

人工智能激励型立法审思：以激励规制理论为视角

袁秀挺, 严 驰

(同济大学 法学院, 上海 200092)

摘要:人工智能引领的科技创新和产业创新是加快培育形成新质生产力之关键。激励规制理论有效克服了信息不对称背景下传统规制经济学理论的局限性,为中国构建助力人工智能技术和产业创新发展的法律制度提供了思路参考。在理论层面,激励规制与法律激励和人工智能均具有内在契合性。人工智能法作为中国人工智能治理领域的基础性法律,应立足国情实际,通过引入激励规制理论,在传统立法模式的基础上有所扬弃和创新,考虑采用激励型立法,选择减少成本型和增加收益型的激励方式,以基金激励机制为镜鉴,既维护人工智能企业的利益,又促进社会公共利益的实现。

关键词:新规制经济学;激励规制;人工智能法;激励型立法;基金激励机制

中图分类号:D922.17 **文献标志码:**A **文章编号:**2096-028X(2025)04-0032-09

一、问题的提出:为何需要人工智能立法?

人工智能被认为是21世纪世界三大尖端技术之一。自2022年底ChatGPT横空出世以来,大量类ChatGPT产品不断被开发运用,通过集成至大量的应用程序及广泛的应用场景,新一代生成式人工智能技术正深度影响生产、分配、交换、消费等各个环节,成为人类社会智能化迭代发展的重要动力。中国的人工智能发展起步较晚,改革开放后才逐渐开启相关研究,产业发展周期较短。近年来,中国已初步建立起人工智能产业政策支持体系,发布了超千项人工智能产业政策,^①尽管中央和地方层面鼓励人工智能产业发展与创新的政策举措正密集出台,但是由于信息不对称的存在,产业政策可能会引起资源配置扭曲并催生寻租行为,不利于人工智能产业有效运行。^②同时,在当前政策执行过程中,存在政策协作难度大、监督保证体系缺失、制度落实执行不到位等问题,制约着中国人工智能产业的持续高质量发展。

人工智能作为发展新质生产力的重要引擎,能够引领生产要素创新配置、推进产业升级转型,给传统生产力带来整体性变革与跃迁。科学技术的进步和应用离不开市场机制的有效运作,人工智能产业的蓬勃发展及其带动的智能经济增长也离不开相应的政策和法律保障。企业是重要的创新主体,企业创新水平是评估人工智能产业发展情况的关键指标。制度经济学理论认为,正式制度尤其是法律制度,对企业创新能力建设发挥着关键性作用。^③人工智能产业作为战略性新兴产业的典型代表,其发展事关政府与市场、经营者与消费者等多重关系处理,涉及效率与公平、自由与秩序等多种价值权衡。^④为保障人工智能产业政策的稳定性和有效实施,发挥政策与法律的协同共振效用,夯实人工智能高质量发展的产业基础,调控和规制人工智

基金项目:2025年度中国版权保护中心版权研究重点课题“《全面与进步跨太平洋伙伴关系协定》视角下的人工智能生成内容版权问题研究”(BQ2025008),2025年度中共浙江省委政法委员会 浙江省法学会法学研究重点课题“可信数据空间视域下大模型训练的著作权制度优化研究”(2025NA29)

作者简介:袁秀挺,男,法学博士,同济大学法学院教授、博士生导师;严驰,男,同济大学法学院知识产权法专业博士研究生。

^① 参见贾怡伟、戚湧、孙明汉:《新一代人工智能产业政策及治理体系现代化研究》,载《中国科技论坛》2023年第12期,第53页。

^② 参见陈艳霞、张鹏:《人工智能产业政策的创新促进效应——来自企业专利数据的证据》,载《现代经济探讨》2024年第3期,第69页。

^③ 参见阳镇、凌鸿程、陈劲:《数智化时代非正式制度对创新的激励效应——基于人工智能企业的微观检验》,载《上海对外经贸大学学报》2021年第5期,第35页。

^④ 参见张守文:《人工智能产业发展的经济法规制》,载《网络信息法学研究》2019年第1期,第69页。

能技术可能产生的负面影响,加快形成新质生产力,推进国家层面人工智能立法具有充分的必要性。

现今中国人工智能专项立法规制体系尚付阙如,缺乏顶层设计和统筹规划的问题较为突出,^①立法数量、质量和层级上都存在明显不足。科技立法相对传统法律部门立法具有更强的时效性,鉴于人工智能立法在人工智能产业发展和治理领域的基础性、引领性和支撑性作用,近年来世界范围内掀起了人工智能大规模集中立法热潮。2024年8月1日,欧盟《人工智能法案》(*Artificial Intelligence Act*)正式生效。作为世界首部全面监管人工智能的法律,欧盟《人工智能法案》在全球人工智能发展与治理中占据了先发优势。中国人工智能立法应借鉴国际经验,在广泛社会实践基础上凝聚社会共识,制定既能引领国际规则又具有中国特色的法律制度。目前,中国人工智能法治建设中存在的立法层次较低且缺乏系统性、与国际规则接轨不够等问题,已成为制约人工智能产业转型升级、保障公民权益和社会公正的制度性障碍,人工智能统一立法的需求日益迫切。尽管人工智能法草案曾连续两年列入国务院立法计划“预备提请全国人大常委会审议”项目,但是整体上呈现为一种分散式立法方式。国务院2025年的年度立法工作计划仅强调“推进人工智能健康发展立法工作”,《全国人大常委会2025年度立法工作计划》中也未明确提及人工智能法。由此观之,中国的人工智能法从制度设计到出台实施还要经历一段相对漫长的过程。在此期间,理论研究应持续推进,让人工智能在发展与安全之间取得最佳平衡,持续增进人类共同福祉。以往学界针对如何进行人工智能立法作出了深入研究,但仍缺乏对基于中国实践的人工智能法立法模式的探讨。在构建现代化、本土化的人工智能治理体系的关键时期,基于“激励规制”理论的人工智能激励型立法为相关立法提供了一条新的路径。

二、破题之道:激励规制理论的引入

(一) 规制经济学的形成与演进

古典经济学和新古典经济学均认为,市场这只“看不见的手”是调节经济活动的最佳方式。^②福利经济学第一定理表明,完全竞争市场经济的一般均衡是存在且帕累托最优的。但是,市场势力的存在使得市场无法充分发挥资源配置的功能。随着20世纪30年代凯恩斯主义经济学的创立,政府这只“有形之手”逐渐成为影响和调整经济活动的重要力量。规制经济学的发展与市场紧密相连,当自运行的市场机制存在失灵或不完善等问题,无法实现提升生产效率和促进技术革新时,就需要政府介入进行直接规制以维护市场稳定和平衡,推动经济发展重回正轨。^③1970年,美国经济学家阿尔弗雷德·卡恩(Alfred E. Kahn)在著作《规制经济学:原理与制度》中首次提出“规制经济学”的概念,后逐渐形成了遵循完全理性假设的传统规制经济学理论体系。传统规制经济学理论假设政府规制机构与被规制企业的信息资源获取和利用处于同一水平线上,即双方之间进行的是一种对称信息博弈,侧重于关注规制的原因与效果,使得垄断企业缺乏降低成本的动力,可能会导致资本过度集约化的低效后果,且无法有效克服事前的逆向选择和事后的道德风险等非效率损失。随着20世纪70年代以来西方放松规制的改革之风的兴起,石油、电信、铁路等自然垄断行业陆续面向国有企业和私营企业开放。垄断和竞争是市场经济的两个极端,受限于行业的规模经济特征,在缺乏充分市场竞争的前提下,放松规制无法解决寡头企业垄断经营及服务效率偏低、产品供应不足等问题。为克服市场失灵带来的效率损失而对相关产业进行再管制,新规制经济学理论应运而生。

1986年,拉丰(Jean-Jacques Laffont)和梯若尔(Jean Tirole)联合发表了《运用成本观察进行企业规制》一文,^④初步奠定了新规制经济学的理论基础。1993年,二人合著的《政府采购与规制中的激励理论》出版,正式标志着新规制经济学的诞生。传统规制经济学理论仅把政府规制作为一种外生性制度来考虑,未有效考虑到受规制企业的激励问题。新规制经济学的核心即“激励规制”,强调规制制度须遵循激励相容原则,^⑤其

① 参见严驰:《GPT4发展中的若干问题及其规制方案》,载《信息安全研究》2023年第6期,第515页。

② 参见余东华:《政府与市场:一个管制经济学的视角》,载《经济体制改革》2004年第1期,第44页。

③ 规制作为学术概念在经济学、管理学、法学等多个学科中被广泛使用。在后文的法学视角下,规制主要是指广义视角下基于法律规范对市场和社会的干预和控制。

④ Jean-Jacques Laffont & Jean Tirole, *Using Cost Observation to Regulate Firms*, *Journal of Political Economy*, Vol.94: 614, p.614-641(1986).

⑤ 根据美国经济家里奥尼德·赫维茨(Leonid Hurwicz)提出的“机制设计理论”,在市场经济中,每个理性经济人都会追求自身利益,其个人行为会按自利的规则行动。若存在一种制度安排能使行为人追求个人利益的行为与企业实现集体价值最大化的目标相吻合,这种制度安排就被视为激励相容。

致力于解决的关键问题是如何立足于实践中非对称信息博弈及规制双方行为目标存在差异的现状,通过“委托—代理”论、信息经济学及机制设计等现代经济学研究方法的引入,设计出最优的激励规制方案,充分发挥政府指导作用和市场竞争优势,扼制代理人的机会主义行为,从而提高生产经营效率,推动资源配置最优化。

传统政府规制具有响应滞后、程序冗杂、执法成本高等局限性,难以在不确定的变革社会中进行精确分析和快速回应。^①发展新质生产力要求深化科技体制改革,形成鼓励创新的激励机制。如今在人工智能领域,政府与企业间存在高度的信息不对称。由于算法黑箱的客观存在,人工智能系统缺乏透明度和可解释性,无论是作为规制主体的政府,还是作为被规制主体的企业,对于人工智能的认知都相当有限。如何建立合理的激励规制制度,在满足企业利益诉求的同时实现人工智能治理目标、促进社会公共利益最大化是一个亟待解决的现实难题。激励规制理论颠覆了传统规制经济学理论中的信息对称性假设,强调规制者与被规制者之间激励相容的重要性,希望通过设计最优的激励规制机制以提升规制质量和效果,在中国人工智能产业发展与法律规制中有着广阔的应用前景,也为新质生产力的形成开拓了创新思路。需要注意的是,激励规制理论的建立主要是基于发达国家的社会条件和经济特点。发展中国家的规制理论设计须充分考虑与发达国家在规制能力、承诺能力、法律制度等方面的不同。今后应立足本土资源禀赋和制度环境,分析人工智能的内生性问题,而非简单地外生给定。通过将激励规制理论与中国人工智能发展现状相结合,秉持可信、安全、公平的基本原则,以法律形式进一步厘清政府和市场的关系,构建更科学有效的人工智能规制制度安排,推进国家治理体系和治理能力现代化。

(二)激励规制理论的融合发展

第一,激励规制与法律激励。法律激励的思想古已有之,《史记》载秦昭王“欲以激励应候”,意在激发和鼓励范雎。基于管理学视角,激励不仅可以视为个人的一种静态状态,也可以理解为管理者调动被管理者状态的动态过程,法律激励更接近于后者。法律的主要调控对象是个体行为,法律对个体行为的激励功能体现为鼓励个体作出法律所要求和期望的行为,最终形成理想的法律秩序。^②法律激励存在于法律关系形成、变更和消灭的整个过程中,既包括奖励性的正向激励和惩罚性的反向激励,也涵盖经济学中的激励理论,^③具有权利义务规范内容与激励制度设计方法两层含义。传统法学领域研究以法解释学为核心,较少直接关注法律激励问题。如今法学界关于激励问题的探讨主要从权利、义务和责任视角出发,侧重于研究实施某种正外部性行为后取得的收益及可能产生的负外部性影响,^④希望通过制度激励促进实现立法目标。其研究范围多为交叉学科,覆盖政府治理、食品安全、知识产权、个人信息保护等诸多领域。

个人利益与社会利益常处于相互对立的状态,法律作为社会制度的一部分,不仅应发挥其强制性,调整和保证社会财富的公平分配,还应通过非强制性的法律激励,保证个人利益与社会利益的一致性,使每个人都有遵守激励制度的积极性。法律的首要目的即通过提供一种激励机制,引导个体作出从社会角度来看最优的行动。法律的激励功能体现为经济学中追求自身利益最大化的行为,激励功能的强弱取决于当事人接受某项法律时机成本的大小。^⑤激励规制与法律激励虽然基于不同学科视角,但是在“尽可能实现社会效益最大化并作出公平分配”的最终目的上殊途同归,前者更关注社会效益增加,后者更注重效益公平分配。^⑥法律制定后的实施并非理所当然,形式上有效的法律仍然有很高的非执行比例。以往法律实施上的“强化惩罚论”在提高法律实施效果上收效甚微,通过将激励规制理论应用于法律领域,可以更好地解释法律监管规则实施成本高昂、法律责任规则威慑作用有限等问题,促使法律实施行为为外部性内部化,有助于解决信息不对称背景下“委托—代理”关系中的道德风险和逆向选择问题,构造激励相容的法律实施机制,将被规制主体行为调整至与立法目标一致。^⑦

① 参见黄新华、赵荷花:《制度型开放中政府规制变革的动因、挑战与路径》,载《北京社会科学》2022年第3期,第119页。

② 参见付子堂:《法律的行为激励功能论析》,载《法律科学》1999年第6期,第22页。

③ 参见丰霏:《法律治理中的激励模式》,载《法制与社会发展》2012年第2期,第151页。

④ 参见胡元聪:《人工智能产品发展风险抗辩制度的理论冲突与平衡协调》,载《中南大学学报(社会科学版)》2020年第6期,第75页。

⑤ 参见冯玉军:《经济立法的效益——功能分析》,载《兰州学刊》1997年第6期,第36页。

⑥ 参见陈彩虹:《法律:一种激励机制》,载《书屋》2005年第5期,第58页。

⑦ 参见董淳铎:《法律实施激励机制的基本原理及立法构造》,载《法学》2023年第9期,第158页。

第二,激励规制与人工智能。新规制经济学发展至今已有三十余年,在此期间,国内学者探讨了将激励规制理论应用于商业银行、农地配置、政府治理等领域的可能,但对激励规制理论与人工智能的结合研究较少。理性经济人假设是西方主流经济学的逻辑起点,即每个人在进行经济活动时都会追求自身利益最大化。激励规制理论秉承着“堵不如疏”的原则,在承认被规制主体自利行为的同时,通过激励相容的制度设计对其进行积极引导。以经济利益为核心的激励机制乃提升数字经济活力、推动数字技术发展的重要助力,作为支撑数字货币交易的底层技术,激励机制是引领区块链健康运行和相关应用良性发展的核心,^①不论是比特币的工作量证明机制,还是 Filecoin 项目的时空证明机制,^②其中都包含了分布式共识的经济激励机制,通过构建激励相容的开放式环境,无需第三方机构介入即可实现资产确权、交易和转移。

人工智能领域同样渗透着激励规制的制度设计。首先,人工智能技术呈现出双重信息不对称性,正契合激励规制理论信息不对称的前提。其中既包括开发者与应用者之间的不对称,又包括应用者与用户之间的不对称,从而引致诸多规制困境。^③从“开发者—应用者”角度来看,OpenAI、Google 等大型科技公司在面对政府部门时具有明显的信息不对称优势,通常居于“委托—代理”关系中的“代理人”位置,缺乏信息自主权的政府居于“委托人”位置。从“应用者—用户”^④角度来看,ChatGPT 等人工智能模型使用前的用户协议或隐私政策中往往包含“一揽子同意”条款,用户若不同意其个人信息和对话数据被收集、使用,将无法使用相关模型。用户被迫承担的信息损失和应用者利用信息优势获取的商业收益间亦存在高度不对称性。在巨大的经济利益驱使之下,不完全竞争可能会加剧,对市场经济下的公平和效率产生负面影响。其次,生成式对抗网络(Generative Adversarial Networks,简称 GAN)作为生成式人工智能的底层支撑技术之一,亦深度契合激励相容的理念,其本质上即在求解一个多方博弈的纳什均衡问题。GAN 主要由生成网络和判别网络组成,整个训练过程中贯彻着博弈论中的零和博弈思想,博弈双方的利益之和是一个常数,二者通过互相对抗、共同学习,交替提升各自的性能,希望最终将网络优化至一个确定的均衡状态,即纳什均衡点。最后,强化学习算法中的启发式奖赏函数也蕴含着激励机制设计的理念。英文中 agent 意为代理人,既包括授权为他人或代替他人行事的人,也包括自动执行某些任务的计算机应用程序。若暂不考虑主体性因素的影响,人类和人工智能之间亦可以视为一种新型的“委托—代理”关系。强化学习的核心原理即基于奖赏函数的试错,通过智能体(agent)^⑤与环境的连续交互以确定最佳行为策略,实现帕累托最优。当智能体根据当前状态作出行动后,会对环境产生影响并生成一个新的状态,环境会给予智能体相应的正面或负面回馈,然后进入下一个循环。强化学习的最终目标是尽可能多地获得正面奖励、减少负面惩罚,最大化累计回报。通过和变分自编码器、循环神经网络、Transformer 模型等技术相结合,生成式人工智能大模型表现出更接近于人类的认知、理解和思考方式,在计算机视觉、自然语言处理等领域取得了显著突破。

三、人工智能激励型立法的提出与证立

2017 年印发的《新一代人工智能发展规划》确立了人工智能立法和产业化的“三步走”战略目标:到 2020 年,人工智能产业成为新的重要经济增长点,部分领域的人工智能法规初步建立;到 2025 年,人工智能成为带动产业升级和经济转型的主要动力,初步建立人工智能法律法规体系;到 2030 年,人工智能产业竞争力达到国际领先水平,建成更加完善的人工智能法律法规体系。由此观之,在法治轨道上促进人工智能产业健康有序发展将是今后一段时间内的工作重点。人工智能法作为基本法,于中国人工智能发展和治理中起着固根本、稳预期、利长远的核心作用。在技术路线并未确定、产业发展无法预见的情况下,短期内出台人工智能法的可行性和必要性还有待进一步论证。今后应结合中国人工智能技术发展现状,以现实问题为导向,妥善回应产业发展需求,循序渐进地推动人工智能立法工作发展。概言之,人工智能立法是一个复杂且漫长

① 参见胡元聪:《区块链技术激励机制的制度价值考察》,载《现代法学》2021 年第 2 期,第 137 页。

② 参见王晓丽、严驰:《IPFS 未来规制进路探寻——基于相关裁判文书的分析》,载《电子科技大学学报(社科版)》2023 年第 5 期,第 101 页。

③ 参见杨建:《认真对待个人信息保护的社会结构:AI 科技时代的不对称性挑战》,载《环球法律评论》2022 年第 2 期,第 28 页。

④ 不管是面对人工智能技术的开发者还是应用者,用户都处于相对弱势地位。开发者亦可成为应用者,在此处不作明确区分。

⑤ 人工智能中 agent 的中文译名尚未形成共识,通常被译为“行为体”。参见寿步:《人工智能术语 agent 的精准译解及其哲学意义》,载《哲学分析》2023 年第 3 期,第 130 页。

的过程,现阶段关于人工智能法具体规则的探讨很难面面俱到,当务之急是阐释人工智能法的整体规制思路,明确人工智能法应采取的立法模式。

(一)传统立法模式的批判与扬弃

“立法模式”的概念有广义和狭义之分,广义的立法模式是指国家创制法律的惯常套路、基本体制和运作程式,包括立法主体、立法程序、立法权限等内容。狭义的立法模式则是指法律的外部表现形式,包括一国立法所采取的方法、结构、体例及形态等,下文主要聚焦于狭义的立法模式。陈昶屹教授指出,当代中国立法模式的演进历程大致可分为设范型立法、管理型立法和促进型立法三个阶段。^① 广义上来看,一切立法均属于国家干预的范畴。^② 由于国家干预通常由政府具体实施,因此又被称为政府干预。政府干预在经济领域体现为宏观经济领域的调控及微观经济领域的规制,并逐渐发展出激励、引导、促进等措施。^③ 人工智能具有技术属性和社会属性的双重面向,人工智能立法须及时回应技术迭代升级和产业快速增长的需求。^④ 人工智能治理是一项复杂的系统工程,其发展过程中涉及诸多权利义务关系,并伴随着算法歧视、数据泄露、版权侵权等技术风险。在探讨中国人工智能法的立法模式时,应重视产业迭代带来社会关系的不确定性,贯彻发展与安全并重的宗旨,考虑人工智能产业发展与社会治理的现实需求,极端的经济自由主义或国家干预主义均不可取。人工智能法需要具有高度国际竞争视野,在立足本土实践的同时汲取域外经验,以最大程度地实现激励相容。

设范型立法和管理型立法以义务性规范和禁止性规范为主,规定了严格的法律责任。近年出台的《互联网信息服务深度合成管理规定》《生成式人工智能服务管理暂行办法》等,都属于具有国家强制力的“硬法”。在信息不对称、目标不一致的背景下,简单的令行禁止可能会引发不同激励机制间的冲突,挤压市场机制发挥作用的空间,使法律规范难以得到有效实施和执行。促进型立法作为对市场失灵进行补正的新型立法模式,采用了大量授权性规范和任意性规范,更契合人工智能产业发展需求,近年来,人工智能治理实践中亦出现了《上海市促进人工智能产业发展条例》《深圳经济特区人工智能产业促进条例》等促进型地方性法规。但是,促进型立法的实施效果亦引起了诸多质疑。就促进型法律规范的具体规范类型来看,软法性质的指导性规范、鼓励性规范较多,硬法性质的强制性规范较少,^⑤ 总体上表现为一种概括性制度宣示。批评观点认为,促进型法律规范内容较为模糊,没有规定具体的行为模式和明确的法律责任,缺乏实践可操作性。^⑥ 法律的生命力在于实施,为实现人工智能发展与安全的平衡,人工智能法中需包含紧密相连的风险规制和产业促进两组规范类型,^⑦ 若缺乏刚性约束、仅依靠软法治理难免会流于形式,立法的激励效果与保障力度将大打折扣,可能会导致不确定和不可预见的法律风险。如上所述,今后人工智能法应在传统立法模式的基础上有所扬弃和创新,通过引入激励规制理论,实现软法引导与硬法规制的有效结合。^⑧

(二)激励型立法的理论证立

激励型立法是指对“人”的特定行为实施激励的法律。中国立法实践中至今尚未出现名称含有“激励”的法律,尽管域外诸如美国《拜杜法》(Bayh-Dole Act)等激励型立法普遍存在,但激励法作为按照法律基本功能划分的重要类别之一,却鲜少获得关注,学界关于激励法的研究深度和广度亦较为有限。

1.人工智能激励型立法之应然性

只有准确反映和体现所调整社会关系客观规律的法律,才能经得起实践和历史的检验。以往的立法更多体现出管理甚至管制的思想,过分强调政府权力,对市场主体的权利关注不够。^⑨ 当下全球人工智能正处于加速发展期,人工智能产业化进程明显提速。由于构建先进的深度学习模型和人工智能系统需要大量数

① 参见陈昶屹:《论“促进型立法”的形成背景》,载《北京行政学院学报》2005年第1期,第67页。

② 参见张乃根:《论市场经济的立法取向》,载《法学》1993年第5期,第3页。

③ 参见卢代富:《经济法中的国家干预解读》,载《现代法学》2019年第4期,第121页。

④ 参见张凌寒:《中国需要一部怎样的〈人工智能法〉?——中国人工智能立法的基本逻辑与制度架构》,载《法律科学》2024年第3期,第5页。

⑤ 参见焦海涛:《论“促进型”经济法的功能与结构》,载《政治与法律》2009年第8期,第81页。

⑥ 参见傅爱竹:《数字新兴议题专门立法热之反思》,载《法商研究》2023年第5期,第50页。

⑦ 参见郑志峰:《人工智能立法的一般范畴》,载《数字法治》2023年第6期,第55页。

⑧ 参见严驰:《论人工智能的激励型立法——基于〈拜杜法〉的思考》,载《人工智能》2024年第2期,第91页。

⑨ 参见张守文:《政府与市场关系的法律调整》,载《中国法学》2014年第5期,第69页。

据、算力和资金的支持,如今产业界的人工智能发展步伐已远超学术界和非营利组织。在信息不对称的背景下,仅依靠作为被规制主体的人工智能企业的内部自律规范来建立行之有效的激励约束机制显然并不现实。根据政府规制俘获理论,在经济利益驱使下,规制主体可能会为被规制主体所支配和控制,制定出不是为实现公共利益最大化而是为少数人利益服务的规制方案。因此,有必要对规制主体的权利进行适当的限制。在公共管理领域,西方国家已开始探索通过法律助力由管制型政府向服务型政府的定位转变,化刚性管理为柔性管理。^① 根据“守夜人”理论,政府无须对经济运行作过多干预,只需维护和保障市场安全稳定,并对不足之处加以补救即可。中国未来发展道路的选择不仅要符合现实国情,也须与国际上政府改革的普遍经验相契合,通过持续深化简政放权,保护市场主体的经济自由。激励规制理论对于中国人工智能立法模式选择具有重要的参考价值,人工智能法在借鉴外来制度的同时,也需要扎根本土实践,体现中国本土特色的制度属性,回应保障和激励人工智能产业高质量发展的现实需求。

公共领域的软硬法混合治理是现阶段普遍存在且应当采取和推广的规范模式。人工智能技术具有复杂性、动态性、隐性、不可预见性等特征,^②在进行人工智能领域立法时,应保持公权力的审慎性和谦抑性。传统自上而下的回应性监管和主要基于硬法规范的刚性治理无法实现与人工智能技术的同频更新,可能会产生阻滞创新或强化风险的负面影响。中国式现代化背景下的人工智能法采用激励规制方式而非命令控制方式,有助于继承和发挥促进型立法中的积极因素,明确激励人工智能产业发展的主体思想。通过软硬法混合治理,在软法规范的基础上加入一定比例的管理、监管与惩戒内容,更有利于实现“约束公共权力、保障人民权益”的法治基本价值。从规制经济学的发展可以看出,规制与竞争的关系逐渐从二元对立转向了互补融合。激励规制理论看似是国家干预主义的体现,本质上仍是在新古典经济学框架下的展开,反对凯恩斯主义下政府职能的过度扩张,最终目的是实现个人价值与社会价值相统一,和人工智能法之间存在高度的理论契合性。为妥善处理内在激励机制与强制性外部要求间的冲突,弥补理性经济人决策逻辑在具体情景下的不足,贯彻激励相容的规制理念,人工智能法采用激励型立法模式具有逻辑上的应然性。

2. 人工智能激励型立法之必然性

激励规制的思想本质是基于公共利益理论,可以将其主旨解读为调整市场行为所带来的经济无效或社会不公,在追求经济效率的同时实现社会效率。法律与利益之间存在天然联系,利益是法律规范产生的根本动因,^③人工智能产业发展过程中不可避免地会涉及利益分配问题。激励法具有鲜明的利益属性,充分发挥了法律的“权利—利益”调节机制,不仅会进行利益衡量和价值评估,还能通过设计合理的激励制度以平衡和协调不同群体间的利益冲突。在探索人工智能法的立法模式时,强政府干预性质的设范型立法和管理型立法可能会给人工智能产业发展套上沉重的枷锁,促进型立法相较而言更契合建设公共服务型政府所需。激励型立法是在促进型立法基础上的进一步变革和调整。一方面,其与促进型立法一样,本质上是激励性政策、规范性文件等的法律化,顺应了社会经济发展需求,在更接地气的同时提高了稳定性、可预期性和政府承诺能力。另一方面,激励型立法相较于促进型立法更为明确、具体,加入了更多涉及权利、义务、责任的规范,具备更强的可实施性和可操作性,更符合促进人工智能安全发展的要求。再者,激励法立足于激励契约即合同的基础之上,具有鲜明的契约性,而激励契约的形成需要充分尊重要约人和承诺人双方的意愿,这样才能建立绝大多数社会成员认同的激励制度和运行机制,提高激励相容度。激励型立法并非简单地尊重契约自由即可,而是需要同时保障社会公共利益,对被规制主体作出必要限制,综合考量规制主体和被规制主体背后的利益权衡和博弈情况。激励法符合人类本性中向上、向善的一面,因此其运行阻力最小,也最容易受到欢迎。^④ 对人工智能企业等被规制主体而言,激励法表现为正面的激励效果,不会直接强制其接受或放弃激励。为保护被规制主体的利益,激励法中还包含对规制主体不实施激励行为或激励行为实施不当时应采取的强制性措施,更好地保护了个人利益乃至社会公共利益。

人工智能是中国当下发展最快、创新最活跃的重点领域和新兴领域之一。《2025 人工智能发展白皮书》

① 参见倪正茂:《加强激励法学研究 完善法治国家建设》,载《上海大学学报(社会科学版)》2014年第2期,第93页。

② 参见严驰:《生成式AI大模型的全球治理方案:何以可能与何以可为》,载《科技与法律(中英文)》2024年第1期,第92页。

③ 参见吕世伦、孙文凯:《赫克的利益法学》,载《求是学刊》2000年第6期,第62页。

④ 参见倪正茂:《激励法学与法制发展》,载《人民法院报》2012年10月31日,第5版。

显示,中国已经形成了相对完整的人工智能产业链条。截至2024年底,中国人工智能相关企业数量达到13 069家,排名世界第二,仅次于美国。尽管人工智能商业化应用的场景正不断丰富,但是现阶段中国人工智能产业尚存在发展不平衡、不充分的问题,从事人工智能基础设施和核心技术研究的企业数量较少,一定程度上阻碍了中国人工智能产业向全球高端产业链、价值链的跃升。人工智能激励型立法作为一种软硬结合的创新规制理论和科学制度设计,是破解产业发展难题、持续赋能新质生产力的必然选择。人工智能激励型立法所体现出的中国的法律多元主义和以产业需求为导向的发展理念,能够更好地扶植产业上游的基础层人工智能企业发展、提升核心技术层人工智能企业的市场竞争力、拓展应用层人工智能企业与各领域的深度融合,积极壮大和培育中国人工智能产业的整体创新能力和核心技术竞争力。

四、人工智能激励型立法方案的延伸探讨

(一) 激励方式:减少成本型和增加收益型激励

法律激励规范具有奖惩一体性,今后人工智能法可以考虑采取的激励方式不仅包括授予上线许可、允许扩大运营范围等正面激励,也包括强制下线、删除算法等负面惩罚。中国法律中现有的激励方式可以概括为三大类:收益、成本配置;权利、义务、责任分配;荣誉、待遇给予。其中,荣誉、待遇给予类激励多面向特殊主体,如《中华人民共和国科学技术进步法》(简称《科学技术进步法》)规定的提高工资水平、给予优厚待遇和荣誉激励等。下文重点讨论前两类激励方式,其中,收益、成本配置类激励又可分为增加收益型激励和减少成本型激励,具体情形包括《科学技术进步法》规定的通过“设立自然科学基金”进行经费资助,及《中华人民共和国企业所得税法》规定的对小型微利企业和高新技术企业以优惠比例征收企业所得税等。权利、义务、责任分配类激励则可分为赋予权利型激励、减免义务型激励和减免责任型激励。例如,《中华人民共和国专利法》就是通过赋予专利发明者垄断性专有权的方式实现激励;减免义务型激励和减免责任型激励则体现为《中华人民共和国劳动合同法》的“不得要求劳动者提供担保”、《中华人民共和国刑法》中关于自首、立功的规定等。

规制影响评价是提升规制质量的重要路径,既包括采用统计学、计量经济学等方法的定量评价,也包括从宏观角度和制度层面进行的定性评价。鉴于定量评价对数据依赖度过高、适用范围有限,在人工智能法正式实施前,定性分析更有助于从整体上把握法律规制给被规制者和社会整体利益所带来的影响,提升规制方案的科学性和有效性。人工智能法中激励规范的设计和激励方式的选择将直接影响法律的实施效果。为此,可以引入维克托·弗鲁姆(Victor H. Vroom)在其著作《工作与激励》中提出的期望理论,通过确定影响激励效果的因素,及时调整激励规制方案,推动人工智能法中设计的激励机制达到激励相容的理想状态。期望理论用公式表达,即激励取决于行动结果的价值评价和与其相对应的期望值二者之乘积。其中,价值评价是指某项结果对于满足个体期望和需求的价值,数值介于-1和1之间,期望值则是对某种激励效能的预测,数值介于0和1之间,相应的激励范围也位于-1和1之间。为提升人工智能法激励效果,一方面,要尽可能地满足合理范围内的需求,不同个体或群体的需求内容和程度有所不同,当法律中的激励措施确能促进各类需求实现时,就会提升人们对法律的整体评价。只有法律令人满意、被普遍接受时,才会表现为正效价。另一方面,要提高激励机制和奖励措施的客观可行性,让人们相信预期目标可以实现,并且在达到预期目标后能得到相应奖励,尽可能地使期望值接近于1,从而使人工智能法产生更好的激励效果。

根据期望理论,为提高社会公众对于激励规制措施的期望值和信任度,强化法律规范的激励力量,人工智能法中宜采取的激励方式应是减少成本型激励和增加收益型激励。二者的共同特点是主要针对一般法律主体,规定内容较为简单且激励效果明显。具言之,人工智能法构建时应考虑纳入税收优惠、费用减免、设立专项资金、财政经费补助等直接助力人工智能产业发展的规定。生成式人工智能的重要特征之一是资本投入的规模收益递增,在高投资和市场的相互刺激下,模型的技术进步可以在不断扩大的投资 and 市场中螺旋上升。^① 人工智能产业研发周期长、投资风险大,人工智能技术的应用和创新需要大量资金支持才能产生规模效应和产业集群。实践中各地政府正在持续加大资金投入力度,例如,北京市拟在人工智能领域设立百亿级

^① 参见李寅:《智能革命还是“规模”革命?——历史视野下的人工智能技术与产业演化》,载《文化纵横》2024年第2期,第27页。

的政府投资基金,以更好地发挥政府资金的产业助推作用。为保障激励方式的可行性,推动人工智能激励型立法提质增效,今后应树立开放创新意识,制定实施更加积极的法律法规和财政支持政策,基于经济法的视域,中立、审慎地应对人工智能发展中产生的负外部性影响对现有社会经济秩序之冲击。^①

(二)激励机制:知识产权制度之外的创新激励思路

知识产权作为人们对其智力创造所享有的专有权利,是基于创造性智力成果和工商业标记依法产生的权利的统称。人工智能法作为国家层面的基础性法律,须观察技术和产业的走向,及时回应社会关切,妥善处理其与知识产权法尤其是著作权法的关系问题,考虑如何与现有法律规定相衔接。同时,知识产权制度与新质生产力也有着天然的联系。创新是引领发展的第一动力,保护知识产权就是保护创新。新质生产力具有突出的创新驱动特性,加强知识产权保护是发展新质生产力的内在要求和重要保障。长久以来,人们习惯于将知识产权制度的本质解释为一种激励机制。^② 知识产权制度具有内在的利益平衡机制,实现激励知识创造和对知识产品需求的社会利益间的平衡是知识产权制度的最终价值旨归。在法社会学语境下,经济功能是知识产权法律制度的核心功能。从理论视角来看,激励机制学说是经济分析方法运用在产权有关问题上的结果。知识产权制度体系的内容除了激励个体创新之外,还包括保护完整市场。正如波斯纳(Richard A. Posner)法官所言,传统知识产权领域中经济分析的焦点在于激励和接触间的交换。^③ 从市场角度而言,赋予知识财产创造者垄断性权利的目的在于给广大社会公众提供接触的可能,以享受知识财产更新带来的便利和效益,更好地平衡个人激励与社会发展。

人工智能时代知识产权的内涵和外延都得到了拓展,生成式人工智能大模型输出的人工智能生成内容对著作权、专利、商标等传统知识产权领域产生了极大的冲击,立足于机械功利主义的知识产权激励理论不管是在激励对象还是激励效果上,都越发显得难以为继。^④ 对此,李石教授直言,知识产权激励说正面临着“囚徒困境”,提出基金激励机制是一种不同于知识产权制度的创新保护方案。^⑤ 21世纪伊始提出的健康影响力基金计划对全球药品研发和推广工作产生了积极而广泛的影响,其以影响力为核心、重视社会理性的基金激励机制设计,不同于传统金融领域追求经济理性的基金激励机制,在最大限度地促进创新和传播的同时,更好地维系了个人利益与公共福利的平衡。健康影响力基金有以下三个特征:第一,基于政府、慈善基金会或国际税收的资助运行;第二,向健康影响力基金注册新药品创新者须放弃申请专利,注册后的创新者将在奖励期内每年获得一定报酬,健康影响力基金会设立专门的评估委员会进行“全球健康影响评估”,注册者获得的报酬份额与注册药品的健康影响力在所有注册药品的全球健康贡献中所占份额相对应;第三,药品注册者应允许其注册药品以不超过最低可行生产和销售分配成本的价格在全球范围内销售,并在奖励期结束后提供免费许可,以实现产品的批量制造和销售。^⑥ 健康影响力基金可以被视为一项全面的预先市场承诺,在其激励机制下,新药研发者的经济回报直接取决于药品销量和健康水平促进程度。健康影响力基金提供了和药品专利制度不同的激励制度选择,新药产品可能同时适用于两种激励机制或是只能择一适用。健康影响力基金不会对药品专利制度构成直接威胁,制药公司可以从开发治疗原本被忽视疾病的药物中获得新收入来源,制药创新者则可以根据自身情况综合考量竞争情况、预期收益等因素在二者中进行抉择。概言之,以健康影响力基金为代表的基金激励机制是一种兼具经济可行性和人道主义精神的创新激励路径,符合加速发展新质生产力的时代要求,今后可以考虑将其应用于更广泛的知识创新领域。

以健康影响力基金为代表的基金激励机制在中国尚处于萌芽阶段,其落地应用面临着一定阻碍。如基金激励机制的运转主要依靠政府出资,可能会因资金投入大、建设周期长、收益见效慢而无法获得足够的政策支持。此外,如何建立科学合理的影响力评估方案以证明和衡量经注册知识产权客体的影响力及贡献所占份额仍有待明确。基金激励机制在有效破除上述阻碍前,还无法成为学者所构想的“知识产权制度的替

① 参见胡元聪、廖娟:《人工智能的负外部性及其经济法规制》,载《大连理工大学学报(社会科学版)》2020年第3期,第71页。

② 参见杨明:《知识产权制度与知识财产创造者的行为选择》,载《中外法学》2012年第4期,第742页。

③ Richard A. Posner, *Intellectual Property: The Law and Economics Approach*, Journal of Economic Perspectives, Vol.19: 57, p.58(2005).

④ 参见林秀芹、郭壬癸:《人工智能对知识产权正当性理论的挑战与应对》,载《知识产权》2023年第11期,第92页。

⑤ 参见李石:《“知识产权制度”的哲学反思》,载《哲学研究》2019年第8期,第120页。

⑥ Amitava Banerjee, Aidan Hollis & Thomas Pogge, *The Health Impact Fund: Incentives for Improving Access to Medicines*, Lancet, Vol.375:166, p.168(2010).

代方案”。^①以新促质是新质生产力的核心要义。知识产权是随着社会发展、应社会需求而创设的法定权利,为实现新质生产力的培育、壮大和提升,将基金激励机制引入知识产权领域将会是一次有益尝试。基金激励机制提出了别具匠心的创新融资激励思路,在激励创新者追逐个人利益的同时,亦能促进社会公共福利持续增长,有助于营造勇于创新、尊重创新的人工智能产业发展环境。新质生产力与中国的创新驱动战略、建设创新型国家目标相辅相成。中国正处于向创新型国家前列迈进的关键时期,今后在构建人工智能立法方案时,应高度重视和发挥知识产权制度激发创新动力之功能,持续深入探索现有激励理论和基金激励机制之间双向衔接、深度协同的可能。^②

五、结语

当下飞速发展的人工智能技术正衍生出层出不穷的新模式、新场景和新业态,激发了新兴产业崛起及传统产业转型升级的蓬勃活力。生产力具有鲜明的时代特征,立足于人工智能时代,发展新质生产力是实现中国式现代化的必由之路。加快形成新质生产力需要发挥中国的独特优势,大力推进人工智能技术产业化。今后须更好地落实总体国家安全观,以制度创新为核心,建立健全中国特色人工智能治理体系,在法治道路上统筹实现人工智能发展和安全之间的最佳平衡。新规制经济学引入了信息不对称因素和激励相容的理念,克服了传统规制经济学理论的局限性,对中国探索制定最优的激励规制规则、营造公平有序的人工智能产业发展环境具有重要的借鉴意义。在国家层面的人工智能法中采用激励型立法模式具有可能性和必要性。有别于传统的设范型立法、管理型立法及促进型立法,激励型立法不仅体现为立法模式的革新,也是刚柔并济治理理念及软硬法结合治理方式的体现,深入贯彻了激励相容的规制理念,在维护人工智能企业利益的同时,亦能有效促进社会公共利益的实现,更契合中国人工智能产业发展新阶段的战略需求,有助于加快形成新质生产力,增强发展新动能,是未来人工智能法构建时的题中应有之义。

Consideration of Incentive Legislation of Artificial Intelligence: From the Perspective of Incentive Regulation Theory

YUAN Xiuting, YAN Chi

(Law School, Tongji University, Shanghai 200092, China)

Abstract: Scientific and technological innovation and industrial innovation led by artificial intelligence are the key to accelerating the cultivation and formation of new quality productive forces. The incentive regulation theory effectively overcomes the limitations of the traditional regulation economics theory against the background of asymmetric information, and provides a reference for constructing the legal system to help artificial intelligence technology and industrial innovation development. At the theoretical level, incentive regulation and legal incentives are intrinsically compatible with artificial intelligence. As the basic law in the field of artificial intelligence governance in China, artificial intelligence law should be based on the actual conditions of the country, introduce incentive regulation theory, sublate and innovate on the basis of the traditional legislative model, consider the adoption of incentive legislation, choose the incentive mode of reducing costs and increasing income, and take the fund incentive mechanism as a reference point, maintaining the interests of artificial intelligence enterprises and promoting the realization of social public interests.

Key words: new regulatory economics; incentive regulation; artificial intelligence law; incentive legislation; fund incentive mechanism

^① 参见李石:《基金激励机制:知识产权制度的替代方案》,载《读书》2020年第1期,第142页。

^② 参见严驰、沈汪成:《知识产权制度赋能新质生产力发展的优化路径——以“基金激励机制”为视角》,载《电子知识产权》2024年第12期,第4页。