普遍地依赖利益集团提供的信息。^③其次，具有利益偏好的集团将通过这一过程向政府有选择性地输出符合自身偏好的信息，以争取对自身有利的决策。对这一信息输送过程的研究指出，企业为政府部门提供决策所需的信息企业使用的信息策略往往比信息内容更重要，而信息策略包括内部信息策略（与欧盟机构内部官员和政治人物直接互动）与外部信息策略（通过公众或媒体来影响决策过程），^④外部策略的逻辑是压力（Pressure），而内部策略的逻辑本质是说服（Persuasion），^⑤基于压力的外部信息策略被认为更加具有成功潜力。科技企业常常交替使用这两种策略以达到游说目的。以欧盟为例，Google、Facebook 和 Microsoft 的年游说支出超过了 500 万欧元，同时大规模雇佣前政治

^① Tamar Sharon and Raphaël Gellert, "Regulating Big Tech Expansionism? Sphere Transgressions and the Limits of Europe's Digital Regulatory Strategy," *Information, Communication & Society*, 2023, pp.1-18.

^② [美]海伦·米尔纳：《利益、制度与信息：国内政治与国际关系》，曲博译，王正毅校，上海：上海人民出版社 2010 年版，第 18-21 页。

^③ Tarrant Andy, Tim Cowen, "Big Tech Lobbying in the EU", pp.218-219.

^④ Iskander De Bruycker and Jan Beyers, "Lobbying Strategies and Success: Inside and Outside Lobbying in European Union Legislative Politics," *European Political Science Review*, Vol.11, No.1, 2019, pp.57-74.

^⑤ Iskander De Bruycker, "Pressure and Expertise: Explaining the Information Supply of Interest Groups in EU Legislative Lobbying," *JCMS: Journal of Common Market Studies*, Vol.54, No.3, 2016, pp.599-616.

家、公务员与技术专家作为政治顾问，同时利用媒体为自身的观点进行背书与宣传，以期发挥更大的政治影响力。^①

上述研究多聚焦于大型企业，杰弗里·德罗普（Jeffery Drope）等人指出，以往对美国公司政治参与的研究中，学者们一直依靠大部分或完全由大型企业组成的样本，样本上的不平衡可能导致研究结论上的偏见。^②从参与度与参与方式来看，这一倾向具有其客观原因：经典经济学路径的研究指出，由于在政策过程中拥有的更庞大的经济利益及更复杂的业务形态，大型公司有更大的动力和资源通过直接游说影响政策制定的过程。^③而中小企业受限于游说资源与资金规模的约束，加之对披露企业专有信息的天然抵触，其游说活动的发生频次显著低于大型企业，且在组织形态上更倾向于通过行业协会进行集体游说，或者完全不进行游说，和有多种组织选择的大型企业存在显著区别。^④

然而，中小企业在游说活动中并非完全处于劣势地位。基于公共物品理论，大卫·莱维（David Levine）等人的研究构建了检验大型团体和小型团体的游说竞争力的理论模型，并发现小型团体在游说竞争中的优势可能源于组织、监督成员所需的总成本较低，能够减少克服公共品问题的成本，从而提高游说活动的成功率。^⑤何塞·阿隆索（José Alonso）等人对英国政府合同游说的研究发现，游说“深度”策略带来的好处对于小企业来说更为显著，与关键政治人物的重复互动可能特别有助于小型公司在政府合同竞争中取得优势。^⑥总的来说，现有研究发现，虽然小企业参与和支出游说的比例较低，但它们在面对特定议题（如影响盈利能力）时会积极游说，并且可以从某些特定类型的游说策略（如与关键人物建立深层关系）中获得比大企业更显著的好处。

^① Corporate Europe Observatory: “The lobby network: Big Tech’s web of influence in the EU”, August 2021, <https://corporateeurope.org/en/2021/08/lobby-network-big-techs-web-influence-eu>, 2025-05-12. And Corporate Europe Observatory: “The lobbying ghost in the machine: Big Tech’s covert defanging of Europe’s AI Act”, February 2023, <https://corporateeurope.org/en/2023/02/lobbying-ghost-machine>, 2025-05-12.

^② Jeffrey M. Drope and Wendy L. Hansen, “Does Firm Size Matter? Analyzing Business Lobbying in the United States,” *Business and Politics*, Vol.8, No.2, 2006, pp.1–17.

^③ Holly Brasher and David Lowery, “The Corporate Context of Lobbying Activity,” *Business and Politics*, Vol.8, No.1, 2006, pp.1–23.

^④ John M. De Figueiredo and Emerson H. Tiller, “The Structure and Conduct of Corporate Lobbying: How Firms Lobby the Federal Communications Commission,” *Journal of Economics & Management Strategy*, Vol.10, No.1, 2001, pp.91–122.

^⑤ David K. Levine and Salvatore Modica, “Size, Fungibility, and the Strength of Lobbying Organizations,” *European Journal of Political Economy*, Vol.49, 2017, pp.71–83.

^⑥ José M. Alonso and Rhys Andrews, “Insider Lobbying and Government Contracts: The Moderating Role of Firm Size,” *European Management Review*, Vol.19, No.3, 2022, pp.462–475.

对跨国公司与游说的研究已经较为成熟，许多学者指出了大型跨国科技公司的技术地位转化为政治话语权的机制，并对其游说策略的有效性进行了检验。虽然一部分学者关注了中小企业的游说活动，并指出了中小企业影响中小企业游说有效性的因素，但对中小企业相对特殊的自身性质、游说渠道等缺乏具体的研究；反过来，对中小企业的政治参与也常因为数据的数量、可见性和可触达性等原因存在困难，因此缺少被进一步探索的机会。

三、技术监管相关研究

最后，我们将回顾既往文献中对技术监管这一话题的研究。新兴技术向来被视为一种双刃剑：一方面，其强大的创新潜力能够带动现代产业发展、促进技术革新；另一方面，其迅速发展带来的技术陷阱、伦理问题与滥用风险也不容小觑。基于此，从监管与治理的相对概念出发，本节对以人工智能为代表的新型技术监管研究现状做了小结。

本文讨论的监管（regulation）被视为治理（governance）的一个子集，其核心在于引导事件和行为的流程，而更少涉及提供和分配。^①传统上，监管往往与国家制定的法律及其强制执行相关联，而治理的概念则涵盖了国家、市场、公民社会及其他非国家行为者。新兴的“监管治理”（regulatory governance）概念将上述国家、市场等不同部门视为一个综合的、相互关联的系统，在这种系统中，私营主体不再仅仅是监管的对象，更是参与规范制定的主体之一。^②上述研究反映出，即使是传统上与统一规则和强制执行相关的监管研究，也逐渐吸纳了以私营企业为代表的非国家行为体作为研究对象之一。

从新型技术治理的现状研究来看，人工智能监管模式的比较研究是目前研究的主流方向。例如，王彦雨等人从治理理念、治理主体与治理困境等方面比较了欧盟和美国的人工治理模式，指出相比美国，欧盟在人工治理中倾向于使用伦理优先与治理领导策略弥补其技术竞争力劣势，同时质疑了欧盟治理模式在技术创新与强治理间的平衡能力。^③詹姆斯·布彻（James Butcher）等人则比较了目前全球私营部门、公共部门、研究机构以及联合国在人工智能治理中的现有角色，展示了私营部门在自我监管方面的努力与目前人工智能领域无组织与复杂的博

^① John Braithwaite, Cary Coglianese and David Levi-Faur, “Can Regulation and Governance Make a Difference?,” *Regulation & Governance*, Vol.1, No.1, 2007, pp.1-7.

^② David Levi-Faur: “Regulation and Regulatory Governance”, in David Levi-Faur: *Handbook on the Politics of Regulation*, Edward Elgar Publishing, 2013, pp.3-17.

^③ 王彦雨、李正风、高芳：《欧美人工智能治理模式比较研究》，《科学学研究》2024年第3期，第460-468页。

弈动态。^①从全球视野看，拉杜（Radu）分析了2016-2019年间12个国家发布的人工智能战略，重点研究了国家和跨国公司专家合作的混合治理模式下，国家如何设计人工治理机制；研究发现，这些战略都侧重于伦理导向而非规则导向，且未明确人工智能发展中的公共利益与私人利益。^②娜塔莉·斯穆哈（Nathalie Smuha）则考察了全球人工智能监管竞赛这一概念：人工智能监管竞赛是人工智能竞赛的产物，先发国家试图通过率先确立监管规范来成为这一领域的规范领导者，虽然这一监管存在“竞底”的风险，但“信任”作为一种竞争优势，加之民间力量的影响，正在促使各国在AI监管方面趋于融合。^③

如上文所述，监管者必须在促进创新与基本权利保护等公共利益中的相互冲突的目标之间进行权衡，换句话说，监管者最优先的愿望也是在部署新技术的收益和风险之间建立最佳平衡，以避免监管过度或监管不足。^④阿拉兹·泰伊哈格（Araz Taeihagh）等人的研究指出，人工智能作为一种“破坏性创新”，使技术监管面临着技术信息不对称、发展方向难以预测、利益影响复杂等挑战。^⑤雅各布·哈塞尔巴赫（Jacob Hasselbalch）对破坏性监管中的监管评估的研究指出，由于以人工智能为代表的技术创新发展过于迅速，政策制定者发现自己常常被迫对这些变化做出仓促反应，而不是从一开始就加以塑造。^⑥

在这一背景下，监管方应该使用什么手段完成上述重要的平衡？洪胜勋（Seung-Hun Hong）等人衡量并对比了英国与韩国的自动驾驶汽车监管政策，基于规则结构，执行结构和法规反馈建立了政策灵活性指标，并指出灵活监管作为破坏性创新的一种对策的潜力。^⑦马克·芬威克（Mark Fenwick）等人提出了类似的“动态监管”框架，认为监管决策不应被视为“最终事件”，而应该被看作是

^① James Butcher and Irakli Beridze, “What Is the State of Artificial Intelligence Governance Globally?,” *The RUSI Journal*, Vol.164, Nos.5–6, 2019, pp.88–96.

^② Roxana Radu, “Steering the Governance of Artificial Intelligence: National Strategies in Perspective,” *Policy and Society*, Vol.40, No.2, 2021, pp.178–193.

^③ Nathalie A. Smuha, “From a ‘Race to AI’ to a ‘Race to AI Regulation’: Regulatory Competition for Artificial Intelligence,” *Law, Innovation and Technology*, Vol.13, No.1, 2021, pp.57–84.

^④ Araz Taeihagh, “Governance of Artificial Intelligence,” *Policy and Society*, Vol.40, No.2, 2021, pp.137–157.

^⑤ Araz Taeihagh, M. Ramesh and Michael Howlett, “Assessing the Regulatory Challenges of Emerging Disruptive Technologies,” *Regulation & Governance*, Vol.15, No.4, 2021, pp.1009–1019.

^⑥ Jacob A. Hasselbalch, “Innovation Assessment: Governing through Periods of Disruptive Technological Change,” *Journal of European Public Policy*, Vol.25, No.12, 2018, pp.1855–1873.

^⑦ Seung-Hun Hong et al., “Making Regulation Flexible for the Governance of Disruptive Innovation: A Comparative Study of AVs Regulation in the United Kingdom and South Korea,” *Journal of European Public Policy*, Vol.29, No.11, 2022, pp.1845–1865.

一种“权衡性决策”，应保持高度的开放性以培养负责的创新。^①王（Wang）等人则详细说明了生成式人工智能的破坏性，并通过对美国人工智能监管政策的案例研究，检验并说明了国际合作、积极立法以及伦理整合在破坏性创新监管中的有效作用。^②总而言之，主动且积极的事前监管逻辑被认为是应对破坏性创新与有效处理技术创新-基本权利保护这一平衡的有效手段。

从上述观察来看，学者们描述并讨论了创新—监管这一新兴技术监管中的困境，但学者们在现状的比较之外，多是对监管方应该使用何种监管手段达成高效监管的应然分析，在实然层面的本质讨论仍有缺乏；同时，现有文献虽然提出了诸多应对监管争议的方法，但对应对在政策的制定、监管争议的形成、最终的妥协与调解等领域的多元利益相关者的动态互动的分析上仍有欠缺。

第一章 人工智能法案与通用人工智能监管

第一节 人工智能法案的起源、方法与挑战

人工智能法案是欧盟对近年来迅速发展的人工智能带来的技术与社会风险的监管回应。为全面讨论人工智能法案的决策过程。本节将首先从技术风险和规范意愿两个方面讨论人工智能法案的决策起源，再由其使用的“基于风险的方法”进入通用人工智能监管挑战的讨论。

一、技术风险与规范意愿

经合组织（OECD）将人工智能系统定义为“一种基于机器的系统，它能够根据收到的输入推断如何生成输出，例如预测、建议或决策，以实现明确或隐含的目标。”^③随着机器学习和深度学习技术的不断发展，人工智能技术从 20 世纪 50 年代开始不断发展迭代，从自然语言处理、计算机视觉到金融科技与智能制造，人工智能技术已经成为目前最具潜力的通用技术之一。2022 年以来以 ChatGPT 为代表的生成式人工智能的发展进一步掀起了全球 AI 技术发展的热潮。

^① Mark Fenwick, Erik P. M. Vermeulen et al., “Business and Regulatory Responses to Artificial Intelligence: Dynamic Regulation, Innovation Ecosystems and the Strategic Management of Disruptive Technology”, in Marcelo Corrales et.al, *Robotics, AI and the Future of Law*, Springer, 2018, pp.81-103.

^② Wang Xukang, Ying Cheng Wu, “Balancing Innovation and Regulation in the Age of Generative Artificial Intelligence”, *Journal of Information Policy*, Vol.14, 2024, pp.385–416.

^③ Marko Grobelnik, Karine Perset, Stuart Russell, “What is AI? Can you make a clear distinction between AI and non-AI systems?”, March 6, 2024, <https://oecd.ai/en/wonk/definition>, 2025-05-12

据 Statista 统计，人工智能市场规模在 2020-2025 年间增长了 97.3%，用户规模也增长了 3 倍以上，达到约 3.78 亿。^①

人工智能技术在现代的发展与普及不仅显著推动了相关产业的革新与现代生活方式的转变，也引发了全球范围内对其潜在风险的高度关注和广泛讨论。欧盟在 2020 年的人工智能白皮书中，将人工智能的风险分为物质与非物质两类，物质风险包括个人安全和健康，例如生命损失和财产损失等，而非物质风险包括隐私权与言论自由权的侵犯，以及就业等方面的歧视等。^②具体而言，物质风险主要存在于以自动驾驶为例的人机交互领域：当人工智能被整合进物质产品与服务时，可能会给使用者带来新的安全风险。例如，由于物体识别中整合的人工智能技术存在缺陷，自动驾驶汽车可能会错误识别道路上的物体，从而导致人员伤亡与财产损失，这些缺陷可能存在于人工智能本身的设计中，也可能存在于其使用的数据可用性与质量上。^③非物质风险则主要指向对个人基本权利的侵犯。例如，利用人工智能进行的远程生物特征识别与预测性警务将侵犯个人隐私权，而未经适当审查与调整的训练数据集可能会使模型产出结果存在基于性别、种族或年龄的歧视。^④物质上的风险对应主权国家对公民生命与财产的保护义务，而非物质的风险则挑战了民主与人权等重要价值观。

除上述风险以外，人工智能技术应用的不可预测性（unpredictability）也对监管造成了巨大挑战，加之通用人工智能的井喷发展与迅速迭代，人工智能系统在训练中形成的决策逻辑越来越难以观测。以 IBM 开发的国际象棋人工智能深蓝（deep blue）为例，开发者无法预测它在对局的每个转折中会作出什么决定，只能预测的是它将采取可能的手段获取胜利这一目的。^⑤人工智能开发者常以此作为辩护依据，致使监管方难以通过具体条款对人工智能行为形成有效规范。

从 2018 年开始，欧盟已经着手处理人工智能带来的诸多风险。2018 年 4 月，欧盟委员会在提交给其他欧盟机构的一份名为“欧洲的人工智能”（Artificial

^① Statista, “Artificial Intelligence - Worldwide”, <https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide#market-size>, 2025-05-12.

^② European Commission: “White Paper on Artificial Intelligence - A European approach to excellence and trust”, February 19, 2020, pp.10-15.

^③ Ibid, pp.10-12.

^④ Songül Tolan et al., “Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia,” in *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*, Montreal QC Canada: ACM, 2019, pp.83-92.

^⑤ Roman V. Yampolskiy, “Unpredictability of AI: On the Impossibility of Accurately Predicting All Actions of a Smarter Agent,” *Journal of Artificial Intelligence and Consciousness*, Vol.7, No.01, 2020, pp.109-118.

Intelligence for Europe) 的通讯中, 表达了对人工智能创新的重视与对其建立适当的道德与法律监管框架的愿望。^①2020年2月, 欧盟委员会发布了《人工智能白皮书: 欧洲追求卓越与信任的方法》(后简称白皮书), ^②阐明欧盟人工智能政策的目标: 维护欧盟的技术领先地位, 并确保新技术服务于所有欧洲人——在尊重他们权利的同时改善他们的生活。在白皮书中, 欧盟委员会识别了诸如基本权利风险、安全责任风险等具体的人工智能风险点, 并确认了未来欧盟人工智能监管框架的范围与核心关切。^③2021年4月, 欧盟委员会发布了人工智能法案的立法提案。

在2021年初该立法提案发布之际, 全球主要国家尚未针对人工智能出台专门的监管立法安排, 其关注重心多以国家人工智能发展战略为载体, 主要聚焦于提升技术竞争力, 而对人工智能监管框架与方法论的系统性构建关注有限。^④欧盟迅速布局人工智能监管, 首先有通过创建可信人工智能框架、创建统一的人工智能市场, 以此增强欧盟在人工时代的竞争力的目的。在立法提案中, 欧盟委员会表示, “(法案) 避免单一市场进一步分裂成可能相互矛盾的国家框架, 从而阻碍包含人工智能的商品和服务的自由流通... 一个稳固的欧洲可信人工智能监管框架也将... 增强欧洲在人工智能领域的竞争力和产业基础。”^⑤人工智能法案搭建的可信人工智能框架与统一人工智能市场被预期将为欧盟人工智能产业创造有利的发展环境。

除了促进本土人工智能发展外, 法案的另一重目的是通过率先立法设置立法标杆, 从而实现其监管的全球影响与输出,^⑥即发挥“布鲁塞尔效应”(brussels effect)。布鲁塞尔效应最早由阿努·布拉福德(Anu Bradford)提出, 用于描述

^① European Commission: “Communication from the Commission to The European Parliament, the European Council, the Council, the European Economic and Social Committee and The Committee of The Regions: Artificial Intelligence for Europe”, Apr 2018, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52018DC0237>, 2025-05-12.

^② European Commission, “White Paper on artificial intelligence – a European approach to excellence and trust”, February 29, 2020.

^③ European Commission: “White Paper on Artificial Intelligence - A European approach to excellence and trust”, p.10.

^④ 例如, 中国国务院在2017年印发的《新一代人工智能发展规划》、美国政府2019年2月发布的“行政命令13859: 保持美国在人工智能领域的领导地位”(EO 13859: Maintaining American Leadership in AI)、英国政府在2021年9月发布的《国家人工智能战略》(National AI Strategy)等, 均基本仅以增强人工智能竞争力为主要目标进行了方向规划。

^⑤ European Commission, “Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts”, April 21, 2021, p.6.

^⑥ Sean Musch, Michael Borrelli and Charles Kerrigan, “The EU AI Act As Global Artificial Intelligence Regulation,” *SSRN Electronic Journal*, 2023, pp.4-5.

欧盟凭借其市场力量在自身边界之外施加监管影响的能力。^①夏洛特·西格曼（Charlotte Siegmann）等人认为，在过去数据保护指令（DPD）、通用数据保护条例等法律成功显著影响他国立法的情况下，欧盟希望继续利用其巨大的人工智能市场规模影响在欧盟经营的诸多人工智能跨国公司，使遵守欧盟标准的跨国公司游说其本国政府采纳类似的法规，以避免在国内市场面临不遵守欧盟标准从而成本较低的竞争对手。^②通过这一影响路径，欧盟试图通过先发优势设定全球标准与规范，巩固欧盟在人工智能时代的全球监管领导者地位，传播欧盟价值观并增强欧盟在全球人工智能（监管）竞赛中的竞争力。^③

二、“基于风险的方法”

欧盟在人工智能监管决策上具有培育产业发展、扩大全球影响等宏大的目标，然而面对高度不确定与迅速发展的人工智能技术，传统的监管方法显然无法帮助欧盟建立成比例与高效的监管框架。在决策中，欧盟最终决定利用“基于风险的方法”搭建监管框架，这一方法基于对潜在不利影响发生概率和影响程度的评估，提议制定与风险水平相匹配的不同干预措施，涵盖从标准制定到执行等各个方面。^④人工智能法案确定了四种人工智能风险类别，从大到小分别是“不可接受风险”（unacceptable risk）、“高风险”（high risk）、“有限风险”（limited risk）和“极小或无风险”（minimal or no risk）。^⑤

其中，具有“不可接受风险”的人工智能均被完全禁止，这些人工智能被认为对人类安全、升级和权利构成明显威胁，例如社会评分、情绪识别等。“高风险”人工智能则是可能对健康、安全或基本权利构成严重风险的类别，主要包括在教育机构、交通工具、司法与警务当中应用的人工智能。这些人工智能在投放市场前需遵循严格的义务，例如充分的风险评估、详细的技术文档、提供高质量

^① Anu Bradford, “The Brussels Effect,” in Anu Bradford, *The Brussels Effect*, Oxford University Press, 2020, pp.25–66.

^② Charlotte Siegmann, Markus Anderljung, “The Brussels Effect and Artificial Intelligence: How EU regulation will impact the global AI market”, pp.4-5.

^③ Shu Li, Béatrice Schütte and Suvi Sankari: “The ongoing AI-regulation debate in the EU and its influence on the emerging economies: a new case for the ‘Brussels Effect’?” in Mark Findlay, Li Min Ong and Wenxi Zhang, eds., *Elgar Companion to Regulating AI and Big Data in Emerging Economies*, Edward Elgar Publishing, 2023.

^④ Regine Paul, “European Artificial Intelligence ‘Trusted throughout the World’: Risk-based Regulation and the Fashioning of a Competitive Common AI Market,” *Regulation & Governance*, Vol.18, No.4, 2024, pp.1065–1082.

^⑤ European Union, “laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008 (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)”, June 13, 2024, pp.7-8.

的数据集等。“有限风险”类别仅需遵守特定透明度义务，“风险极小或无风险”的人工智能则未被要求遵循特定义务。^①

表 1 人工智能法案的风险分级

风险类型	定义	用例	规定义务
不可接受风险	对人类安全、生计和权利构成明显威胁的人工智能	社会评分、工作场所和教育机构中的情绪识别、个人犯罪风险评估或预测	被完全禁止
高风险	可能对健康、安全或基本权利构成严重风险的人工智能	教育机构使用的人工智能解决方案、关键基础设施(例如交通运输)中的人工智能安全组件	严格的合规义务，例如充分的风险评估、为系统提供高质量的数据集、详细的技术文件
有限风险	旨在与个人互动或生成内容的人工智能系统	聊天机器人和数字助理、生成合成音频、图像、视频或文本的系统	有限的透明度义务，例如声明内容是由人工智能生成等
极小或无风险	被认为风险极低或无风险的人工智能	支持人工智能的视频游戏、垃圾邮件过滤器	无

(来源：作者根据法案内容自制)

欧盟并非第一次在立法中使用基于风险的方法，早在通用数据保护条例和数字服务法（DSA）的立法实践中，这一方法就已经得到了应用。但与前述两法案不同的是，人工智能法案采用了自上而下的管理方法，即由监管方全权定义风险类别并加以管理，被监管方无法自行决定风险评估与管理措施。^②然而，基于风险的方法并非人工智能法案设想中唯一的路径，欧盟委员会在 2020 年 7 月发布的立法初步影响评估（Inception Impact Assessment）中提出了四种达到欧盟人工智能治理目标的路径：无正式法律的软约束方法、企业自愿参加的“可信”标签计划立法、使用基于风险的方法对所有或某些类型的人工智能建立强制要求，以及上述选项的集合。^③在 2021 年发布的后续影响评估中，欧盟委员会将第三选项，即使用基于风险的方法对高风险人工智能应用制定强制性要求（并对非高风险的人工智能制定资源行为准则）认定为最佳选项，因为这一方法“提供了针对人工智能带来的风险成比例的保障措施，同时将行政和合规成本降至最低”。^④

^① 均基于人工智能法案最终版本列出。

^② Pietro Dunn and Giovanni De Gregorio, “The Ambiguous Risk-Based Approach of the Artificial Intelligence Act: Links and Discrepancies with Other Union Strategies”, CEUR Workshop Proceedings, Vol.3221, June 2022, pp.3-9.

^③ European Commission, “Inception Impact Assessment”, July 23, 2020, Ref. Ares(2020)3896535, pp.4-5.

^④ European Commission, “Commission Staff Working Document: Executive Summary of The

然而，利用基于风险的方法进行人工智能监管也存在缺点。首先，法案中基于风险的方法被批评过于侧重监管，忽略技术带来的积极贡献，这将使某些高潜力的人工智能因其风险而失去促进社会福利的机会；其次，法案中的风险分类被认为缺乏实证证据，由于分类基于抽象描述而非具体用例，人工智能很有可能进入错误的风险分类。^①另外，由于法案对风险的划分是基于人工智能的“预期目的”（intended purpose）进行，即通过其用途或应用场景确定其风险等级，以通用人工智能为代表的无特定预期目的的人工智能将难以被整合进这一监管逻辑，甚至对这一方法造成冲击。最后，瑞吉·保罗（Regine Paul）从文化政治经济学的角度指出，在人工智能法案中基于风险的方法只是表面上符合既往研究中对这一方法理性的假定，其本质是欧盟在新兴技术领域寻求监管合法性的修辞工具。^②在通用人工智能兴起并进入法案的监管视野中后，这一方法的过度监管、标准模糊和评估困难问题进一步显现，最终成为了后续通用人工智能监管争议中最重要的议题之一。

三、通用人工智能的突现性风险与制度反应

欧盟委员会在 2021 年 4 月提出立法提案时，机器学习与自然语言处理等人工智能的基础技术已经发展得比较成熟，也出现了 Alpha Go 等引发国际关注的人工智能应用，但其普及度、生成与互动能力仍然受限，今天具有强大文本与图片生成能力、逻辑推理能力与互动能力的通用人工智能尚未兴起。因此，在最初的立法提案中，通用人工智能并未被纳入监管文本。2022 年初，能够与用户对话、根据用户指令生成文本的 ChatGPT 在上线五天内收获了逾百万用户，揭开了以其为代表的通用人工智能兴起的序幕，同时引发了全球范围内对其潜在风险和影响的广泛关注。

人工智能法案将通用人工智能定义为“包括使用大规模自我监督用的大量数据进行训练的，无论该模型在市场上的投放方式如何，该模型都具有显著的通用性，能够胜任执行各种不同的任务，并且可以集成到各种下游系统或应用程序中”的人工智能模型。通用人工智能的核心优势在于其能够通过大规模数据训练，执

Impact Assessment Report”, April 21, 2021, pp.1-2.

^① Martin Ebers, “European Union Law Working Papers: Truly Risk-Based Regulation of Artificial Intelligence: How to Implement the EU’s AI Act”, EU Law Working Papers No. 101, Stanford-Vienna Transatlantic Technology Law Forum, October 2024, pp.29-33.

^② Regine Paul, “European Artificial Intelligence ‘Trusted throughout the World’: Risk-based Regulation and the Fashioning of a Competitive Common AI Market”, p.1068.

行多种复杂任务，并集成到各种下游系统和应用中。斯坦福基础模型研究中心的报告指出，通用人工智能在视觉、语言与机器人上均有高度潜力，能够被应用在医学与教育领域，发展疾病辅助诊断、教育个性化计划等专业功能，在促进社会福利上有巨大潜力。^①

然而，通用人工智能的井喷发展也对人工智能法案的制定提出了巨大的挑战。首先，通用人工智能具有高度通用性与广泛的应用范围，这使通用人工智能提供商难以预见、评估和缓解其具体应用场景下的所有风险，从而冲击了人工智能法案基于风险的核心方法。^②提供商可能无法预测其通用人工智能将被下游用户如何具体使用，以及在这些具体用途中可能产生哪些特定风险，监管者因此无法将通用人工智能有效地置入任何预定义的风险层级中。于是，在人工智能法案的立法讨论过程中，出现了将通用人工智能整体作为高风险人工智能监管、反对对通用人工智能的任何监管以及对通用人工智能实施分层监管的不同声音，前两种监管方法将冲击基于风险的方法，而最后一种分层方法则面对着分层标准难以划分或模糊等问题，无论哪种解决方案都面临着困境。因此，通用人工智能的风险分类成为立法过程中最大的难题之一。

除此以外，通用人工智能的其他风险也不容忽视。其中最突出的是其生成能力造成的风险，在未受到有效监管的情况下，它们可能生成具有冒犯性或有害的内容，也可能成为制造与传播虚假信息的工具，从而对社会稳定、民主进程和知识产权保护构成威胁。^③另外，通用人工智能日益增长的自主性也带来了风险，虽然监管者倾向于将人工智能视为一种工具而非自主智能体，但通用人工智能已被证实存在某些潜在的危险目标，例如，测试显示模型可能表现出对“权力”或“财富”的“渴望”。更先进的通用人工智能甚至具备制定多步计划并执行与进行欺骗行为的能力。^④最后，通用 AI 模型常常作为“构建模块”被整合到各种下游系统或应用中，因此起源于基础模型设计和开发阶段的风险将可能被被众多下游应用

^① Rishi Bommasani, Drew A. Hudson, Ehsan Adeli, et al., “On the Opportunities and Risks of Foundation Models.” Research Paper by Center for Research on Foundation Models (CRFM), August 16, 2021, pp.21-67.

^② Pietro Dunn and Giovanni De Gregorio, “The Ambiguous Risk-Based Approach of the Artificial Intelligence Act: Links and Discrepancies with Other Union Strategies.”, pp.3-5.

^③ Pegah Maham and Sabrina Küspert, “Governing General Purpose AI,” July 2023, pp2-6.

^④ Scott McLean et al., “The Risks Associated with Artificial General Intelligence: A Systematic Review,” *Journal of Experimental & Theoretical Artificial Intelligence*, Vol.35, No.5, 2023, pp.649–663.

所继承，从而带来潜在的系统性风险，对公共健康、安全、基本权利乃至整个社会产生重大影响。^①

第二节 通用人工智能监管的主要争议场域

通用人工智能的兴起带来了巨大的创新机遇，也对人工智能法案的立法决策提出了方法和内容层面的重大挑战，并最终成为了后续立法争议的焦点。在这一认知的基础上，本节将具体讨论通用人工智能的监管争议，展现主要的争议话题、争议原因与冲突焦点，以期对通用人工智能的自身特点与监管重点形成较为全面的认识，为后文探索决策过程中人工智能监管具体条文内容的变动提供认知基础。与传统的人工智能系统不同，通用人工智能不仅具有高度的任务通用性，还能作为基础模型广泛应用于各种下游任务，使其风险特征更具复杂性。这不仅冲击了现有的基于风险的方法，还对立法机构提出了更高的监管要求。在法案的决策过程中，围绕通用人工智能监管的讨论逐步形成了多个争议场域，其中既包括如何准确定位通用人工智能的风险级别，也涉及监管公平、产业影响等多方面的问题。

一、风险定位与监管方式

正如上文对通用人工智能对基于风险的方法的挑战所述，对在通用人工智能的监管过程中，如何准确界定其风险级别已经成为决策争议的关键争议之一。这不仅涉及技术本身的复杂性，还反映了不同利益方对通用人工智能潜在社会影响的不同预期。

支持严格监管的一方主张，通用人工智能具备广泛的任务适应性和自主学习能力，可能在多种情境下带来重大风险。例如，欧盟基本权利署（FRA）等研究机构认为，通用人工智能不仅可能生成反民主言论、仇恨言论，还存在侵犯个人隐私等系统性风险。^②同时，虽然通用人工智能普遍存在内置的伦理与法律保护措施，提示注入(prompt injection)等技术已被证明可以规避这些措施，从而导致了人工智能系统自我监管无法控制的新风险领域。^③这些特性使其在研发阶段即具备高风险潜力，因此应将其纳入高风险系统的监管范畴。2023年3月，欧洲

^① Pegah Maham and Sabrina Küspert, “Governing General Purpose AI,” July 2023, pp.34-42.

^② Shaleen Khanal, Hongzhou Zhang and Araz Taeiagh, “Why and How Is the Power of Big Tech Increasing in the Policy Process?” European Agency for Fundamental Rights (FRA): “Bia in Algorithms-Artificial Intelligence and Discrimination”, December 2022, pp.21-22.

^③ Oskar J. Gstrein, Noman Haleem and Andrej Zwitter, “General-Purpose AI Regulation and the European Union AI Act,” *Internet Policy Review*, Vol.13, No.3, August 2024, p.3.

议会的联合报告员曾提出妥协文本（compromised text），^①建议将无人监督的复杂文本生成系统（如 ChatGPT）视为高风险，并要求其提供商和部署者承担更高的合规义务。这一主张反映了对通用模型潜在负面外部性的高度警惕，强调应在模型开发之初就对其进行风险控制。^②

然而，反对方则认为这种一刀切的监管模式缺乏针对性，可能对技术创新产生压制效应。批评者指出，通用人工智能不仅包含潜在高风险的应用，也涵盖了大量对社会有积极作用的低风险场景，如听障人士的实时字幕生成、文本朗读辅助等。如果将所有通用模型统一视为高风险，不仅会大幅增加合规成本，还可能阻碍这些有益创新的广泛应用。^③与此同时，许多欧洲中小企业已经将通用人工智能整合入自身在市场提供的服务中，对通用人工智能的过度监管将损害欧盟诸多中小企业的利益。另外，这种分类方法在最初的委员会提案中并未得到充分讨论和评估，因此缺乏明确的技术风险标准支持，不应在法案中被应用。^④

从上述观点可以看出，这一争议的核心在于如何在充分防范系统性风险的同时，避免对创新能力造成不必要的抑制，以及如何在高度不确定的技术演化过程中设定合适的监管边界，商业利益与风险控制的冲突在这一议题上被进一步放大。

二、开源与研究型通用人工智能的豁免

如何处理开源与研究用途的基础模型，是通用人工智能监管中的另一大争议焦点。开源模型不仅是前沿技术创新的重要推动力量，也是学术研究和中小企业创新的重要基础。然而，这类模型的开放性也带来了显著的监管挑战。

支持开源与研究型通用人工智能豁免的一方从两方面发展论点，一是创新生态与技术发展，二是监管公平。在创新生态和技术发展上，以 Mistral AI 为代表

^① 妥协文本指在欧盟立法中，各机构为吸纳修正案、弥合各方意见产出的非正式文件，这种文本通常是为了在各方立场之间找到平衡，并形成可被多数成员国和议员接受的法律草案；虽然是非正式文件，但这些文件依旧能够反映出各立法机构的意图与动向。See James P. Cross and Henrik Hermansson, “Legislative Amendments and Informal Politics in the European Union: A Text Reuse Approach,” *European Union Politics*, Vol.18, No.4, 2017, pp.5-6.

^② Luca Bertuzzi, “Leading EU lawmakers propose obligations for General Purpose AI”, March 14, 2023, <https://www.euractiv.com/section/artificial-intelligence/news/leading-eu-lawmakers-propose-obligations-for-general-purpose-ai/>, 2025-05-13.

^③ Polish Confederation Lewiatan, “GPAI Considerations for EU AI Act High-Risk Requirements”, March 2023, <https://lewiatan.org/wp-content/uploads/2023/03/GPAI-Considerations-for-EU-AI-Act-High-Risk-Requirements.pdf>, 2025-05-13.

^④ KI BUNDESVERBAND (German AI Association), “Position Paper on the EU AI Act Remaining Issues and Current Discussions in the European Parliament”, April 2023, p.3.

的欧洲中小企业强调，严格的前置合规要求如果不区分开源与闭源、商业与非商业用途，可能对开源社区形成极大冲击，阻碍技术协作与创新，并认为围绕创新社区建立强大模型并创建开源资产才是创建人工智能最安全的方式。^①创新组织“数据创新中心”（Center for Data Innovation）同样表示，许多开源人工智能创新已经进入多个商业应用领域，这些开源技术不仅能够促进竞争，降低新进入人工智能领域的门槛，还能提高系统的透明度和稳健性评估。^②在监管公平方面，全球最大代码托管平台 GitHub 的首席执行官认为，人工智能法案应该对开源者宽容，因为开源软件开发者通常只是志愿者而非专业开发者，没有任何合规预算应对开源人工智能的监管，监管应该对开源创新社区“公平公正”。^③这些企业认为，欧盟应充分保护并平等对待开源通用人工智能及其开发者，避免对其进行“一刀切”式监管。

然而，反对方则指出，对开源模型的监管豁免并不可靠。从利好来看，在大型技术公司通过控制数据、计算资源和市场渠道，仍然能够主导人工智能技术的发展和应用的背景下，一味鼓励开源、豁免其监管并不能真正培育一个可信的、多样化的人工智能生态系统。另外，在人工智能技术领域是一个模糊的概念，人工智能系统的开放性需要考虑模型、数据、劳动、框架和计算能力等多个维度。仅仅开放其中一个或几个方面，并不能真正实现系统的开放性。^④从危害来看，部分人工智能开发者利用开源概念的模糊性，对自身的人工智能模型实施“开放洗白”（Open-washing），通过开放部分非核心的训练内容获得开源的标签，从而逃避监管。^⑤这将导致强大人工智能的风险脱离监管范围。另外，随着模型功能越来越强大，开发人员可能在有意或无意中发布不安全的模型（例如，具有间谍软件生成功能的模型），从而导致模型扩散，风险随之增加，且无需承担任何

^① Natasha Lomas: “France’s Mistral dials up call for EU AI Act to fix rules for apps, not model makers”, November 16, 2023. <https://techcrunch.com/2023/11/16/mistral-eu-ai-act/>, 2025-05-13.

^② Aswin Prabhakar, “The EU AI Act Is a Cautionary Tale in Open-Source AI Regulation”, November 20, 2023 <https://datainnovation.org/2023/11/the-eu-ai-act-is-a-cautionary-tale-in-open-source-ai-regulation/>, 2025-05-13.

^③ Jude Karabus: “GitHub CEO says EU AI Act shouldn't apply to open source devs”, Feb 7, 2023, https://www.theregister.com/2023/02/07/github_ceo_ai_act/, 2025-05-13.

^④ David Gray Widder, Meredith Whittaker and Sarah Myers West, “Why ‘Open’ AI Systems Are Actually Closed, and Why This Matters,” *Nature*, Vol.635, No.8040, November 2024, pp.827-833.

^⑤ Andreas Liesenfeld and Mark Dingemans, “Rethinking Open Source Generative AI: Open Washing and the EU AI Act,” in *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, Rio de Janeiro Brazil: ACM, June 2024, pp.1774-1787.

责任或问责。如果这些模型可以被自由下载和修改，可能会形成巨大的安全漏洞，例如被用于恶意软件生成或恐怖组织的信息战。^①

综上所述，这一争议的核心在于如何在鼓励技术开放与防范滥用之间找到合理的监管平衡，并确保监管体系对不同技术路径具有足够的包容性。开源社区的创新文化与需求、大型跨国科技公司的逃避监管意图、欧盟监管方的风险控制意愿在这一争议上不断碰撞。

第二章 决策互动结构与企业决策介入分析

第一节 监管决策互动结构与过程

一、主要利益攸关者

在人工智能法案的立法过程中，主要利益攸关者主要以监管方—受监管方的组合呈现，组合双方内部的不同主体也拥有各自偏好。监管方包括欧盟委员会、欧盟理事会及欧洲议会，受监管方按企业规模分，主要分为大型跨国科技企业（Big Tech）和中小科技企业（Small and Medium-sized Enterprises, SMEs）。

在监管方中，欧盟委员会主要负责问题识别与立法议程设定，在立法过程中提供提案以供欧盟决策机构表决，在通用人工智能监管立法决策中，由欧盟委员会提出立法提案并进行了后期调解工作；欧洲议会（European Parliament）是欧盟唯一直接选举产生的机构，具有较强的超国家性，是欧盟的监督、咨询机构，并在日益增多的政策领域与欧盟理事会分享立法权，欧盟的立法均需经议会一读与二读程序审议，在立法决策中偏好强监管路线。欧盟理事会（Council of the EU）是欧盟的决策中枢，由成员国各政策领域的部长级任务组成，是欧盟成员国意见表达的重要渠道，欧盟的立法均需理事会最终通过才能生效；欧盟委员会在监管决策中注重创新和监管的平衡，其立场整体上略比议会保守。

另一方面，受监管方存在大型跨国科技企业与本土中小企业的区别，受人工智能领域企业的既有分布特征影响，大型跨国科技企业主要指向来自美国的大型跨国科技企业，例如 OpenAI、苹果公司、微软、谷歌等。^②而中小企业以欧洲本

^① Max Langenkamp and Daniel N. Yue, “How Open Source Machine Learning Software Shapes AI” in *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, Oxford United Kingdom: ACM, July 2022, pp.385-395.

^② Shaleen Khanal, Hongzhou Zhang and Araz Taeihagh, “Why and How Is the Power of Big Tech Increasing in the Policy Process?”, *Policy and Society*, Vol. 44, No. 1, January 2025, pp.53-56.

土的人工智能企业为主，由于欧洲人工智能产业起步相对较晚、科研与资金分布较为分散等原因，欧洲的人工智能企业规模与科技水平虽然在近年来发展较为迅速，但仍未达到美国同行的水准。基于以上分类，本研究总结了各利益相关者在人工智能法案的立法过程中的核心诉求、主要行动机制以及对政策演化的结构性影响如下表。

表 2 立法决策利益攸关者分析

利益攸关者	核心诉求	主要行动路径/机制	对政策演化的结构性影响
欧盟委员会	平衡创新与风险、推动统一监管	初期主导议程，后期协调妥协角色	从议程设定者逐步转向妥协推动者
欧洲议会	强监管路线、强化 GPAI 监管责任	激进文本推动者、议会程序主导	初期试图设定高标准，但最终立场受制于理事会妥协
欧盟理事会	平衡创新与监管	审慎推进监管，更新完善妥协文本	为成员国提高在最终决策讨价还价中的地位提供通道
欧盟主要国家（德、法）	灵活监管，保护本国创新能力	通过理事会谈判向立法施加压力	对妥协起到重要推动作用
大型跨国科技企业	自我规范、反对一刀切监管	游说、提交立场文件、旋转门机制	初期主攻阻断监管，后期立场弱化，转向规范监管手段
本土中小企业	争取开源豁免、反对过高合规门槛	借助成员国力量、塑造政策共识	后期政策转向的关键推动力量，影响欧盟治理动力结构

（来源：作者自制）

二、立法决策过程分析

欧盟的立法过程被称为“共同决定程序”（codecision），首先由委员会提出立法提案，期间理事会对与欧洲议会各自对提案内容进行审议并提出修正案。后取决于议会与理事会的意见分歧情况进行一读、二读乃至三读程序，或涉及调解委员会的参与，如议会与理事会最终达成一致则立法通过，而如果议会与理事会无法在法案上达成一致，则立法将宣告失败。^①

^① 朱仁显、唐哲文：《欧盟决策机制与欧洲一体化》，《厦门大学学报（哲学社会科学版）》2002年第6期，第81-88页。

人工智能法案的立法程序与上述一般程序稍显不同，因其涉及了大篇幅的“三方会谈”程序。这一制度并非《里斯本条约》设定的官方制度，而是为了更好地沟通和达成协议而举行的非正式的涉及委员会、议会和理事会三方的会谈。^①虽然正式的立法程序包含了一读、二读、甚至可能的调解程序，但实际上，为了提高效率并尽早达成一致，非正式的三方会谈已成为常态。在第七届欧洲议会任期(2009-2014)中，几乎90%的共同决定立法议案都是通过第一读协议获得通过的，而促成这一现象的关键要素正是机构间的三方会谈，这一程序已经成为立法议案事实上主要的协议场所。^②

在人工智能法案的立法过程中，欧盟委员会于2021年4月正式提出了立法提案，欧盟理事会在2022年12月形成了自身立场，欧洲议会随后在2023年6月通过了自身的谈判立场，在欧盟三方机构均形成自身立场后，三方于2023年6月开始至2023年12月进行了五次三方会谈，以期在三方会谈中解决主要问题以迅速推进立法。在三方会谈中，各方就政治妥协与技术细节进行多层次的谈判，主要着眼于高风险人工智能分类、执法与规范权力分配与通用人工智能监管问题。^③

三方会谈在各种突发状况中最终于2023年12月达成了临时协议。三方会谈的临时协议达成后，虽然理论上议会与理事会仍然能够选择拒绝这一协议，但由于三方会谈自身已致力于弥合各方矛盾，并已经充分考虑到各机构内部对协议的支持，^④三方会谈达成的协议在提交给欧洲议会全体会议和理事会时，极少会被修改或否决。例如，欧盟1999年至2007年的数据显示，96%的三方会谈达成的妥协文本在欧洲议会投票时未做任何修改就获得通过，并被理事会直接接受。^⑤因此，2023年12月临时协议的达成既宣告了此次立法成功，也基本上代表了立法文本的最终形态，研究对立法过程的后续追踪亦截止于此。

^① Gijs Jan Brandsma, “Transparency of EU Informal Trilogues through Public Feedback in the European Parliament: Promise Unfulfilled,” *Journal of European Public Policy*, Vol.26, No.10, 2019, pp.1464–1483.

^② Christilla Roederer-Rynning and Justin Greenwood, “The Culture of Trilogues,” *Journal of European Public Policy*, Vol.22, No.8, 2015, pp.1148–1165.

^③ Clara Hainsdorf, Tim Hickman, Dr. Sylvia Lorenz, et.al, “Dawn of the EU's AI Act: political agreement reached on world's first comprehensive horizontal AI regulation”, December 14, 2023, <https://www.whitecase.com/insight-alert/dawn-eus-ai-act-political-agreement-reached-worlds-first-comprehensive-horizontal-ai>, 2025-05-13.

^④ Tom Delreux and Thomas Laloux, “Concluding Early Agreements in the EU: A Double Principal-Agent Analysis of Trilogue Negotiations,” *JCMS: Journal of Common Market Studies*, Vol.56, No.2, 2018, pp.300–317.

^⑤ Raya Kardasheva, “Trilogues in the EU Legislature”, Reaserch Paper of Department of European and International Studies, King’s College London, July 2012, pp.13-21.

值得一提的是，12月的最后一场三方会谈经历了长达连续三天的谈判，三方都致力于在最后一次三方会谈中取得成果。如果12月仍未达成协议，当时即将卸任的欧盟轮值主席国西班牙将没有动力继续推动立法，而即将上任的欧盟轮值主席国比利时也只有几周的时间来处理如此复杂的问题，因为2024年6月欧洲议会将解散并重新举行欧盟选举。^①这意味着立法达成将更加遥遥无期，难以取得欧盟预想中的立法先发优势。这也在一定程度上造成了立法后期的紧迫形势。

表 3 人工智能法案立法决策主要事件时间节点

时间	主体	事件
2020.2.19	欧盟委员会	发布人工智能白皮书
2021.4.21	欧盟委员会	发布了人工智能法案立法提案（proposal）
2022.12.6	欧盟理事会	通过了对人工智能法案的立场文件，即“一般方法”（general approach）
2023.6.14	欧洲议会	通过议会对人工智能法案的立场文件
2023.12.9	三方机构 ^②	对人工智能法案达成了临时协议
2024.2.13	欧洲议会	投票通过三方会谈后的协议
2024.5.21	欧盟理事会	最终通过人工智能法案

（来源：作者自制）

第二节 企业决策介入分析

一、游说频率与游说手段

人工智能法案野心勃勃地希望成为世界第一部全面监管人工智能的企业的同时，人工智能在谷歌、微软和 Meta 等大型跨国科技企业的商业模式中越来越重要。谷歌在与欧盟委员会的私人会议上将自己描述为“人工智能第一导向公司”，声称其所有产品都由人工智能驱动。^③因此，欧盟的监管努力在决策过程的每个阶段都面临着大型跨国科技企业强烈的游说企图。与此同时，在欧盟“数字主权”

^① Luca Bertuzzi: “EU’s AI Act negotiations hit the brakes over foundation models”, November 10, 2023, <https://www.euractiv.com/section/tech/news/eus-ai-act-negotiations-hit-the-brakes-over-foundation-models/>, 2025-05-13.

^② 指欧盟委员会、欧盟理事会与欧洲议会三方机构

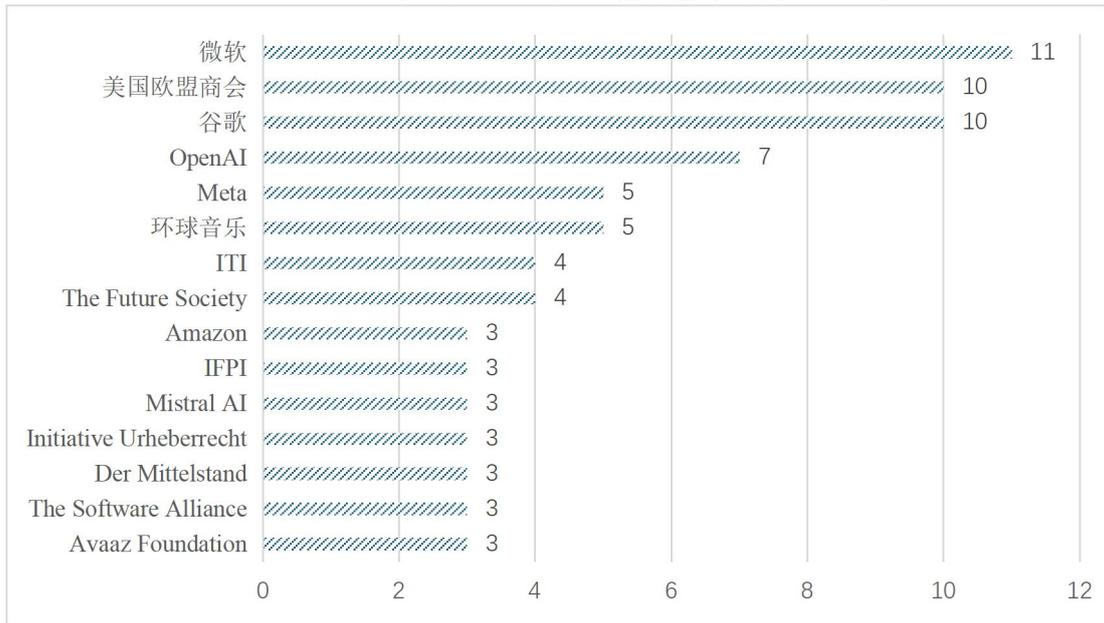
^③ “European Meeting with Google on the AI Act”, 28 June, 2022, Ref. Ares(2024)6547475, pp.1-2.

与“人工智能全球竞争”的话语中，欧盟本土的人工智能中小企业也逐渐深入地介入了立法决策的过程。

大型跨国科技企业与行业协会在立法决策过程中不遗余力地在人工智能法案决策上施加影响：根据欧洲议会欧洲企业天文台的报告与议会公开数据，2023年欧洲议会议员登记了287场有关人工智能法案的会议，其中三分之二的会议与行业与贸易协会举行。在会议频率上，谷歌、微软和欧盟美国商会依旧位列前三，而OpenAI作为顶级游说者中的新人以7次会议场次紧跟其后。欧洲本土的人工智能中小企业中，除已登记3场会议的Mistral AI与德国中小企业协会Der Mittelstand外，其余鲜见于会议记录中。^①

传统上，大型跨国科技企业凭借资源优势在立法进程中占据更大话语权，从而在影响立法决策走向上获得了更大的机会。^②但在欧盟人工智能法案立法过程中，虽然本土科技中小企业对欧盟机构的直接游说较少，但其仍然通过中小企业—归属国家—欧盟机构这一影响路径发挥了超越其规模的影响力，后文将讨论欧盟人工智能法案立法过程中的这一大特点。

表 4 2023 年主要企业欧洲议会游说会议登记数



(来源：欧洲议会游说会议公开数据^③)

^① Corporate Europe Observatory, “Byte by byte: How Big Tech undermined the AI Act”, November 2023, <https://corporateeurope.org/en/2023/11/byte-byte>, 2025-05-13.

^② Shaleen Khanal, Hongzhou Zhang and Araz Taeihagh, “Why and How Is the Power of Big Tech Increasing in the Policy Process?”, pp.53-55.

^③ Data provided by European Parliament: Search MEP meetings, <https://www.europarl.europa.eu/meps/en/search-meetings?textualSearch=&procedureReferences=2021%2F0106%28COD%29&fromDate=&toDate=>, 2025-05-13.

二、游说的竞争力逻辑

在人工智能法案的立法过程中，企业游说不仅依赖于高频的直接接触，还通过信息策略构建了复杂的利益表达框架。竞争力逻辑正是这一框架中的核心要素之一。它不仅能够影响立法机构对技术风险与经济效益的权衡，还影响了不同企业对监管议程的参与路径。^①与简单的利益表达不同，基于竞争力逻辑的游说不仅在商业逻辑上强调严格监管的潜在负面影响，还试图在地缘政治竞争的大背景下将宽松监管视为符合公共利益的合理选择。这种策略传达的信息也在人工智能法案的最后决策关头起到了重要作用。

竞争力逻辑从人工智能时代欧盟的竞争力出发，基于欧盟在全球人工智能竞赛中的相对劣势，将监管视作对市场的干预，认为对通用人工智能的监管将压制技术创新、扼杀初创企业并降低欧盟在人工智能赛道的国际地位。例如，2021年4月微软在与欧盟委员会的一封往来邮件中提及，由于中国对人工智能的监管较小，而投资不断增长，这将威胁欧盟在人工智能全球竞赛中的地位；^②2023年6月一份来自欧洲150多个科技企业高管的联合声明声称“此类监管可能导致高度创新的公司将业务转移到海外，投资者将从欧洲基础模型乃至整个欧洲人工智能的开发中撤资”；^③Mistral AI的首席执行官在2023年10月表示，将要出台的人工智能法案“可能会杀死我们公司”。^④竞争力逻辑不仅严格地将当前阶段的人工智能监管严格置于技术创新的对立面，还将这一叙事放置在了更大的地缘政治竞争语境中；这实际上是在重新定义“合理监管”的边界，将“宽松”视为符合公共利益的默认路径。总而言之，基于其与立法背景的契合性与这一策略被预期的有效性，利用竞争力逻辑向监管方提供信息成为了后续过程追踪中企业影响立法决策的重要手段之一。

^① Pepper D. Culpepper and Kathleen Thelen, “Are We All Amazon Primed? Consumers and the Politics of Platform Power,” *Comparative Political Studies*, Vol.53, No.2, 2020, pp.288–318.

^② “Email exchange between Microsoft and European Commission, 26 April 2021”, in Corporate Europe Observatory: “The lobbying ghost in the machine: Big Tech’s covert defanging of Europe’s AI Act”, Feb 2023, <https://corporateeurope.org/en/2023/02/lobbying-ghost-machine>, 2025-05-12.

^③ “Open letter to the representatives of the European Commission, the European Council and the European Parliament; Artificial intelligence: Europe’s chance to rejoin the technological avantgarde”, signed by German AI Association CEO, Siemens Energy CEO, etc., [https://uploads-ssl.webflow.com/6034e2e35a869ae02e67f55f/649eadd757bee709f4fe1f49_Brandbrief%20en-US%20\(2\).pdf](https://uploads-ssl.webflow.com/6034e2e35a869ae02e67f55f/649eadd757bee709f4fe1f49_Brandbrief%20en-US%20(2).pdf), 2025-05-13.

^④ Zosia Wanat: “‘EU’s AI act could kill our company,’ says Mistral’s Cédric O”, October 27, 2023, <https://sifted.eu/articles/eu-ai-act-kill-mistral-cedric-o>, 2025-05-13.

使用竞争力逻辑的潜在有效性根植于欧盟增强技术竞争力的立法目标，这一目标在预期上与传播欧盟价值观应是一体两面。^①然而，从后续的实践来看，虽然欧盟机构致力于将增强欧盟人工智能技术实力这一实然层面的目标与维护欧盟公民基本权利与价值观这一应然层面的目标结合，但在决策时仍然无可避免地面临上述两个目标的冲突。在立法决策过程中，通用人工智能监管的条款的不断变化即是上述两种逻辑内在动态平衡的体现：一方面，必须确保人工智能的开发和部署不危及数据权利、公平性与透明度等基本社会价值；另一方面，又需避免设置过高的制度门槛，以至于将本已处于全球竞争劣势的欧洲人工智能生态进一步边缘化，在后续的过程追踪中，我们将在对立法后期妥协转向的原因的探索中继续观察这一平衡。

第三章 监管决策过程追踪

第一节 过程追踪节点划分

欧盟 AI Act 的立法过程历经多轮谈判，在前文提到的关于通用人工智能的定义、风险分类与监管方式、中小企业的适应性这三大议题的争议中，科技公司与欧盟机构展开了多轮磋商与博弈。为了分析这些议题如何在政策形成过程中演变，以及不同利益相关方如何影响最终立法结果，下文将对 2022 年 6 月至 2023 年 12 月的立法进程进行追踪，重点关注政策文本的变化、关键行动者的立场转变及其影响路径。

表 5 通用人工智能监管适用条款整体演变

时间	主体	内容
2021 年 11 月	欧盟理事会 斯洛文尼亚 轮值主席	对于 GPAI 的开发和使用，不自动适用法案。只有当 GPAI 具有法案所定义的预期用途，或者被集成到受法案约束的人工智能系统中时，法案才会适用。 ^②

^① European Commission: “Impact Assessment Accompanying the Proposal for a Regulation of The European Parliament and of The Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts”, April 2021, p.30.

^② Council of the EU, “Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Presidency compromise text”, November 2, 2021, p.17.

2022年6月	欧盟理事会 法国 轮值主席	拓宽了法案对 GPAI 的管辖适用，取消了认为 GPAI 本身并无“特定目的(intended purpose)”的表述，默认 GPAI 可能的高风险用途，其在进入市场或投入使用前就需要满足相关法规要求。 ^①
2022年12月	欧盟理事会 捷克 轮值主席	将法案中原本直接适用的高风险人工智能系统要求替换为未来可能实施的法案(future implementing acts)，以“确保成员国充分参与，并对如何在此背景下应用这些要求拥有最终决定权”。 ^②
2023年6月	欧洲议会	在通用人工智能的基础上引入了“基础模型”的概念，并对基础模型的提供商提出了特殊约束，且无法获得开源模型的监管豁免，加强了对通用人工智能的监管。 ^③
2023年12月	欧盟理事会 西班牙 轮值主席	剔除了仅用于内部研究的 GPAI 的监管，但进一步限缩了开源模型的豁免范围。对 GPAI 采取分层监管，被认为具有“系统性风险”的 GPAI 将必须遵循额外义务，判断标准为计算资源使用量等。 ^④

(来源：作者据法案各版本文本整理)^⑤

本研究的时间划分基于上述人工智能法案自 2022 年 6 月起发布的四份关键立法文本。选取的四份立法文本均在决策中扮演了关键节点，例如 2022 年 6 月的理事会版本首次将通用人工智能纳入监管框架中，2022 年 12 月理事会形成了其谈判立场“一般方法”，2023 年 6 月欧洲议会也形成了自身相对强硬的谈判立场，最终在艰难的三方会谈后，立法机构最终在 2023 年 12 月达成了最终的

^① Council of the EU, “Proposition de Règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union - Text de compromis de la présidence - Version consolidée”, June 15, 2022, p.13, 30-31.

^② Council of the EU, “Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Presidency compromise text”, November 25, 2021, pp.4-6.

^③ European Parliament, “Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts”, June 14, 2023, pp.102-103, 115-116.

^④ European Parliament, “Provisional Agreement Resulting from Interinstitutional Negotiations”, February 2, 2024, pp.48-57, 148-150.

^⑤ 为方便说明，本文对不同时间、不同主题发布的法案文本均以“版本”称呼，实际上分析的法案文本存在立法提案、“一般方法”、临时协议及妥协文本等不同类别。

临时协议。这一时间划分旨在通过关键节点的对比，揭示政策制定中动态变化的参与者、事件及引发关键条款变动的逻辑链条。

2022年6月的法国轮值理事国版本初步将通用人工智能纳入法案的监管视野，反映了欧盟理事会监管新兴技术的强烈意愿。2022年12月的捷克轮值理事国版本虽然保留了对通用人工智能的监管措施，但在具体措施与实施时间上做出了让步。2023年6月的议会版本面对企业界的积极游说，反而进行了较为激进的转向，不仅加强了前序文本对通用人工智能的监管，还取消了某些监管豁免，直指功能最强大的通用人工智能。

最后2023年12月的临时协议则呈现为在外部冲击后高度妥协的决策，立法机构在风险分类、开源豁免等方面做出了关键性让步，在立法机构监管意愿、大型跨国科技企业反对意见、中小企业适应性之间进行了微妙平衡。这三个阶段反映出了欧盟人工智能法案立法中的戏剧性过程：从2022年6月到2023年6月，欧洲议会与欧盟理事会仍然旗帜鲜明地试图严格监管通用人工智能，甚至在外部密集的游说攻势下仍然不为所动；而在2023年12月的三方会谈中，立法机构却在诸多方面放弃了严格监管的措施走向妥协，下文的过程追踪将详细梳理这一决策过程，以揭示欧盟这一立法决策的内部动态。

第二节 机制探索与有限妥协：2022年6月—2022年12月

在2022年6月，欧盟理事会法国主席提出了一项远比先前版本严格的通用人工智能监管标准，试图将其纳入高风险系统框架。这一提案迅速引发了欧盟内部及外部科技公司与商业协会的强烈反应，以大型跨国科技公司为代表的行为者展开了高强度的游说，以阻止该版本的严格监管条款延续到后续版本中。

表 6 2022年6月理事会版本

主要争议	2022年6月理事会法国轮值主席版本的回应
风险定位与监管方式	在一定程度上视为高风险，对通用人工智能进行普遍监管，并直接应用部分高风险人工智能的义务。
开源与研究型通用人工智能的豁免	存在“使用说明或通用人工智能系统随附的信息中明确排除了所有高风险用途”的提供者的监管豁免条款，但条件苛刻。

（来源：作者据法案文本整理）

2022年4月，在政治新闻公司 Politico 主办的“人工智能与技术峰会”（AI & Tech Summit）上，谷歌副总裁玛丽安·克罗克（Marian Croak）公开反对该提案，强调不应将通用人工智能模型纳入高风险系统框架。^① 6月，谷歌在与欧盟委员会的一场会议中，明确表态愿意接受技术文档存档与对价值链下游的透明性义务，但强烈反对欧盟将监管负担集中施加于开发者（developer），并特别呼吁对开源模型与研究用途模型提供监管豁免。^②到了7月，谷歌向欧盟委员会提交的立场文件则更为强硬，称高风险框架下的透明度、人类监督和数据治理要求“在实践中难以实现，甚至不可行”，并批评法国轮值主席的提案，指责其过度集中监管负担于通用人工智能提供商，而忽略了价值链中其他主体的责任。谷歌主张，监管应当仅限于旨在用于高风险用途的系统，否则将削弱欧洲本土企业的竞争力，最终不利于欧洲人工智能产业生态的发展。^③

不仅是谷歌，其他科技企业与产业联盟也迅速跟进。2022年11月，包括微软在内的23家研究机构和企业联合致信欧盟轮值主席国捷克，呼吁取消《人工智能法案》中关于通用人工智能的单独章节，并强调“在无法预知通用工具的具体用途时，任何高风险要求都难以得到有效落实”。^④美国政府在10月的一份非正式文件中，也公开支持这一立场，强调通用人工智能提供商对其技术的具体应用场景控制能力有限，因此“要求所有通用人工智能提供商遵守《人工智能法案》的风险管理义务，将导致极高的合规负担，甚至在某些情况下难以执行”。这一表态反映了美国政府对欧盟人工智能监管方向的直接干预，并强化了美国企业的游说立场。^⑤

与此同时，欧盟内部的行业联盟也积极推动调整政策，以保护本国科技企业利益。9月，来自法国、希腊、捷克的科技产业联盟联合发表公开信，警告该提

^① Politico: “Panel discussion - The AI package, one year on: from theory to practice | POLITICO AI & Tech Summit”, April 21, 2022, https://www.youtube.com/watch?v=uvaUQg_q2Ho, 2025-05-13.

^② “BTO Meeting with Google on the AI Act”, June 28, 2022, Ref. Ares(2024)6547475 - 16/09/2024, pp.1-2.

^③ “Google Feedback on general purpose AI systems”, July 17, 2022, Ref. Ares(2022)5217601-18/07/2022, pp.1-3.

^④ Confederation of Industry of the Czech Republic, “Open letter on the proposed regulation of artificial intelligence”, 7 November 2022, pp.3-5.

^⑤ “U.S. Non-Paper regarding the September 2022 Revisions to the Draft EU Artificial Intelligence Act Proposed by the Czech Presidency”, Provided by Open Future Foundation, October 22, 2022, p.4.

案将抑制创新，并呼吁将监管义务从开发者转移至部署者（**deployer**），认为将通用人工智能整体纳入监管将冲击原立法提案中的“基于风险的方法”。^① 瑞典人工智能研究机构 **AI Sweden** 在 9 月的立场文件中也表达了相似意见，但原因是对于开发者的监管可能损害中小企业的利益。^② 商业欧洲（**BusinessEurope**）在 11 月的立场文件中指出，“对通用人工智能系统实施全体监管，与欧盟‘基于风险的方法’相悖”，主张应重新调整 AI 价值链中的责任分配。^③

面对科技企业和商业联盟的持续压力，欧盟理事会最终做出调整。在 9 月末发布的第三版妥协文本中，虽然仍维持了对通用人工智能的监管要求，但决定将具体监管措施推迟，由法案生效 18 个月后通过执行法案（**implementing act**）进行细化。这一改动实质上是一种妥协策略，通过延迟决策来缓和科技企业的反对意见，同时也避免了欧盟内部因政策分歧而陷入立法僵局。

这一调整最终被保留至 11 月 25 日理事会发布的“一般方法”文件（**General Approach**），确立了理事会在立法谈判中的核心立场。然而，这一版本并未最终解决通用人工智能的监管问题，其风险分级仍保持模糊态度，政策制定者刻意回避了关于高风险分类的争议。12 月，欧洲议会在其修正案中更是明确将高风险人工智能限定在“有预期目的”的范围内，从而暂时将通用人工智能排除在高风险人工智能的分类外。文本边缘的注释甚至直接指出：“在讨论之后，通用人工智能将被单独处理。”^④ 虽然科技企业和商业协会的核心诉求在立法文本中得到了部分体现，但监管框架并未根本改变，而是采取了推迟决策、弱化监管义务的方式进行调整。价值链责任分配仍在争议之中，具体监管措施的执行细节依然悬而未决，科技公司的政策干预仍在持续推进。

^① Business Software Alliance (BSA), “Business Joint Industry Statement on the EU Artificial Intelligence Act and High-Risk Obligations for General Purpose AI”, September 27, 2022, p.2-5.

^② AI Sweden, “position paper – AI Act”, September 16, 2022, pp.3-4.

^③ BusinessEurope, “Statement on Potential AI Act General Approach”, November 30, 2022.

^④ Luca Bertuzzi, “Leading MEPs exclude general-purpose AI from high-risk categories – for now”, December 12, 2022, <https://www.euractiv.com/section/tech/news/leading-meps-exclude-general-purpose-ai-from-high-risk-categories-for-now/>, 2025-05-13.

表 7 2022 年 12 月理事会 “一般方法”

主要争议	2022 年 12 月理事会 “一般方法” 的回应
风险定位与监管方式	模糊风险分级，依旧进行普遍监管，但不再直接应用某些高风险人工智能的义务，具体监管措施将在法案生效后 18 个月内由委员会决定。
开源与研究型通用人工智能的豁免	未给予豁免，但延续了对 “使用说明或通用人工智能系统随附的信息中明确排除了所有高风险用途” 的提供商的监管豁免条款。

(来源：作者据法案文本整理)

第三节 游说高潮与文本严格转向：2022 年 12 月-2023 年 6 月

在欧盟理事会捷克轮值主席提出“一般方法”后，随着生成式人工智能的迅速发展及监管方的进一步行动，通用人工智能的监管措施继续成为讨论的中心议题。2023 年 2 月，人工智能法案的主要立法机构提议将使用大型语言模型在无人工监督情况下生成文本输出的人工智能平台归类为高风险，这一提议直接针对以 ChatGPT 为代表的通用人工智能。^①同年 3 月，意大利政府以隐私与网络安全问题为由暂时禁用了 ChatGPT，凸显了欧盟内部对通用人工智能的安全性和合规性日益增长的关注。^②在同月的一份欧洲议会研究 (European Parliamentary Research Service, EPRS) 报告中，研究者援引专家意见呼吁通过评估和测试机制对通用人工智能进行监督和监控，强调让这些工具只掌握在少数公司手中的危险，并指出开发通用人工智能的公司与其部署者之间复杂的依赖关系需要进一步审慎评估。^③

③

^① Luca Bertuzzi, “Leading EU lawmakers propose obligations for General Purpose AI”, March 14, 2023,

<https://www.euractiv.com/section/artificial-intelligence/news/leading-eu-lawmakers-propose-obligations-for-general-purpose-ai/>, 2025-05-13.

^② Shiona McCallum, “ChatGPT banned in Italy over privacy concerns”, April 1, 2023, <https://www.bbc.co.uk/news/technology-65139406>, 2025-05-13.

^③ European Parliamentary Research Service, “General-purpose artificial intelligence”, March 2023, PE 745.708, p.2.

欧洲本土企业联盟在这一阶段继续进行联合行动。波兰私营企业家联合会 Lewiatan 于 3 月发表声明，反对将通用人工智能普遍归类为高风险，主张法律责任应主要由应用而非开发人工智能的企业承担。^①法国 AI 联合会 Hub France IA 在同期发布的立场文件中指出，部分通用人工智能预训练模型难以预见后续具体用途，因此难以满足法案中的合规要求，并警告此类监管可能抑制欧洲中小企业的创新能力，主张放松对开源和研究型人工智能的监管，以促进技术发展。^②

面对监管趋严的趋势，来自美国的大型跨国科技企业也迅速展开新一轮游说，试图影响即将出台的监管框架。2023 年 3 月也成为游说活动最密集的时期，共记录了 67 场相关会议。仅在这一月，美国商会就与欧洲议会议员举行了 7 场会议，而谷歌、微软和 OpenAI 各自举行了 4 场。^③与此同时，公民组织在议会外部会议的占比下降到十分之一。据公开会议纪要，谷歌在 3 月 6 日与欧盟委员会的会面中对理事会提出的“一般方法”表示担忧，并重申新增责任条款可能损害欧洲的人工智能创新。^④OpenAI 在 3 月 31 日的会议中向委员会展示了 ChatGPT 的安全性，试图证明通用人工智能的自我监管能力，并在 5 月提交白皮书，强调通用人工智能整体风险较低，且能够自行监督高风险应用，建议委员会放宽豁免条款，减少通用人工智能提供商的监管负担，转向正面激励机制。^⑤5 月 24 日，OpenAI 的首席执行官萨姆·奥特曼（Sam Altman）表示，如果其公司无法遵守人工智能法案的义务，则可能会退出欧洲。

这一表态遭到了来自欧盟的强烈批评，欧盟内部市场专员蒂埃里·布雷顿（Thierry Breton）强硬批评称这一退出威胁毫无意义，欧盟的人工智能“不可讨价还价”；^⑥而参与起草欧盟法律的荷兰议会议员金·范·斯帕伦泰克（Kim van

^① Polish Confederation Lewiatan, “GPAI Considerations for EU AI Act High-Risk Requirements”, March 2023, <https://lewiatan.org/wp-content/uploads/2023/03/GPAI-Considerations-for-EU-AI-Act-High-Risk-Requirements.pdf>, 2025-05-13.

^② Hub France IA, “Position Paper AI Act”, March 2023, p.6.

^③ Data provided by European Parliament: Search MEP meetings, <https://www.europarl.europa.eu/meps/en/search-meetings?textualSearch=&procedureReferences=2021%2F0106%28COD%29&fromDate=&toDate=>, 2025-05-13.

^④ “Meeting with ... Google’s ...”, March 6, 2023, Ref. Ares(2025)15904 - 03/01/2025, p.2.

^⑤ “OpenAI White Paper on the European Union’s Artificial Intelligence Act”, May 2, 2023, Ref. Ares(2023)3076642, p.2.

^⑥ Foo Yun Chee, “Exclusive: AI rules 'cannot be bargained', EU's Breton says after OpenAI CEO threat”, May 25, 2023, <https://www.reuters.com/technology/eus-breton-slams-openai-ceos-comments-blocs-draft-ai-rules-2023-05-25/>, 2023-05-13.

Sparrentak) 表示, 她和同事“不应该让自己受到美国公司的勒索”, 如果 OpenAI 的产品不能遵守基础的数据治理与透明度要求, 则他们的产品可能并不适合欧洲市场。^① 奥特曼最终于 5 月 26 日收回了其表态, 但取消了后续原定的布鲁塞尔之行, 表现了其对欧盟监管思路仍存的不满。负责起草欧盟提案的罗马尼亚欧洲议会议员德拉戈斯·图多拉切 (Dragos Tudorache) 在路透社的采访中表示, 短期内人工智能法案不会出现任何削弱的情况。^②

与上述外部压力与内部强硬表态并行, 欧洲议会在 2023 年上半年逐步确立了自身的谈判立场。4 月, 议会立法委员会提出在通用人工智能内部采用分层监管模式, 将以 ChatGPT 为代表的一类通用人工智能归类为“基础模型”(foundation models), 并对其施加更严格的质量管理和外部审计要求。^③ 这一概念最早由斯坦福大学研究团队于 2022 年提出, 指“任何在大规模数据上进行训练, 并能够适配多种下游任务的模型”。^④ 虽然基础模型有时被认为与通用人工智能是相通的概念, 但它实际上是通用人工智能中最强大的一类, 不包括其他用于模式识别或翻译的基础类型。^⑤ 无论如何, 欧洲议会试图通过这一概念拓宽对强大通用人工智能的监管空间, 同时尽量减少对其他人工智能系统施加不必要的合规负担。这一策略最终体现在 2023 年 6 月 14 日通过的谈判立场文本中, 其中原有针对通用人工智能的条款被大幅削减至 3 条, 而基础模型的监管条款则增至 16 条, 议

^① Julia Tar, “OpenAI CEO’s backtracks on quitting EU threat”, May 26, 2023, <https://www.euractiv.com/section/artificial-intelligence/news/openai-ceos-backtracks-on-quitting-eu-threat/>, 2025-05-13.

^② Supantha Mukherjee and Martin Coulter: “ChatGPT-maker OpenAI says has no plans to leave Europe”, May 27, 2023, <https://www.reuters.com/technology/openai-has-no-plans-leave-europe-ceo-2023-05-26/>, 2025-05-13.

^③ Luca Bertuzzi: “AI Act: MEPs close in on rules for general purpose AI, foundation models”, April 20, 2023, <https://www.euractiv.com/section/tech/news/ai-act-meps-close-in-on-rules-for-general-purpose-ai-foundation-models/>, 2025-05-13. 需要指出的是, 欧盟后续的表述并未明显区分此二者, 例如, 在 2023 年 3 月的研究文件与 6 月的新闻稿中, 通用人工智能与基础模型概念均被混合使用, 出于研究考虑, 本文也并未对此二者概念做过多区分, 特说明。European Parliamentary Research Service, “General-purpose artificial intelligence”, March 2023, PE 745.708. And European Parliament, “MEPs ready to negotiate first-ever rules for safe and transparent AI”, June 14, 2023, <https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai>, 2025-05-13.

^④ Rishi Bommasani, Drew A. Hudson, Ehsan Adeli, et.al, “On the Opportunities and Risks of Foundation Models”, pp.6-11.

^⑤ Computer & Communication Association(CCIA): “Understanding Artificial Intelligence Ai Foundation Models – Explained”, September 2023, pp.1-3.

会版本较理事会 2022 年 6 月的版本进行了显著调整，几乎直接将文中的“通用人工智能”直接替换为“基础模型”。^①

与前述各方积极的游说活动形成对比的是，议会在 2023 年 6 月通过的这一版本虽然在中小企业保护等特定方面有所妥协，但在整体上则呈现了十分严格的监管态度。但在议会通过的文件中，虽然基础模型（通用人工智能）未被整体归类为高风险，并成为了一个监管的中间类别，但是对通用人工智能开发前和开发过程中的风险控制与信息透明要求增强了，通用人工智能提供商现在需不仅在风险管控上增加了“通过适当的测试、设计和分析”以证明风险已被识别和减轻的义务、证明仅使用符合基础模型数据治理规范数据集的义务，还被要求准备“大量（extensive）”准备大量技术文档和清晰易懂的使用说明，以便利下游供应商能够履行各自义务。同时，能够生成文本、图像、音频或视频等复杂内容的通用人工智能提供商被添加了使用符合数据治理标准的数据集的义务、公开受版权保护数据的使用情况的额外透明度义务与标注人工智能合成内容的义务。一系列新的严格义务被分配给了通用人工智能基础模型。^②除此之外，在 2022 年 6 月理事会版本中出现的例外豁免条款被删除，即使提供商已在使用说明或随附信息中明确排除该通用人工智能系统的任何高风险用途，亦不可得到监管豁免。不仅如此，对开源模型或研究型模型的豁免意见同样没有得到议会的通过。

下表展现了 2022 年 6 月理事会版本至 2023 年 6 月议会版本大型跨国科技企业的主要诉求与议会版本的回应情况。

表 8 2023 年 6 月议会版本

主要争议	2023 年 6 月议会立场的回应
风险定位与监管方式	采取了分层方法， ^③ 对能力较强的基础模型提出了更严格风险控制、信息透明及数据合规的义务，这些模型

^① European Parliament, “Artificial Intelligence Act: Amendments Adopted by the European Parliament on 14 June 2023 on the Proposal for a Regulation of the European Parliament and of the Council on Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts”, June 14, 2023, pp.8-9, 70-71.

^② European Parliament, “Artificial Intelligence Act: Amendments Adopted by the European Parliament on 14 June 2023 on the Proposal for a Regulation of the European Parliament and of the Council on Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts”, pp.115-117.

^③ 需要说明的是，议会的这一处理本质上是将通用人工智能划分为一类单独的监管对象，从而试图规避风险分级上的争议，但这一划分显然也带来了新的争议。

主要由大型跨国科技企业提供。

开源与研究型通用人工智能监管豁免 未给予豁免，同时删除了原有的可通过“证明排除所有高风险用途”来获得豁免的条款。

（来源：作者据法案文本整理）

不难发现，议会并未在两个主要争议点上作出有效让步。虽然对通用人工智能监管应用了分层而非一视同仁的原则，但对基础模型提出的更严格的义务无疑进一步增加了以大型跨国科技企业为主的提供者的合规成本，应用的基础权利评估、投放市场前的一致性测试等义务也远比之前严格和全面。在这一版本被议会通过后，来自科技企业的反对声再次出现，这些论点强调了科技企业在议会立场中未得到满足的诉求，并被整合至最后一轮游说的过程中。

第四节 本土企业介入与最终让步：2023年6月-2023年12月

在议会通过其谈判立场后，法案就进入了立法进程的最后跨机构谈判阶段，即三方会谈。在人工智能法案上，欧洲议会、欧盟理事会与委员会在2023年6月至12月展开了五次主要会议，并最终在12月9日达成了临时协议。这一阶段的企业游说出现了新特点，也产生了不同于前序游说的效果。

这一时期，大型跨国科技企业的立场进行了微妙的转变，比起前述完全拒绝对通用人工智能的监管，他们转而更加关注触发监管的阈值。谷歌在7月与欧洲议会议员的会议中强调，通用人工智能不应被默认为高风险而受到整体严格监管；其在9月提交的一份立场文件中继续表示，对监管应仅关注“被应用于高风险环境中的最有能力的基础模型”，而非所有通用人工智能或所有应用场景，这一表述结合了议会和理事会各自的监管倾向。^①微软同样表达了担忧，并在其10月13日更新的针对三方会谈的文件中建议，监管方应依据每秒浮点运算次数（FLOPs）的计算阈值进一步区分普通基础模型与“高级基础模型（advanced foundation models）”，进而对强大的高级基础模型提出更严格的合规要求。然而，上述两者基本认可了议会立场中的分层监管思路，仅在进一步细化监管、提高标准随技术进步而变化的灵活性上提出了意见。^②

^① Corporate Europe Observatory, “GPAI, Foundation Models and Generative AI in EU AI Act”, November 2023, pp.2-4.

^② Microsoft, “Microsoft Trilogue Recommendations”, October 13, 2023, pp.1-2.

在以基础模型为主的通用人工智能义务上，微软对提供技术文档和使用说明、建立管理系统监督合规情况等要求表示认可，仅对公布具有版权的训练数据摘要一条表示了反对。^①Meta 同样反对发布训练数据摘要的义务，声称人工智能法案并不应该被用于版权管理。^②主要由美国大型跨国科技企业组成的商业软件联盟在 7 月底提出的三方会谈意见同样认可了提供技术文档与风险识别信息等措施，但依旧对外部专家评估、公布有版权的训练数据摘要提出了反对意见，同时延续了对开源模型的监管豁免需求。^③这一阶段，仅有谷歌对依旧反对对通用人工智能提供商的任何监管，即使是可能的分层监管模式，其声称“（监管框架）缺乏明确的概念来区分各个层级和阈值...考虑到（通用人工智能）领域还处于萌芽阶段，这些概念可能并非衡量人工智能风险的最佳指标。”^④

整体来看，美国大型跨国科技企业的话语逐渐从“是否监管通用人工智能”转向了“如何有效监管人工智能”的话题，在监管范围和分层方法也呈现延续立法机构思路的趋势，话语走向谨慎。

然而，在这一阶段，来自欧洲的人工智能初创公司在反对监管上发挥了令人意想不到的作用。6 月底，超过 150 个欧洲大型跨国科技企业与商业联盟高管共同签署了一份联名信，指出在议会最新通过的版本中，基础模型无论其使用情况如何，都将受到严格监管，开发和实施此类系统的公司都将面临不成比例的合规成本和不成比例的责任风险，从而导致科技企业与创新投资外流的风险，最终损害欧洲的竞争力。^⑤7 月德国人工智能协会（German AI Association）针对三方会谈方向的一份报告认为，额外合规义务模糊且难以达成，而开源基础模型的开发

^① Ibid. pp.1-2.

^② Meta: “Meta’s AI Act Position - 4 Column Document”, November 2023, pp.1-2, 6.

^③ BSA: “Joint Industry Statement on the EU Artificial Intelligence Act and High-Risk Obligations for General Purpose AI”, September 27, 2022, <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>, 2025-05-13.

^④ Google, “Considerations regarding a tiered approach for foundation models and general purpose AI”, October 2023, pp.1-2.

^⑤ “Open letter to the representatives of the European Commission, the European Council and the European Parliament; Artificial intelligence: Europe’s chance to rejoin the technological avantgarde”, signed by German AI Association CEO, Siemens Energy CEO, etc., [https://uploads-ssl.webflow.com/6034e2e35a869ae02e67f55f/649eadd757bee709f4fe1f49_Brandbrief%20en-US%20\(2\).pdf](https://uploads-ssl.webflow.com/6034e2e35a869ae02e67f55f/649eadd757bee709f4fe1f49_Brandbrief%20en-US%20(2).pdf), 2025-05-13.

者必须承担商业化应用的合规责任这一点也与欧洲一贯支持开创新的立场相悖。^①

更引人注目的是法国与德国本土人工智能中小企业在人工智能法案立法过程中的活跃参与,尤其是 Mistral AI 与 Aleph Alpha 两家公司。以 Mistral AI 为例,该公司成立于 2023 年 4 月,尽管仍处于初创阶段,却在欧盟决策体系中获得了显著影响力。2023 年 7 月, Mistral AI 即与欧盟通信网络、内容和技术总司 (DG Connect) 负责人会面,讨论欧洲在通用人工智能领域的竞争力,并强调构建本土 AI 能力的战略必要性。^②随后在 10 月,该公司又与多位欧洲议会议员举行了三场闭门会议,重点围绕开源人工智能的监管标准展开游说。^③在国家层面, Mistral AI 也获得了高规格的政治支持,其与另一家法国 AI 公司 Dust 在 2023 年受邀参与法国总统马克龙主持的晚宴。此类场合不仅象征着政策接近性,也体现了法国政府在 AI 政策制定中对本土企业的倾斜;有证据表明,以 Mistral AI 为代表的与会人工智能公司的观点均受到了法国决策层的讨论,并可能通过法院等官方渠道转化为国家意志的表达。^④

更具代表性的是 Mistral AI 通过“旋转门”机制构建的制度性通道。其股东兼公关主管塞德里克·奥 (Cédric O) 曾任法国数字化转型事务国务秘书,离任后直接进入企业管理层。在其任内,奥支持加强人工智能监管,而在加入 Mistral AI 后则公开反对人工智能法案对通用人工智能的过多限制,认为过度监管将削弱欧洲人工智能企业的全球竞争力,使中美受益。此种立场转变清晰反映了企业利益如何通过个人轨迹影响政策态度,也揭示了中小企业与政府之间的密切协商关系在塑造人工智能监管路径中的关键作用。^⑤

^① KI BUNDESVERBAND (German AI Association), “Position Paper on the EU AI Act Remaining Issues and Current Discussions in the European Parliament”, pp.3-4.

^② “Briefing – Meeting with ... on the AI Act”, July 13, 2023, Ref. Ares(2024)4108481, p.4.会议提到,欧洲在获取通用人工智能方面越来越依靠非欧洲公司,73% 的 AI 基础模型来自美国,15% 来自中国,这限制了欧盟对技术的掌控并使其 AI 能力面临风险。

^③ LobbyFacts: “Mistral AI”, <https://www.lobbyfacts.eu/datacard/mistral-ai?rid=539437350687-42#data-card-data-meetings>, 2025-05-13.

^④ Stanislas Polu: “And the state is not kidding about that...”, June 15, 2023, <https://x.com/spolu/status/1669303713440079873>, 2025-05-13.

^⑤ Zosia Wanat: “‘EU’s AI act could kill our company,’ says Mistral’s Cédric O”, October 27, 2023, <https://sifted.eu/articles/eu-ai-act-kill-mistral-cedric-o>, 2025-05-13.

中小企业在人工智能监管立法过程中的影响不仅限于法国。Mistral AI 联合创始人亚瑟·门奇（Arthur Mensch）在 2023 年 11 月的采访中，明确反对 AI Act 中要求基础模型提供商披露训练数据的透明度义务，同时要求监管者仅关注最强大的基础模型。^①在德国，Aleph Alpha 也展现出高度的政策接近性。该公司在 2023 年 6 月至 11 月期间与德国政界高层，包括总理、经济事务部长与数字事务部长等，共举行了 12 场会议。同年 8 月，其创始人乔纳斯·安德鲁里斯（Jonas Andrusis）受邀出席德国内阁会议，与整个联邦政府直接对话。根据其提交的政策文件，Aleph Alpha 与 Mistral AI 立场一致，反对对基础模型进行直接监管，主张应将监管重点放在应用层面。^② Lobby Control 对 Aleph Alpha 的游说活动进行的调查显示，德国部长们逐一讨论了 Aleph Alpha 提出的反对对通用人工智能模型进行过度严格监管的论点，这些观点成功进入了德国最高的决策层，并获得了充分的注意。^③

这类立场通过成员国间的协调机制进一步上升为国家层面的政策诉求。在 2023 年 10 月举行的法德意三国人工智能产业部长会议上，三国的经济、工业与数字事务部长与多家本国科技企业代表展开磋商。会后，德国经济部长罗伯特·哈贝克（Robert Habeck）公开表示，欧洲应制定支持创新的人工智能监管框架，并明确指出通用人工智能不应被纳入过于严格的监管中。这一表态不仅呼应了企业关于以基础模型为代表的通用人工智能不应被强监管的主张，也反映了产业政策在成员国国家利益框架下的重新定位。^④

10 月 17 日，欧盟理事会西班牙轮值主席在《人工智能法案》谈判中为通用人工智能制定了更加细致的监管条款，首次推出了“高能力基础模型”分类，针

^① No Priors: AI, Machine Learning, Tech, & Startups, “No Priors Ep. 40 | With Arthur Mensch, CEO Mistral AI”, November 9, 2023, <https://www.youtube.com/watch?v=EMOFRDOMiU>, 2025-05-13.

^② FragDenStaat, “Stellungnahmen von Aleph Alpha an das BMWK”, December 14, 2023, <https://fragdenstaat.de/anfrage/stellungnahmen-von-aleph-alpha-an-das-bmwk/>, 2025-05-13.

^③ Olivier Petitjean: “AI Act: le troublant lobbying des « champions » européens, Mistral AI et Aleph Alpha”, Observatoire des multinationales, Mars 2024, <https://multinationales.org/fr/enquetes/intelligence-artificielle-lobbying-et-conflits-d-interets/mistral-ai-aleph-alpha-gafam-ai-europe>, 2025-05-11.

^④ Ministry of Enterprises and Made in Italy, “Italy, Germany and France agree on strengthening their cooperation on Artificial Intelligence”, November 2023, <https://www.mimit.gov.it/en/media-tools/news/italy-germany-and-france-agree-on-strengthening-their-cooperation-on-artificial-intelligence#:~:text=Such%20a%20strengthened%20cooperation%20in,start-ups%20and%20SMEs%20through>, 2025-05-13.

对“能力超群、具备广泛应用潜力的模型”提出了更为复杂且严格的合规要求。然而，这一进展并未如预期顺利推进。^① 11月初，法国、德国和意大利三国联合在11月9日的电信工作组（Telecom Working Party）会议上明确反对推动对基础模型实施任何强制监管措施，认为这不仅会威胁到欧洲在全球技术创新中的竞争力，还违反了欧盟一贯推崇的技术中立原则。三国代表表示，基础模型的监管应当通过行为准则（Code of Conduct）进行自我监管，而非直接通过法律法规进行强制性管理。三国代表于此次会议上立场的陡然变化，致使其他与会人员对人工智能法案迅速达成协议的可能性表示悲观。^②

这一立场的提出，反映了成员国间在人工智能监管政策上的深刻分歧，尤其是关于如何处理新兴技术的监管问题。虽然欧盟委员会和议会早已在推动对通用人工智能的监管框架设定中取得共识，但三国的强硬态度让整个立法进程面临严峻考验。在这一关键时刻，欧盟的决策面临了前所未有的压力。

11月19日，法国、德国和意大利联合发布的非正式文件进一步加剧了紧张局势。三国在文件中再次强调，反对对通用人工智能基础模型进行任何强制性的监管，而是支持通过非强制性行为准则来引导模型开发者自律。^③此举直接挑战了欧洲议会的立场，后者认为，若放弃对基础模型的监管，将为无序竞争埋下隐患，并放任技术风险在欧盟范围内蔓延。随着各方分歧的加剧，立法进程陷入了严重的僵局。三方会谈的前景变得异常复杂，部分观察者甚至提出了法案可能面临停滞的风险，给整个立法进程增添了巨大的不确定性。特别是在2023年12月即将举行的最后一轮三方会谈前，欧盟委员会与欧洲议会逐渐意识到，若不能迅速达成共识，法案的顺利通过将面临巨大的挑战，甚至可能导致立法工作完全中断，迫使立法机构必须在创新与风险防控之间找到更加微妙的平衡。

在12月三方会谈前，欧盟立法机构面临前所未有的制度性压力。一方面，法国、德国与意大利在11月的联合反对立场，使得议会原本推动的分层监管框

^① Luca Bertuzzi, “AI Act: Leading MEPs revise high-risk classification, ignoring negative legal opinion”, October 23, 2023, <https://www.euractiv.com/section/artificial-intelligence/news/ai-act-leading-meps-revise-high-risk-classification-ignoring-negative-legal-opinion/>, 2025-05-13.

^② Luca Bertuzzi: “EU’s AI Act negotiations hit the brakes over foundation models”, November 10, 2023, <https://www.euractiv.com/section/tech/news/eus-ai-act-negotiations-hit-the-brakes-over-foundation-models/>, 2025-05-13.

^③ “An innovation-friendly approach based on European values for the AI Act - Joint Non-paper by IT, FR and DE -”, November 2023, pp.1-2.

架几近搁浅；欧洲议会联合报告员在 2023 年 11 月表示，围绕是否应对基础模型实施直接监管的争论已经成为欧盟立法机构之间的重大分歧，其背后是大量针对成员国政府的企业游说压力。^①另一方面，以 Mistral AI、Aleph Alpha 为代表的本土中小企业，联合大型跨国科技企业持续向成员国政府施压，强调严格监管将损害欧洲 AI 产业生态。这一双重压力最终迫使欧盟委员会与议会在 11 月 20 日提出新的妥协文本。^②

该文本最显著的变化之一，是在“高影响力基础模型”的判定标准中引入了定量阈值：每秒浮点运算次数（FLOPs）高于 10^{26} 的模型将被视为“高影响力”，并适用高风险人工智能的监管义务。这一门槛极高，在技术上几乎将绝大多数通用人工智能模型排除在监管之外，也同样远远高于议会的最初预期，^③这意味着欧盟立法机构在核心监管对象的划定上作出了实质性退让。此外，来自德、法、意非正式文件中提出的“模型卡”机制，也被纳入妥协方案中，替代原本议会版本中的强制基本权利影响评估，监管在较大程度上出现了妥协。^④

在 12 月 7 日最后一次三方会谈中，妥协进一步深化。立法机构将“高影响力基础模型”的每秒浮点运算次数门槛降低至 10^{25} ，虽然在技术标准上有所调整，但仍显著高于大多数模型当前能力水平，依然将除 GPT-4 之外的模型排除出了强监管范围。同时，在开源模型监管上，新妥协文本虽未完全豁免监管责任，但明确指出“鼓励其遵守透明度义务”，实质上在监管实践中构成豁免。^⑤同时，最终达成的临时协议也将使用的术语从基础模型回退到了通用人工智能，表现出欧洲议会主导的向更强大的基础模型附加更多义务的监管思路被排除。

^① Natasha Lomas: “Europe’s AI Act talks head for crunch point”, November 14, 2023, <https://techcrunch.com/2023/11/14/eu-ai-act-trilogue-crunch/>, 2025-05-13.

^② Luca Bertuzzi, “AI Act: EU Commission attempts to revive tiered approach shifting to General Purpose AI”, November 20, 2023, <https://www.euractiv.com/section/tech/news/ai-act-eu-commission-attempts-to-revive-tiered-approach-shifting-to-general-purpose-ai/>, 2025-05-13.

^③ Council of the European Union, “Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Analysis of the final compromise text with a view to agreement”, January 26, 2024, p.5.

^④ Council of the European Union, “Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Analysis of the final compromise text with a view to agreement”, January 26, 2024, p.55.

^⑤ European Parliament, “Provisional Agreement Resulting from Interinstitutional Negotiations”, February 2, 2024, pp.45, 49.

更为关键的是，议会此前推动的若干约束性条款——如强制基本权利评估、外部审计义务等——在最终版本中多数被削弱或删除。最终达成的临时协议，只保留了一套适用于所有基础模型的通用透明度要求。^①这一调整不仅体现了立法机构在通用人工智能监管上的重大让步，也标志着科技公司与成员国多重压力在关键立法节点上的集中释放。

表 9 2023 年 12 月临时协议的回应情况

主要争议	2023 年 12 月临时协议的回应
风险定位与监管方式	保留分层监管，但通过触发监管的高门槛将绝大多数通用人工智能排除在监管之外，这些人工智能仅需履行基础的透明度义务，基本实现行业主张。
开源与研究型通用人工智能监管豁免	对已公开模型权重、结构信息并获得开源许可的开源通用人工智能提供监管豁免。

（来源：作者据法案文本整理）

这一轮妥协，不仅结束了三方僵局，也标志着人工智能法案在通用人工智能监管上的最终制度性降调。虽然法案并未完全放弃对以基础模型的监管尝试，但最终所采纳的制度架构，已基本实现了科技企业和部分成员国对灵活性与自主性的关键诉求。尽管临时协议保留了对基础模型的形式监管，但其实际效果与企业所主张的“灵活治理”已非常接近。欧盟最终在成员国压力与企业游说下，进行了通用人工智能监管中的决策回撤。

这一转变同时表明，中小企业的政策主张通过成员国机制实现上升，并进一步影响到欧盟层级的谈判进程。这一现象表明，在人工智能立法博弈中，中小企业不仅通过技术立场介入政策争论，更通过国家与欧盟的欧盟多互动介入政策制定过程，对监管议题的重心与优先顺序产生影响。

^① Ibid, pp.49-51.

第四章 讨论与解释

第一节 本土中小企业与立法的妥协转向

上述过程追踪呈现了人工智能法案立法决策中显著的妥协转向：在数个监管争议场域中，立法机构原本雄心勃勃的严格监管措施在立法后期转向了宽松与豁免。在过程追踪中，我们能够识别欧盟内部中小企业，特别是来自法国的 Mistral AI 与德国的 Aleph Alpha 对立法决策的介入，是这一妥协转向中的关键变量。这与一般游说理论的预期并不相符：具有强大人脉网络、游说资金与渠道的大型跨国科技企业并未发挥与其规模相当的影响力，反而是规模较小的中小企业发挥了与其规模不相符的影响力，其介入成功促成了立法机构在关键条款上的妥协。^①基于这一认知，下文将讨论这一关键变量的因果机制与作用路径。

基于在过程追踪中的观察以及对既往研究的总结，本文对这一转向的原因提出了三个假设：

- (1) 本土中小企业在母国拥有排他且具有特权性质的游说渠道。
- (2) 本土中小企业在游说中采取了与大型跨国科技企业不同的信息策略或信息内容。
- (3) 本土中小企业提供的技术信息被认为更加可信。^②

一、独家游说渠道

在 2022 年 12 月理事会版本（“一般方法”）发布后，企业层面的游说主体仍以美国大型跨国科技企业为主导。典型如谷歌对通用人工智能监管条款的担忧表达，以及 OpenAI 发布的辩护白皮书。2023 年 3 月成为立法提案发布以来外部游说会议最多的一个月份——众多跨国科技企业密集约见议会议员，系统性表达反对立场；与此同时，欧盟本土受监管主体则主要通过行业协会渠道进行意见整合，例如以 France Hub IA 为代表的行业组织。然而，2022 年 12 月至 2023 年 6 月的密集游说并未成功地对欧洲议会的决策施加决定性的压力，在欧盟议会谈判立场正式通过前，以共同报告员布兰多·博尼菲（Brando Bonifei）为代表的议员群体已对外部利益集团的游说压力展现出极强的抗压立场。议会最终通过

^① Juho Lindman, Jukka Makinen and Eero Kasanen, “Big Tech’s Power, Political Corporate Social Responsibility and Regulation,” *Journal of Information Technology*, Vol.38, No.2, 2023, pp.144–159.

^② David Coen, et.al, *Business Lobbying in the European Union*, Oxford University Press, 2021, pp.58-82.

的谈判文本不仅对大型基础模型增设了更多合规义务，更直接废止了欧盟理事会版本中部分已通过的豁免条款。

当立法决策进入三方会谈阶段后，尽管大型跨国科技企业的关注更多集中于严格监管触发阈值的设定，但仍可观察到其对立法进程的持续性干预。从行为模式看，企业在该阶段的游说活动与前一阶段呈现出显著的延续性。然而，不同于前序阶段欧洲本土公司主要以行业集团的形式展开游说介入决策的方式，这一时期以 Mistral AI 和 Aleph Alpha 为代表的欧洲本土独立人工智能中小企业开始迅速发挥影响力。值得注意的是，上述两家企业在三方会谈之前在游说领域仍然处于籍籍无名的状态：Mistral AI 在 2023 年 4 月刚刚创立，在 2023 年 6 月才发布第一批人工智能模型；Aleph Alpha 虽然创立较早，但在这一时期前几乎没有游说方面的努力，创始人乔纳斯·安德鲁里斯（Jonas Andrusis）曾在 2023 年 6 月前表示，作为创新领域的领导者，他们仍未进行任何游说活动。^①

从 2023 年 6 月开始，Mistral AI 和 Aleph Alpha 的母国政府确实给予了二者排外的特权渠道，使其能够直接与母国政府决策层交流。如上文所述，从 2023 年 6 月的 Vivatech 贸易展览会到最近的达沃斯论坛，Mistral AI 的领导层在许多场合都受邀与马克龙（Macron）或其他法国领导人同台亮相，包括参与在公共会议期间在爱丽舍宫举行的晚宴。2023 年 6 月至 11 月期间，Aleph AI 与联邦政府成员之间举行了不少于 12 次游说会议，甚至有一次与德国总理朔尔茨（Scholz）本人会面。^② Aleph AI 提出的反对监管的论点还得到了德国各部门部长的仔细讨论。虽然如此，但大型跨国科技企业显然也具有接触欧盟成员国决策层的权力。罗伯特·戈尔瓦（Robert Gorwa）等人对欧盟数字服务法立法决策的研究显示，与欧盟成员国直接接触是大型跨国科技企业游说欧盟层面立法的重要策略之一。^③然而，在人工智能法案的立法决策中，大型跨国科技企业直接游说成员国政府的记录十分有限，仅存的公开记录包括谷歌在 2023 年 6 月曾邀请法国议员与高

^① Olivier Petitjean: “AI Act: le troublant lobbying des « champions » européens, Mistral AI et Aleph Alpha”, Observatoire des multinationales, Mars 2024, <https://multinationales.org/fr/enquetes/intelligence-artificielle-lobbying-et-conflits-d-interets/mistral-ai-aleph-alpha-gafam-ai-europe>, 2025-05-11.

^② Felix Duffy: “AI Act: Von der KI-Industrie in die Zange genommen”, LobbyControl, Mars 2024, <https://www.lobbycontrol.de/macht-der-digitalkonzerne/ai-act-von-der-ki-industrie-in-die-zange-genommen-114508/>, 2025-05-11.

^③ Robert Gorwa, Grzegorz Lechowski et al., "Platform Lobbying: Policy Influence Strategies and the EU's Digital Services Act", *Internet Policy Review*, Vol.13, No.2, 2024, pp.3-17.

级官员参加其副总裁参加午餐会，以共同商讨人工智能的挑战等。^①在与成员国政府的直接接触渠道方面，大型跨国科技企业显然缺乏此类独家路径。

在人工智能法案的立法决策过程中，我们需承认法德两国政府确实为上述中小企业创造了大型跨国科技企业无法利用的特权渠道，在三方会谈期间为这些中小企业提供了显著的政策接触优势，这种官方渠道不仅增强了本土中小企业话语的政治正当性，也增强了其话语的传递力度。不同于谷歌、微软等公司在人工智能法案的立法决策中并无过多直接接触成员国政府进行游说的记录，这些初创企业通过“高层互动”与“国家背书”的形式直接参与了立法机构案的讨论，甚至部分主张在随后被法德政府转化为成员国立场，在三方谈判中施加了实际影响。同时，在时间节点上，由于在 2023 年 6 月后上述两企业才开始通过游说对立法决策施加影响力，这也为我们建立因果联系提供了一些依据。因此，游说手段相对大型跨国科技企业的异质为中小企业成功促成立法决策妥协提供了一个解释。这一假设进一步验证了企业影响力并不完全等同其资源规模的理论，企业对决策的影响能力在这一场景中更受其构建对自身有利的游说通道的能力影响。

二、信息策略与信息内容

接下来，我们将继续比较二者在游说过程中使用的游说逻辑。我们在前文中介绍了企业在人工智能法案游说中主要使用的竞争力逻辑。在检验了来自大型跨国科技企业与本土中小企业不同的立场文件与公开表态后，我们认为信息策略或信息内容的不同并不能构成对中小企业成功促成立法决策妥协的一个解释：从信息策略上看，两者均使用了类似的内部与外部相结合的信息策略；从信息内容上看，两者的在游说中提出的信息均未偏离竞争力逻辑。

既有文献肯定了信息策略对游说效果的重要影响：游说中的信息策略甚至比信息类型更加重要，它直接关系到游说信息的影响力与可见性。^②在上述立法过程中，大型跨国科技企业与本土中小企业均被观察到使用这两种策略。首先，从内部信息策略来看，本土中小企业通过直接与本国决策层以及欧盟立法机构会面传达信息，例如 Mistral AI 参加爱丽舍宫晚宴、Aleph Alpha 面见德国总理等；大型跨国科技企业则更多通过内部咨询与会议安排，例如发送对人工智能白皮书与

^① Octave Bonnaud, Elsa Trujillo, “Google prospecte les parlementaires pour un déjeuner autour de l'IA”, La Lettre, June 2023, https://www.lalettre.fr/fr/action-publique_lobbying/2023/06/13/google-prospecte-les-parlementaires-pour-un-dejeuner-autour-de-l-ia,109993537-art, 2025-05-11.

^② Iskander De Bruycker and Jan Beyers, “Lobbying Strategies and Success.”, *European Political Science Review*, Vol.11, No.1, February 2019, pp.57-74.

三方会谈阶段文本的反馈意见、密集安排与欧洲议会议员的会议等等。其次，从外部信息策略来看，双方都利用媒体为自身观点造势：OpenAI 对欧盟的退出威胁引起了轩然大波，Mistral AI 和 Aleph Alpha 也在媒体表态称人工智能法案的严格监管可能使其在欧洲难以生存，以至于失去人工智能时代的“欧洲冠军”。当然，二者在外部策略施用频率上也存在区别，大型跨国科技企业在内部会议与咨询意见上的输出内容显著多于本土中小公司，但这同样削弱了信息策略在此次监管妥协中的解释力：更频繁施用的信息策略并未起到相应的干预效果。值得一提的是，虽然美国的大型跨国科技企业在对欧盟的游说领域常被认为“固守美国的思路”，对法案内容进行了错误的指摘与威胁（而非讨价还价与提出替代方案），^①但在此次立法过程中，本土的中小企业同样被观察到了使用威胁策略，二者在信息策略方面的差距实际上不如预期中大。

从信息内容来看，二者提供的信息均未跳出竞争力逻辑的框架。虽然大型跨国科技企业的反监管立场显然有相当部分是出于维护其在欧洲的技术地位与市场利益的目的，但在游说过程中，其仍然使用竞争力逻辑来说服立法机构。例如，谷歌 2022 年 7 月针对通用人工智能的立场文件提到，对通用人工智能应用某些高风险人工智能的要求将“阻碍欧洲各地寻求使用这些系统的企业满足其特定需求的合作和创新的需求”；^②大型跨国科技企业主导的商业软件联盟在 2022 年 9 月发表的意见声称，对通用人工智能新的义务要求将“破坏欧洲人工智能的使用、创新和数字化转型”。^③对本土中小企业而言，他们更加顺理成章地使用竞争力逻辑来正当化自身反监管的立场。2023 年 11 月，Mistral AI 的领导者门奇在社交平台上表示，目前通过的通用人工智能监管思路“对正在崛起的欧洲人工智能生态系统来说是极其不利的...我们只希望制定不会给现有企业（这些企业恰好都不是欧洲企业）带来不公平优势的规则”。^④另一方面，虽然 Aleph Alpha 采用了一贯的欢迎监管的叙事，但仍然在其对人工智能法案的回应中强调，其介入的

^① Andy Tarrant and Tim Cowen, “Big Tech Lobbying in the EU”, *The Political Quarterly*, Vol. 93, No. 2, June 2022, pp.219-222.

^② “Google Feedback on general purpose AI systems”, Ref. Ares(2022)5217601, July 18, 2022.

^③ BSA: “Joint Industry Statement on the EU Artificial Intelligence Act and High-Risk Obligations for General Purpose AI”, September 27, 2022, <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>, 2025-05-11.

^④ Arthur Mensch, “Mistral AI’s position on the AI Act”, November 16, 2023, <https://x.com/arthurmensch/status/1725076260827566562>, 2025-05-11.

重点目的是“让监管有效...以进一步促进欧洲负责任的人工智能的创新、研究和发 展。”^①

综上所述，无论是在信息策略的运用上，还是在信息内容的叙事上，大型跨国科技企业与本土中小企业在人工智能法案游说过程中的表现都展现出高度相似性，均采用内部外部相结合的策略，并围绕“欧洲竞争力”构建其主张。然而，这种高度一致的信息实践却带来了截然不同的游说效果——大型跨国科技企业未能推动其关键诉求进入最终法案，而本土中小企业却在关键条款上促成了监管的实质性妥协。这表明，游说中使用的信息策略与信息内容的差异本身并不足以解释政策结果的差异。

三、技术信息的可信性

在上一段，我们讨论了信息策略与信息内容对立法妥协的潜在影响，接下来我们将讨论企业提供的技术信息在立法机构视角下的可信性是否对立法妥协产生了影响。这一思路受到欧盟游说研究中的“信息换取准入 (information in exchange for access)”框架启发，这一框架描述了企业通过向欧盟立法决策部门提供决策所需的技术信息以换取在立法决策中的内部地位的现象。^②在人工智能等新兴高技术领域，立法机构对技术原理、模型能力及风险评估普遍不具备内部专业知识，因此外部技术信息在立法过程中的地位更加关键。

然而，在本次人工智能法案的立法中，我们并未观察到 Mistral AI 或 Aleph Alpha 提供了显著高于其他企业的信息增量——相比之下，拥有雄厚研究资源的大型跨国科技企业，如谷歌、OpenAI 等，在立法前期就已频繁提交详尽的技术意见与评估材料。在 Mistral AI 对过度监管的批评与 Aleph Alpha 对创新空间的呼吁中，其所提供的信息多集中在本土生态系统发展受限的产业叙事上，技术细节与风险框架贡献相对有限。因此，如果仅从技术信息供给能力来看，中小企业并不具备更强的可信性基础。

大卫·科恩 (David Coen) 等人的研究指出，政策制定者在接收和使用企业提供的信息时，会严格地审查其可信性，他们更倾向于从那些被认为具有合法性和代表性的利益集团获取信息，这些利益集团不仅需要提供高质量的信息，还需

^① Natasha Lomas, “France’s Mistral dials up call for EU AI Act to fix rules for apps, not model makers”, Tech Crunch, November 16, 2023, <https://techcrunch.com/2023/11/16/mistral-eu-ai-act/>, 2025-05-11.

^② Pieter Bouwen, “Corporate Lobbying in the European Union: The Logic of Access,” *Journal of European Public Policy*, Vol.9, No.3, 2002, pp.365–390.

要在政策制定过程中表现出透明度和责任感。^①既有的对欧盟数字市场法案的研究为大型跨国科技企业的游说失败做出了基于信息可信性的解释：由于大型跨国科技企业市场行为和商业模式已经被视为潜在的垄断者，它们在欧盟的声誉受到了损害，提供的信息因此被认为更加不可信。本土中小企业和行业参与者则通常被视为更“中立”的信息来源，因为他们往往与大型跨国科技企业存在直接的竞争关系，并且更加了解本地市场与监管环境。^②这与 Mistral AI 和 Aleph Alpha 的态度十分吻合：Mistral AI 面对大型跨国科技公司，声称“我们与所有人竞争”，直指非欧盟的人工智能巨头；^③而 Aleph Alpha 也被认为是欧洲对于美国在人工智能领域压倒性地位的有力回应。^④

一些证据显示来自本土中小企业在特定政策窗口中被赋予了更高的认可与信任。例如，法国与德国在 2023 年联合提出修改议案，主张削弱对基础模型的监管强度，其背后的立场与中小企业表达的关切高度一致。在这一过程中，Mistral AI 和 Aleph Alpha 的观点被多次在媒体与政策讨论中引用为“代表欧洲创新利益”的声音。^⑤同时，在 Mistral AI 频繁参与法国决策层活动后，法国数字部长于 11 月公开表示，“我们必须避免用过于严格的监管压垮欧洲创新者”。^⑥德国经济部长在 11 月的法、德意三国部长会议中也表达了类似的观点，这些表态均被认为受到了来自本土中小企业的强力推动。^⑦这种信息被采纳的现象，可能更多源于其身份所承载的象征性政治价值——即代表本土竞争力的技术叙事——而非技术信息本身的专业可信性，这些因素使得本土中小企业提供的技术信息天然地比大型跨国科技企业提供的信息“更可信”。

^① David Coen, et.al, *Business Lobbying in the European Union*, pp.58-82.

^② Andy Tarrant and Tim Cowen, “Big Tech Lobbying in the EU”, pp.216-222.

^③ Tim Bradshaw, et.al, “‘We compete with everybody’: French AI start-up Mistral takes on Silicon Valley”, November 7, 2023, <https://www.ft.com/content/387eeeab-1f95-4e3b-9217-6f69aeeb5399>, 2025-05-13

^④ Ola Adach, “Aleph Alpha's OS LLM might be the answer to US AI dominance: German OS AI startup raises €486 million in Series B financing”, November 8, 2023, <https://interoperable-europe.ec.europa.eu/collection/open-source-observatory-osor/news/aleph-alpha-has-os-llm-might-be-answer-us-ai-dominance>, 2025-05-13.

^⑤ Morgan Heaker, “Meet Aleph Alpha, Europe’s Answer to OpenAI”, Wired, August 30, 2023, <https://www.wired.com/story/aleph-alpha-europe-openai/>, 2025-05-11.

^⑥ Olivier Petitjean: “AI Act: le troublant lobbying des « champions » européens, Mistral AI et Aleph Alpha”, Observatoire des multinationales, Mars 2024, <https://multinationales.org/fr/enquetes/intelligence-artificielle-lobbying-et-conflits-d-interets/mistral-ai-aleph-alpha-gafam-ai-europe>, 2025-05-11.

^⑦ Corporate Europe Observatory, “Trojan horses: how European startups teamed up with Big Tech to gut the AI Act”, March 11, 2024, <https://corporateeurope.org/en/2024/03/trojan-horses-how-european-startups-teamed-big-tech-gut-ai-act>, 2025-05-11.

当然，限于研究设计与可获取材料的限制，研究难以直接识别立法机构在信息采纳过程中的具体偏好及其对不同信息源“可信性”的主观判断。因此，技术信息可信性在此次立法妥协中所起的实际作用仍然更多依赖间接迹象与理论预设。未来研究可借助更系统的文本追踪方法或对欧盟决策人员的深度访谈，进一步揭示信息可信性在新兴技术治理过程中的真实机制。

第二节 欧洲竞争力与企业—成员国意见传导

上文讨论了可能引起立法过程中的重大妥协的诸多因素，下文将进一步讨论欧洲竞争力这一话题与中小企业—成员国的意见与压力传导机制。从欧盟委员会最初的立法提案中不能看出，人工智能法案在最初底层就存在着一组创新与竞争力—价值观与基本权利的矛盾，人工智能法案面临的首要挑战也正是在促进创新和有效监管间达成平衡。^①发布立法提案的委员会专员玛格丽特·维斯塔格（Margret Vestager）表示，“...欧盟正在引领新的全球规范的制定...我们可以为全球范围内的道德技术铺平道路，并确保欧盟在此过程中保持竞争力。我们的规则面向未来且有利于创新，将在欧盟公民的安全和基本权利受到威胁时，在绝对必要的情况下进行干预。”^②一方面，欧盟试图通过一贯的高标准严监管巩固自身在新兴技术上的全球监管领导者地位；另一方面，欧盟也试图在此基础上构建自身的创新生态与在技术上的全球竞争力。与其他国家统一时期发布的人工智能战略相比，欧盟人工智能法案的一大特点就是其对欧盟价值观的强调与整合。^③然而对通用数据保护条例的研究表明，侧重价值观的法规在针对 AI 等快速发展技术实施时，会面临来自技术、法律、商业利益和监管能力等多方面的更大的阻力。^④这种阻力常常被描述为一种“创新的刹车”（a brake on innovation），体现了技术创新与价值观与基本权利保护间仍旧紧张的关系。在人工智能法案的立法决策过程中，法案诸多版本的变动也能够体现出此两者的此消彼长。

^① Lorenzo Ricciardi Celsi, “The Dilemma of Rapid AI Advancements: Striking a Balance between Innovation and Regulation by Pursuing Risk-Aware Value Creation,” *Information*, Vol.14, No.12, 2023, pp.4-8.

^② European Commission, “Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence”, April 21, 2021, https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682, 2025-05-11.

^③ Xukang Wang and Ying Cheng Wu, “Balancing Innovation and Regulation in the Age of Generative Artificial Intelligence.”, *Journal of Information Policy*, Vol.14, pp.387-395.

^④ Araz Taeihagh, M. Ramesh and Michael Howlett, “Assessing the Regulatory Challenges of Emerging Disruptive Technologies.”, *Regulation & Governance*, Vol.15, No.4, October 2021, pp.1009-1019.

基于此，本文进一步认为，人工智能法案立法决策后期的显著妥协实质上是技术创新逻辑权重逐渐大于价值观与基本权利保护逻辑产生的压力的结果，这一压力经由欧盟特殊的多层次治理结构最终从成员国内部层面传导至欧盟立法机构层面。这一转变表面上是国家自主性的恢复，实质上则体现了成员国在受到本国科技企业利益驱动下，主动构建起政策联盟的策略性介入。在欧盟，成员国之间的分歧常常成为其监管理念落实到现实的重要阻力，^①这一转变由欧盟内部人工智能产业最为发达的德国和法国主导也说明了这一点。

欧盟在世界人工智能技术竞赛中的相对落后并非秘密——2019年麦肯锡撰写了一份欧盟人工智能技术发展报告，其中明确指出欧盟在人工智能研发投入、市场份额等层面与主要竞争对手中美之间的差距。^②而作为“欧洲明星”，Mistral AI与Aleph Alpha的迅速崛起无疑为其母国与欧盟注入了相当的信心。建立于2023年4月的Mistral AI在短短两个月中筹集了超过1亿欧元的融资，Aleph Alpha也在11月宣布共获得了超过5亿欧元的投资，在某些美国媒体的叙述中，Mistral AI已经成为OpenAI未来的主要竞争对手之一。^③欧盟委员会在2023年6月与Mistral AI的会议中承认：“73%的人工智能基础模型来自美国，15%来自中国，对外部公司的依赖限制了欧盟对其技术的控制，并使人工智能能力面临风险”，因此，欧盟需要开发欧洲的大型人工智能模型，以确保其人工智能“主权”与不依赖非欧盟企业的人工智能技术能力。^④

这类被寄予厚望的本土企业，逐步被建构为欧洲“科技主权”的象征性代表，而企业自身也表达了巩固欧洲“科技主权”或“人工智能主权”的意愿。^⑤在这种象征框架下，Mistral AI与Aleph Alpha的崛起不仅是产业发展的自然结果，也被赋予了在中美技术竞争中重塑欧洲位置的战略使命。这种政治赋权机制使得

^① Julia Rone, “‘The Sovereign Cloud’ in Europe: Diverging Nation State Preferences and Disputed Institutional Competences in the Context of Limited Technological Capabilities,” *Journal of European Public Policy*, Vol.31, No.8, 2024, pp.2343–2369.

^② McKinsey Global Institute: Notes from the AI Frontier, Tackling Europe’s Gap in Digital and AI, February 2019, pp.5-14.

^③ France 24: “French start-up Mistral AI emerges as a leader in European artificial intelligence”, December 11, 2023, <https://www.france24.com/en/europe/20231211-french-start-up-mistral-ai-emerges-as-leading-force-in-european-artificial-intelligence>, 2025-05-11.

^④ European Commission, “Briefing – Meeting With (Mistral AI) on the AI Act.”, Ref. Ares(2024)3108481, June 7, 2024.

^⑤ Kyle Wiggers, Devin Coldewey: “This week in AI: Mistral and the EU’s fight for AI sovereignty”, December 9, 2023, <https://techcrunch.com/2023/12/09/this-week-in-ai-mistral-and-the-eus-fight-for-ai-sovereignty/>, 2025-05-11. And Aleph Alpha: “Luminous: European AI closes gap to world leaders”, Feb 2023, <https://aleph-alpha.com/luminous-european-ai-closes-gap-to-world-leaders/>, 2023-05-11

原本处于政策边缘的中小企业主张，获得了超出其市场规模与技术成熟度的关注度与影响力。在欧盟治理结构中，这种国家层面对企业声音的放大，最终在立法协商阶段显著改变了原有的监管平衡。

在立法决策过程中，虽然法德两国在公共论述中仍然保持对“欧洲价值观”与“基本权利保护”的持续承诺，^①但在具体立法技术上，却明显做出了重要让步。原方案中对高风险系统的强制性审查机制被调整为更具灵活性的行业自律，基础模型的透明度与解释义务亦被弱化为更低程度的风险披露。从结果上看，这种重大的妥协不仅是企业影响的结果，更体现了这些国家在产业现实面前的价值排序逻辑，当本土技术冠军的潜力与全球竞争态势产生冲突时，部分成员国选择了推动在实质的制度设计上降低监管强度。

这进一步说明，在欧盟的多层治理结构中，掌握技术资源的行为体能够在决策中占据优势，其相同的技术利益不仅被本土中小企业自身代表，还在成员国与欧盟层面被再次代表。^②欧盟内部的成员国扮演了某种技术企业的代理人角色，将后者的政策诉求政治化、制度化，并将技术合规争议上升为主权议题，形成了一种本土中小企业—欧盟成员国的压力与意见传导机制与政治联系。

国家级行为体的立场转换成为了本土中小企业影响最终政策结果的关键机制，而这一机制受到实际经济利益与产业竞争力的强大驱动，这与欧洲近年来强调全球技术竞争，逐渐地缘政治化的倾向相吻合。^③同时，在欧盟层面，我们能够观察到委员会与议会作为的超国家机构，更倾向于强调保护基本权利与统一市场标准；而成员国政府，尤其是产业基础较强的国家（如法国、德国），则更敏感于国内企业的竞争处境，更易接受以“激励而非限制”为导向的监管安排。在此过程中，企业信息策略与产业利益通过影响成员国立场，间接塑造了最终的监管妥协结构。

^① Héloïse Fayet: “French thinking on AI integration and interaction with nuclear command and control, force structure, and decision-making”, European Leadership Network, November 2023, https://www.europeanleadershipnetwork.org/wp-content/uploads/2023/11/French-bibliography_AI_Nuclear_Final.pdf, 2025-05-11.

^② [德]贝娅特·科勒-科赫等：《欧洲一体化与欧盟治理》，顾俊礼等译、刘立群校，中国社会科学出版社 2004 年，第 258-265 页。Also see Otto Holman, “Asymmetrical Regulation and Multidimensional Governance in the European Union,” *Review of International Political Economy*, Vol.11, No.4, 2004, pp.714-735.

^③ 解楠楠、张晓通：《“地缘政治欧洲”：欧盟力量的地缘政治转向？》，《欧洲研究》2020 年第 2 期，第 26-33 页。

第三节 技术风险界定与偏好表达

最后，本节讨论将讨论与贯彻全文的监管概念相对立的风险概念，在监管的框架之外通过技术风险的角度审视研究对象与研究结果。在安全研究领域，作为一种客观因素的“风险”逐渐取代了基于主观认知的“威胁”，成为了反向定义安全的新依据，具有高度不可预测性的通用人工智能技术作为技术领域的新风险因素，必然会引起监管方的风险焦虑，从而试图迅速构建相应的监管框架。需要指出的是，通用人工智能技术的高度不可预测性暗合了风险研究中风险概念由“可计算的不确定性”到“不可接受的不确定性”的转变，这一转变也使风险真正从被认为可知的客观事实转化为一种社会政治事实，^①故而，对通用人工智能监管的讨论并不拘泥于具体技术细节，而关注作为政治概念的监管与风险本身。

监管（或治理）作为风险控制的一种手段出现，不仅是对将风险治理到何种程度的选择，也对将风险治理容忍到何种程度的选择。^②换句话说，上文反复出现的创新—监管平衡实质上是监管方在治理—容忍的光谱中将决策向治理一端推动与新兴科技企业将决策向容忍一方推动间的拉扯。然而上述治理—容忍光谱的冲突立足于一个关键问题，即欧盟在人工智能立法中治理或容忍的风险究竟为何物。这一问题固然有其技术的面向，但作为具有高度不可预测性的通用人工智能，在技术上如何界定和监管通用人工智能的风险本质上是政治的，反过来说，人工智能法案的决策过程实际上是一个风险界定的过程，科技企业—立法机构的冲突本质上是风险界定的冲突。前文反复提及了通用人工智能的两大争议场域，而之所以称之为“争议场域”而非“风险点”，是因为风险在这一决策过程中是尚未确定的，通用人工智能的风险在决策过程中既是监管对象，也是双方争夺界定的对象。

我们已经知道，欧盟的人工智能立法本质上是一种风险驱动的政治，虽然我们发发现欧盟在监管上同样存在经济利益与技术竞争力的考量，但就其立法动因而言，控制通用人工智能技术在两个层面上的风险是重要的：首先是欧盟过分依赖外部人工智能提供商的风险，欧盟希望通过人工智能法案培育负责任的本土创新生态环境，以捍卫其技术主权；其次是具有高度不确定性的通用人工智能本身的滥用、误用与系统性风险等。通用人工智能技术的风险并非指向其今时今地造成

^① 庞珣：《反思与重构：全球安全和全球治理的风险政治理论及方法》，《国际政治研究》2024年第2期，第14页。

^② 唐士其、庞珣：《综合安全论：风险的反向界定和政治逻辑》，《国际政治研究》2022年第6期，第20页。

的实际损害，而是对其造成损害的潜在能力的判断。^①因而，使用“基于风险的方法”的人工智能法案不仅承担着监管这些风险的任务，也在底层上起着判断、规定这些风险的作用。在仔细观察上文提到的两个主要争议场域后，我们能够发现，它们本质上从属于风险界定的议题，而非只是单纯的技术细节：第一个争议场域关注通用人工智能的整体风险在欧盟的风险分级光谱中位于什么位置，第二个争议场域则关注具体的开源与研究型人工智能的风险表达，这些风险是否能够被容忍到豁免监管。

欧盟关于风险的认知全系于其内部行为体的风险认知，内部行为体的风险认知与分布通过制度结构被最终表达，而行为体自身对风险的界定既受知识和信息的影响，又受价值观与文化传统的影响。^②首先，对风险的界定需要科学技术与专业知识的背书，这是欧盟在决策过程中开展社会咨询、接受游说会议的主要原因。首先，在科学技术与专业知识的角度上，人工智能产业中的企业是优势方，他们能够通过自身较监管方掌握的更多信息来使自身的风险界定正当化，而作为依靠通用人工智能技术获取技术优势地位与经济利益的科技企业，他们自然地偏好将通用人工智能的风险界定标准提高，这也是我们最后看到的妥协形态（例如将判定“高影响力”通用人工智能的每秒浮点运算次数的阈值提高到几乎没有企业在现阶段能达到的水平），而作为相对缺少专业知识的监管方，虽然其具有扩大风险界定范围的本能，但其专业信息的来源仍需依靠技术企业，这导致了欧盟立法机构在通过知识和信息塑造监管框架中的弱势地位，而为对抗这一弱势地位，欧盟立法者也将表达出强硬的斗争态度。这一基于风险界定的逻辑解释了前文中的创新—监管之争，立法方与科技公司两者具有内生的不同的风险界定偏好，这在技术层面表现为通用人工智能的风险定位与监管手段、开源与研究型人工智能的监管争议上。

前述分析遵循了科技企业—监管方的二分关系，然而，在价值观与文化传统方面，欧盟（及其内部行为体）—欧盟外部行为体的关系在解释上发挥了主导作用。价值观决定了人们对待风险分配（或风险容忍）的组合方案的态度。^③贾斯卡·贝里斯(Jasca Bareis)指出，人工智能政策作为一种“述行性政治”(performative politics)，价值观不仅是其文本中的静态原则，更是塑造技术治理路径、定义市

^① Ortwin Renn and Christina Benighaus, "Perception of Technological Risk: Insights from Research and Lessons for Risk Communication and Management," *Journal of Risk Research*, Vol.16, No.3-4, 2013, pp.293-313.

^② 唐士其、庞珣：《综合安全论：风险的反向界定和政治逻辑》，第 23 页。

^③ 同上，第 23 页。

场规则的重要力量。^①，虽然二者在技术经济效应的利益政治上有所分野，但欧盟与本土中小企业仍然共享同一套价值观，即欧盟技术主权与技术创新。如果说前文描述的风险界定中的科技企业—监管方的二分关系解释了为何在通用人工智能监管出现巨大争端以及为何出现妥协，那么此处欧盟（及其内部行为体）—欧盟外部行为体基于价值观与传统文化的二分关系则解释了为何中小企业的干预比跨国大型科技企业更加奏效，虽然本土中小企业作为技术提供商与大型跨国科技公司共享某些偏好，但本土中小企业与欧盟立法机构共享的价值观与传统文化使欧盟立法机构作为监管方更容易出现与内部行为体一致的风险界定偏好，它们在风险定义上的立场更符合欧盟长期推崇的技术主权和本土创新价值。这种共同的政治价值和文化认同，使得中小企业在立法博弈中能够更有效地推动自身主张，其对风险界定的偏好在欧盟被多次代表，并最终通过制度结构表达了出来，本土中小企业从而在风险定义的争议中赢得了关键性突破。

总而言之，我们在前述决策争端中看到的争夺与拉扯过程在实质上是介入决策的行为体在风险界定上的冲突，相异的背景条件塑造了行为体对风险的不同界定偏好与容忍程度，这些风险界定上的偏好在决策过程中不断融合、分离、重组，最终形塑了人工智能法案的风险界定方式与监管方式。

^① Jascha Bareis and Christian Katzenbach, “Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics,” *Science, Technology, & Human Values*, Vol.47, No.5, 2022, pp.855–881.

结语

本研究聚焦通用人工智能监管这一国际政治中的新兴议题，从国际政治经济学视角出发，以欧盟人工智能法案中通用人工智能监管决策为研究案例。文章基于欧盟机构、成员国及专家发布的政策文件、磋商意见与立法草案，以及参与决策过程的企业提交的立场文件等一手资料，结合相关学术研究与媒体报道，运用过程追踪与一致性检验相结合的方法，细致剖析了欧盟在立法初期主张强监管，而在后期却出现重大妥协的决策转向。研究揭示了这一重大转向背后的政治动态与逻辑，旨在为理解新兴技术条件下的全球治理演进提供理论与经验启示。

研究跳出既有技术监管研究局限于法律或技术视角的传统框架，转而从政治维度剖析通用人工智能带来的新风险与监管对策。研究首先关注了欧盟人工智能法案的立法目的与背景，指出通用人工智能的突现性风险对立法理念与进程的挑战及随之而来的主要监管争议；随后梳理了监管决策中主要的行为体及其行为模式与偏好，特别关注了其中科技企业对法案的游说情况与策略；在此基础上细致构建了2022年6月至2023年12月间决策主要事件的时间序列，确定了决策中各行为体的行为及相应节点的立法文本反馈，逐步厘清其中的因果线索，在最后进行系统归纳与总结。分析发现，本土中小企业与欧盟人工智能竞争力考量在决策的重大妥协中扮演了重要角色。在欧盟层面与欧盟成员国层面对培育本土人工智能竞争力的偏好下，本土中小企业通过渠道与信息可信性上的优势影响了本国立场，从而对此次决策中的重大妥协产生关键影响。另外，决策过程中各行为体意见的竞合归根结底是在风险界定上的冲突，不同主体基于各自不同的背景条件，形成了对风险界定的独特偏好。在决策过程中，这些风险界定偏好不断相互作用，经历融合、舍弃与重组的过程，最终经由欧盟制度结构被表达出来。

首先，对企业决策介入的分析发现，本土中小企业利用在排他游说渠道与信息可信度上的优势，在决策上获得了大于跨国大型科技企业的影响力。作为构建欧洲主权人工智能的关键力量，欧洲本土中小企业被赋予了直接接触母国决策层的渠道。在成立初期，它们即可参与国际人工智能论坛、出席政府高层研讨会，并与最高决策者直接沟通。通过这些渠道，中小企业能够有效向母国决策层输出通用人工智能监管意见，从而塑造有利于自身的监管决策；大型跨国科技企业则一般只能依靠公开游说渠道进行观点输出，对决策的影响力大打折扣。另外，欧洲本土中小企业在地理上归属欧盟，且与大型跨国科技企业在人工智能上具有直接的竞争关系，还承载了欧盟技术主权的竞争力叙事，这使他们向立法机构提供的信息更倾向于被认为中立与可信，因而更易被采纳，从而增强其对决策的影响力；而大型跨国科技企业在欧盟则常被视为市场垄断者或对欧盟技术主权的潜在威胁，其提供的信息因此在欧盟立法机构处更值得怀疑、缺乏可信度。利用上述机制，本土中小企业有力地推动了法案监管决策中的重大妥协。

其次，研究还探索了“本土中小企业—成员国—欧盟”多层次互动中的压力传导机制，聚焦决策转变中的经济利益与竞争力因素。在推动决策妥协的过程中，中小企业将自身政策诉求充分传递至母国决策层，再由母国决策层传导至欧盟层面，最终推动了最终文本在监管强度上的大幅弱化。与此同时，以法国、德国为代表的产人工智能基础较好的欧盟强国，在欧洲人工智能技术竞争力逻辑的推动下，逐渐弱化了其在基本权利保护上的承诺，最终推动监管决策走向妥协。这些国家在一定程度上充当了技术企业的代言人角色，将这些企业的政策诉求转化为政治议程，并通过制度渠道将技术合规争议提升至欧盟层面，最终使用基于本土中小企业诉求的国家意志影响了欧盟层面的决策。这揭示了在欧洲技术主权话语下，中小企业作为“本土技术代表”的重要象征性角色，其不仅代表着欧洲在全球技术竞赛中的竞争力诉求，还通过各种手段塑造了成员国在立法谈判中的立场；其通过与母国的政治—经济联盟，利用欧盟多层治理的制度结构传递并表达了自身诉求与偏好，形成了其对决策结果的影响力。

最后，研究试图解释人工智能立法过程中围绕通用人工智能监管的本质矛盾，即不同行为体在风险界定上的深层冲突，这一分析也揭示了在价值观与传统文化上的一致性对中小企业游说成果的显著影响。虽然这一决策过程同样存在对经济利益、技术竞争力以及全球规范设定等其他复杂因素的考量，但其核心始终围绕着对风险的不同理解与界定。具有高度不可预测性的通用人工智能挑战了以往的

风险治理手段，使其风险界定方式与适用监管手段成为重要的政治议题。这表明新型技术监管，相较于一般社会经济领域的监管，因其无先例可参照，所以对社会价值与监管偏好具有更强的反映与表达能力。在本案例中，风险界定的过程不仅涉及技术判断，还受到政治价值与制度结构等重要影响，大型跨国科技企业由于掌握更多的信息和技术优势，往往试图将通用人工智能的风险界定为可控的、可管理的，以维持自身的商业利益和技术主导地位，而本土中小企业在风险界定上虽然与大型跨国科技企业在提高界定阈值等方面存在类似之处，但根源于其政治归属与文化传统，其在风险界定上的偏好更符合欧盟长期以来所倡导的技术主权与本土创新的政治价值。这种共享的价值观和文化传统，使得中小企业能够在立法过程中更加有效地影响监管框架，并在风险定义的拉锯战中取得关键突破。从本质上看，这些争议不仅是对具体技术细节的讨论，更是对风险如何界定、管理与容忍的深层分歧。风险不仅是监管的对象，更是各方争夺定义的核心领域。正是在这一风险建构的过程中，大型跨国公司、立法机构和本土中小企业之间的拉锯与博弈，共同确定了最终的决策妥协与监管路径。这不仅揭示了新兴技术监管的复杂性，也突显了监管政策背后的深层风险政治逻辑。

本研究仍然存在一定的局限性。首先，从材料获取来看，由于欧盟的信息公开限制，建立因果联系所需的文件无法被全面获取，使用的一手材料在完整性和代表性上有所不足，导致事件梳理和因果链条构建上存在一些薄弱之处，对中小企业决策影响机制的验证也略显缺乏；其次，虽然研究指出了风险界定中的技术信息与价值观问题，但对通用人工智能技术本身的不确定性讨论较为有限。由于人工智能技术的发展路径与速度高度不确定，技术创新对监管措施的反馈可能会进一步改变现有的监管逻辑，研究在此类动态因素的捕捉上仍有改进空间。因此，未来的研究可以基于更丰富与全面的材料，对进一步细分下的行为体在决策中的具体影响路径与效果进行分析。此外，随着人工智能技术的继续发展，技术对监管的反馈效应将变得更加突出与可观察，未来的研究可以基于这些反馈，构建评估技术变迁对监管决策的长期影响框架。

总体而言，本研究致力于深度剖析欧盟人工智能法案的通用人工智能监管决策进程，旨在精准揭示法案决策过程中关键妥协转向的内在逻辑；同时从学术层面为理解新兴技术治理领域内多元利益主体的互动模式及决策机制提供了全新研究案例，以期为相关领域的理论拓展与实践探索提供有益借鉴。作为全球首部全面监管人工智能的法律，人工智能法案不仅是欧盟在新兴技术治理中的里程碑，

也在全球范围内发布了一个高影响的政策框架，其立法路径与成果将对其他地区的技术监管实践将产生重要影响。理解欧盟的立法过程，有助于理解欧盟内部行为体的风险界定与监管偏好在其内部的表现与分布，以及这些复杂因素通过欧盟制度进行加总与表达的逻辑。这一过程能够帮助中国更全面地把握欧盟在中美竞争格局中的角色，特别是其在人工智能时代对风险的独特理解。这不仅有助于厘清中欧在人工智能领域可能的合作与分歧，还能为中国在全球技术治理中的战略选择提供重要参考。随着人工智能技术的持续演进，相关的监管立法和政策动向将继续成为国际政治经济学研究的重要课题。

参考文献

一、中文文献

(一) 专著

- [1] 王正毅：《国际政治经济学通论》，北京：北京大学出版社 2010 年版。
- [2] [美] 海伦·米尔纳：《利益、制度与信息：国内政治与国际关系》，曲博译，王正毅校，上海：上海人民出版社 2010 年版。
- [3] [德] 贝娅特·科勒-科赫等：《欧洲一体化与欧盟治理》，顾俊礼等译、刘立群校，北京：中国社会科学出版社 2004 年版。

(二) 论文

- [1] 崔宏伟：《“数字技术政治化”与中欧关系未来发展》，《国际关系研究》2020 年第 5 期。
- [2] 宫云牧：《数字时代主权概念的回归与欧盟数字治理》，《欧洲研究》2022 年第 3 期。
- [3] 庞珣：《反思与重构：全球安全和全球治理的风险政治理论及方法》，《国际政治研究》2024 年第 2 期。
- [4] 戚凯、崔莹佳、田燕飞：《美欧英人工智能竞逐及其前景》，《现代国际关系》2024 年第 5 期。
- [5] 唐士其、庞珣：《综合安全论：风险的反向界定和政治逻辑》，《国际政治研究》2022 年第 6 期。
- [6] 王彦雨、李正风、高芳：《欧美人工智能治理模式比较研究》，《科学学研究》2024 年第 3 期。
- [7] 吴桐、刘宏松：《地缘经济转向、数字主权与欧盟人工智能治理》，《国际安全研究》2024 年第 5 期。
- [8] 解楠楠、张晓通：《“地缘政治欧洲”：欧盟力量的地缘政治转向？》，《欧洲研究》2020 年第 2 期。
- [9] 殷佳章、房乐宪：《欧盟人工智能战略框架下的伦理准则及其国际含义》，《国际论坛》2020 年第 2 期。
- [10] 朱仁显、唐哲文：《欧盟决策机制与欧洲一体化》，《厦门大学学报(哲学社会科学版)》2002 年第 6 期。

二、英文文献

(一) 专著

- [1] Blatter, Joachim, Haverland, Markus. *Designing Case Studies*. Palgrave Macmillan UK, 2012.
- [2] Braford, Anu, *The Brussels Effect*. Oxford University Press, 2020.
- [3] Coen, David, et al. *Business Lobbying in the European Union*. Oxford University Press, 2021.
- [4] Fuchs, Doris A., *Business Power in Global Governance*, Lynne Rinner Publishers, 2007.
- [5] Zekos, Georgios I., *Political, Economic and Legal Effects of Artificial Intelligence: Governance, Digital Economy and Society*, Cham: Springer International Publishing, 2022.

(二) 论文

- [1] Ahmed, Nur, Wahed, Muntasir, and Thompson, Neil C., "The Growing Influence of Industry in AI Research," *Science*, Vol.379, No.6635, 2023.
- [2] Alonso, José M., and Andrews, Rhys, "Insider Lobbying and Government Contracts: The Moderating Role of Firm Size," *European Management Review*, Vol.19, No.3, 2022.
- [3] Anamaria Nastasa, Maer Matei Monica Mihaela et al., "Artificial Intelligence: Friend or Foe? Experts' Concerns on European AI Act", *Journal of Economic Computation and Economic Cybernetics Studies and Research*, Vol.57, No.3/2023, 2024
- [4] Bareis, Jascha, and Katzenbach, Christian, "Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics," *Science, Technology, & Human Values*, Vol.47, No.5, 2022.
- [5] Blatter, Joachim, and Blume, Till, "In Search of Co-variance, Causal Mechanisms or Congruence? Towards a Plural Understanding of Case Studies," *Swiss Political Science Review*, Vol.14, No.2, 2008.
- [6] Bommasani, Rishi et al., "On the Opportunities and Risks of Foundation Models."
- [7] Bouwen, Pieter, "Corporate Lobbying in the European Union: The Logic of Access," *Journal of European Public Policy*, Vol.9, No.3, 2002.
- [8] Braithwaite, John, Coglianese, Cary, and Levi-Faur, David, "Can Regulation and Governance Make a Difference?," *Regulation & Governance*, Vol.1, No.1, 2007.
- [9] Brandsma, Gijs Jan, "Transparency of EU Informal Trilogues through Public Feedback in the European Parliament: Promise Unfulfilled," *Journal of European Public Policy*, Vol.26, No.10, 2019
- [10] Brasher, Holly, and Lowery, David, "The Corporate Context of Lobbying Activity," *Business and Politics*, Vol.8, No.1, 2006.
- [11] Broeders, Dennis, Cristiano, Fabio, and Kaminska, Monica, "In Search of Digital Sovereignty and Strategic Autonomy: Normative Power Europe to the Test of Its Geopolitical Ambitions," *JCMS: Journal of Common Market Studies*, Vol.61, No.5, 2023
- [12] Butcher, James, and Beridze, Irakli, "What Is the State of Artificial Intelligence Governance Globally?," *The RUSI Journal*, Vol.164, Nos.5–6, 2019
- [13] Cancela-Outeda, Celso, "The EU's AI Act: A Framework for Collaborative Governance," *Internet of Things*, Vol.27, 2024.
- [14] Cantero Gamito, Marta, and Marsden, Christopher T., "Artificial Intelligence Co-Regulation? The Role of Standards in the EU AI Act," *International Journal of Law and Information Technology*, Vol.32, 2024.
- [15] Cioffi, John W., Kenney, Martin F., and Zysman, John, "Platform Power and Regulatory

- Politics: Polanyi for the Twenty-First Century,” *New Political Economy*, Vol.27, No.5, 2022.
- [16] Cross, James P., and Hermansson, Henrik, “Legislative Amendments and Informal Politics in the European Union: A Text Reuse Approach,” *European Union Politics*, Vol.18, No.4, 2017.
- [17] Culpepper, Pepper D., and Thelen, Kathleen, “Are We All Amazon Primed? Consumers and the Politics of Platform Power,” *Comparative Political Studies*, Vol.53, No.2, 2020.
- [18] De Bruycker, Iskander, “Pressure and Expertise: Explaining the Information Supply of Interest Groups in EU Legislative Lobbying,” *JCMS: Journal of Common Market Studies*, Vol.54, No.3, 2016.
- [19] De Bruycker, Iskander, and Beyers, Jan, “Lobbying Strategies and Success: Inside and Outside Lobbying in European Union Legislative Politics,” *European Political Science Review*, Vol.11, No.1, 2019.
- [20] De Figueiredo, John M., and Tiller, Emerson H., “The Structure and Conduct of Corporate Lobbying: How Firms Lobby the Federal Communications Commission,” *Journal of Economics & Management Strategy*, Vol.10, No.1, 2001.
- [21] Delreux, Tom, and Laloux, Thomas, “Concluding Early Agreements in the EU: A Double Principal-Agent Analysis of Trilogue Negotiations,” *JCMS: Journal of Common Market Studies*, Vol.56, No.2, 2018.
- [22] Drope, Jeffrey M., and Hansen, Wendy L., “Does Firm Size Matter? Analyzing Business Lobbying in the United States,” *Business and Politics*, Vol.8, No.2, 2006.
- [23] Dunn, Pietro, and Gregorio, Giovanni De, “The Ambiguous Risk-Based Approach of the Artificial Intelligence Act: Links and Discrepancies with Other Union Strategies”, *CEUR Workshop Proceedings*, Vol.3221.
- [24] Dyer-Witheyford, Nick, and Mularoni, Alessandra, "Framing Big Tech: News Media, Digital Capital and the Antitrust Movement", *The Political Economy of Communication*, Vol.9, No.12, 2021.
- [25] Eke, Damian, and Stahl, Bernd, “Ethics in the Governance of Data and Digital Technology: An Analysis of European Data Regulations and Policies,” *Digital Society*, Vol.3, No.1, 2024.
- [26] Falletti, Tullia G., and Lynch, Julia F., "Context and Causal Mechanisms in Political Analysis", *Comparative Political Studies*, Vol.42, No.9, 2009.
- [27] Fenwick, Mark, Vermeulen, Erik P. M., and Compagnucci, Marcelo Corrales, “Business and Regulatory Responses to Artificial Intelligence: Dynamic Regulation, Innovation Ecosystems and the Strategic Management of Disruptive Technology”, in Marcelo Corrales et.al, *Robotics, AI and the Future of Law*, Springer, 2018.
- [28] Gorwa, Robert, Lechowski, Grzegorz, and Schneiß, Daniel, “Platform Lobbying: Policy Influence Strategies and the EU’s Digital Services Act,” *Internet Policy Review*, Vol.13, No.2, 2024.
- [29] Gstrein, Oskar J., Haleem, Noman, and Zwitter, Andrej, “General-Purpose AI Regulation and the European Union AI Act,” *Internet Policy Review*, Vol.13, No.3, 2024.
- [30] Hasselbalch, Jacob A., “Innovation Assessment: Governing through Periods of Disruptive Technological Change,” *Journal of European Public Policy*, Vol.25, No.12, 2018..
- [31] Helberger, Natali, and Diakopoulos, Nicholas, “ChatGPT and the AI Act,” *Internet Policy Review*, Vol.12, No.1, 2023.
- [32] Holman, Otto, “Asymmetrical Regulation and Multidimensional Governance in the European Union1,” *Review of International Political Economy*, Vol.11, No.4, 2004.

- [33] Hong, Seung-Hun et al., "Making Regulation Flexible for the Governance of Disruptive Innovation: A Comparative Study of AVs Regulation in the United Kingdom and South Korea," *Journal of European Public Policy*, Vol.29, No.11, 2022.
- [34] Jurowetzki, Roman et al., "'The Privatization of AI Research(-Ers): Causes and Potential Consequences", published in Arxiv, 2021.
- [35] Justo-Hanani, Ronit, "The Politics of Artificial Intelligence Regulation and Governance Reform in the European Union," *Policy Sciences*, Vol.55, No.1, 2022.
- [36] Kardasheva, Raya, Raya Kardasheva, "Trilogues in the EU Legislature", Research Paper of Department of European and International Studies, King's College London, July 2012.
- [37] Khanal, Shaleen, Zhang, Hongzhou, and Taihagh, Araz, "Why and How Is the Power of Big Tech Increasing in the Policy Process? The Case of Generative AI," *Policy and Society*, 2024.
- [38] Kusche, Isabel, "Possible Harms of Artificial Intelligence and the EU AI Act: Fundamental Rights and Risk," *Journal of Risk Research*, 2024.
- [39] Lake, David A., "Open Economy Politics: A Critical Review," *The Review of International Organizations*, Vol.4, No.3, 2009.
- [40] Langenkamp, Max, and Yue, Daniel N., "How Open Source Machine Learning Software Shapes AI," in *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, Oxford United Kingdom: ACM, 2022.
- [41] Levine, David K., and Modica, Salvatore, "Size, Fungibility, and the Strength of Lobbying Organizations," *European Journal of Political Economy*, Vol.49, 2017
- [42] Liesenfeld, Andreas, and Dingemans, Mark, "Rethinking Open Source Generative AI: Open Washing and the EU AI Act," in *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, Rio de Janeiro Brazil: ACM, 2024.
- [43] Lindman, Juho, Makinen, Jukka, and Kasanen, Eero, "Big Tech's Power, Political Corporate Social Responsibility and Regulation," *Journal of Information Technology*, Vol.38, No.2, 2023.
- [44] Li, Shu and Schütte, Béatrice and Sankari, Suvi: "The ongoing AI-regulation debate in the EU and its influence on the emerging economies: a new case for the 'Brussels Effect'?", in Mark Findlay, Li Min Ong, Wenxi Zhang, *Elgar Companion to Regulating AI and Big Data in Emerging Economies*, Edward Elgar Publishing, 2023
- [45] McLean, Scott et al., "The Risks Associated with Artificial General Intelligence: A Systematic Review," *Journal of Experimental & Theoretical Artificial Intelligence*, Vol.35, No.5, 2023.
- [46] Morgan, Kimberly J., "Process Tracing and the Causal Identification Revolution," *New Political Economy*, Vol.21, No.5, 2016.
- [47] Mügge, Daniel, "EU AI Sovereignty: For Whom, to What End, and to Whose Benefit?," *Journal of European Public Policy*, Vol.31, No.8, 2024.
- [48] Musch, Sean, Borrelli, Michael, and Kerrigan, Charles, "The EU AI Act As Global Artificial Intelligence Regulation," *SSRN Electronic Journal*, 2023.
- [49] Pagallo, Ugo, "Why the AI Act Won't Trigger a Brussels Effect", in *AI Approaches to the Complexity of Legal Systems*, Springer, forthcoming, 2024.
- [50] Palmiotto, Francesca, "The AI Act Roller Coaster: The Evolution of Fundamental Rights Protection in the Legislative Process and the Future of the Regulation," *European Journal of*

Risk Regulation, 2025.

- [51] Paul, Regine, “European ARTIFICIAL INTELLIGENCE ‘Trusted throughout the World’: Risk-based Regulation and the Fashioning of a Competitive Common AI Market,” *Regulation & Governance*, Vol.18, No.4, 2024.
- [52] Radu, Roxana, “Steering the Governance of Artificial Intelligence: National Strategies in Perspective,” *Policy and Society*, Vol.40, No.2, 2021.
- [53] Ricciardi Celsi, Lorenzo, “The Dilemma of Rapid AI Advancements: Striking a Balance between Innovation and Regulation by Pursuing Risk-Aware Value Creation,” *Information*, Vol.14, No.12, 2023.
- [54] Rikap, Cecilia, “Varieties of Corporate Innovation Systems and Their Interplay with Global and National Systems: Amazon, Facebook, Google and Microsoft’s Strategies to Produce and Appropriate Artificial Intelligence,” *Review of International Political Economy*, Vol.31, No.6, 2024.
- [55] Roederer-Rynning, Christilla, and Greenwood, Justin, “The Culture of Trilogues,” *Journal of European Public Policy*, Vol.22, No.8, 2015.
- [56] Rone, Julia, “‘The Sovereign Cloud’ in Europe: Diverging Nation State Preferences and Disputed Institutional Competences in the Context of Limited Technological Capabilities,” *Journal of European Public Policy*, Vol.31, No.8, 2024.
- [57] Sharon, Tamar, and Gellert, Raphaël, “Regulating Big Tech Expansionism? Sphere Transgressions and the Limits of Europe’s Digital Regulatory Strategy,” *Information, Communication & Society*, 2023.
- [58] Smuha, Nathalie A., “From a ‘Race to AI’ to a ‘Race to AI Regulation’: Regulatory Competition for Artificial Intelligence,” *Law, Innovation and Technology*, Vol.13, No.1, 2021.
- [59] Taeihagh, Araz, “Governance of Artificial Intelligence,” *Policy and Society*, Vol.40, No.2, 2021.
- [60] Taeihagh, Araz, Ramesh, M., and Howlett, Michael, “Assessing the Regulatory Challenges of Emerging Disruptive Technologies,” *Regulation & Governance*, Vol.15, No.4, 2021.
- [61] Tallberg, Jonas, Lundgren, Magnus, and Geith, Johannes, “AI Regulation in the European Union: Examining Non-State Actor Preferences,” *Business and Politics*, Vol.26, No.2, 2024.
- [62] Tarrant, Andy, and Cowen, Tim, “Big Tech Lobbying in the EU,” *The Political Quarterly*, Vol.93, No.2, 2022.
- [63] Tolan, Songül et al., “Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia,” in *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*, Montreal QC Canada: ACM, 2019.
- [64] Verdegem, Pieter, “Dismantling AI Capitalism: The Commons as an Alternative to the Power Concentration of Big Tech,” *AI & SOCIETY*, Vol.39, No.2, 2024.
- [65] Wang, Xukang, and Wu, Ying Cheng, “Balancing Innovation and Regulation in the Age of Generative Artificial Intelligence,” *Journal of Information Policy*, Vol.14, 2024.
- [66] Widder, David Gray, Whittaker, Meredith, and West, Sarah Myers, “Why ‘Open’ AI Systems Are Actually Closed, and Why This Matters,” *Nature*, Vol.635, No.8040, 2024.
- [67] Yampolskiy, Roman V., “Unpredictability of AI: On the Impossibility of Accurately Predicting All Actions of a Smarter Agent,” *Journal of Artificial Intelligence and*

(三) 欧盟官方文件

- [1] Council of the EU: “Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Presidency compromise text”, 2 November, 2021.
- [2] Council of the EU: “Proposition de Règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union - Text de compromis de la présidence - Version consolidée”, June 15, 2022.
- [3] Council of the EU: “Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Presidency compromise text”, November 25, 2021.
- [4] Council of the European Union: “Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Analysis of the final compromise text with a view to agreement”, January 26, 2024.
- [5] European Commission: “European Approach to artificial intelligence: AI Act”, <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>, 2025-05-12.
- [6] European Commission: “White Paper on Artificial Intelligence - A European approach to excellence and trust”, February 19, 2020.
- [7] European Commission: “Communication from the Commission to The European Parliament, the European Council, the Council, the European Economic and Social Committee and The Committee of The Regions: Artificial Intelligence for Europe”, Apr 2018.
- [8] European Commission: “Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts”, April 21, 2021.
- [9] European Commission: “Inception Impact Assessment”, July 23, 2020.
- [10] European Commission: “Commission Staff Working Document: Executive Summary of The Impact Assessment Report”, April 21, 2021.
- [11] European Commission: “Impact Assessment Accompanying the Proposal for a Regulation of The European Parliament and of The Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts”, April 2021.
- [12] European Parliament: “Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts”, June 14, 2023.
- [13] European Parliament: “Provisional Agreement Resulting from Interinstitutional Negotiations”, February 2, 2024.
- [14] European Parliament: “Artificial Intelligence Act: Amendments Adopted by the European Parliament on 14 June 2023 on the Proposal for a Regulation of the European Parliament and of the Council on Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts”, June 14, 2023.
- [15] European Parliament: “Results of roll - call votes: Artificial Intelligence Act, by Political

Group, all countries”, March 13, 2024.

- [16] European Commission, "Inception Impact Assessment", July 23, 2020, Ref. Ares(2020)3896535.
- [17] European Commission: "Briefing - Meeting With (Mistral AI) on the AI Act.", June 7, 2024, Ref. Ares(2024)3108481.

(四) 企业立场文件、会议纪要

- [1] AI Sweden, "position paper - AI Act", September 16, 2022, p.3-4.
- [2] "Briefing - Meeting with ... on the AI Act", July 13, 2023, Ref. Ares(2024)4108481.
- [3] BSA: "Joint Industry Statement on the EU Artificial Intelligence Act and High-Risk Obligations for General Purpose AI", September 27, 2022
- [4] BusinessEurope, "Statement on Potential AI Act General Approach", November 30, 2022
- [5] "BTO Meeting with Google on the AI Act", June 28, 2022, Ref. Ares(2024)6547475
- [6] Google, "Considerations regarding a tiered approach for foundation models and general purpose AI", October 2023
- [7] Google Feedback on general purpose AI systems", July 17, 2022, Ref. Ares(2022)5217601
- [8] Google: "Recommendations of Regulating AI", May 2023.
- [9] Hub France IA, "Position Paper AI Act", March 2023.
- [10] KI BUNDESVERBAND (German AI Association), "Position Paper on the EU AI Act Remaining Issues and Current Discussions in the European Parliament"
- [11] "European Meeting with Google on the AI Act", June 28, 2022, Ref. Ares(2024)6547475.
- [12] Mckinsey Global Institute: Notes from the AI Frontier, Tackling Europe's Gap in Digital and AI, February 2019
- [13] "OpenAI White Paper on the European Union's Artificial Intelligence Act", May 2, 2023, Ref. Ares(2023)3076642.
- [14] "Open letter to the representatives of the European Commission, the European Council and the European Parliament: Artificial intelligence: Europe's chance to rejoin the technological avantgarde", signed by German AI Association CEO, Siemens Energy CEO, etc.
- [15] Polish Confederation Lewiatan, "GPAI Considerations for EU AI Act High-Risk Requirements", March 2023
- [16] "Meeting with ... Google's ...", March 6, 2023, Ref. Ares(2025)15904 - 03/01/2025,

三、网络资源

- [1] AccessNow: "The EU AI Act: a failure for human rights, a victory for industry and law enforcement", March 2024, <https://www.accessnow.org/press-release/ai-act-failure-for-human-rights-victory-for-industry-and-law-enforcement/>, 2025-05-13.
- [2] Adach, Ola, "Aleph Alpha's OS LLM might be the answer to US AI dominance: German OS AI startup raises €486 million in Series B financing", November 8, 2023, <https://interoperable-europe.ec.europa.eu/collection/open-source-observatory-osor/news/aleph-alpha-os-llm-might-be-answer-us-ai-dominance>, 2025-05-13.
- [3] Aleph Alpha: "Luminous: European AI closes gap to world leaders", Feb 2023, <https://aleph-alpha.com/luminous-european-ai-closes-gap-to-world-leaders/>, 2025-05-13.
- [4] Aswin Prabhakar, "The EU AI Act Is a Cautionary Tale in Open-Source AI Regulation", November 20, 2023

- <https://datainnovation.org/2023/11/the-eu-ai-act-is-a-cautionary-tale-in-open-source-ai-regulation/>, 2025-05-13.
- [5] Arthur Mensch, “Mistral AI’s position on the AI Act” , November 16, 2023, <https://x.com/arthurmensch/status/1725076260827566562>, 2025-05-13.
- [6] Bradshaw, Tim, et.al, “‘We compete with everybody’: French AI start-up Mistral takes on Silicon Valley”, November 7, 2023, <https://www.ft.com/content/387eeeab-1f95-4e3b-9217-6f69aeeb5399>, 2025-05-13
- [7] Clara Hainsdorf, Tim Hickman, Dr. Sylvia Lorenz, et.al, “Dawn of the EU’s AI Act: political agreement reached on world’s first comprehensive horizontal AI regulation”, December 14, 2023, <https://www.whitecase.com/insight-alert/dawn-eu>, 2025-05-13.
- [8] Corporate Europe Observatory: “The lobby network: Big Tech’s web of influence in the EU”, Aug 2021, <https://corporateeurope.org/en/2021/08/lobby-network-big-techs-web-influence-eu>, 2025-05-13.
- [9] Corporate Europe Observatory: “The lobbying ghost in the machine: Big Tech’s covert defanging of Europe’s AI Act”, Feb 2023, <https://corporateeurope.org/en/2023/02/lobbying-ghost-machine>, 2025-05-13.
- [10] Corporate Europe Observatory, “Byte by byte: How Big Tech undermined the AI Act”, November 2023, <https://corporateeurope.org/en/2023/11/byte-byte>, 2025-05-13.
- [11] Corporate Europe Observatory, “Trojan horses: how European startups teamed up with Big Tech to gut the AI Act”, March 11, 2024, <https://corporateeurope.org/en/2024/03/trojan-horses-how-european-startups-teamed-big-tech-gut-ai-act>, 2025-05-13.
- [12] European Center for Not-for-Profit Law (ECNL): “Packed with loopholes: Why the AI Act fails to protect civic space and the rule of law”, April 2024, <https://edri.org/our-work/packed-with-loopholes-why-the-ai-act-fails-to-protect-civic-space-and-the-rule-of-law/>, 2025-05-13.
- [13] European Commission, “Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence”, April 21, 2021, https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682, 2025-05-13.
- [14] European Parliament, “MEPs ready to negotiate first-ever rules for safe and transparent AI”, June 14, 2023, <https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai>, 2025-05-13.
- [15] Felix Duffy: “AI Act: Von der KI-Industrie in die Zange genommen” , LobbyControl, Mars 2024, <https://www.lobbycontrol.de/macht-der-digitalkonzerne/ai-act-von-der-ki-industrie-in-die-zange-genommen-114508/>, 2025-05-13.
- [16] Foo Yun Chee, “Exclusive: AI rules 'cannot be bargained', EU's Breton says after OpenAI CEO threat”, May 25, 2023, <https://www.reuters.com/technology/eus-breton-slams-openai-ceos-comments-blocs-draft-ai-rules-2023-05-25/>, 2025-05-13.
- [17] FragDenStaat, “Stellungnahmen von Aleph Alpha an das BMWK”, December 14, 2023, <https://fragdenstaat.de/anfrage/stellungnahmen-von-aleph-alpha-an-das-bmwk/>, 2025-05-13.

- [18] France 24: “French start-up Mistral AI emerges as a leader in European artificial intelligence”, December 11, 2023, <https://www.france24.com/en/europe/20231211-french-start-up-mistral-ai-emerges-as-leading-force-in-european-artificial-intelligence>, 2025-05-13.
- [19] Killian Vieth-Ditlmann, Nikolett Aszodi: “EU Parliament votes on AI Act; member states will have to plug surveillance loopholes”, Algorithm Watch, March 2024. <https://algorithmwatch.org/en/eu-parliament-votes-on-ai-act/>, 2025-05-13.
- [20] Kyle Wiggers, Devin Coldewey: “This week in AI: Mistral and the EU’s fight for AI sovereignty”, December 9, 2023, <https://techcrunch.com/2023/12/09/this-week-in-ai-mistral-and-the-eus-fight-for-ai-sovereignty/>, 2025-05-13.
- [21] LobbyFacts: “Mistral AI”, <https://www.lobbyfacts.eu/datacard/mistral-ai?rid=539437350687-42#data-card-data-meetings>, 2025-05-13.
- [22] Luca Bertuzzi, “AI Act: EU Commission attempts to revive tiered approach shifting to General Purpose AI”, November 20, 2023, <https://www.euractiv.com/section/tech/news/ai-act-eu-commission-attempts-to-revive-tiered-approach-shifting-to-general-purpose-ai/>, 2025-05-13.
- [23] Luca Bertuzzi, “AI Act: Leading MEPs revise high-risk classification, ignoring negative legal opinion”, October 23, 2023, <https://www.euractiv.com/section/artificial-intelligence/news/ai-act-leading-meps-revise-high-risk-classification-ignoring-negative-legal-opinion/>, 2025-05-13.
- [24] Luca Bertuzzi, “AI Act: MEPs close in on rules for general purpose AI, foundation models”, April 20, 2023, <https://www.euractiv.com/section/tech/news/ai-act-meps-close-in-on-rules-for-general-purpose-ai-foundation-models/>, 2025-05-13.
- [25] Luca Bertuzzi, “EU’s AI Act negotiations hit the brakes over foundation models”, November 10, 2023, <https://www.euractiv.com/section/tech/news/eus-ai-act-negotiations-hit-the-brakes-over-foundation-models/>, 2025-05-13.
- [26] Luca Bertuzzi, “Leading EU lawmakers propose obligations for General Purpose AI”, March 14, 2023, <https://www.euractiv.com/section/artificial-intelligence/news/leading-eu-lawmakers-propose-obligations-for-general-purpose-ai/>
- [27] Luca Bertuzzi, “Leading MEPs exclude general-purpose AI from high-risk categories – for now”, December 12, 2022, <https://www.euractiv.com/section/tech/news/leading-meps-exclude-general-purpose-ai-from-high-risk-categories-for-now/>, 2025-05-13.
- [28] Melissa Heikkilä: “Dutch scandal serves as a warning for Europe over risks of using algorithms”, March 29, 2022, <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>, 2025-05-12.
- [29] Ministry of Enterprises and Made in Italy, “Italy, Germany and France agree on strengthening their cooperation on Artificial Intelligence”, November 2023,

<https://www.mimit.gov.it/en/media-tools/news/italy-germany-and-france-agree-on-strengthening-their-cooperation-on-artificial-intelligence#:~:text=Such%20a%20strengthened%20cooperation%20in,start-ups%20and%20SMEs%20through,> 2025-05-13.

- [30] Microsoft: “Microsoft’s AI Access Principles: Our commitments to promote innovation and competition in the new AI economy”, February 2024, <https://blogs.microsoft.com/on-the-issues/2024/02/26/microsoft-ai-access-principles-responsible-mobile-world-congress/>, 2025-05-13.
- [31] Morgan Heaker, “Meet Aleph Alpha, Europe’s Answer to OpenAI”, Wired, August 30, 2023, <https://www.wired.com/story/aleph-alpha-europe-openai/>, 2025-05-13.
- [32] Natasha Lomas: “Europe’s AI Act talks head for crunch point”, November 14, 2023, <https://techcrunch.com/2023/11/14/eu-ai-act-trilogue-crunch/>, 2025-05-13.
- [33] Natasha Lomas: “France’s Mistral dials up call for EU AI Act to fix rules for apps, not model makers”, November 16, 2023. <https://techcrunch.com/2023/11/16/mistral-eu-ai-act/>, 2025-05-13.
- [34] No Priors: AI, Machine Learning, Tech, & Startups, “No Priors Ep. 40 | With Arthur Mensch, CEO Mistral AI”, November 9, 2023, <https://www.youtube.com/watch?v=EMOFRDOMiIU>, 2025-05-13.
- [35] Octave Bonnaud, Elsa Trujillo, “Google prospecte les parlementaires pour un déjeuner autour de l’IA”, La Lettre, June 2023, https://www.lalettre.fr/fr/action-publique_lobbying/2023/06/13/google-prospecte-les-parlementaires-pour-un-dejeuner-autour-de-l-ia,109993537-art, 2025-05-13.
- [36] Olivier Petitjean: “AI Act: le troublant lobbying des « champions » européens, Mistral AI et Aleph Alpha”, Observatoire des multinationales, Mars 2024, <https://multinationales.org/fr/enquetes/intelligence-artificielle-lobbying-et-conflits-d-interets/mistralai-alephalpha-gafam-ai-europe>, 2025-05-13.
- [37] OpenAI: “Frontier AI regulation: Managing emerging risks to public safety”, July 2023, <https://openai.com/index/frontier-ai-regulation/>, 2025-05-13.
- [38] Polish Confederation Lewiatan, “GPAI Considerations for EU AI Act High-Risk Requirements”, March 2023, <https://lewiatan.org/wp-content/uploads/2023/03/GPAI-Considerations-for-EU-AI-Act-High-Risk-Requirements.pdf>, 2025-05-13.
- [39] Shiona McCallum, “ChatGPT banned in Italy over privacy concerns”, April 1, 2023, <https://www.bbc.co.uk/news/technology-65139406>, 2025-05-13.
- [40] Statista, “Artificial Intelligence - Worldwide”, <https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide#market-size>, 2025-05-13.
- [41] Tanislas Polu:” And the state is not kidding about that...”, June 15, 2023, <https://x.com/spolu/status/1669303713440079873>, 2025-05-13.
- [42] Politico: “Panel discussion - The AI package, one year on: from theory to practice | POLITICO AI & Tech Summit”, April 21, 2022, https://www.youtube.com/watch?v=uvaUQg_q2Ho, 2025-05-13.
- [43] Supantha Mukherjee and Martin Coulter: “ChatGPT-maker OpenAI says has no plans to leave Europe”, May 27, 2023, <https://www.reuters.com/technology/openai-has-no-plans-leave-europe-ceo-2023-05-26/>,

2025-05-13.

[44] Julia Tar, “OpenAI CEO’s backtracks on quitting EU threat”, May 26, 2023, <https://www.euractiv.com/section/artificial-intelligence/news/openai-ceos-backtracks-on-quitting-eu-threat/>, 2025-05-13.

[45] Zosia Wanat: “‘EU’s AI act could kill our company,’ says Mistral’s Cédric O”, October 27, 2023, <https://sifted.eu/articles/eu-ai-act-kill-mistral-cedric-o>, 2025-05-13.

仅供学习交流

附录

2023 年欧洲议会人工智能法案游说会议数量（清洗后， ≥ 2 场）

Attendees	Meeting Count
Microsoft Corporation	11
Google	10

American Chamber of Commerce to the European Union	10
OpenAI	7
Universal Music Group International	5
Meta Platforms Ireland Limited and its various subsidiaries	5
The Future Society	4
ITI _ The Information Technology Industry Council	4
Amazon	3
Avaaz Foundation	3
BSA \ The Software Alliance	3
Der Mittelstand, BVMW e.V.	3
IFPI Representing recording industry worldwide	3
Initiative Urheberrecht	3
Mistral AI	3
Adobe	2
Apple Inc.	2
AT&T, Inc.	2
Bundesdruckerei GmbH	2
Bundesverband Musikindustrie e.V.	2
Cohere Inc.	2
Creative Commons	2
European AI forum	2
European Disability Forum	2
European Internet Forum	2
EUROPEAN TRADE UNION CONFEDERATION	2
Friends of Europe	2
General Catalyst	2
Handelsverband Deutschland	2
OECD	2
Palo Alto Networks Inc.	2
Snap, Inc.	2
Spotify	2

Stability AI Ltd	2
The German Marshall Fund of the United States - The Transatlantic Foundation	2
Twilio Inc.	2
Vereniging VNO-NCW	2
Wirtschaftsrat der CDU e.V.	2

Source: European Parliament: Search MEP meetings,

<https://www.europarl.europa.eu/meps/en/searchmeetings?textualSearch=&procedureReferences=2021%2F0106%28COD%29&fromDate=&toDate=, 2025-05-12.>

仅供学习交流

致谢

想象了很多次自己写致谢的场景，在论文基本完成的当下，自己也突然有了强烈的即将毕业的实感。回忆起自己的本科生活，总觉得自己在四处奔跑，从北京到深圳再到牛津，从一个公司到另一个公司，从宿舍到教学楼再到咖啡厅，一篇论文从零到数万字，几个月后又要从颐和园路走向双清路，虽然每段路程都难

免辛苦，但这些奔跑里蕴含着意义，一路上的所有片段与画面构筑了我的生活，以及我对生活意义的认知。

这篇论文今近完成，首先要感谢我的指导老师董昭华老师，在《全球治理的政治经济分析》课上，我第一次系统地学习 IPE 研究中的利益攸关者分析，从希斯考克斯到卡赞斯坦，前人的 IPE 理论在课程学习中闪闪发光，让我领略到 IPE 研究的精妙之处。同时，还十分感谢董老师愿意接下我的毕业论文指导，从选题、研究问题与对象的确定、结构的添补到引言、结论的更新，董老师一如既往地鼓励我，在百忙之中仍然优先处理学生论文，撰写修改意见，与我一同探索研究内容，促使我在研究上不断精进、不断成长，让这篇毕业论文的撰写成为了一段充满意义的经历，董老师对学术的认真与热爱让我真切地认识到学术作为一种志业的分量。一直记得第一次和董老师讨论毕业论文时，董老师让我思考自己学习四年之后，到底想要为国关这个领域做出什么 contribution，回看这篇论文，觉得至少做出了一些不愧对自己本科学习、不白费自己撰写过程中努力的结果。

感谢向我传道授业的老师们。感谢王联老师、雷少华老师、节大磊老师、钱雪梅老师、李扬帆老师、王正毅老师、项佐涛老师，在国际政治的学习过程中，老师们的课程使我逐渐搭建起自己的知识体系，其中涉猎的不同政治研究方向也不断刺激着我的思考，让我真切领略到国际关系研究的魅力，在课堂之外，老师们的为人处事与治学态度同样对我启发良多。感谢 Young Joon Park 老师、史蛟老师、David Lander 老师和 Jonathan Pearson 老师，在汇丰的日子里得到了诸位老师在学业上的帮助与照顾，不仅大大帮助了我对国际金融知识的顺利探索，让我从一个商科小白不断成长到今天，也让我在南国燕园留下了温暖的回忆。

感谢一路陪伴的朋友们。感谢 IPE 幼儿园的范钊焯、高丽星、余嘉欣，在无数失意与迷茫的时刻，你们是最坚强的后盾；在我们诸多命运的十字路口，相互扶持让我们最后坚强地成长为今天的我们，我对此无比感激。感谢 2025 迎来好人生的黄淇鑫、朱霖儿，我们在生活中彼此感知、彼此支持，又在各自的磕磕绊绊里走向我们期待的未来，在共同经历的无数黑夜里，我们总是能够找到彼此的光芒。感谢北上港成功人士的邬锦鑫、彭开欣，有时候会惊奇地发现我们已经认识十年之久，赣州、深圳、上海、北京、香港、北海道，我们成长的印记落在世界的每个角落。至今还记得高中和邬锦鑫用支付宝聊天的日子，怎么一眨眼我们都毕业了，不过我们好像都还成为了不错的大人，为我们感到高兴！感谢肖睿，在生活与智识上都给了我十分大的帮助，谢谢音乐、羽毛球和网球。感谢生煎包

的陈思凝和余静怡，我们在上海的雨中手忙脚乱地交换相机，在石舫上看新一天的太阳升起，外滩的夜晚将成为我最珍贵的回忆之一。感谢比特币的朋友们，我们一起创造了许多想到还会笑出声的难忘事迹。

感谢在汇丰遇见的好朋友们，感谢 SISDF 的（排名不分先后）陈相宇、许大为、林欣韵、陈欣仪、潘仕刚、胡佳禾、洪劲恒、孔智裕、尹颂怡、成超涵、胡齐齐、方睿，在汇丰度过的一年因为有你们变得无比快乐，我们在 Cambridge River 泛舟，一起去 City Center 喝奶茶，一起熬夜做头疼的小组作业，一起通宵聊天，学困难的博弈论和 M&A，无论身在在世界何处，都祝愿大家一切顺利。感谢博雅的朋友们，感谢李祥宇常常成为我的海底捞搭子，分享我们的生活。感谢国辩，感谢国辩的学长学姐和所有朋友们，让我在偌大的学校里仍然感觉有一个归属的港湾。感谢在实习和生活中遇到的朋友们，我们同行的路，我不会忘记。

感谢永远支持我的家人们。谢谢爸爸妈妈对我无条件的支持，总是让我不要担心各种事情又总是在担心我，给我物质和精神上的鼓励和支持，让我勇敢地去看看外面更大的世界，督促我养成各种好习惯（虽然结果存疑），容忍我的各种懒惰和脾气，感谢你们给予我的一切。谢谢总是在等我回家的爷爷奶奶，总是惭愧常常没有时间陪老人家，但好在老人们总也不嫌我空不出时间和他们聊天，希望长大的我没有让你们失望。谢谢姑姑姑父对我的照顾，知道我考上北大之后姑姑比我还要激动，特别感谢姑姑从小到大对我的关心。谢谢外公外婆、大姨、舅舅、舅妈、小姨，总是很照顾我们几个小辈；谢谢大姨总是关照我，谢谢舅舅舅妈在我每次到访的时候都不嫌打扰。谢谢总是支持我、帮助我的家人们，这一切都是我生命中最宝贵的财富。

谢谢欧盟的透明度部门，你们闪烁着民主精神的信息公开框架使这篇论文成为可能。谢谢音乐，在无数日子里与我为伴。谢谢殷祥昶带给我的一切帮助和成长，我会继续努力成为更强大的人，I promise I'll always have this faith in me.

谢谢国际关系学院和北京大学对我的教育和滋养，谢谢某些精神依旧在课堂上流淌，在学术之外教会我诚实、正直、严谨、勇敢等等珍贵的品质，让我在诸多重要的思辨中不断成长，这一切我将受用终身。

王子健

2025 年 5 月 19 日于燕园

仅供学习交流

北京大学学位论文原创性声明和使用授权说明

原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品或成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

论文查重网站： 中国知网 论文查重率： 0.4%

论文作者签名：

日期：2025 年 5 月 23 日

学位论文使用授权说明

本人完全了解北京大学关于收集、保存、使用学位论文的规定，即：

- 按照学校要求提交学位论文的印刷本和电子版本；
- 学校有权保留学位论文的印刷本和电子版，并提供目录检索与阅览服务，在校园网上提供服务；
- 学校可以采用影印、缩印、数字化或其它复制手段保存论文；

论文作者签名：

导师签名：

日期：2025 年 5 月 23 日