

# 欧盟人工智能治理方案 会产生“布鲁塞尔效应”吗？

## ——基于欧盟《人工智能法》的分析\*

苏可桢 沈 伟

**摘 要：**欧盟率先制定了全球第一部人工智能(AI)治理专门法律——《人工智能法》(简称《AI法》),该法以安全性、创新促进性、尊重基本权利、维护内部市场等为规制目标。作为欧盟AI治理的基本法,该法以风险规制为治理框架,对高风险AI系统提出诸多要求,设置了以监管沙盒为代表的监管工具。虽然欧盟AI的发展与投入稍逊美中,但继在数据治理、环境保护等领域实现“布鲁塞尔效应”后,欧盟试图在AI领域复制规则引领者的成功经验。当前,全球已经形成AI治理的原则性共识,AI治理的“最大公约数”轮廓逐渐清晰。中国的AI规制方案包括国际倡议、法律和部门规章,基于产业优势的监管竞争可能会阻碍AI领域的“布鲁塞尔效应”。中国的AI规制体系应契合全球共识,兼顾一般法与特别法,争取在全球AI治理竞争中的优势地位。

**关键词：**AI治理；《人工智能法》；欧盟；布鲁塞尔效应；风险防范

**作者简介：**上海交通大学 凯原法学院 博士研究生 上海 200030

上海交通大学 凯原法学院 教授 博士生导师 上海 200030

**中图分类号：**D814.1；D99

**文献标识码：**A

**文章编号：**1005-4871(2024)02-0066-23

---

\* 本文系国家社科基金重大项目“美国全球单边经济制裁中涉华制裁案例分析与对策研究”(项目编号:21&ZD208)的阶段性成果。

## 引 言

近年来,人工智能(Artificial Intelligence,以下简称 AI)以其快速发展和广泛应用成为全球科技领域的焦点。一系列值得关注的问题随着 AI 的广泛应用不断涌现,隐私保护、伦理准则、失业风险和安全威胁等问题已经出现在不同的场景之中,AI 技术与应用的复杂性和风险性逐渐成为全球共同关注的规制重点。习近平总书记在 2021 年 11 月主持召开的中央政治局会议上强调加快提升生物安全、网络安全、数据安全、AI 安全等领域的治理能力。<sup>①</sup>探索 AI 治理路径迫在眉睫,无论是国家、国际组织等公主体,还是高校、企业、社会组织等非公主体,都在密切关注 AI 治理问题。<sup>②</sup>AI 治理路径的探索面临诸多困难。一方面,AI 本质上被视为不同技术的集合,技术领域缺乏对 AI 的明确定义,AI 在不同行业和应用场景中可能会带来不同的风险特征,因此定义 AI 治理的适用范围、设置统一的监管路径难度较高。<sup>③</sup>另一方面,传统的技术治理经验与政策工具难以匹配应用于 AI 技术,监管者与开发者对于 AI 存在诸多信息不对称情况,导致各利益相关方难以对监管范式形成共识。<sup>④</sup>因此,不同主体从不同角度提出的 AI 治理规范存在着排他、笼统、矛盾、碎片化、脱离实践等缺陷。<sup>⑤</sup>

不同国家和地区在 AI 治理上的路径和措施各异,反映了各国技术实力、社会需求和价值取向的差异。2021 年 4 月,欧盟发布《人工智能法》(以下简称《AI 法》)提案,<sup>⑥</sup>此后,欧盟成员国又数次对法案文本进行修订,于 2022 年 12 月发布《AI 法》

<sup>①</sup> 《中共中央政治局召开会议 审议〈国家安全战略(二〇二一—二〇二五年)〉〈军队功勋荣誉表彰条例〉和〈国家科技咨询委员会二〇二一年咨询报告〉 中共中央总书记习近平主持会议》,2021 年 11 月 19 日, <http://cpc.people.com.cn/n1/2021/1119/c64036-32286407.html>, 访问日期:2023-12-12。

<sup>②</sup> 周慎、朱旭峰、梁正:《全球可持续发展视域下的人工智能国际治理》,载《中国科技论坛》,2022 年第 9 期,第 163-169 页,这里第 164 页。

<sup>③</sup> Peter Stone et al., “Artificial Intelligence and Life in 2030: The One Hundred Year Study on Artificial Intelligence”, Report of the 2015 Study Panel, September 2016, [https://ai100.stanford.edu/sites/g/files/sbiy-bj18871/files/media/file/ai100report10032016fml\\_singles.pdf](https://ai100.stanford.edu/sites/g/files/sbiy-bj18871/files/media/file/ai100report10032016fml_singles.pdf), 访问日期:2024-02-04; Peter Cihon/Matthijs M. Maas/Luke Kemp, “Should Artificial Intelligence Governance be Centralised? Design Lessons from History”, in Annette Markham et al. (eds.), *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, New York: Association for Computing Machinery, 2020, pp. 228-234, here p. 229.

<sup>④</sup> Urs Gasser/Virgilio A. F. Almeida, “A Layered Model for AI Governance”, *IEEE Internet Computing*, Vol. 21:6, 2017, pp. 58-62, here p. 58.

<sup>⑤</sup> Brent Mittelstadt, “Principles Alone Cannot Guarantee Ethical AI”, *Nature Machine Intelligence*, Vol. 1:11, 2019, pp. 501-507, here p. 503.

<sup>⑥</sup> European Commission, “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”, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>, 访问日期:2024-01-03。

折中修订草案。<sup>①</sup> 2023年6月,欧洲议会通过了《AI法》折中修订草案,<sup>②</sup>又于2023年12月与欧盟理事会就该法案达成临时协议。<sup>③</sup> 2024年3月,欧洲议会正式通过了《AI法》,该法预计将于2024年年中生效。<sup>④</sup> 欧盟基于风险预防的理念,为AI应用制定了覆盖全过程的风险规制体系,<sup>⑤</sup>试图实现“布鲁塞尔效应”(The Brussels Effect),通过颁布法规提高相关全球标准,使该商业领域“欧洲化”(Europeanization)。<sup>⑥</sup> 中国在AI法律规制方面采取的是更多基于特定应用场景的针对性立法,这一路径在多样性、灵活性方面具有一定的优势,但也面临缺少更为全面统一的顶层法律设计的挑战。

本文以欧盟《AI法》为切入点,第一部分分析世界首部全面监管AI之法律的监管原则、框架和规则。第二部分考察欧盟AI立法能否实现“布鲁塞尔效应”并引领全球AI监管。欧盟利用市场优势在多个领域推行规则之治、实现规则引领,这种“布鲁塞尔效应”的实现既需要形式上的多重内部因素集合作用,又取决于欧盟能否平衡域内AI立法与全球相关治理规范,契合全球AI治理的最大公约数,这是研判AI“布鲁塞尔效应”能否实现的两个重要维度。第三部分从比较视角出发,分析中国关于AI规制的相关法律及其特点,讨论不同规制方案的利弊和未来中国进行AI治理的路径设计。最后是结语。

## 一、欧盟《AI法》的规制路径:目标、框架与工具

《AI法》对AI系统(Artificial Intelligence System)的定义是:“采用附件I所列的一种或多种技术和方法开发的软件,可以针对人类确定的特定目标,生成内

<sup>①</sup> 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 — General Approach”, December 6, 2022, <https://data.consilium.europa.eu/doc/document/ST-15698-2022-INIT/EN/pdf>, 访问日期:2024-01-03。

<sup>②</sup> Yasmina Yakimova/Janne Ojamo, “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>, 访问日期:2024-01-03。

<sup>③</sup> Dimosthenis Mammonas, “Artificial Intelligence Act: Council and Parliament Strike a Deal on the First Rules for AI in the World”, December 9, 2023, <https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/>, 访问日期:2024-01-04。

<sup>④</sup> Yasmina Yakimova/Janne Ojamo, “Artificial Intelligence Act: MEPs Adopt Landmark Law”, March 13, 2024, <https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law>, 访问日期:2024-02-09。

<sup>⑤</sup> 曾雄、梁正、张辉:《欧盟人工智能的规制路径及其对我国的启示——以〈人工智能法案〉为分析对象》,载《电子政务》,2022年第9期,第63-72页,这里第65页。

<sup>⑥</sup> Anu Bradford, *The Brussels Effect: How the European Union Rules the World*, New York: Oxford University Press, 2020, pp. xiii-xix.

容、预测、建议或决定等输出结果，对与之互动的环境产生影响。”<sup>①</sup>欧盟高度重视包括 AI 在内的数据和信息领域安全风险，试图推出 AI 领域的《通用数据保护条例》(General Data Protection Regulation, 以下简称 GDPR)，<sup>②</sup>对 AI 系统进行风险评估与定级，提出具有较强针对性的规制路径。

### (一) 规制目标

《AI 法》旨在实现一系列具体的规制目标。首先，该法将确保投放到欧盟市场的 AI 系统是安全的并尊重现有的欧盟法律。其次，《AI 法》的内容需具备规则确定性，发挥促进相关投资和创新的功能。再次，该法注重对公民基本权利的保护，减小 AI 系统对基本权利的影响、加强对相关欧盟法律的执行，这是该法的重点内容。最后，该法还将促进欧盟内部 AI 单一市场发展，防止市场分裂。<sup>③</sup>

#### 1. 安全性和可靠性

安全可靠是《AI 法》对 AI 系统最基本的要求。AI 的使用具有不透明性、复杂性、数据依赖性、行为自主性等特征，可能对 AI 系统的部署者、用户及使用对象的基本权利产生不利影响。为此，《AI 法》以风险为起点，区分 AI 系统的不同风险类别并予以法律干预。但该法也保证 AI 应用程序只会在必要情况下受到监管以解决特定级别风险。<sup>④</sup>此外，为确保 AI 系统在其生命周期内的安全可靠、减少可能威胁公共安全的风险，要求 AI 系统保障透明度成为该法的立法目的之一。<sup>⑤</sup>

#### 2. 规则确定性与创新促进性

提供法律上的确定性也是《AI 法》的立法目标之一。<sup>⑥</sup>有观点认为，《AI 法》可能会对欧洲 AI 投资产生寒蝉效应，阻碍中小企业开发高风险 AI 系统，使欧洲 AI

---

① Article 3 “Definitions”, the AI Act. “附件 I”中所列举的技术包括：1. 机器学习方法，包括监督学习、无监督学习和强化学习，使用包括深度学习在内的各种方法；2. 基于逻辑和知识的方法，包括知识表示、归纳（逻辑）编程、知识库 (knowledge bases)、推理和演绎引擎、(符号)推理和专家系统 (expert systems)；3. 统计方法、贝叶斯估计、搜索和优化方法。

② Ursula von der Leyen, *Political Guidelines for the Next European Commission 2019 - 2024*, Luxembourg: Publications Office of the European Union, 2020, p. 39.

③ Part 1 “Context of the Proposal”, Explanatory Memorandum, European Commission Proposal for the AI Act.

④ European Parliamentary Research Service, “EU Legislation in Progress: Artificial Intelligence Act”, June 2023, [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS\\_BRI\(2021\)698792\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf), 访问日期:2023-12-10。

⑤ Article 1 “Subject matter”, the AI Act.

⑥ Jonas Schuett, “Risk Management in the Artificial Intelligence Act”, *European Journal of Risk Regulation*, February 8, 2023, <https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/risk-management-in-the-artificial-intelligence-act/2E4D5707E65EFB3251A76E288BA74068>, 访问日期:2023-12-11。

投资减少近 20%，并在五年内给欧洲造成 310 亿欧元的损失。<sup>①</sup>但也有经济学家对前述估算有所质疑，认为只要相关商业主体满足《AI 法》的部分要求，遵循常见最佳实践，企业合规成本就可以降低。<sup>②</sup>总之，《AI 法》针对系统开发者设置的规则需要具备确定性，并设置促进 AI 领域研发活动的制度。为了促进 AI 创新，《AI 法》中设置了对研究活动以及根据开源许可提供 AI 组件的豁免，提倡由公共机构建立监管沙盒 (Regulatory Sandbox) 或受控环境，以便在应用 AI 前对技术或系统进行测试。<sup>③</sup>

### 3. “负责任创新”与尊重基本权利

“负责任的研究与创新” (Responsible Research and Innovation, 以下简称 RRI) 是欧盟在科技发展中提倡的重要理念，也被贯彻于 AI 相关领域。<sup>④</sup> RRI 是“一个透明、互动的过程，通过这一过程，社会参与者和创新者相互响应，共同关注创新过程及其市场产品的道德可接受性、可持续性和社会可取性，以便将科技进步恰当地融入社会”。<sup>⑤</sup> 在《AI 法》的立法过程中，立法者认为对 AI 的使用以及 AI 系统本身的不透明性、复杂性、数据依赖性、行为自主性等特征，可能会对《欧盟基本权利宪章》(The EU Charter of Fundamental Rights) 中规定的许多基本权利产生不利影响。例如，用于移民、庇护和边境管制的 AI 系统，就可能会对使用对象的自由行动权、不受歧视权、保护私人生活和个人数据权等权利产生威胁。<sup>⑥</sup> 《AI 法》

<sup>①</sup> Benjamin Mueller, “How Much Will the Artificial Intelligence Act Cost Europe?”, July 2021, <https://www2.datainnovation.org/2021-ai-a-costs.pdf>, 访问日期:2023-12-08。

<sup>②</sup> Moritz Laurer/Andrea Renda/Timothy Yeung, “Clarifying the Costs for the EU’s AI Act”, September 24, 2021, [https://www.ceps.eu/clarifying-the-costs-for-the-eus-ai-act/?\\_mc\\_cid=1b1e61c5af&-mc\\_eid=9a740783cd](https://www.ceps.eu/clarifying-the-costs-for-the-eus-ai-act/?_mc_cid=1b1e61c5af&-mc_eid=9a740783cd), 访问日期:2023-12-08; Meeri Haataja/Joanna J. Bryson, “What Costs Should We Expect from the EU’s AI Act?”, August 20, 2020, <https://osf.io/preprints/socarxiv/8nzb4>, 访问日期:2023-12-08。

<sup>③</sup> Yasmina Yakimova/Janne Ojamo, “AI Act: A Step Closer to the First Rules on Artificial Intelligence”, May 11, 2023, <https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence>, 访问日期:2023-12-08。

<sup>④</sup> 例如欧洲议会研究服务处 (European Parliamentary Research Service) 于 2019 年发布的一份报告就要求在算法开发与应用的过程中发挥 RRI 的作用，通过算法促进公平。参见 European Parliamentary Research Service, “A Governance Framework for Algorithmic Accountability and Transparency”, April 2019, [https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS\\_STU\(2019\)624262\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS_STU(2019)624262_EN.pdf), 访问日期:2023-12-11。

<sup>⑤</sup> René von Schomberg, *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields*, Luxembourg: Publications Office of the European Union, 2011, p. 9.

<sup>⑥</sup> Para 39, Foreword of the AI Act.

将确保对这些基本权利的高水平保护，通过监管来解决威胁基本权利的风险来源。<sup>①</sup>为此，在《AI法》的折中修订草案中，立法者明确禁止六种 AI 应用场景。<sup>②</sup>但是，该法也对经营自由、艺术与科学自由等部分基本权利施加限制，以确保高风险 AI 技术的开发者和使用者可以进行“负责任创新”，维护公民健康、安全、消费者保护等方面的公共利益。<sup>③</sup>

#### 4. 对单一市场的维护

欧盟之所以能够扮演全球监管规则供给者的角色，其整体的市场规模是其中的重要因素，欧盟同时还有能力通过法律体制架构将市场规模优势转化为监管影响力。<sup>④</sup>《AI法》的总体目标是，通过为在欧盟开发和可信 AI 系统创造条件，确保单一市场正常运作。<sup>⑤</sup>《AI法》是根据《欧盟运作条约》(Treaty on the Functioning of the European Union)第 114 条制定的，后者强调以欧盟成员国国内法律间的相似性促进欧盟统一内部市场的运作。<sup>⑥</sup>因此，维护欧盟市场规模、防止市场分割，促进合法、安全和值得信赖的 AI 应用单一市场发展，<sup>⑦</sup>既是《AI法》的规制目的，又是该法产生“布鲁塞尔效应”的基础。

#### (二) 欧盟 AI 规制框架

首先，《AI法》在欧盟 AI 治理的整体框架中居于基本法地位。“防止成员国采取可能导致市场碎片化的单边行为(unilateral actions)，以及防止对开发或使用 AI 系统的运营商施加更多的监管负担”是《AI法》的重要目标。<sup>⑧</sup>实现这一目标的方式是通过《AI法》对相关规制框架进行“最大程度的协调”(maximum harmonisation)，削弱成员国在该领域采取行动的能力，迫使其取消与该法冲突的国内法。因此，该法采取统一通用的立法模式，使其成为欧盟范围内适用于所有 AI 系统的

<sup>①</sup> Part 3 “Results of Ex-Post Evaluations, Stakeholder Consultations and Impact Assessments”, Explanatory Memorandum, European Commission proposal for the AI act.

<sup>②</sup> Yasmína Yakimova/Janne Ojamo, “Artificial Intelligence Act: Deal on Comprehensive Rules for Trustworthy AI”, September 12, 2023, <https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>, 访问日期:2023-12-10。

<sup>③</sup> 同注①。

<sup>④</sup> Anu Bradford, *The Brussels Effect: How the European Union Rules the World*, p. 25.

<sup>⑤</sup> European Parliamentary Research Service, “EU Legislation in Progress: Artificial Intelligence Act”.

<sup>⑥</sup> Michael Veale/Frederik Zuiderveen Borgesius, “Demystifying the Draft EU Artificial Intelligence Act — Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach”, *Computer Law Review International*, Vol. 22:4, 2021, pp. 97-112, here p. 98.

<sup>⑦</sup> Part 1 “Context of the Proposal”, Explanatory Memorandum, European Commission Proposal for the AI act.

<sup>⑧</sup> European Commission, “Impact Assessment Accompanying the AI Act”, April 21, 2021, <https://ec.europa.eu/newsroom/dae/redirection/document/75792>, 访问日期:2023-12-11。

“基本法”。<sup>①</sup>这一立法路径将允许相关主体利用欧盟法律,挑战国内法对 AI 系统的使用限制。<sup>②</sup>

其次,在《AI法》中,欧盟还搭建了基于全过程风险划分体系的监管框架。为配合落实《AI法》的基本法地位,欧盟在《AI法》中设置了针对 AI 系统开发、入市、使用的全过程风险划分体系,将几乎所有场景都纳入其监管范围。在《AI法》的修订过程中,欧盟采用基于风险的规制思路,重点引入了以风险为导向、分类分级监管 AI 系统的制度,将 AI 系统划分为不可接受的风险(unacceptable risk)、高风险(high risk)、有限风险(limited risk)、低风险或极小风险(low or minimal risk)四类,并针对每一类别分别采取不同的监管措施。<sup>③</sup>不可接受风险涉及违反基本权利的 AI 系统应用,它们可能对基本权利和自由构成威胁,有违欧盟的价值观,例如操纵人类行为以造成人员伤害,或利用儿童、残疾人等特定弱势群体弱点的应用。<sup>④</sup>这些应用被视为不可接受的风险,被欧盟禁止使用。<sup>⑤</sup>高风险主要出现在关键基础设施、就业、私人和公共服务、执法等领域的 AI 应用中,这些应用必须在投入市场前经过严格的合规性评估。<sup>⑥</sup>《AI法》要求对这些应用进行严格评估和持续监管,确保其在整个生命周期中的安全和合规性,该法附录三列出了八个可能存在高风险 AI 系统应用的领域。<sup>⑦</sup>有限风险则常见于与人类互动的系统(即聊天机器人)、情感识别系统、生物识别分类系统,以及生成或处理图像、音频、视频内容(即深度伪造)的 AI 系统。这些 AI 系统将受到一系列有限的透明度义务的约束。<sup>⑧</sup>例如对于聊天机器人,《AI法》要求相关系统保证一定程度的透明度,确保用户知道他们正在与 AI 系统进行交互。<sup>⑨</sup>对于由 AI 系统驱动的视频游戏或垃圾邮件过滤器等低风险或极小风险的 AI 应用,该法规定的限制相对较少,但仍鼓励这类应用遵守最佳实践。这一类别的 AI 系统并无详细的专门条款,它们将在该

① Michael Veale/Frederik Zuiderveen Borgesius, “Demystifying the Draft EU Artificial Intelligence Act -Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach”, p. 98.

② Case C-148/78, Criminal proceedings against Tullio Ratti, ECLI:EU:C:1979:110, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A61978CJ0148>, 访问日期:2023-12-11.

③ European Parliamentary Research Service, “EU Legislation in Progress: Artificial Intelligence Act”.

④ Part 5 “Other Elements” of Explanatory Memorandum, 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 (COM/2021/206 final).

⑤ Article 5, Title II “Prohibited Artificial Intelligence Practices” the AI Act.

⑥ Article 6 “Classification Rules for High-Risk AI Systems” and Article 7 “Amendments to Annex III”, the AI Act.

⑦ Annex III “High-risk AI Systems Referred to in Article 6(2)”, the AI Act.

⑧ 同注③。

⑨ Article 52 “Transparency Obligations for Certain AI Systems”, the AI Act.

法总体指导原则下运作。

### (三) 规制工具

#### 1. 义务设置

基于风险分级路径,《AI法》将高风险 AI 系统供应者和使用者作为最主要的监管对象。<sup>①</sup> 对于高风险 AI 系统,该法施加了整体性的合规要求,<sup>②</sup>以及构建风险管理系统(Risk management system)、<sup>③</sup>数据治理(Data and data Governance)、<sup>④</sup>记录保留(Record-keeping)、<sup>⑤</sup>用户透明度与知情(Transparency and provision of information to users)、<sup>⑥</sup>人类监管(Human oversight)<sup>⑦</sup>等具体义务。例如,“记录保留”义务要求高风险 AI 系统在系统生命周期内自动记录事件(即“日志”,log)。<sup>⑧</sup>“用户透明度与知情”义务要求相关系统的设计和开发确保其运行足够透明,并遵守相关义务,使用户能够正确理解和使用该系统。<sup>⑨</sup>再以“人类监管”义务为例,该法要求高风险 AI 系统在设计 and 开发中设置适当的人机界面工具,以便在系统使用期间能够受到自然人的有效监督。<sup>⑩</sup>此外,高风险 AI 系统的提供者还需要承担署名、设置质量管理体系(quality management system)、保留日志、知会所在国监管机构等义务。<sup>⑪</sup>

针对 ChatGPT 等通用 AI 系统(General Purpose AI Systems),欧盟理事会在最初的法案文本中没有定义该术语,<sup>⑫</sup>直到 2022 年的折中草案修订中才提供了对于通用 AI 系统的定义。根据《AI法》的规定,通用 AI 系统系旨在执行普遍适用功能的 AI 系统,不论该系统以何种方式投放市场或投入使用(包括作为开

---

<sup>①</sup> Nikos Th. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies — The AI Act*, Cham: Springer Nature Switzerland, pp. 327 – 741.

<sup>②</sup> Article 8 “Compliance with the Requirements”, Council’s compromise text on the proposed AI Act, 7 March 2022.

<sup>③</sup> Article 9 “Risk Management System”, Council’s compromise text on the proposed AI Act, 7 March 2022.

<sup>④</sup> Article 10 “Data and Data Governance”, Council’s compromise text on the proposed AI Act, 7 March 2022.

<sup>⑤</sup> Article 12 “Record-Keeping”, Council’s compromise text on the proposed AI Act, 7 March 2022.

<sup>⑥</sup> Article 13 “Transparency and Provision of Information to Users”, Council’s compromise text on the proposed AI Act, 7 March 2022.

<sup>⑦</sup> Article 14 “Human Oversight”, Council’s compromise text on the proposed AI Act, 7 March 2022.

<sup>⑧</sup> 同注<sup>⑤</sup>, para 1.

<sup>⑨</sup> 同注<sup>⑥</sup>, para 1.

<sup>⑩</sup> 同注<sup>⑦</sup>, para 1.

<sup>⑪</sup> Article 16 “Obligations of Providers of High-Risk AI Systems”, Council’s compromise text on the proposed AI Act, 7 March 2022.

<sup>⑫</sup> 同注<sup>①</sup>.



源软件),如图像和语音识别、音频和视频生成、模式检测、问答、翻译等;通用 AI 系统可以在多种环境中使用,并集成到多个其他 AI 系统中。<sup>①</sup> 该文本还明确了通用 AI 系统提供商须承担的义务,<sup>②</sup>要求这类基础模型提供者确保系统对基本权利、健康、安全、环境、民主和法治的有力保护。通用 AI 系统提供商须评估和减轻其模型所带来的风险,相关系统的运营受到数据治理规定和登记义务的约束。通用 AI 系统供应者有义务确保系统生成内容的可预测性、安全性,并向用户披露其内容是由 AI 而非人类生成,还须公布与模型训练数据相关的知识产权等信息。

## 2. 监管沙盒

《AI 法》引入 AI 监管沙盒制度,“建立受控环境,在有限时间内测试创新技术、接入数字创新中心、使用测试和实验设施,这将有助于创新公司、中小型企业 and 初创企业在遵守新法规草案的前提下继续创新”。<sup>③</sup> 在该制度下,成员国或欧洲的数据保护监管机构必须建立 AI 监管沙盒,即在主管当局监督和指导下的受控环境,鼓励开发者根据与主管当局商定的计划,在技术投放市场前对其进行测试和验证,以期开发者遵守《AI 法》的有关规定。<sup>④</sup> 在沙盒中处理的任何个人数据均位于参与者控制下的独立、隔离和受保护的数据处理环境中,只有获授权人员才能访问该数据。<sup>⑤</sup> 监管部门对在沙盒中处理的任何个人数据均应通过适当的技术和措施进行保护,在沙盒终止或个人数据保留期结束后删除。监管沙盒能够使参与者通过使用个人数据来促进基于公共利益目的的 AI 创新,且不会影响相关市场主体遵守 GDPR 的合规要求。<sup>⑥</sup> 监管机构应在沙盒内提供指导和监督以确定风险,特别是测试针对基本权利、民主和法治、健康和环境等已确定风险的缓解措施及其有效性,确保 AI 系统符合《AI 法》以及其他相关立法的要求。<sup>⑦</sup> 《AI 法》折中版本要求,在监管沙盒中,AI 系统的潜在提供者须按照应适用的欧盟和成员国法律,对于沙盒实验造成的第三方损害承担责任。沙盒的设立机构还应当每年向欧盟的

① Article 3, point (1b), Council's compromise text on the proposed AI Act, 13 May 2022.

② Nikos Th. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies — The AI Act*, pp. 327 – 741.

③ European Commission Communication, “Fostering a European Approach to Artificial Intelligence”, COM (2021) 205 final, Brussels, April 21, 2021, p. 7.

④ Article 53 “AI Regulatory Sandboxes”, the AI Act.

⑤ Article 54 “Further Processing of Personal Data for Developing Certain AI Systems in the Public Interest in the AI Regulatory Sandbox”, Council's compromise text on the proposed AI Act, 7 March 2022.

⑥ 同上。

⑦ 同注④。

AI 办公室(AI Board)提交年度报告,自沙盒设立一年后开始,直到 AI 系统项目终止。<sup>①</sup>

## 二、AI 治理“布鲁塞尔效应”的实现及其挑战

“布鲁塞尔效应”是欧盟依托统一市场的规模优势,通过颁布塑造国际商业环境的法规提高相关全球标准,制定并推广数据隐私、消费者权益、环境保护、反垄断等领域的欧盟规则,导致全球商业许多重要方面显著“欧洲化”的一种现象。<sup>②</sup> 欧盟提出只有在第三国提供与欧盟规范同等水平的保护后,这些国家才能在相关领域与欧盟进行交流。<sup>③</sup> 由此,欧盟将内部立法与标准推向世界,在相关领域的治理中获得优势。如今,欧盟试图在 AI 监管领域形成新的“布鲁塞尔效应”,这一目标的实现可以从其产生动因、实现机制与可行性挑战三个方面展开分析。

### (一) 欧盟推动 AI“布鲁塞尔效应”的现实动因

欧盟已成功将其监管标准输出到世界其他地区的多个领域,数据治理方面的 GDPR 便是典型例证。自 2018 年实施以来,GDPR 已成为数据保护的黄金标准,影响着亚洲、非洲和美洲各国和地区的相关立法。来自全球各地的公司都不得不遵守其对数据处理和隐私的严格规定,因为该法规可适用于处理欧盟境内数据但在欧盟境内未设立实体机构的主体,<sup>④</sup>这就相当于欧盟通过提高跨境数据输出的审查标准及“长臂管辖”制度,在世界范围内推广 GDPR 建立的“欧盟数字标准”,<sup>⑤</sup>呈现出明显的“布鲁塞尔效应”。在环境规制方面,欧盟在碳排放标准制定方面的领导力同样体现了“布鲁塞尔效应”。欧盟排放交易体系(EU Emissions Trading System,以下简称 EU ETS)是世界上第一个也是规模最大的碳市场,欧盟通过 EU ETS 影响了其他国家在碳定价机制方面的决策,为全球碳市场设定了基准。在欧盟规则下,相关进口产品在进口许可方面是否需要支出额外成本、成本数额多少,都取决于 EU ETS 规则;进口产品可否免除进口许可成本,取决于出口国是否建立起和 EU ETS 相同或相通的碳市场。<sup>⑥</sup> 依托市场规模优势与先发立法优势,欧盟

① Article 53 “AI Regulatory Sandboxes”, para 5, the AI Act.

② Anu Bradford, *The Brussels Effect: How the European Union Rules the World*, pp. xiii-xix.

③ European Commission, “Adequacy Decisions”, December 17, 2021, [https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions\\_en](https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en), 访问日期: 2023-12-27; 靳思远、沈伟:《〈数字经济伙伴关系协定〉的科技战略导向和中国选择进路》,载《中国科技论坛》,2023 年第 2 期,第 122-134 页,这里第 124 页。

④ Article 3.2, GDPR.

⑤ 靳思远、沈伟:《〈数字经济伙伴关系协定〉的科技战略导向和中国选择进路》,第 124 页。

⑥ 陈红彦:《欧盟碳边境调整机制的合法性考辨及因应》,载《法学》,2021 年第 12 期,第 177-192 页,这里第 185 页。

将内部标准推上世界舞台,使自身在相关领域治理中获得一席之地。

随着欧盟立法向新兴的 AI 监管进军,欧盟是否计划在 AI 治理中复制其在监管方面的成功经验,在该领域重现“布鲁塞尔效应”? 尽管《AI 法》雄心勃勃,但在世界范围内,包括中美两国在内的主要经济体都在密切关注 AI 产业的发展并在该领域大规模投入。与此形成鲜明对比的是,欧盟在 AI 的研发、创新、生态、数据及技术应用等方面都难以取得相对于中美等竞争对手的优势。<sup>①</sup> 在发展现状方面,美国和英国拥有全世界最前沿的 AI 实验室和模型,<sup>②</sup>中国则在相关领域的科学出版物数量和 AI 专利申请数量方面领先世界。<sup>③</sup> 在研发投入方面,2022 年的数据显示,美国的 AI 私人投资总额居世界首位,达到 474 亿美元,投资额大约是排名第二的中国(134 亿美元)的 3.5 倍。<sup>④</sup> 欧盟在 AI 领域的产业发展与研发投入都落后于中美两国。但是,中美两国所采用的监管策略都显著有别于《AI 法》,例如,美国的《关于安全、可靠和可信赖的人工智能发展和使用的行政命令》(Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence,以下简称《行政命令》)就缺少关于基本权利保障的内容。<sup>⑤</sup> 但前述 RRI 理念在《AI 法》中体现为“对公共利益的高度保护”,这种保护又具体落实为对欧盟公民的基本权利(fundamental rights)的保护。《AI 法》强调对 AI 系统的开发和使用应尊重人类的基本权利,这一原则在不同系统使用场景中都应得到落实。欧盟的严格监管有机会影响中美以及其他国家和地区在监管 AI 方面的策略,尤其是当来自这些经济体的公司发现需要遵守欧盟标准以保持对欧盟市场的通行时,这便是“布鲁塞尔效应”发挥作用的时刻。但是,欧洲数字产业组织“欧洲数字中小企业联盟”(European Digital SME Alliance)警告立法者不应提出过于严格的合规要求,而是应允许中小企业参与相关行业和法律标准制定过程,要求在所有欧盟成员国

---

① Daniel Castro/Michael McLaughlin/Eline Chivot, “Who is Winning the AI Race: China, the EU or the United States?”, August 19, 2019, <https://datainnovation.org/2019/08/who-is-winning-the-ai-race-china-the-eu-or-the-united-states/>, 访问日期:2024-02-04。

② Nestor Maslej et al., “Artificial Intelligence Index Report 2023”, [https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI\\_AI-Index-Report\\_2023.pdf](https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf), 访问日期:2023-12-27。

③ Lizka, “The State of AI in Different Countries—An Overview”, September 14, 2023, <https://forum.effectivealtruism.org/posts/Lb2TjSsjpqA8rQ7dP/the-state-of-ai-in-different-countries-an-overview>, 访问日期:2023-12-27。

④ 同注②。

⑤ The White House, “Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence”, October 30, 2023, <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>, 访问日期:2023-12-11。

强制实施监管沙盒制度，<sup>①</sup>以提升《AI法》的规则确定性以及企业合规便利性，促进相关领域的创新活动。

在产业发展与研发投入落后于其他主要经济体的情况下，<sup>②</sup>如果欧盟成功地将自己确立为AI监管的规范设定者，其国际谈判中的地位便可得到增强，还可为其他国家设立先例。倘若《AI法》中的限制标准与产业促进措施实现了良好平衡，则该法可能会成为增强欧盟相关竞争力和影响力的利器。“布鲁塞尔效应”是理解欧盟全球监管影响力的重要框架，可以应用于AI等欧盟监管涉足的新兴领域。随着欧盟在AI立法道路上继续前进，全世界都在关注或者试图避免“布鲁塞尔效应”在这一新兴关键领域的重现。欧盟输出AI监管标准的成功与否，取决于欧盟的战略视野、立法技术及其与该领域中其他参与者的互动情况。

## （二）AI“布鲁塞尔效应”赖以实现的核心机制

“布鲁塞尔效应”是欧盟在新兴领域实现监管领先和监管辐射的一种话语和规则形成机制，是欧盟“单边监管全球市场能力”（unilateral ability to regulate the global marketplace）的表现。<sup>③</sup>从已经形成“布鲁塞尔效应”的新兴领域看，监管主体的市场规模（Market Size）或市场权力（Market Power）、监管能力（Regulate Capacity）、严格规则（Stringent Regulations）偏好、非弹性监管对象（Inelastic Targets）偏好以及采用不可切割标准是形成“布鲁塞尔效应”不可缺少的五项条件，它们共同构成实现“布鲁塞尔效应”的核心要素和机制。

**1. 市场规模或市场权力。**在市场的语境下，监管者的权力通常与市场规模密切相关，因为跨国企业将为确保其进入具有较大规模的市场而倾向于采用这些市场的标准。<sup>④</sup>市场规模的重要性在于，相较于监管环境更为宽松的出口国，采取严格监管的进口国市场规模越大、向该国出口的商品比例越高，“布鲁塞尔效应”就越有可能发生。<sup>⑤</sup>

**2. 监管能力。**并非每一个具有相当市场规模的国家都能成为监管大国（regulatory power）和全球标准制定者，后者还必须具备将其市场力量转化为实际监管影响力的监管能力。可制裁违规行为的权威性是监管能力的重要组成部分

<sup>①</sup> European Digital SME Alliance, “DIGITAL SME Position Paper on the EU AI Act”, September 2021, <https://www.digitalsme.eu/digital/uploads/DIGITAL-SME-Position-Paper-AI-Act-FINAL-DRAFT-1.pdf>, 访问日期:2023-12-08。

<sup>②</sup> Nestor Maslej et al., “Artificial Intelligence Index Report 2023”.

<sup>③</sup> Anu Bradford, *The Brussels Effect: How the European Union Rules the World*, p. 1.

<sup>④</sup> Daniel W. Drezner, “Globalization, Harmonization, and Competition: The Different Pathways to Policy Convergence”, *Journal of European Public Policy*, Vol. 12:5, 2005, pp. 841-859, here p. 843.

<sup>⑤</sup> David Vogel/Robert A. Kaga (eds.), *Dynamics of Regulatory Change: How Globalization Affects National Regulatory Policies*, Oakland: University of California Press, 2004, pp. 4-5.

分,只有市场规模足够大的司法管辖区,才有能力通过将不合规公司排除在市场之外使其承担巨大成本,并在此基础上强制进行监管调整(regulatory adjustment),<sup>①</sup>使“布鲁塞尔效应”得以实现。监管者的监管经验与执行能力是监管能力的直接体现。当前,仅有美国、欧盟等少数司法管辖区因监管体系的精巧设计以及自身市场规模而取得了相应的监管能力和全球影响力,其中欧盟在贸易、竞争政策等建立和加强单一市场的关键领域获得了全球监管权(global regulatory power)。<sup>②</sup>

**3. 严格规则。**该要求是指监管者行使其监管能力的意愿,即监管者颁布严格监管标准的倾向。有研究认为,高收入司法管辖区更倾向于实行严格监管,因为新兴市场经济体可能不愿负担消费者保护所带来的本国企业盈利能力下降的后果。<sup>③</sup>高收入经济体就全球标准设置权开展竞争的关键在于,谁能在本国推行更为严格的标准。美国与欧盟之间的监管竞争便沿着这一路径展开,标准设置权在二者推进本国标准不断收紧的过程中交替。<sup>④</sup>这些严格的标准会使企业付出合规成本,企业也更有动力促进这些标准在全球推广,以确保企业在其他业务区的公平竞争环境与自身竞争力。

**4. 非弹性监管对象。**所谓“非弹性”,是指监管对象无法随着监管方式变化做出具体反应,故监管者得以将其与特定的监管制度绑定的特征。欧盟选择非弹性对象进行监管,以防止监管效果受到监管套利行为的不利影响。如果监管对象是公司税、金融交易税,企业就可以轻松地通过转移公司注册地与金融交易开展地来规避高额企业所得税或金融交易相关税收,甚至还可能引发监管者之间的逐底竞赛。<sup>⑤</sup>欧盟将监管重点置于消费者市场,意图避免这一问题,因为消费者正是企业服务无从绕开或改变的目标,除非该企业放弃欧盟市场。

**5. 标准的不可分割性。**该条件又可以根据具体内容的差异区分为“技术不可分割性”“经济不可分割性”以及“法律不可分割性”三类。“技术不可分割性”是指在技术上无法将投放于严格司法管辖区的产品或服务区别于其他产品或服务的情形,例如,由于无法将在欧盟收集的信息独立储存,谷歌被迫对全球业务进行调整

---

<sup>①</sup> David Bach/Abraham L. Newman, “The European Regulatory State and Global Public Policy: Micro-Institutions, Macro-Influence”, *Journal of European Public Policy*, Vol. 14:6, 2007, pp. 827 – 846, here pp. 831 – 832.

<sup>②</sup> Anu Bradford, “The Brussels Effect”, *Northwestern University Law Review*, Vol. 107:1, 2012, pp. 1 – 68, here pp. 13 – 14.

<sup>③</sup> J. Luis Guasch/Robert W. Hahn, “The Costs and Benefits of Regulation: Implications for Developing Countries”, *The World Bank Research Observer*, Vol. 14:1, 1999, pp. 137 – 158, here p. 138.

<sup>④</sup> 同注<sup>②</sup>, p. 15.

<sup>⑤</sup> 同注<sup>②</sup>, pp. 16 – 17.

以符合欧盟的严格监管标准。<sup>①</sup>“经济不可分割性”所指向的是企业将欧盟业务从其他业务中拆分的高昂经济成本，为不同消费市场开发不同产品的成本将逼迫生产者将其产品全部依照欧盟标准进行生产。“法律不可分割性”的最典型例证为欧盟、美国监管者对于兼并行为的审查权力，最严格的反垄断司法管辖区可以决定兼并交易的最终结果，<sup>②</sup>相关企业将因此被迫遵守最严格的反垄断要求。若相关规则中采用了不可分割标准，即便生产者在全球不同地区所提供的产品或服务标准不一，采取最严格监管的司法辖区也可以要求企业按照本辖区要求提供产品或服务。

在上述的五项条件中，前三者是欧盟出于自身市场规模与监管意愿设计的监管要求，后两者更多的是监管技术的体现。在市场规模、监管能力与严格监管偏好的基础上，欧盟采用了针对非弹性对象的不可分割标准，这是实现“布鲁塞尔效应”的根本要求。我们可以对标上述“布鲁塞尔效应”的核心要素，检验《AI法》相关规则产生“布鲁塞尔效应”的可能性。

### （三）AI“布鲁塞尔效应”的可行性挑战

欧盟《AI法》作为国内法具有一定的领先性，兼顾了全球 AI 治理安全目标和价值的“最大公约数”。欧盟采用的全场景与风险分级规制确保了《AI法》对于产业链的全覆盖，基本权利、内容治理、产品安全甚至劳动保护都是该法涉及的议题，大而全的严格法律体系与对外效力都彰显出欧盟企图复制 GDPR 成功、重现“布鲁塞尔效应”的野心。<sup>③</sup>但在“布鲁塞尔效应”严格标准的要求下，法律规则与现实需求的适配与否会成为掣肘“布鲁塞尔效应”的因素。<sup>④</sup>核心机制与需求适配两个角度是考察《AI法》相关规则能否产生“布鲁塞尔效应”的重要维度。

一方面，基于“布鲁塞尔效应”形成的核心机制，在面对其他法域产业发展优势与规则竞争的当下，实现《AI法》的“布鲁塞尔效应”道阻且长。

第一，欧盟的市场规模巨大，这在消费者保护、环境保护和数据治理等诸多新兴领域的重要性已经得到验证，欧盟的巨大消费人口规模对于企业具有无可比拟的吸引力。

<sup>①</sup> Ryan Singel, “EU Tells Search Engines to Stop Creating Tracking Databases”, April 8, 2008, <https://www.wired.com/2008/04/eu-tells-search/>, 访问日期:2024-02-16。

<sup>②</sup> Anu Bradford, “Antitrust Law in Global Markets”, in Einer R. Elhauge (ed.), *Research Handbook on The Economics of Antitrust Law*, Northampton: Edward Elgar Publishing, 2012, pp. 283-326, here p. 310.

<sup>③</sup> 王融:《AI时代,欧盟会复制下一个 GDPR 吗?》,2023年12月21日, <https://www.tisi.org/27214>, 访问日期:2023-12-27。

<sup>④</sup> 赵海乐:《欧盟数字市场治理规则“布鲁塞尔效应”的限度与我国因应》,载《德国研究》,2023年第6期,第78-95页,这里第89页。

第二,欧盟具有将其市场规模转化为监管能力的监管经验与执行能力,在竞争法这一影响共同市场形成的关键领域,欧盟频频“重拳出击”,在2017年、2018年、2019年连续就谷歌的垄断、不正当竞争等行为开出巨额罚单,总罚款金额近百亿美元。欧盟的相关执法行为常常指向在相关领域处于领先地位的美国公司,使这些公司在欧盟、美国或其他第三国的行为产生显著改变。<sup>①</sup>《AI法》的“布鲁塞尔效应”也须基于欧盟的自由流动能力(EU free movement competences)才能实现。欧盟法院曾指出,如果成员国对特定技术或设施的使用限制措施阻碍了市场准入,则该措施可构成相当于贸易数量限制的措施,因为这种限制“对消费者的行为有相当大的影响”,使消费者“对购买该产品的兴趣有限”。<sup>②</sup>可见,《AI法》的基本法性质与欧盟内部市场的统一性相辅相成,前者确保欧盟成员国不会采取单边措施导致欧盟的AI统一市场受到干预,后者又构成了前者具备基本法地位的基础,因为只有商品自由流通的统一内部市场,讨论《AI法》的基本法地位才有意义。这种关系在AI领域得到了直接体现,在多个涉及AI系统应用的案件中,欧盟法院已经利用自由流通相关规则来进行裁判。

第三,欧盟对于严格标准有着显著偏好。2003年颁布的《限制有害物质指令》(Restriction of Hazardous Substances Directive)禁止包括家用电器和电脑等在内的所有电子产品在使用寿命结束时向环境中排放有害物质,这是最为严格的电子废物监管规则,包括中国、日本、韩国和美国加州等国家和地区在内的司法管辖区都选择响应该规则,将严格的欧盟标准纳入本土立法,<sup>③</sup>体现了“布鲁塞尔效应”在环保领域的作用。《AI法》所设置的监管规则是目前最为严格的AI相关规则。

第四,在监管对象的非弹性方面,《AI法》强调AI系统在使用时对消费者和统一市场的影响,将风险监管的对象分布于AI系统的整个产业链,通过立法技术限缩相关企业进行监管套利的空间。

第五,由于AI系统开发与模型训练的高昂成本,<sup>④</sup>开发者与供应者或将受制

① Anu Bradford, *The Brussels Effect: How the European Union Rules the World*, p. 99.

② Case C-110/05, *Commission of the European Communities v Italian Republic*, ECLI:EU:C:2009:66, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELLAR:6fc44dd2-3e99-47a3-b97a-01b3999dfde7&from=FR>, 访问日期:2023-12-11.

③ Henrik Selin/Stacy D. VanDeveer, “Raising Global Standards: Hazardous Substances and E-Waste Management in the European Union”, *Environment: Science and Policy for Sustainable Development*, Vol. 48:10, 2006, pp. 6-18, here pp. 14-15.

④ AI训练是教授AI系统感知、解释和学习数据的过程,在这一过程中通过输入大量数据和算法,AI模型将学会识别和生成规律,只有经过训练的AI才能够进行推理并根据得到的信息做出决策。参见KJ Jacoby, “Tech Explainer: What is AI Training?”, March 10, 2023, <https://www.performance-intensive-computing.com/objectives/tech-explainer-what-is-ai-training>, 访问日期:2024-02-17.

于“经济不可分割性”而放弃重新开发和训练新模型，在全球范围内遵守《AI法》的严格标准以期进入消费人口规模庞大的欧盟市场。但是由于数字产品的无形性、可复制性等特征，数字企业有可能通过业务拆分、法人人格分割、法律选择、差异化隐私政策等策略剥离欧盟业务，在不同司法管辖区适用不同技术标准。<sup>①</sup> 囿于欧盟在 AI 研发与投入方面的落后，欧盟立法者可能无法很好地掌握 AI 发展现状，立法盲区随之产生。一个典型案例是，欧盟直到 ChatGPT“出圈”后才将通用 AI 系统纳入《AI法》规制范围，<sup>②</sup>且该法直至 2024 年年中才会逐步落地生效。因此，欧盟实际上无法确保其严格标准的不可分割性。倘若因技术发展，AI 系统的开发者在技术上能够以较低成本实现 AI 系统的“欧盟特供”，或许“布鲁塞尔效应”在新兴领域的“攻城略地”会在 AI 产业暂告一段落。此外，作为 AI 开发的领先者，美国也越发重视对于 AI 的监管，于 2024 年 2 月 8 日成立了 AI 安全研究所联盟 (Artificial Intelligence Safety Institute Consortium)，“制定新的指南、工具、方法、协议和最佳实践，以促进行业标准的发展，以安全、可靠和值得信赖的方式开发或部署 AI”是该联盟的重要工作内容。<sup>③</sup> 面对行业领先者，欧盟《AI法》中规则与标准的领先性与严格性势必受到挑战和质疑。因此，AI 治理的“布鲁塞尔效应”能否经由《AI法》实现，在很大程度上将取决于欧盟在 AI 领域的未来发展水平。

另一方面，在欧盟之外，世界其他主要国家的不同主体也在单边和多边层面尝试推进 AI 规制，制定、出台 AI 治理规范，我们可以从中窥见寻求全球 AI 治理的共识或“最大公约数”的可能性。由于欧盟是世界上最大、最具吸引力的市场之一，跨国公司严格遵守欧盟标准就是相关领域的“最佳实践”，这导致了欧盟标准在全球范围内的推广，此谓“事实上的布鲁塞尔效应”(de facto Brussels Effect)。在此基础上，跨国公司还将更具动力在全球各市场推动当地立法向欧盟标准看齐，当地政府受此影响的相关立法行为形成了“法律上的布鲁塞尔效应”(de jure Brussels Effect)。<sup>④</sup> 由此，“布鲁塞尔效应”的限度来自其规则对于其他国家乃至跨国企业的接合成本。因此，我们应基于全球 AI 治理共识分析《AI法》内容，以明确 AI“布鲁塞尔效应”的可行性。

在单边层面，美国作为当前科技竞争力和产业竞争力双重领先的国家，<sup>⑤</sup>在

① 赵海乐：《欧盟数字市场治理规则“布鲁塞尔效应”的限度与我国因应》，第 80 页。

② Nikos Th. Nikolopoulos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies — The AI Act*, pp. 327 - 741.

③ U. S. Artificial Intelligence Safety Institute, <https://www.nist.gov/artificial-intelligence/artificial-intelligence-safety-institute>, 访问日期：2024 - 02 - 16。

④ Anu Bradford, *The Brussels Effect: How the European Union Rules the World*, p. 2.

⑤ 王彦雨、李正风、高芳：《欧美人工智能治理模式比较研究》，载《科学学研究》，2023 年第 3 期，第 460 - 468 页，这里第 461 页。



2019年就提出了美国AI政策的十项原则,并以公众参与、科学诚信和促进“可靠、有弹性和值得信赖的AI应用”作为其中的重要内容。<sup>①</sup>2023年10月,美国白宫发布了《行政命令》,旨在确保美国AI技术的发展与应用安全、可靠、可信,强调AI技术在促进经济增长、增强国家安全、改善政府服务以及提升公民生活质量方面的重要作用,指出AI技术发展带来的挑战和风险。《行政命令》的核心内容包括:第一,为确保AI的安全可靠,所有相关联邦机构必须确保其AI系统的安全性和可靠性,具体包括保护数据隐私、确保算法的公平性和透明度以及提高AI系统的可解释性。第二,为促进AI技术的创新与发展,该命令鼓励联邦机构支持相关研究与创新,包括投资与AI相关的教育和培训项目以培养人才,加强与私营部门的合作。第三,关于国家安全和防御,该命令强调AI技术在国家安全和防御方面的重要性,指出相关部门应从保护国家安全的角度识别和降低使用AI技术的风险。第四,关于国际合作,该命令呼吁在国际层面开展合作以促进全球治理,包括制定国际标准和规范、加强国际伙伴关系。第五,在监管和政策框架方面,该命令要求制定一套全面的监管框架,以指导相关技术的发展和应用,并确保其符合联邦政府的政策和标准。<sup>②</sup>

在多边层面,包括中国、美国、欧盟在内的国家和地区于全球首届AI安全峰会签署了AI安全的纲领性文件《布莱切利宣言》(Bletchley Declaration),力求推动AI的安全、负责任和伦理性的发展,体现了参与国家和地区对在该领域合作的承诺。该宣言强调了五个方面的主要内容:第一,该宣言强调AI系统的安全性和可靠性,呼吁采取必要措施确保其安全可靠运行,防止AI系统的误用或滥用。第二,该宣言强调任何AI技术的开发和应用都应遵循国际人权法和伦理原则,以保护个人隐私、防止算法歧视。第三,该宣言还强调透明度和问责制,主张增强AI决策过程的透明度,建立有效的问责机制,以确保其使用符合公众利益。第四,该宣言同时还强调国际合作的重要性,主张加强世界各国和地区在AI领域的合作,包括共享最佳实践、协调政策和标准,以解决跨境技术应用带来的挑战。第五,该宣言也认识到AI技术的创新发展对社会和经济的重要性,提倡在确保安全和伦理的基础上推动技术进步。<sup>③</sup>

① The White House, “Maintaining American Leadership in Artificial Intelligence”, February 14, 2019, <https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence>, 访问日期:2023-12-11。

② The White House, “Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence”.

③ “The Bletchley Declaration by Countries Attending the AI Safety Summit”, November 1, 2023, <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>, 访问日期:2023-12-11。

《行政命令》《布莱切利宣言》等不同国家和地区的 AI 治理政策文件中体现了一些共识性内容，AI 治理“最大公约数”已经有所显现。这些共识的内容主要集中在安全性和可靠性、伦理先行和人权保护、透明度和问责制、国际合作、促进技术创新和发展五个方面，具体包括：（1）安全性和可靠性。几乎所有的 AI 治理文件都强调了安全性和可靠性的重要性，包括确保 AI 系统在操作中的安全性、防止 AI 系统被滥用，以及确保 AI 系统在预期的应用领域内可靠运行。（2）伦理先行和人权保护。AI 治理政策强调必须尊重人权和伦理标准，包括保护个人隐私、防止算法歧视和确保数据的透明使用。（3）透明度和问责制。这一理念要求各相关主体提升在 AI 应用和决策过程中的透明度，建立问责机制，确保在出现问题时可以溯源并采取相应措施。（4）国际合作和标准制定。这包括在国际层面建立共同的标准和原则，以及通过跨国合作解决 AI 发展和应用中的共同挑战。（5）促进技术创新和发展。除了关注风险管理和合规性要求，各国和地区的政策文件还强调支持 AI 技术的创新和发展，认为 AI 技术的发展对于经济增长和社会福祉至关重要。

“最大公约数”共识为全球 AI 治理提供了基础，有助于促进 AI 负责任发展。欧盟《AI 法》作为国内法，其文本具有一定的领先性，兼顾了全球 AI 治理安全目标和价值的“最大公约数”。其中，安全性和可靠性、伦理先行和人权保护、透明度和问责制、促进技术创新和发展都是《AI 法》中的重要内容，其中关于安全可靠、伦理保护的内容更是领先全球。《AI 法》设置了风险分级治理、沙盒监管等监管制度，对 AI 开发者、供应者和使用者都课以诸多义务，可谓当前最为严格的 AI 监管框架。相较于国际通行的 AI 治理原则，《AI 法》唯一欠缺的是有关国际合作机制与标准制定的内容。欧盟在环境保护、消费者保护等新兴领域的严格监管偏好体现了政策制定者对于相关风险的厌恶，这既是消费者（即欧盟公民）本身的要求，又得益于欧洲政策制定者在采取严格监管策略上的意识形态统一，有利于政策响应民众对于更领先、更严格监管的需求。<sup>①</sup> 因此，在包含必要议题的基础上，欧盟所采用的规制路径确保了《AI 法》对于相关产业的广泛覆盖，其对该法体系结构与域外效力的设计，都彰显出欧盟在 AI 治理中重现“布鲁塞尔效应”的目标。但是，过于严苛的标准会对“布鲁塞尔效应”的实现造成负面影响：欧盟的标准如果被视为过于繁重或滞后，则可能会被忽视或绕过。欧盟标准限制产业发展的情况已经在 GDPR 所处的数据领域出现。在 GDPR 生效后的半年间，欧盟的

---

<sup>①</sup> David Vogel, *The Politics of Precaution: Regulating Health, Safety, and Environmental Risks in Europe and the United States*, New Jersey: Princeton University Press, 2012, pp. 31–32.

相关中小企业在融资方面陷入了困境,大型企业却获得了合规优势。<sup>①</sup>因此,若要实现AI的“布鲁塞尔效应”,欧盟还需基于国际通行的治理原则,确保该法推行与合规的成本可控,防止主权国家回应寥寥、相关主体切割市场,掣肘《AI法》的全球推广。

### 三、AI规制的“布鲁塞尔效应”和中国因应

在AI规制方面,我国以纲领性文件为方向,在国际倡议、法律及标准等层面构建起AI监管的中国体系。我们以“布鲁塞尔效应”为视角,对我国的国际倡议、法律法规和行业标准等AI治理相关文本进行归纳分析,尝试基于中欧差异探究我国AI治理路径的发展方向与未来方案。

#### (一)AI规制的中国体系构成

##### 1. 国际倡议

《全球人工智能治理倡议》(以下简称《治理倡议》)是中国在国际上提出的一项重要倡议,目标为推动全球AI技术的负责任发展和治理。这一倡议强调了在全球化背景下,AI技术的发展应遵循公平正义、安全透明、符合伦理等六项原则,具体包括:第一,促进公平正义,反对技术霸权和单边主义,提倡通过国际合作确保各国平等参与相关技术的研发和利用,帮助发展中国家增强AI能力。第二,保障数据安全和个人隐私。鉴于AI技术对数据的依赖,《治理倡议》指出必须加强对数据的保护,避免数据滥用和个人隐私泄露,确保数据处理的透明度和可追溯性。第三,提高AI系统的安全性和可靠性。该倡议要求提高相关系统的安全性和可靠性,防止相关技术被用于非法和不道德目的。第四,推动伦理规范建设,尊重人类尊严和社会价值,提倡在全球建立相关伦理规范,引导AI技术的健康发展。第五,加强国际合作,呼吁国际社会共同制定全球AI治理规则,加强跨国界、跨学科的交流与合作,共同应对AI技术带来的挑战。第六,推动技术创新和普惠发展,该倡议主张相关技术应用应有利于社会发展和人类福祉。<sup>②</sup>

##### 2. 法律法规

在我国现行法律中,《数据安全法》明确了开展数据处理活动的原则,要求技术促进经济社会发展,增进人民福祉,符合社会公德和伦理。<sup>③</sup>《个人信息保护法》从

<sup>①</sup> Jian Jia/Ginger Zhe Jin/Liad Wagman, “The Short-Run Effects of GDPR on Technology Venture Investment”, November 2018, [https://www.nber.org/system/files/working\\_papers/w25248/w25248.pdf](https://www.nber.org/system/files/working_papers/w25248/w25248.pdf), 访问日期:2023-12-27。

<sup>②</sup> 《全球人工智能治理倡议》,2023年10月20日,[https://www.mfa.gov.cn/web/ziliao\\_674904/1179\\_674909/202310/t20231020\\_11164831.shtml](https://www.mfa.gov.cn/web/ziliao_674904/1179_674909/202310/t20231020_11164831.shtml), 访问日期:2023-12-13。

<sup>③</sup> 《数据安全法》第二十八条。

个人信息保护角度规制自动化决策，要求使用个人信息进行自动化决策者确保决策的透明度和结果公平、公正，保障相关个人的知情权与拒绝权。<sup>①</sup>此外，还有诸多聚焦于电子商务、金融科技等特定领域的与算法治理相关的规定，散见于《电子商务法》《反垄断法》等法律法规中。

在部门规章中，《互联网信息服务算法推荐管理规定》重点治理以“大数据杀熟”为代表的算法歧视性决策，对生成合成类、个性化推送类、排序精选类、检索过滤类、调度决策类等多种算法技术进行全面规制。<sup>②</sup>《互联网信息服务深度合成管理规定》则秉持“辨伪求真”的愿景，以利用深度学习、虚拟现实等生成合成类算法制作文本、图像、音频、视频、虚拟场景等网络信息的深度合成技术作为规制客体，<sup>③</sup>规定深度合成技术的适用边界、规制主体和法律责任。与AI规制最直接相关的《生成式人工智能服务管理暂行办法》（以下简称《办法》）则在整体上聚焦生成式AI的规范应用与发展，坚持发展和安全兼顾、促进创新和依法治理相结合等原则，既鼓励生成式AI创新发展，又对生成式AI服务实行包容审慎和分类分级监管。<sup>④</sup>在机制设计、算法开发、内容生成与用户引导等不同的规制主题下，《办法》对生成式AI系统的开发者、提供者、使用者设置了权利与义务，促进生成式AI服务全链条的合法、合理、合情发展。

### 3. 行业标准

除立法之外，我国还重视AI领域的标准体系设计。2020年，国家标准化管理委员会联合中央网信办、国家发展改革委、工业和信息化部等五部门，印发了《国家新一代人工智能标准体系建设指南》，试图加强AI领域标准化顶层设计，推动AI产业技术研发和标准制定，促进产业可持续发展。<sup>⑤</sup>2021年，全国信息安全标准化技术委员会秘书处发布的《人工智能伦理安全风险防范指引》则力求防范AI伦理安全风险，针对潜在的伦理安全风险问题，为AI产业相关组织或个人开展研究开发、设计制造、部署应用等相关活动提供指引。<sup>⑥</sup>此外，在《办法》发布后，全国信息

① 《个人信息保护法》第二十四条。

② 《互联网信息服务算法推荐管理规定》第二条第二款。

③ 《互联网信息服务深度合成管理规定》第二十三条。

④ 《国家互联网信息办公室有关负责人就〈生成式人工智能服务管理暂行办法〉答记者问》，2023年7月15日，[https://www.gov.cn/zhengce/202307/content\\_6892001.htm](https://www.gov.cn/zhengce/202307/content_6892001.htm)，访问日期：2023-12-13。

⑤ 《国家标准化管理委员会 中央网信办 国家发展改革委 科技部 工业和信息化部关于印发〈国家新一代人工智能标准体系建设指南〉的通知》，2020年7月27日，[https://www.gov.cn/zhengce/zhengceku/2020-08/09/content\\_5533454.htm](https://www.gov.cn/zhengce/zhengceku/2020-08/09/content_5533454.htm)，访问日期：2023-12-13。

⑥ 全国信息安全标准化技术委员会秘书处：《网络安全标准实践指南——人工智能伦理安全风险防范指引》，2021年1月，<https://www.tc260.org.cn/upload/2021-01-05/1609818449720076535.pdf>，访问日期：2023-12-13。

安全标准化技术委员会相继提出将构建拟作为其配套标准的《生成式人工智能预训练和优化训练数据安全规范》《信息安全技术生成式人工智能人工标注安全规范》等标准,对生成式 AI 产品的训练数据合法性与可追责性提出量化要求。<sup>①</sup>

## (二) 中国与“布鲁塞尔效应”: AI 治理的竞争与未来

欧盟的 AI 法律治理体系被认为具有自上而下的结构和强约束的特征,这些特征具有避免规则碎片化、降低相关主体合规成本等优势,但也导致了技术创新与严格治理、规则相同与资源差异等方面的矛盾。<sup>②</sup> 相较我国当前所采用的以场景规制为主的 AI 治理路径,欧盟的一体式立法和风险分级治理具有通用性更高、合规成本更低等优势。当前,世界上已经形成了 AI 治理的价值理念共识。《AI 法》的既有内容涵盖了“最大公约数”的大部分内容,只有关于国际合作的部分未在该法中体现。在此方面,《AI 法》可谓提纲挈领,以一部全面监管 AI 各领域的通用立法占据了相关技术国内法监管的领先地位。

但是, AI 领域的通用立法饱受争议,原因在于各个行业应用 AI 的风险差异较大,在缺乏广泛调研、未深入了解各行业 AI 技术应用差异时,尝试用普遍规则来规制多行业应用,有可能掣肘特定行业 AI 的发展。<sup>③</sup> AI 技术在众多产品、服务中均有使用,在法律规范中追求单一、宽泛的定义,或许会使相关技术的研究者、供应者、使用者等都不得不面对高额合规成本。<sup>④</sup> 例如,有观点认为《AI 法》对 AI 技术施以严格的监管要求,将使开发者面临高昂的合规成本和不成比例的责任风险,导致创新企业和投资者将业务和资金转移至欧盟以外地区,<sup>⑤</sup>这无疑与“布鲁塞尔效应”的形成路径背道而驰。立法者之所以采用既粗放又严苛的立法路径,是因为其对 AI 技术具体实现路径和未来发展趋势的把握能力有限,无法预先精准规制尚未成型的 AI 应用。虽然欧盟《AI 法》的输出规则与标准来势汹汹,但仍有学者指出欧盟的 AI 监管标准还面临《美日数字贸易协定》《数字经济伙伴关系协定》等先进国际数字经济规则的监管竞争,<sup>⑥</sup>在产业发展上的落后

① 《附件 1:2023 年度第一批网络安全国家标准需求清单》,2023 年 5 月 6 日, <https://www.tc260.org.cn/upload/2023-05-06/1683366666445047167.pdf>, 访问日期:2023-12-13。

② 王彦雨、李正风、高芳:《欧美人工智能治理模式比较研究》,第 464-465 页。

③ 宁宣凤、吴涵:《“不要温和地走进那良夜”——对〈生成式人工智能服务管理办法〉的思考》,2023 年 4 月 16 日, <https://www.kwm.com/cn/zh/insights/latest-thinking/reflections-on-draft-regulation-on-management-of-generative-artificial-intelligence-services.html>, 访问日期:2023-12-13。

④ 张凌寒:《深度合成治理的逻辑更新与体系迭代——ChatGPT 等生成型人工智能治理的中国路径》,载《法律科学(西北政法大学学报)》,2023 年第 3 期,第 38-51 页,这里第 45 页。

⑤ Supantha Mukherjee, “Draft EU Artificial Intelligence Rules Could Hurt Europe, Executives Say”, July 1, 2023, <https://www.reuters.com/technology/draft-eu-artificial-intelligence-rules-could-hurt-europe-open-letter-2023-06-30/>, 访问日期:2024-01-03。

⑥ 赵海乐:《欧盟数字市场治理规则“布鲁塞尔效应”的限度与我国因应》,第 88 页。

身位还可能引发欧盟在 AI 监管竞争中的视野受限，最终导致其在竞争中落败。中国应当把握自身在 AI 开发中的领先优势，利用技术发展的领先身位，研究、分析、归纳 AI 系统在不同使用场景、不同发展阶段、不同技术基础中的不同规制重点，在自身技术优势的基础上实现立法超越，以更为合理、严格的立法标准反哺 AI 领域技术发展，推动相关国际标准的制定，争取在 AI 领域实现中国版“布鲁塞尔效应”。

我国未来的 AI 治理体系，可以兼顾场景化、分散式立法与统一化、专门式立法，<sup>①</sup>既通过基本法勾画 AI 治理的基本原则和规制工具，又利用特别法灵活应对 AI 系统的技术发展。首先，AI 基本法应当将前述“最大公约数”的相关原则确立为原则性规范。其次，基于这些原则，基本法应确定 AI 系统的开发者、供应者和使用者等 AI 系统生命周期中的各阶段参与者需要承担的根本性义务。再次，鉴于该法对支持 AI 创新与发展的作用，基本法中不应设置过于苛刻的 AI 研发限制，而是应该鼓励 AI 创新发展，并引入监管沙盒等包容性的监管工具。最后，对于未尽义务的相关主体，基本法也要对法律责任的承担方式有所设计和反映。在特别法方面，AI 包含各种元素、技术和场景，在研发、生产、应用和产生影响的过程中必然涉及 AI 算法和应用的各种特殊问题，因此势必要将 AI 治理的对象明细化。<sup>②</sup>我国可以沿用现有的规制方式，结合“布鲁塞尔效应”产生的必要条件，立足我国市场规模，针对应用频率较高的 AI 系统出台特别规范。我国应以严格监管规则对特定 AI 技术和应用进行风险防范，通过选定非弹性监管对象、适用不可分割性标准等立法技术，填补“粗线条”基本法的立法留白。由此，我国或可兼顾欧盟方案与中国模式的优点，既避免欧盟立法中 AI 定义过宽的缺陷，又探索不同 AI 规制场景下的最优方案，争取实现我国对 AI 立法的规则外溢。<sup>③</sup>

## 结 语

考察 AI 治理的当前发展，AI 可能对人类造成的潜在威胁绝非科幻臆想，实乃 AI 在其生命周期的各环节、各场景对人类根本价值和基本权利带来的现实挑战。规制 AI 应用的意义便在于维护人类社会的既有秩序，防止技术应用对人类造成不可承受的风险。从欧盟立法方案与中国规制路径的分殊，到全球 AI 治理理念“最

<sup>①</sup> 史凤林、张志远：《论人工智能的公法规制：美欧模式与中国路径》，载《理论月刊》，2023 年第 8 期，第 127—139 页，这里第 134 页。

<sup>②</sup> 张辉、曾雄、刘鹏：《中国式治理现代化视域下的人工智能治理》，载《中国科学基金》，2023 年第 4 期，第 632—639 页，这里第 636 页。

<sup>③</sup> 张凌寒：《深度合成治理的逻辑更新与体系迭代——ChatGPT 等生成型人工智能治理的中国路径》，第 47 页。

大公约数”的浮现,背后折射的是对 AI 规制的必要性与可行性。

各国应当为 AI 设计合理的治理体系和规则,既防范风险又避免对其创新与发展构成限制,以实现技术与风险治理的平衡。作为负责任大国,我国应在国内法层面建立完善的 AI 国内规制体系,以欧盟方案为参照设置高水平的规范标准,立足我国处于领先地位的 AI 产业发展与投入,争取在 AI 领域实现中国版“布鲁塞尔效应”。在此基础上,我国可在国际层面积极参与 AI 治理规范的国际法制定,如国际软法、行业规范以及伦理共识等,<sup>①</sup>促进我国国内法治、涉外法治与国际法治有机结合,以国际规则推动 AI 领域朝着安全可靠、平等共益的方向发展。

责任编辑:朱苗苗

---

<sup>①</sup> 沈伟、赵尔雅:《数字经济背景下的人工智能国际法规制》,载《上海财经大学学报》,2022年第5期,第123-137页,这里第136页。