

# 人工智能法律主体地位的困境与出路

曹子豪

青岛科技大学法学院, 山东 青岛

收稿日期: 2025年9月10日; 录用日期: 2025年10月9日; 发布日期: 2025年10月20日

## 摘要

近年来人工智能技术呈加速演进态势,明确其法律主体定位成为破解复杂人工智能法律议题的关键所在。本文探讨了人工智能法律主体地位的理论基础、现实需求、核心困境及出路。传统法律主体以意识、自由意志等为核心,而人工智能虽具自主性等特征却缺乏真正意识。现实中,人工智能引发的责任归属、权利认定等问题亟需法律回应。其核心困境在于意向性与自由意志缺失、工具理性的片面性、缺乏权利与责任能力。为此,本文提出构建法律拟制主体理论,建立责任承担机制并完善配套制度,主张采取渐进式路径,根据人工智能功能配置法律地位,以平衡技术发展与法律稳定。

## 关键词

人工智能, 法律主体地位, 法律拟制主体, 工具理性, 责任承担

# The Dilemma and Way out of the Legal Subject Status of Artificial Intelligence

Zihao Cao

Law School, Qingdao University of Science and Technology, Qingdao Shandong

Received: September 10, 2025; accepted: October 9, 2025; published: October 20, 2025

## Abstract

In recent years, artificial intelligence technology has accelerated its development trend, and clarifying its legal subject status has become the key to solving many legal problems of artificial intelligence. This paper discusses the theoretical basis, practical needs, core dilemma and way out of the legal subject status of artificial intelligence. Traditional legal subjects take consciousness and free will as the core, while artificial intelligence has the characteristics of autonomy but lacks real consciousness. In reality, issues such as responsibility attribution and rights identification caused by AI urgently need legal response. Its core dilemma lies in the lack of intention and free will, one-sided

文章引用: 曹子豪. 人工智能法律主体地位的困境与出路[J]. 争议解决, 2025, 11(10): 165-172.

DOI: 10.12677/ds.2025.1110320

instrumental rationality, and lack of rights and responsibilities. Therefore, this paper proposes to construct the theory of legal fiction subject, establish the responsibility mechanism and improve the supporting system, and advocate the gradual path to allocate the legal status according to the AI function to balance the technological development and legal stability.

## Keywords

Artificial Intelligence, Legal Subject Status, Instrumental Rationality, Legal Fiction Subject, Responsibility

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 问题的提出

人工智能技术的迅猛发展正在深刻改变人类社会的运行模式，其在自主决策、创造性活动和社会交互等方面展现出的能力，对传统法律体系构成了前所未有的挑战。从 AlphaGo 击败人类围棋冠军到 ChatGPT 引发的生成式人工智能浪潮，从自动驾驶汽车的规模化应用到智能投资顾问的普及，人工智能正在越来越多地参与原本由人类主导的社会活动。这种技术变革不仅带来了效率提升和创新机遇，更引发了关于人工智能法律地位的思考。

现行法律主体制度建立在康德式的理性主体观之上，以自由意志和道德主体能力为核心要素。该制度预设只有具备意识、理性和自主决策能力的实体才能成为法律关系的承载者。然而人工智能的出现打破了这种预设——它展现出类似人类的智能行为，却缺乏传统意义上的意识和自由意志。这种“似人非人”的特性使法律陷入两难：承认其主体地位可能颠覆整个法律体系的伦理基础，否认其主体地位则无法有效应对人工智能引发的法律问题。

理论界与实务界对人工智能法律地位的回应呈现出明显的滞后性和分歧性。2023 年北京互联网法院的“人工智能文生图”一案首次认定人工智能生成内容具有作品属性<sup>1</sup>，但回避了创作主体问题；广州互联网法院的“奥特曼案”则将侵权责任直接归于开发者<sup>2</sup>。这种司法实践的不一致反映出制度供给的严重不足。国务院《新一代人工智能发展规划》明确要求“明确人工智能法律主体以及相关权利、义务和责任”，但具体路径仍然模糊。面对人工智能带来的颠覆性变革，法律必须在坚守人类中心主义与适应技术发展之间寻找平衡点。

## 2. 人工智能法律主体地位的理论基础

### (一) 法律主体的传统理论

法律主体作为法学的基础概念，其理论演进经历了漫长的历史过程。在传统法学理论中，自然人的主体地位源于其作为理性存在者的本质属性。如德国法学家萨维尼所言：“所有的权利都是为了道德的、内在于每个人的自由而存在的[1]。”可见，自然人主体地位的构成要素包括：意识能力、自由意志、道德责任能力以及社会交往能力。这些要素共同构成了自然人享有权利和承担义务的基础。

法人主体地位的确立则经历了从否定到承认的演变过程。拟制说认为法人是法律的创造物，其主体资格完全源于法律的赋予；实在说则主张法人具有独立于成员的现实存在基础。现代法学理论普遍接受

<sup>1</sup>参见北京互联网法院(2023)京 0491 民初 11279 号判决书。

<sup>2</sup>参见广州互联网法院(2024)粤 0192 民初 113 号判决书。

了法人作为独立主体的观点，承认其具有独立的权利能力，但在行为能力方面需要通过自然人代表来实现。这种理论发展为理解人工智能的法律地位提供了重要参照。

权利能力与行为能力的区分是法律主体理论的重要内容。权利能力是成为权利义务主体的资格，而行为能力则是通过自己的行为取得权利、承担义务的能力。这种区分在人工智能语境下具有特殊意义，因为人工智能可能在某些方面展现出行为能力的特征，但是否具备完整的权利能力仍存在争议。

## (二) 人工智能的技术特征

当代人工智能技术呈现出显著的自主性、学习能力和决策能力特征。深度学习算法使人工智能系统能够从海量数据中自主学习规律，做出超越预设程序的决策。人工智能的自主性表现为其能够在没有人类直接干预的情况下，根据环境变化调整行为策略，这种自主性使得人工智能的行为结果往往超出设计者和使用者的预期范围。

强人工智能与弱人工智能的区分对于法律主体地位的讨论至关重要。弱人工智能专注于特定任务的智能化处理，如图像识别、语言翻译等；强人工智能则追求达到或超越人类的通用智能水平。目前的技术发展仍处于弱人工智能阶段，但其展现出的能力已经在某些领域接近甚至超越人类水平。但是现阶段的人工智能仍存在不容忽视的技术局限性，尽管在特定领域表现出色，但人工智能仍缺乏真正的理解能力、情感体验和道德判断能力。当前的人工智能系统本质上是复杂的模式识别工具，距离真正的智能还有很长的路要走。这些局限性构成了赋予人工智能完整法律主体地位的重要障碍。

## (三) 人工智能法律主体地位的理论争议

支持赋予人工智能法律主体地位的学者主要基于功能主义立场进行论证。他们认为，如果人工智能能够履行法律主体的功能，就应当被赋予相应的法律地位。这种观点强调结果导向，认为法律主体资格的认定应当基于实际功能而非本体论属性。支持者认为，既然公司等非自然实体能够获得法律人格，人工智能也应当在满足特定条件时获得类似地位。反对论者则从意识和道德主体的角度提出质疑。他们认为，人工智能仅是符号操作系统，缺乏真正的理解和意识，法律主体地位的核心在于道德主体性，而人工智能缺乏道德感知和责任能力，不具备成为法律主体的根本条件。这种观点坚持法律主体必须具备内在的道德维度。折中论者提出了有限主体或特殊主体的构想。比如有学者认为，人工智能所具有的“工具性人格”可以看做一种有限人格，人工智能既不完全是“人”，也不纯粹是“物”。它处于一种“中间状态”——具有某种主体性特征，又保持着客体属性<sup>[2]</sup>。欧洲议会法律事务委员会建议创设“电子人”这一特殊法律类别，赋予最复杂的自主机器人特定的权利和义务。这种方案试图在完全否定和完全承认之间找到平衡点，既承认人工智能的特殊性，又避免将其完全等同于传统法律主体。

人工智能法律主体地位认定的分歧体现了人们对“主体”本质的不同理解。支持者从功能主义出发，认为人工智能已具备独立决策和行为能力，应当赋予其相应的法律地位以明确责任归属<sup>[3]</sup>。反对者则坚持本体论立场，强调法律主体必须具备意识、情感等人类特有属性，人工智能本质上只是工具，人才是也必须是这个世界的本原，无论机器如何被智能化，在法律关系中，其只能是权利义务指向的客体，意义仅限于服务于人类<sup>[4]</sup>。这种分歧的深层原因在于法律主体概念本身的多重性——它既是权利义务的承载者，又是道德责任的承担者，还是社会关系的参与者。人工智能在这三个维度上表现出的不一致性加剧了认定的困难。

人工智能法律主体地位认定的困境源于其自身权利能力与责任能力的“能”与“不能”。人工智能能够创作出具有独创性的作品、进行复杂的决策、甚至表现出某种“创造力”，这似乎具备了权利能力的实质要件。但同时，它缺乏独立财产、无法承担刑事责任、不具备真正的意志自主性，这又否定了其责任能力的可能。更为复杂的是，人工智能的行为往往涉及多方主体——设计者、生产者、训练者、使用者，责任分配面临前所未有的挑战。算法黑箱的存在使得因果关系认定几乎不可能，传统的过错责任

原则失去了适用基础。

### 3. 人工智能法律主体地位的现实需求与国际实践

#### (一) 现实需求分析

人工智能决策造成损害的责任归属问题日益突出。2018 年，优步自动驾驶测试车辆在亚利桑那州造成行人死亡事故，引发了关于责任主体认定的激烈讨论。传统的产品责任和侵权责任理论在面对人工智能自主决策造成的损害时显得力不从心。人工智能的不可预测性和自主性特征使得传统的过错责任和严格责任原则都难以直接适用。

人工智能创作作品的权利归属争议不断涌现。2018 年，佳士得拍卖行以 43.25 万美元的价格售出了人工智能创作的肖像画《Edmond de Belamy》，引发了关于人工智能作品著作权归属的广泛讨论。中国法院在处理人工智能生成内容的著作权案件中，也面临着权利主体认定的困境。北京互联网法院在“菲林诉百度”案中认定，人工智能生成内容不构成作品，但可以作为法人作品给予保护<sup>3</sup>。目前人工智能生成技术方案的可专利性与权属也是学界广泛讨论的话题，在自然人没有作出创造性贡献的前提下，由人工智能生成的技术方案能否被授予专利权，发明人如何认定，专利权归属于谁并没有达成共识。

智能合约的法律效力认定成为区块链时代的新挑战。智能合约通过代码自动执行合同条款，其执行过程完全由算法控制，人类难以干预。这种自动化执行机制对传统合同法理论提出了挑战，特别是在合同解释、违约救济等方面。如何确定智能合约的法律主体地位，关系到整个去中心化金融体系的法律基础。人工智能的引入则进一步增强了智能合约的功能，使其在数据处理、决策优化和自动化执行方面更加高效。明确人工智能法律主体地位有助于解决人工智能在智能合约自主决策上的效力与责任承担问题。

#### (二) 国际立法实践

欧盟在人工智能法律规制方面走在世界前列。2024 年通过的《人工智能法案》采用基于风险的分级规制方法，将人工智能系统分为不可接受风险、高风险、有限风险和最小风险四个等级<sup>4</sup>。虽然该法案没有直接赋予人工智能法律主体地位，但通过详细的合规要求和责任分配机制，为人工智能的法律规制提供了全面框架。欧盟还在探索“算法问责”机制，要求人工智能系统的开发者和使用者承担相应的法律责任。

美国采取分散式立法与判例法相结合的发展路径。联邦层面尚未出台统一的人工智能法律，但各州纷纷制定相关立法。加州的《自动驾驶汽车测试规定》、纽约的《算法问责法案》等都在特定领域对人工智能进行规制。美国法院通过判例逐步明确人工智能相关的法律规则，如在 *Loomis v. Wisconsin* 案中，法院认可了人工智能风险评估工具在刑事司法中的使用，但要求保障正当程序权利<sup>5</sup>。

日本和韩国在人工智能法律探索方面表现积极。日本内阁府发布的《以人为中心的人工智能社会原则》强调人工智能应当服务于人类福祉，提出了人类中心、教育/素养、隐私保护、安全保障、公平竞争、公平性/问责制/透明性、创新等七项原则<sup>6</sup>。韩国则通过《智能信息化基本法》建立了人工智能发展的法律框架，明确了政府、企业和社会在人工智能发展中的责任分工。

国际组织提供了重要的原则性指引。OECD 发布的《人工智能原则》强调包容性增长、可持续发展和福祉，要求人工智能系统透明和可解释<sup>7</sup>。UNESCO 的《人工智能伦理建议书》则从人权、包容、透明等角度提出了全球性的伦理框架。这些国际性文件虽不具有强制力，但为各国立法提供了重要参考。

<sup>3</sup>参见北京互联网法院(2018)京 0491 民初 239 号判决书。

<sup>4</sup>欧盟委员会，人工智能法规(人工智能法案)提案[R]. COM(2021) 206 决赛，2021: 34.

<sup>5</sup>*State v. Loomis*, 881 N.W.2d 749 (Wis. 2016), p.761.

<sup>6</sup>日本内阁府，以人为本的人工智能的社会原则[R]. 2019: 8.

<sup>7</sup>经合组织，人工智能理事会的建议[R]. OECD/LEGAL/0449, 2019: 7.



## 4. 人工智能法律主体地位的核心困境

### (一) 意向性与自由意志的缺失

传统法律主体理论以意识、理性、自由意志等为核心标准，但这些概念本身就存在定义上的模糊性。在人工智能语境下，这种模糊性被进一步放大。我们甚至无法对人类意识给出精确定义，更遑论判断人工智能是否具有意识。功能主义标准虽然提供了替代方案，但“功能”的界定同样面临困难。著名的“中国大脑”思想实验提出，如果 14 亿中国人每人模拟一个神经元的功能，整体是否具有意识？这个问题在人工智能领域同样存在。深度神经网络可能在功能上模拟了大脑的某些运作方式，但是否因此具有意识仍是未解之谜。主观体验如何从物理过程中产生在人工智能意识讨论中尤为关键。有学者认为，意识存在于人工智能“一切皆可算”的逻辑之外，而一旦人工智能产生意识则意味着象征性网络的消亡，这与人工智能的存在本身冲突，因此人工智能很难拥有意识[5]。

哲学意义上的意识包含着现象意识、通达意识和自我意识三个层次，而当前的人工智能即使是 ChatGPT 这样的大模型，其本质仍是基于统计学习的模式匹配系统。它能够模拟人类的语言行为，产生看似有意识的回应，但这种“智能”表现背后缺乏真正的理解和体验。算法驱动的决策过程与人类基于意向性的选择存在本质区别——前者是确定性的计算结果，后者包含着不可还原的主观体验和价值判断。人类的自由意志体现为在多种可能性中进行选择的能力，这种选择包含着道德责任的承担。而人工智能的输出是算法、数据和参数共同作用的过程，其不具有独立完整的主观意识，无法进行自由自觉地实践活动，其“选择”是伪选择，“创造”是伪创造[6]。即使引入随机性因素，也无法产生真正的自主性。深度学习的黑箱特性看似赋予了人工智能某种“不可预测性”，但这种不可预测性源于计算复杂性而非意志自由。缺乏自由意志意味着人工智能无法成为真正的道德主体，也就无法承担法律意义上的责任。

### (二) 工具理性的片面性

人工智能展现的工具理性与人类理性存在差异。工具理性关注手段与目的的匹配，追求效率最大化，这正是人工智能的优势所在。它能够在给定目标下寻找最优路径，进行精确计算和预测。然而，人类理性不仅包括工具理性，更包含价值理性——对目的本身的反思和批判能力。人工智能缺乏对“为什么”的追问能力，无法质疑既定目标的合理性。这种片面的理性使其只能作为达成人类目的的工具，而非独立设定目标的主体。

情感、直觉等非理性因素在人类决策中扮演着不可或缺的角色，而人工智能恰恰缺失这些要素。人类的法律判断往往涉及同情、愤怒、恐惧等情感因素，这些情感不是理性的障碍，而是理性判断的必要组成部分。法官的自由裁量权、陪审团的情感共鸣、立法者的价值选择，都体现着理性与情感的交织。人工智能可以模拟情感表达，但无法真正体验情感。它的“同情”是算法的输出，“愤怒”是参数的调整。这种情感缺失使其无法理解法律背后的人文关怀和价值诉求。

哈贝马斯的交往理性理论强调，主体性是在与他者的交往中建构的。人类通过语言交流、相互承认、共同协商形成主体间的理解和共识。而人工智能虽然能够进行语言交互，但缺乏真正的交往能力。它不能进行视角转换，无法理解他者的内心世界，更不能参与平等的对话。人工智能的“交流”是单向度的信息处理，而非双向的意义建构。缺乏主体间性意味着人工智能无法成为法律共同体的真正成员[7]。

### (三) 缺乏权利能力与责任能力

人工智能权利能力的缺乏主要体现在财产权、人格权与著作权上。财产权的本质是对物的排他性支配，这种支配以主体的独立意志为前提。人工智能本身就是他人的财产，无法拥有独立于其所有者的财产。即使赋予其名义上的财产，这种财产的取得、使用和处分仍然受制于人类的控制。更为关键的是，财产权的正当性基础在于劳动或创造，而人工智能的“创造”实质上是对训练数据的重组和优化，不具

有真正的原创性。缺乏独立财产意味着人工智能无法独立承担民事责任，这从根本上否定了其成为民事主体的可能。人格权保护生命、健康、名誉、隐私等人格利益，这些利益以人的生物性和社会性存在为基础。人工智能既无生命可言，也无健康概念，更谈不上名誉和隐私。虽然有学者提出“电子人格权”概念，但这种权利缺乏实质内容。保护人工智能的“尊严”毫无意义，因为尊严以自我意识和道德地位为前提。人工智能可能需要某种形式的“保护”，但这种保护是为了维护其功能完整性，而非人格利益。人工智能生成的文字、图像、音乐是否构成作品？如果构成，著作权归谁？这一问题在实践中引发了激烈争议。北京互联网法院认定人工智能生成内容具有独创性但否认其作者资格，深圳南山法院则认为机器人作品的著作权归属于使用者。这种矛盾反映出深层困境：承认人工智能的创作能力但是否认其创作主体地位，在逻辑上难以自洽。但赋予人工智能著作权又面临激励机制失效的问题——人工智能不需要经济激励，也无所谓精神权利[8]。

法律责任中通常以过错为责任基础，过错以行为人的主观心理状态为核心，包括故意和过失两种形态。但人工智能不存在真正的心理状态，其“故意”或“过失”无从谈起。算法的输出错误是技术缺陷还是“过失”？深度学习的不可预测性是“故意”还是“意外”？这些问题在现有法律框架内无法回答。即使采用客观过错标准，也面临标准设定的困难——以什么为参照？理性人标准显然不适用，而“理性算法”标准又缺乏共识。责任财产的缺失使人工智能无法独立承担赔偿责任。民事责任的最终实现需要责任财产作为保障，而人工智能既无独立财产，也无获得财产的能力。有学者提出设立人工智能责任基金或强制保险制度，但这实质上是将责任转移给了基金缴纳者或投保人，而非人工智能自身承担。更为根本的问题是，责任承担的目的不仅在于补偿，更在于预防和威慑。但人工智能既无趋利避害的本能，也无接受教训的能力，责任的预防功能完全失效[9]。惩罚措施的无效性否定了人工智能承担刑事责任的可能。刑罚的目的包括报应、威慑、矫正和隔离，但这些目的对人工智能都不适用。人工智能没有痛苦感受，监禁毫无意义；没有生命，死刑无从执行；没有财产，罚金难以实现。即使采用“删除程序”等特殊措施，也只是技术处理而非法律惩罚。更重要的是，刑事责任以道德可责性为基础，而人工智能缺乏道德主体地位，对其进行道德谴责毫无意义。刑法的威慑功能对人工智能完全失效，因为它既不怕惩罚，也不追求自由[10]。

## 5. 人工智能法律主体地位的出路探索

### （一）构建法律拟制主体理论

法律主体制度本质上是为了实现特定社会功能而创设的工具，而非对自然事实的简单反映。法人制度的成功先例表明，即使缺乏自然人格基础，基于社会需要进行法律拟制仍然可能且有效。人工智能虽无意识和意志，但在功能上已经能够独立完成许多原本需要人类才能完成的任务——创作、决策、交易等。从功能等效性出发，赋予其相应的法律地位具有工具理性的合理性。这种拟制不是要承认人工智能是“人”，而是为了更有效地规范人工智能参与的法律关系，明确权利归属和责任承担。

人工智能已经深度嵌入社会生活，其社会影响力不断扩大。ChatGPT 的用户超过 1 亿，每天产生海量的内容<sup>8</sup>；自动驾驶汽车在多个城市试运行；人工智能医疗诊断系统在医院广泛应用。这种广泛的社会应用产生了大量需要法律调整的关系，如果坚持否认人工智能的法律地位，将导致规范真空和责任缺失。法律必须回应社会现实，而非固守抽象原则。正如公司从“法律不承认”到“有限承认”再到“完全承认”的历史进程，人工智能的法律地位也可能经历类似演变。社会实践的需要最终会推动法律的变革。

<sup>8</sup>2023 年 1 月，OpenAI 通过官方渠道确认 ChatGPT 月活跃用户突破 1 亿，成为史上用户增长最快的消费级应用。这一数据在 2023 年 2 月的多份权威媒体报道中被广泛引用，例如《中国信息产业网》明确指出：“2023 年 1 月末，ChatGPT 用户数突破 1 亿，打破 Twitter 保持的 9 个月纪录”。

并非所有人工智能都需要或应当被赋予法律主体地位,拟制应当遵循必要性、适宜性和均衡性原则。必要性要求只有在现有法律框架确实无法有效调整时才进行拟制;适宜性要求拟制的程度与人工智能的智能水平和社会功能相匹配;均衡性要求在技术发展、社会需求和法律稳定性之间保持平衡。这种有限的、渐进的拟制路径既能回应现实需要,又不会对法律体系造成颠覆性冲击。

具体可以采取类型化处理,对不同类型的人工智能赋予不同程度的法律地位。强弱人工智能的区分标准需要从技术能力、应用场景和社会影响三个维度综合判断。技术能力方面,可以参考自主性程度、学习能力、创造能力等指标;应用场景方面,区分高风险应用(如自动驾驶、医疗诊断)和低风险应用(如推荐算法、智能客服);社会影响方面,考虑使用规模、潜在损害、可替代性等因素。这种多维度的评估体系避免了简单的技术决定论,也防止了一刀切的制度安排。欧盟《人工智能法案》采用的风险分级方法提供了有益借鉴,但需要结合中国实际进行本土化改造<sup>4</sup>。差异化的权利能力配置应当与人工智能的类型相匹配。对于弱人工智能,可以维持其客体地位,通过产品责任等传统规则进行调整;对于具有高度自主性的强人工智能,可以赋予其特定领域的有限权利能力,如在著作权领域承认其生成物的作品属性,但不赋予完整的著作人格权;对于未来可能出现的通用人工智能,可以考虑创设特殊的法律主体类型,比如有学者提出“人工类人格”,人工智能享有特定的权利(如署名权),承担有限的义务(如不得侵害他人权益),但不具有完整的人格<sup>[11]</sup>。该概念的价值在于承认了人工智能的特殊性,不强求其符合传统主体的全部要件,同时又赋予其必要的法律地位以解决实践问题。这种阶梯式的制度安排既保持了法律的开放性,又避免了激进变革的风险。关键是建立动态调整机制,根据技术发展和社会需求适时调整权利能力的范围。

## (二) 建立责任承担机制

根据不同情形安排直接责任与间接责任。当人工智能的损害行为可以追溯到特定人类主体的过错时,适用直接责任;当损害源于人工智能的自主决策且无法确定具体过错方时,适用替代责任。关键是建立清晰的责任分配规则:开发者对算法缺陷承担产品责任;生产者对硬件故障承担制造责任;使用者对不当使用承担过错责任;而对于真正的“算法自主”造成的损害,可以由获益最大的一方承担替代责任。这种多元化的责任体系既保护了受害者利益,又不会过度抑制创新。

借鉴公司法人格否认制度的思路。当人工智能被用作规避责任的工具时,应当穿透其“独立性”外观,直接追究背后控制者的责任。具体适用条件包括:恶意利用人工智能从事违法行为;故意设计缺陷以规避责任;明知人工智能存在严重安全隐患仍然使用等。这种穿透不是否定人工智能可能具有的有限法律地位,而是防止其被滥用。实践中可以设置举证责任倒置——当存在规避责任的表面证据时,控制者需要证明其行为的正当性。

人工智能的风险通常具有系统性,单一主体难以承担全部责任。可以借鉴环境侵权、产品质量等领域的经验,建立行业基金、强制保险、政府补偿等多层次的责任分担机制。所有人工智能开发和应用企业按照规模和风险等级缴纳基金,用于补偿人工智能造成的损害;强制高风险应用购买责任保险;对于创新试验造成的损害,政府可以提供一定补偿以鼓励创新。这种集体分担既分散了风险,又保证了受害者能够获得及时、充分的救济。

## (三) 完善配套制度

设置强人工智能准入门槛。准入标准应当包括技术能力评估、安全性测试、伦理审查三个维度。技术能力评估重点考察算法的可解释性、决策的可预测性、系统的稳定性;安全性测试强调对抗样本攻击、数据投毒、模型后门等安全威胁的防御能力;伦理审查确保人工智能的设计和应用符合公平、透明、可问责等伦理原则。这种三维评估体系形成了递进式的过滤机制,只有同时满足技术、安全、伦理要求的人工智能才能获得相应的法律地位。借鉴药品审批的做法,可以设置临床试验环节,在受控环境下测试

人工智能的实际表现。

建立动态评估机制。人工智能系统会通过学习不断进化，其能力和风险可能发生变化。因此需要建立定期评估制度，根据实际运行情况调整管理级别。评估指标包括事故率、投诉量、社会影响等客观数据，也包括专家评审、用户反馈等主观评价。当人工智能表现优异时可以提升级别、扩大权限；出现问题时降级甚至撤销资格。这种动态管理类似于驾驶执照的记分制度，通过正负激励引导人工智能朝着安全、可靠的方向发展。建立溯源机制，确保每个人工智能系统都能追溯到责任主体。

## 6. 结语

人工智能法律主体地位问题具有高度的复杂性和长期性，这种复杂性不仅源于技术本身的不确定性，更源于法律、伦理、哲学等多重维度的交织。人工智能技术仍在快速演进，其最终形态难以预测，因此法律制度设计必须保持开放性和适应性，任何试图一劳永逸解决该问题的方案都可能陷入困境。

考虑到现有法律体系的稳定性需求和人工智能技术的不成熟性，采取循序渐进的改革策略更为稳妥。可通过在特定领域、特定地区开展试点，积累实践经验，逐步完善制度设计。这种渐进式路径既能回应现实需求，又能避免制度变革带来的不确定性风险，中国的改革开放经验为此提供了有益借鉴。

相较于纠结人工智能是否具有意识、是否应享有完整主体地位等形而上学问题，从功能角度出发，根据人工智能在特定领域的实际功能配置相应法律地位更为务实。这种做法既避免了理论争议的泥潭，又能切实解决实际问题。功能主义并非权宜之计，而是符合法律实用主义传统的理性选择。

## 参考文献

- [1] 萨维尼.《当代罗马法体系》(第一卷)[M]. 朱虎, 译. 北京: 中国法制出版社, 2010: 89.
- [2] 许中缘. 论智能机器人的工具性人格[J]. 法学评论, 2018, 36(5): 153-164.
- [3] 袁曾. 生成式人工智能的责任能力研究[J]. 东方法学, 2023(3): 18-33.
- [4] 黎四奇. 对人工智能非法律主体地位的解析[J]. 政法论丛, 2023(5): 117-127.
- [5] 谭勇, 孔明安. 人工智能将会成为一种新主体吗?——齐泽克精神分析视野下的人工智能研究[J]. 国外理论动态, 2021(4): 147-154.
- [6] 洪涛. 哲学视角下生成式人工智能的法律主体证否及其融贯治理[J]. 兰州学刊, 2025(3): 83-96.
- [7] 哈贝马斯. 交往行为理论[M]. 麦卡锡, 译. 北京: 中信出版社, 2018.
- [8] 王迁. 再论人工智能生成的内容在著作权法中的定性[J]. 政法论坛, 2023, 41(4): 16-33.
- [9] 司晓, 曹建峰. 论人工智能的民事责任: 以自动驾驶汽车和智能机器人为切入点[J]. 法律科学(西北政法大學學報), 2017, 35(5): 166-173.
- [10] 刘宪权. ChatGPT 等生成式人工智能的刑事责任问题研究[J]. 现代法学, 2023, 45(4): 110-125.
- [11] 杨立新. 人工类人格: 智能机器人的民法地位——兼论智能机器人致人损害的民事责任[J]. 求是学刊, 2018, 45(4): 84-96.