



西安交通大学学报(社会科学版)

Journal of Xi'an Jiaotong University(Social Sciences)

ISSN 1008-245X,CN 61-1329/C

## 《西安交通大学学报(社会科学版)》网络首发论文

题目： 我国人工智能立法的模式选择与制度展开——兼论领域融贯型立法模式  
作者： 胡铭，洪涛  
收稿日期： 2024-03-26  
网络首发日期： 2024-05-10  
引用格式： 胡铭，洪涛. 我国人工智能立法的模式选择与制度展开——兼论领域融贯型立法模式[J/OL]. 西安交通大学学报(社会科学版).  
<https://link.cnki.net/urlid/61.1329.c.20240510.1430.002>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

# 我国人工智能立法的模式选择与制度展开

## ——兼论领域融贯型立法模式

胡 铭, 洪 涛

浙江大学 光华法学院, 浙江 杭州 310008

**摘要:** 中国人工智能立法已经纳入《国务院 2023 年度立法工作计划》，当前，研究重心应从“应否立法”转向“如何立法”。比较国际人工智能立法实践，可归纳出两种基本模式，即统一垂直型立法模式和分散网格型立法模式，两者在制度设计、价值取向等方面各具特征。在分析研判这两种模式的基础上，以法律 3.0 为指导，提出中国应探索领域融贯型立法。该模式强调人工智能立法系领域法，主要内容应包括风险规制、产业促进、技术标准三方面，形式上要以基本法为主体、产业促进法和技术标准法为分支，旨在实现立法宗旨、人工智能面相及法治系统工程的重叠融贯。在该模式下，《人工智能法》（基本法）应当对基本规定、基本原则、基本制度及责任义务体系等内容分别作统一界定、吸纳转介与制度重塑。

**关键词:** 人工智能立法；智能治理；统一垂直型立法；分散网格型立法；领域融贯型立法

为确保人工智能始终向善发展且致力于增进全人类的福祉，全球智能竞争逐渐从技术之争发展到规则之争、治理之争。2021 年 4 月，欧盟委员会发布《欧洲议会和理事会关于制定人工智能统一规则的立法提案》（以下简称《AI 法案》），意图继续维持规范的布鲁塞尔效应<sup>[1]</sup>；2022 年 10 月，美国发布《人工智能权利法案蓝图》（以下简称《权利法案蓝图》），尝试为人工智能的设计、使用和部署提供指导<sup>[2]</sup>。2023 年 6 月，中国人工智能立法纳入《国务院 2023 年度立法工作计划》，预备提请全国人民代表大会常务委员会审议。然而，《人工智能法》并未列入全国人民代表大会常务委员会未来 5 年的立法规划，这似乎给立法的必要性与正当性打上了问号。但是，该立法规划在第一类立法项目（条件比较成熟、任期内拟提请审议）的部署中明确指出，“对推进科技创新和人工智能健康发展，完善涉外法律体系等，要求制定、修改、废止、解释相关法律，或者需要由全国人大及其常委会作出相关决定的，适时安排审议”，足以说明全国人民代表大会常务委员会对于人工智能立法本身的必要性并无异议。2024 年全国两会期间，多位政协委员递交提案，建议尽快推进人工智能立法，弥补监管空白<sup>[3]</sup>。中国现有智能立法、伦理规范和原则的效力级别较低，智能治理整体呈现出部门法色彩和具象化特征<sup>[4]</sup>，聚焦人工智能法律定位、个人信息保护、平台杀熟等方面，在实践中易产生监管竞合或监管盲区。事实上，正是宏观上的人工智能立法缺位，才使得现有规范内容交叉重叠、适用顺序混乱，“九龙治水”乱象难以避免。倘若采用科学的立法模式，便可在其指引下梳理现有伦理规范和原则，同时有计划地制定效力

【收稿日期】2024-03-26。 【修回日期】2024-04-10。

【基金项目】教育部哲学社会科学重大专项（2023JZDZ012）。

【作者简介】胡铭（1978—），男，浙江大学光华法学院院长，浙江大学数字法治实验室教授，博士生导师；洪涛（1998—），男，通信作者，浙江大学光华法学院博士研究生。

较高的法律、行政法规，加快打造智能治理的中国方案和话语体系。当前，研究重心需要从“应否立法”转向“如何立法”，立法模式的选择成为首要问题。鉴于此，本文拟在分析各国人工智能立法现有模式的特点与优劣基础上，探索基于中国实践的人工智能立法模式，并进一步探讨该模式下的制度展开，尝试为中国人工智能立法提供理论支持和可行方案。

## 一、人工智能立法的两种模式及评析

通过立法规制人工智能已是全球共识。只有采用符合本国国情的立法模式，才可能实现技术向善、规范运行的目的。通常来讲，广义的立法模式指由国家创制法律的惯常套路、基本体制和运作程式等要素构成的整体，包括立法权限、立法程序等<sup>[5]34</sup>。狭义的立法模式指立法最终表现出来的外部表现形式，属于法的外部结构技术，如制定法或判例法等<sup>[6]342</sup>。本文聚焦广义的立法模式，将其理解为一国在一定时间内实施法治的综合设计。纵观各国的人工智能立法实践，总体上可归纳为统一垂直型立法模式和分散网格型立法模式两种。

### （一）统一垂直型立法模式

统一垂直型立法模式是指国家制定一部专门的人工智能法案或法典，通过对智能定义、监管机构等内容作统一规定，确保规范体系内部一致，对人工智能进行垂直性和体系性治理。统一垂直型立法模式的代表是欧盟和加拿大。2021年4月，欧盟委员会发布全球首部针对人工智能的法案，即《AI法案》，开启了人工智能治理的“硬法”道路。2022年6月，加拿大发布了其国内首部人工智能立法，即《人工智能和数据法案》（Artificial Intelligence and Data Act, AIDA）。整体上看，该模式具有以下特点：

其一，对人工智能统一定义，以防概念模糊而导致法律适用偏差。《AI法案》提案最初将人工智能界定为软件，几乎涵盖当时已知的大部分人工智能形式。后经过多轮修改，目前仍保持着统一定义，即“基于机器的系统，旨在不同程度的自主运行，并能为明确或隐含的目标生成诸如预测、建议或决策等输出，以影响物理或虚拟环境”<sup>①</sup>。AIDA对人工智能的定义和经济合作与发展组织（OECD）一致，并与《AI法案》理念相同，其目的正是与其他司法管辖区的法律框架衔接适用<sup>②</sup>。

其二，赋予法案域内和域外效力，以确保其在一国内垂直适用，而不受服务提供者、部署者或营业场域所在地域的限制。举例而言，《AI法案》采取了属人、属地和实质判断的管辖标准，不仅适用于欧盟境内，还可规范位于欧盟之外但输出在欧盟境内的提供者或部署者。《AI法案》第2条规定：“本条例适用于：（a）在欧盟内将人工智能系统投放市场或投入使用的提供者，无论这些提供者是在欧盟内还是第三国设立；（b）在欧盟内设立场所或位于欧盟内的人工智能系统的部署者；（c）人工智能系统的提供者和部署者，或者是其营业场所位于第三国，而第三国因国际公法而适用成员国法律，或者是该系统产生的意图是在欧盟内使用。”<sup>③</sup>

其三，由专门机构统一负责监管执法，尽可能地保证执法专业性并避免执法不公。欧盟为此成立人工智能办公室，负责在成员国统一执行《AI法案》。《AI法案》第3条（42）规定：“人工智能办公室指委员会促进人工智能系统的实施、监测和监督以及人工智能治理的职能部门。”有议员指出，人工智能办公室是实现法案一致适用的关键要素，可为人工智能建立一个单一市场，避免成员国各自为政或

① 如无特殊说明，本文《AI法案》的条文款数均指2024年3月13日欧洲议会通过的官方最终版本。

② The artificial intelligence and data act [EB/OL].(2023-09-27)[2024-01-15].<https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act>.

③ Proposal for artificial intelligence act and amending certain union legislative acts[EB/OL].(2024-03-13)[2024-03-14].<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>.

专业性不足<sup>①</sup>。加拿大也建立了一个独立部门，即创新、科学和经济发展局（ISED）。AIDA 赋予 ISED 部长管理和实施该法中不涉及需经诉讼司法程序解决的事项之外的所有内容。

其四，明确规定相关主体违反法案要承担法律责任，将人工智能安全作为首要考量。《AI 法案》第 71 条规定了不同情形的处罚，如不遵守第 5 条提到的禁止性规定，将被处以最高 3 500 万欧元或企业上一财政年度全球年总营业额 7% 的罚款。AIDA 指出，人工智能的发展已经产生监管空白，必须通过立法填补这些空白来保持加拿大人民对于人工智能技术的信任，提出人工监督、透明、公正和公平等原则，配置了行政罚款、违法诉讼、刑罚三种处罚机制。

## （二）分散网格型立法模式

分散网格型立法模式指国家未制定统一的监管法案，而通过区分智能类型和应用场景形成不同网格节点的个别规范，对人工智能进行灵活性和针对性治理。分散网格型立法模式的代表国家是美国和英国。这些国家的技术水平处于世界前列，国内有多种类型的人工智能且在持续研发新类型，而不同类型的人工智能在监管上亦有不同要求，制定统一监管法案的难度较高。譬如，美国 OpenAI 公司于 2024 年 2 月 16 日发布的“文生视频”模型 Sora，可通过文本指令生成高质量视频，被称为物理世界模拟器<sup>②</sup>。Sora 与 ChatGPT 在著作权方面分别对标视听作品和文字作品，在认定新颖性、独创性时需要适用差异性标准。整体观之，该模式具有以下特点：

其一，中央层面无统一强制性的监管立法。以美国为例，联邦层面并没有像《AI 法案》那样的强制性法案，尽管存在人工智能战略、科技伦理及技术原则，但相关内容并无强制性约束力。例如，2021 年的《国家人工智能倡议法案》属于人工智能的战略布局，与监管尚有距离；2022 年的《权利法案蓝图》为原则性规定，提倡而不是要求服务提供者、部署者以权利保护为目标。同样，英国中央政府层面一方面尚未制定统一适用的人工智能法案，另一方面也没有设立专门监管人工智能的内阁部门。2023 年 3 月发布的《支持创新的人工智能监管方法》明确指出，“不会将跨部门原则最初置于法定基础上，在实施初期将继续与监管机构合作，鼓励其按比例原则优先考虑宽松的监管措施”<sup>[7]</sup>。

其二，对人工智能起到实际监管作用的规范分散在不同网格节点——或是原有的传统立法，或是针对某一场景的智能法案。在美国，分散的联邦机构和各州构成人工智能监管的“双重网格”。具体而言，联邦层面，联邦贸易委员会（FTC）、平等就业机会委员会（EEOC）和国家标准与技术研究院（NIST）等联邦机构通过适用现有法律监管人工智能，而不是制定和应用新的人工智能法。FTC 指出《联邦贸易委员会法》第 5 条“禁止不公平或欺诈行为”完全可用于人工智能，并基于该法对 OpenAI 公司进行了正式调查<sup>[8]</sup>；地方层面，伊利诺伊州颁布《人工智能视频面试法案》，规定企业使用人工智能面试应聘者时，应审查其是否满足告知同意、限制共享、删除权响应等合规义务<sup>③</sup>。《英国人工智能监管》报告

---

① TUDORACHE D. Report on the proposal for a regulation of Artificial Intelligence Act and amending certain Union Legislative Acts[EB/OL].(2023-05-22)[2024-01-18].[https://www.europarl.europa.eu/doceo/document/A-9-2023-0188\\_EN.html](https://www.europarl.europa.eu/doceo/document/A-9-2023-0188_EN.html).

② Video generation models as world simulators[EB/OL].(2024-02-15)[2024-03-12].<https://openai.com/research/video-generation-models-as-world-simulators>.

③ FRIEDLER S, VENKATASUBRAMANIAN S, ENGLER A. How California and other states are tackling AI legislation [EB/OL].(2023-03-22)[2024-04-26]. <https://www.brookings.edu/articles/how-california-and-other-states-are-tackling-ai-legislation/>.

显示，英国同样是立足于现有的、分散的监管机构和法律网络，构建了“部门为基础的监管框架”<sup>①</sup>。

其三，在智能定义、风险认定等方面未能统一，彼此衔接存在障碍。例如，美国不同州法案规定的管辖范围有所差异，加利福尼亚州的智能法案涵盖自动化系统的非政府用途，而康涅狄格州和佛蒙特州的智能法案则仅针对政府中自动化系统的应用。在监管机构方面，OpenAI 公司与微软主张组建一个新的联邦机构，谷歌公司更倾向于由 NIST 带头发布技术指导，再交各部门付诸实施<sup>②</sup>。英国《支持创新的人工智能监管方法》认为统一定义会阻碍创新，选择由监管机构结合技术“适应性”和“自主性”特征进行定义，并基于情景关联的方法评估风险<sup>③</sup>。

其四，监管人工智能的主要目的是促进创新，相关措施基本靠人工智能研发者、服务提供者自愿遵守，未配置明确的法律责任。例如，2021 年 NIST 发布的《人工智能风险管理框架》对人工智能风险作了分类分级并提出指导原则以及配套的实践指南，但并没有规定违反相关内容的法律责任；英国《支持创新的人工智能监管方法》在执行摘要部分写道，为避免阻碍人工智能创新，监管机构在应用原则时不必全部转化为强制性义务，而应利用特定领域的专业知识，并根据使用场景来调整原则的实施，经过最初阶段之后，依据监测情况再考虑是否引入法定责任。

### （三）两种模式的比较分析

统一垂直型立法模式的优势在于集中了国家力量，一方面能够推动规范体系内部协调一致，形成严密的治理法网，实现跨领域综合监管；另一方面，通过合规义务和法律责任进行垂直适用，有利于维护监管执法的公平公正，取得全球智能治理的规范优势。但同时，该模式具有以下局限。第一，人工智能内涵统一定义，在面对技术更迭时易因灵活性和针对性不足而导致规范效果“大打折扣”。《AI 法案》最初未涉及大模型，ChatGPT 的出现引发了后续法案的适应性修正，如第 52 条规定的额外透明度要求<sup>④</sup>。第二，基于风险评估的统一监管在处理不同类型的智能技术和应用场景时，面临标准设定困难及风险动态变化的问题。《AI 法案》主要基于预期用途判断风险等级，但在诸如生成式人工智能这类无固定预期目标的技术面前，难以直接套用既有风险框架，只得将“能被用于高风险系统或其组件的生成式人工智能”归入高风险。AIDA 同样采用了风险监管，重点关注“高风险系统”的义务设置，如风险监督检查、实质损害通知等，但亦因智能类型和应用场景的缘故，将何为“高风险系统”留待未来界定<sup>⑤</sup>。第三，法律责任归属和处罚类型有待完善。《AI 法案》尚未厘清损害责任的具体归属，仅规定了较为单一的行政处罚手段，反映出责任体系的不完善。AIDA 虽包含行政罚款、民事赔偿和刑罚等多种处罚形

---

① Regulating AI in the UK[EB/OL].(2023-07-18)[2024-03-11].<https://www.adalovelaceinstitute.org/report/regulating-ai-in-the-uk/>.

② LIMA C, DIMOLFETTA D. Google bucks calls for a new AI regulator[EB/OL].(2023-06-13)[2024-04-26].<https://www.washingtonpost.com/politics/2023/06/13/google-bucks-calls-new-ai-regulator/>.

③ A pro-innovation approach to AI regulation[EB/OL].(2023-03-29)[2024-03-12].<https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>.

④ 加拿大国际治理创新中心研究员 Scassa 指出，AIDA 的目的是监管高影响的人工智能系统，但未对“高影响”作出界定，严重影响了监管效果。此外，AIDA 对危害的定义是对个人造成的可量化的危害，这个定义过于狭隘，如人工智能系统性歧视问题的影响范围远超出个人和社群。参见 SCASSA T. Canada's draft AI legislation needs important revisions[EB/OL].(2023-08-16)[2024-03-13].<https://www.cigionline.org/articles/canadas-draft-ai-legislation-needs-important-revisions/>.

式，但在适用条件、衔接机制及执法主体（ISED）的公正性方面仍需细致的制度设计<sup>①</sup>。第四，该模式希望通过风险分类分级与监管沙盒的运用，尽可能地平衡安全与发展。但整体上看，安全考量还是占优先地位，可能抑制早期技术发展和本国产业的国际竞争力。正如西门子、雷诺等企业指出的，《AI 法案》使生成式人工智能受到严格监管，运营此类系统的企业将面临高昂的合规成本和不成比例的责任风险，可能迫使企业将运营活动转移到欧盟之外，造成欧洲竞争力和技术主权的削弱<sup>②</sup>。

分散网格型立法模式主要依据不同网格节点的传统立法或特定的智能法案进行治理，可最大限度地维护原有规范体系的稳定性。此外，该模式还具有渊源多元、形式灵活等特点，既可结合新型人工智能做到急法先用，又可留出试错空间供人工智能技术迭代。但该模式亦有如下弊端。第一，分散立法不利于全盘布局和持续发展，可能出现立法冲突、重复立法等问题，损害立法的体系性和权威性。例如，加利福尼亚州、康涅狄格州等州法案均提到算法影响评估，但它们在透明度要求方面有所不同，加利福尼亚州的 AB331 法案进一步规定了私人诉讼权，允许个人在其权利受到侵犯时提起诉讼，在处理跨州侵权上极可能引发冲突<sup>③</sup>。有学者指出，制定者针对不同类型人工智能制定不同法规的做法，可能导致纵横交错的法规范之间出现冲突与抵牾<sup>[10]</sup>。第二，多元执法机构在针对同一违规行为时，因规范适用的差异及权责边界的模糊，可能执法不公或重复处罚，甚至形成执法盲区。美国联邦层面，人工智能监管的职责分散于多个机构之中，如 NIST 负责技术研发标准制定与风险框架，而 FTC 侧重于市场投入与应用。当监管议题横跨不同机构的管辖边界时，“碎片化”执法易引发机构之间的权力竞争或互相推诿。第三，该模式下人工智能监管主要依靠柔性的伦理、原则或指南，由于缺乏强制性的法律责任，实际上很难发挥作用。如《权利法案蓝图》规定了算法歧视保护、数据隐私等原则，但缺乏制度细节和责任机制，企业没有足够压力去遵守，削弱了监管的有效性。第四，该模式在自由原则下主要通过政府与部分企业达成承诺协议来确保伦理、原则或指南的落地，进而维系发展与安全的平衡，但该做法的效果相对有限。拜登在签署《关于安全、可靠和值得信赖的人工智能的行政命令》之前，也只是取得亚马逊、谷歌等七家头部企业的自愿遵守承诺，美国国内市场仍然是以创新发展为主导。

## 二、领域融贯型立法模式的新路径

立法模式的选择是一项系统性工程，需考量一国经济、政策、制度等多重因素，人工智能立法亦是如此。在选择立法模式时，应立足于中国人工智能实践，审慎研判统一垂直型立法模式与分散网格型立法模式适合与否，避免盲目照搬而产生“南橘北枳”的问题。

### （一）立足人工智能实践的模式研判

事实上，统一垂直型立法模式和分散网格型立法模式都是与特定国家或地区相适应的，照搬到中国可能并不适应中国的特点和需求。例如，欧盟之所以选择统一垂直型立法模式，一方面是因为欧盟的科

---

<sup>①</sup> 有研究者指出，ISED 是政府支持工业项目的主要参与者，其中数字技术和人工智能供应链两个集群与 AIDA 监管的技术直接相关。人们有理由发问，一个与人工智能产业合作并为其提供资金的部门，如何能够公正地为其客户设计、开发和使用相同的技术。参见 WITZEL M. A few questions about Canada's artificial intelligence and data act[EB/OL].(2023-08-11)[2024-03-13].<https://www.cigionline.org/articles/a-few-questions-about-canadas-artificial-intelligence-and-data-act/>.

<sup>②</sup> MUKHERJEE S. Draft EU artificial intelligence rules could hurt Europe[EB/OL].(2023-07-01)[2024-04-26]. <https://www.reuters.com/technology/draft-eu-artificial-intelligence-rules-could-hurt-europe-open-letter-2023-06-30/>.

<sup>③</sup> FRIEDLER S, VENKATASUBRAMANIAN S, ENGLER A. How California and other states are tackling AI legislation [EB/OL].(2023-03-22)[2024-04-26]. <https://www.brookings.edu/articles/how-california-and-other-states-are-tackling-ai-legislation/>.

技产业相对落后，短期内无法实现技术赶超；另一方面，欧盟长期以来的实践使得其具备强大的监管框架，以及极具识别性的核心价值观，聚焦“技术安全”的刚性立法更能凝聚成员国间的合力，实现单一市场构建并在人工智能治理新赛道上发挥“布鲁塞尔效应”。相较而言，美国之所以青睐分散网格型立法模式，既有出于避免硬性规范阻碍技术创新主导的考虑，也是受联邦体制下统一立法难度高和需求弱的影响，更离不开其长期以来自由理念的熏陶，聚焦“技术发展”的柔性立法更符合美国国家利益，有利于维持其在全球范围的领导地位。然而，中国在产业发展、制度环境以及网络立法前期建设方面均有自身特点，对人工智能立法产生独特需求。其一，中国人工智能产业取得长足进步，但同时也存在重大原创成果不足、区域差距明显等问题，符合中国国情的立法既要保护国内市场避免西方技术侵入和智能倾销，又要推动本土产业发展和扩大国内市场。其二，中国本土制度环境中人工智能立法需更充分考量国家治理意愿，与国家“安全与发展并重”的战略政策、顶层设计相匹配，而美国自由理念下企业往往通过游说、听证等途径，推动立法朝着利于产业发展的方向前进；欧盟因其独特政治体系，立法进程冗长且易产生分歧，立法推进要依托欧洲核心价值观，以便打造单一市场。其三，中国既往网络立法建设中一直秉持工程化思维，即立法设定总目标，采用“急用先行、务实管用”的思维，同步建设多项制度，这与欧盟网络立法从统筹规划至具体建立制度的思维迥异，故中国人工智能立法还要考虑如何承继与协调实践中运行良好的网络立法<sup>[11]</sup>。有研究指出，国际人工智能监管形成了以欧盟、美国和中国为代表的三条路径，其中欧盟通过立法对不同风险施加有差别的强监管，美国以自愿为原则实施促进企业创新发展的弱监管，中国则采取了一条发展与并重的折中式监管道路<sup>[12]</sup>。

既然现有西方两大立法模式不能适应中国产业发展、制度环境等国情，由此自然产生两个问题：第一，中国人工智能立法究竟应采用什么样的新模式？第二，新模式的具体内涵为何？对于第一个问题，在对现有两种模式的“扬弃”之上可以创新形成中国人工智能立法模式。统一垂直型立法模式的体系性较强，但不区分应用场景的智能定义和风险分级，使得其对于特定类型人工智能的效果“大打折扣”；分散网格型立法模式的针对性较强，但强制性责任的缺失和执法机构的混乱，难以保证统一适用。但从两者同一性切入，彼此相互借鉴，便可在相互作用中达到和谐状态。事实上，采用这两种模式的国家和地区已出现融合趋势，如法德意三国发布非正式联合声明指出，人工智能的风险在于应用而非技术本身，应通过行为准则对基础模型实施强制性自律，且最初不宜实施制裁<sup>①</sup>；美国发布《两党人工智能立法框架》《人工智能研究、创新和问责法案》，主张建构独立监督机构管理的许可制度，实施分类监管、透明度要求和认证标准，并采用水印或其他数据方法识别人类作者、验证内容出处<sup>②</sup>。

立法模式的理论形态包括但不限于统一垂直型立法和分散网格型立法，两者之外还存在其他模式。尽管立法模式的分类尚无定论，但统一垂直型立法和分散网格型立法属于基本类型，几乎存在于所有立法，其他模式则是两者的组合或变种。例如，联合国国际贸易委员会将商事调解立法模式分为六种，其中的调解示范法模式、商事调解示范法模式、法律协调化模式属于统一垂直型立法，均制定了专门的法

---

① RINKE A. Exclusive: Germany, France and Italy reach agreement on future AI regulation[EB/OL].(2023-11-18)[2024-04-26].<https://www.reuters.com/technology/germany-france-italy-reach-agreement-future-ai-regulation-2023-11-18/>.

② The AI research, innovation and accountability act of 2023[EB/OL].(2023-11-15)[2024-03-11].[https://www.thune.senate.gov/public/\\_cache/files/7dea8daa-f6d1-4881-ad21-2381fcbe0785/6362CE1D0A17743166BC170A593B5CDA.ccas.kfall23a15.pdf](https://www.thune.senate.gov/public/_cache/files/7dea8daa-f6d1-4881-ad21-2381fcbe0785/6362CE1D0A17743166BC170A593B5CDA.ccas.kfall23a15.pdf).

律或法典<sup>①</sup>。调解法模式、法律修正模式则是分散网格型立法的变种,即将有关规范渗透到其他法律,以单独章节、法律条文和附带性规定的形式存在<sup>[13]</sup>。又如,经济立法存在集中立法和分散立法两种模式,前者包括法典化和统合立法,后者包括分领域立法和分行业立法。因经济立法的调整范围广泛、调整对象复杂多变,集中立法存在封闭性、滞后性和难度高问题,分散立法存在协调性、体系性差和适用乱问题,越来越多的学者主张统筹统合立法与分散立法,采用多元立法路径<sup>[14]</sup>。人工智能立法面对的困境与经济立法颇为相似,如调整对象有交叉性和复杂性,法律属性横跨公法和私法。但两者也有差异,如人工智能立法包括紧密相连的风险规制、产业促进与技术标准,而经济立法囊括财税、金融、竞争等相对独立的领域。

## (二) 法律 3.0 指导下的新路径

对于第二个问题——新模式的具体内涵为何,首先需要明确指导理论。实际上,统一垂直型立法模式背后的风险规制理论和分散网格型立法模式背后的自由竞争理论,都很难与人工智能实践完美适配。人工智能立法应在发展与安全之间维系平衡,主要用于解决技术性问题的法律 3.0 可以提供参考。法律 3.0 由英国学者布朗斯沃德提出,意在解决技术治理与法治之间的关系<sup>[15]</sup>。布朗斯沃德指出,传统法治包括法律 1.0 (法教义学)——通过规范内部的一致性进行规制,和法律 2.0 (法政策学)——通过政策与法律的互补性实现规制。数字技术的出现造成传统法治的两次颠覆:第一次颠覆是技术引发人们对于法律 1.0 的质疑,揭示出规则的实质内容与规范目的不匹配、规则空白或遗漏、规则与技术脱节等三种不适合性<sup>[15]23</sup>;第二次颠覆是技术作为规制工具的可利用性颠覆了法律 2.0,对监管机构提出规制立场的合法性、规制的关联性与规制行为的有效性等三重挑战<sup>[15]29</sup>。故布朗斯沃德主张以技术视角审视传统法治,使技术成为治理方案的一部分,即法律 3.0。但同时,技术治理可能造成人类道德价值的减损和自主性的丧失,还需用法治的标尺衡量。换言之,法律 3.0 不是对法律 1.0 和法律 2.0 的替代,而是一种与后者对话、共存的升级版。有学者认为法律 3.0 实质上超越了法教义学的融贯性,即法律、政策和技术之间的融贯<sup>[16]</sup>。

以法律 3.0 为指导,以法律、政策和技术三者的融贯为特点,本文提出了新的领域融贯型立法模式。其强调人工智能立法系领域法,内容包括风险规制、产业促进、技术标准三方面,形式上表现为基本法为主体、产业促进法和技术标准法为分支,最终目标是融贯治理<sup>②</sup>。

首先,法律 3.0 强调将技术纳入治理方案并与法律 1.0、法律 2.0 对话、共存,所以领域融贯型立法将人工智能立法界定为领域法,涵盖风险规制(规则)、产业促进(政策)、技术标准(技术)三方面。具言之,第一,风险规制法立足于人工智能的潜在风险,具体包括三个层面。一是法律主体及责任风险的规制,亦即解决人工智能因主体定位而产生的责任问题。欧盟《机器人技术民事法律规则》《人工智能责任指令》便属于此类。二是数据、算法及算力要素对应的风险规制。算力风险相对静态,数据与算法的风险则是动态的,比如数据存在泄露风险、非法收集风险等,算法存在歧视风险、错误风险等。

《中华人民共和国数据安全法》(以下简称《数据安全法》)、《互联网信息服务深度合成管理规定》便是分别规制数据和算法的规范。三是智能应用场景下的风险规制。人工智能存在医疗、司法、自动驾驶等应用场景,不同应用场景下需要进行相应风险规则设计。中国《智能网联汽车道路测试管理规范

---

① UNCITRAL Model law on international commercial mediation and international settlement agreements resulting from mediation [EB/OL].(2018-06-25)[2024-02-12].<https://uncitral.un.org/en/texts/arbitration/modellaw/commercialconciliation/status>.

② 法律意义上的融贯通常指法律规范的概念、原则、制度协同一致,构成一个结构化网络。本文所指的融贯性治理,包括基本法与产业促进法、技术标准法的内部融贯,以及人工智能法与数据安全法、网络安全法等规范的整体融贯。



《试行》》针对自动驾驶场景的风险进行了规制。第二，产业促进法主要通过塑造市场环境和科技促进政策两部分内容，为人工智能产业的发展和竞争提供保障。例如，国务院《关于构建数据基础制度更好发挥数据要素作用的意见》通过确立数据产权制度，为数据要素的市场交易及流通提供指引；《上海市促进人工智能产业发展条例》体现了科技政策的作用，提出要“强化投融资、研发、人才、市场等各种政策的支持”。第三，技术标准法主要通过通过对技术研发、部署或应用设定标准，确保技术安全可靠，进而用于开展技术治理。例如，中国《生成式人工智能服务管理暂行办法》（以下简称《暂行办法》）为实现人们知悉相关信息是由机器生成的而非真实信息，借助标识技术进行处理。为此，全国信息安全标准化技术委员会发布了《网络安全标准实践指南——生成式人工智能服务内容标识方法》，就标识术语、定义、标识方式和标识信息进行了标准设定。

其次，法律 3.0 指出过度依赖技术措施进行治理，可能造成严重伦理和法律问题，仍然需要将法治作为衡量尺度。领域融贯型立法未将风险规制法、产业促进法和技术标准法作为孤立的内容，而是主张三者构成整体的领域法，并在体系层次上提出以基本法为主体、产业促进法和技术标准法为分支。简言之，人工智能法不是传统意义上的一部法律，而是由基本法和子法构成的规范体系，颇似民法通则这一基本法与合同法、物权法等子法形成的民法体系。需要是，风险规制法、产业促进法和技术标准法的划分是从内容角度而言的，基本法和子法的划分是从体系层次而言的，不可混淆。人工智能基本法涵盖风险规制、产业促进和技术标准三方面内容，出于基本法的性质定位，风险规制的有关内容会占较大比例，特别是人工智能的法律定位及责任。至于应用场景下的具体风险规制，并不在基本法范畴内，其只需规定通用的、共性的风险治理条款。此外，产业促进和技术标准在基本法中的比例也不会过大，只要核心原则及制度在其中即可，具体规则还是采用单行法的形式。

为更加具体化，本文以中国学者牵头起草的《人工智能示范法 1.1》（以下简称《示范法》）作为示例，大致描绘人工智能基本法的整体框架<sup>①</sup>。具体而言，《示范法》第一章总则，规定了立法依据、适用范围以及以人为本、安全、公开透明等法律原则；第二章人工智能支持与促进，对应产业促进法的内容，如算力基础设施建设、算法和基础模型创新、数据要素供给等；第三章人工智能管理制度和第五章人工智能综合治理机制，如技术监管沙盒、监管科技、合规科技；第四章人工智能研发者和提供者义务；第五章人工智能综合治理机制和第六章法律责任，分别从合规义务、治理措施及法律责任三方面对风险进行规制；第七章附则，对人工智能、人工智能研发者、人工智能提供者等用语作了统一界定。由此可见，《示范法》一定程度上体现了领域融贯型立法的思路，但还有可以进一步完善之处，这将是后文重点论述的内容。

### 三、领域融贯型立法模式的价值论证

除论述领域融贯型立法模式的内涵与特点外，还需进一步论证其价值，以便为中国人工智能立法的模式选择提供参考。

#### （一）发展与安全并重的宗旨融贯

立法宗旨（或立法目的）是立法活动的起点，代表法律所要实现的最终目标，不但在宏观层面上决定着—部法律的具体内容，而且还统领该法的价值取向<sup>[17]</sup>。当选择一种立法模式时，应先明确相关立法宗旨以及采用哪种模式最有利于实现该宗旨。具体到中国，“发展与安全并重”应成为中国人工智能立法及治理的宗旨。顶层设计上，发展与安全并重已经成为中国人工智能发展的基本原则，除《全球人工智能治理倡议》外，中美欧等 28 个国家和地区签署的《布莱切利人工智能安全宣言》亦指出，人工

<sup>①</sup> 人工智能示范法 1.1（专家意见稿）[EB/OL].(2023-11-22)[2024-03-13].<http://iolaw.cssn.cn/zxzp/202309/W020230907361599893636.pdf>.

智能应该以安全、以人为本、可信赖和负责任的方式设计、开发、部署和使用；规范上，《暂行办法》第3条明确规定，国家坚持发展和安全并重、促进创新和依法治理相结合的原则；学理上，发展与安全并重已成为国际人工智能治理的共识原则，有研究指出人工智能立法应当以推动技术创新和防范技术风险为立法目的<sup>[18]</sup>。基于此，审视领域融贯型立法，可发现其承载着发展与安全的整体融贯。首先，领域融贯型立法对统一垂直型立法进行扬弃，吸收了后者通过强制性法律责任实施“硬法治理”的优势，能够最大程度地防范潜在风险，保障人工智能向善发展；其次，领域融贯型立法对分散网格型立法进行扬弃，吸收了后者通过自愿原则对人工智能实施“软法治理”的优势，可为数字技术留出合法合规的迭代空间，避免发展停滞遭致他国技术威胁；最后，领域融贯型立法以法律3.0为指导，认可数字技术给传统法治带来的冲击和难题，进而基于甚至利用技术特性进行治理。综上，领域融贯型立法通过促进技术发展为人工智能提供生态支撑，并借助国家、地方、行业的技术标准为人工智能安全划定底线，从而实现发展与安全的整体融贯。

### （二）经济、社会与技术三重面相的对象融贯

法律调整对象在法律规范体系中属于基础性构成要素，只有当法律规范与调整对象适配时，规范内部才能实现融贯进而发挥实效。对于人工智能立法来讲，调整对象表面上看是人工智能，但因其内部包括数据训练、模式设计和智能输出三个环节，外部涉及技术研发、技术应用、风险防范三类活动，仍需进一步明确调整对象的具体指向。龙卫球<sup>[19]</sup>将近代以来科技立法模式归纳为“技术—经济”范式，即通过科技强国、科技创新等政策鼓励技术发明和市场应用，强调技术的经济规范意义。但人工智能不仅有市场应用、成果转化等经济面相，还具有社会面相，可能造成主体危机、社会伦理和秩序破坏等负外部影响。因此，龙卫球<sup>[20]</sup>主张人工智能立法采用“技术—社会+经济”新范式，兼顾智能科技的经济意义与社会规范。此外，还需要考虑人工智能的技术面相，即由数据、算法、算力组成的内在结构，以及外在表现出的智能涌现、快速迭代、算法依赖等特性。一方面，技术面相关乎规范条款能否落实的问题，如要求人工智能完全杜绝错误、歧视性生成，理论上正当但技术上不可行；另一方面，技术面相还影响着治理能力能否现代化的问题，忽略技术面相将会造成治理措施总是比技术迭代、技术滥用“慢半拍”，遑论源头治理、敏捷治理。

对比之下，以法律3.0为指导的领域融贯型立法，能够兼顾人工智能经济、社会与技术的多重面相。领域融贯型立法强调技术与法律、政策的对话、并存，将风险规制、产业促进、技术标准均纳入立法，并以基本法为主体、产业促进法和技术标准法为分支，勾勒出一幅领域法的图景。其中，产业促进部分的调整对象正是人工智能的经济面相，技术标准和风险规制分别调整人工智能的技术面相与社会面相，由此实现人工智能经济、社会与技术三大面相的融贯。

### （三）智能立法与现有规范的系统融贯

现代国家法治存在诸多法律规范，但它们并非随意混杂的“大杂烩”，而是一套逻辑严密、内在协调的法治系统工程。法治系统工程要求规范体系内部相互协调、相互配合，共同致力于实现法治整体资源配置优化和法治实效提升<sup>[21]</sup>。进言之，人工智能本身是数据、算法、算力一体的系统工程，涉及诸多权利义务关系，其立法模式更应如此。领域融贯型立法模式以法律3.0为指导，将融贯治理作为行动目标，与法治系统工程的思维不谋而合。法治系统工程与法律体系的融贯构成一体两面的关系：前者是实现后者的可靠路径，后者是前者的建设目标。当判断一种立法模式是否满足法治系统工程的要求时，本质上应当以法律体系的融贯与否作为标准。法律体系的融贯包括逻辑融贯和价值融贯，前者指法律规范的概念、原则、规则可相互支持或证立，后者指法律体系的所有规范在实质价值上彼此一致，可被放入体系中加以理解<sup>[22]</sup>。一方面，领域融贯型立法模式以产业促进、技术标准、风险规制为内容进行统筹设计，力求人工智能基本法与产业促进法（如《上海市促进人工智能产业发展条例》《北京市促进通用人

工智能创新发展的若干措施》等)以及技术标准法(如《数据安全法 数据分类分级规则(GB/T 43697—2024)》《信息技术 人工智能术语(GB/T 41867—2022)》等)保持逻辑连贯,这是法律体系逻辑融贯的体现。另一方面,领域融贯型立法贯彻了发展与安全并重的立法原则,能够实现领域意义上的人工智能立法与《数据安全法》、《中华人民共和国网络安全法》(以下简称《网络安全法》)等规范的系统融贯,共同服务于数字治理,这是对法律体系价值融贯的精准诠释。领域融贯型立法模式可使智能立法与现有规范进行衔接,维护体系融贯的同时降低立法难度。

#### 四、领域融贯型立法模式下的制度展开

对未来立法的制度设计,需要具象地阐释模式意旨,继而指引立法者加快立法工作。承前所述,领域融贯型立法模式下基本法属于主体,产业促进法和技术标准法属于分支。唯有基本法的内容得到明确后,规范体系才有望实现融贯。此外,现有立法中已经存在一些产业促进法和技术标准法,尽管规范层级、数量和质量尚且不足,但与基本法的完全缺位相比,显然后者的推进更为紧迫。国务院立法计划和全国人民代表大会常务委员会立法规划针对的亦是基本法,所以应先围绕基本法——《人工智能法》展开。《人工智能法》不可能面面俱到,应聚焦基本规定、基本原则、基本制度和责任义务等重点内容,各部分要结合特性与现有规范进行衔接。

##### (一) 统一界定:基本规定

基本规定在整个法案中最为基础和抽象,是总则的底基。《人工智能法》的基本规定应包括基础定义、管辖范围及监管机关,这些内容均需作统一界定。

第一,人工智能监管的三部主要规范——《互联网信息服务算法推荐管理规定》(以下简称《推荐管理规定》)、《互联网信息服务深度合成管理规定》(以下简称《合成管理规定》)、《暂行办法》,在基础定义方面尚未统一,应当运用“种差+属”方法在《人工智能法》中明确基础定义,之后将其作为“属”结合推荐算法、深度合成算法、生成式人工智能各自“种差”界定特定人工智能。“种差+属”是一种下定义的常见方法,其中“属”是被定义项所从属的类,这一类概念具有共有特征,被定义项系“属”的种概念。“种差”则是被定义项与其他同属于“属”概念的概念之间的区别特征<sup>[23]</sup><sup>117</sup>。通过该方法定义基础人工智能,可彰显人工智能的核心特征,将该核心特征作为“属”,再结合特定人工智能的“种差”,便可实现兼具系统性和灵活性的智能定义。陈亮等<sup>[24]</sup>通过对立法素材“提取公因式”,从属概念、内涵和外延维度对人工智能进行定义,尽管未抽象提炼出基础定义导致结论冗长,但其思路仍有借鉴意义<sup>①</sup>。《示范法》第71条对人工智能的定义把握“自动化系统”这一“属”特征,但在“种差”方面存在不够精准的问题。

第二,现有规范的管辖范围不利于保护中国管辖权,《人工智能法》应兼具域内和域外效力,增强中国的国际地位和话语权。无论是《推荐管理规定》《合成管理规定》采用的“在中国境内提供”相关服务,或是《暂行办法》采用的“面向境内公众提供”相关服务,均将管辖重心放在中国境内,有意无意地回避了境外问题。相比之下,《AI法案》采用的“长臂管辖”更有利于取得国际竞争优势。中国主张共建网络空间命运共同体,但这不妨碍域外的合理管辖。有必要研究如何合理拓宽域外管辖范围,通过管辖权条款赋予中国执法机构域外管辖的权力<sup>[12]</sup>。《人工智能法》可参考《示范法》第2条的规定,将在中国境内从事人工智能的研发、提供和使用活动,以及在中国境外从事人工智能研发、提供和使用活动,但可能影响中国国家安全、公共利益或个人合法权益的,均纳入管辖。

<sup>①</sup> 该学者将人工智能定义为:依托算力基础设施,通过控制系统的算法处理输入的数据,以软件或硬件等多元集成方式嵌入系统后输出,或者直接在具体场景中输出对人类某种理性功能的模拟状态,在环境中进行交互,并在目标约束下经过反馈修正,最终完成预设任务的信息系统。

第三，治理人工智能要综合运用多领域专业知识，而单一个体的知识储备和应用能力有限，故设立专门监管机构较为合理。该机构还需内设法律规范、科技伦理、技术标准等部门，确保其有能力实施融贯性治理。有学者建议参考《AI 法案》和美国《两党人工智能立法框架》新设专门机构（如“人工智能发展和安全委员会”），统筹负责全国范围内的智能治理<sup>[25]</sup>。也有学者认为统一部门监管难度较大，应分行业分领域由各部门监管，同时与《网络安全法》和《数据安全法》保持一致，由中华人民共和国国家互联网信息办公室统筹协调<sup>①</sup>。鉴于中国已组建新一代人工智能治理专业委员会和科技伦理委员会，两者均发布过有关规范且同属科学技术部，可考虑将两者合并，作为专门监管机构。

## （二）吸纳转介：基本原则和基本制度

《人工智能法》作为基本法，自然要对基本原则和基本制度作出安排。现有规范对相关内容已有规定且经过了实践检验，故可将其中成熟部分纳入基本法（吸纳），并借助委任性或准用性规范将细节交单行法或下位法负责（转介）。

首先，可将《新一代人工智能治理原则》《新一代人工智能伦理规范》等规范中的伦理和原则吸纳为基本原则。《示范法》中的法律原则也是如此而来的，如以人为本、公平公正、安全可控、透明可解释等，笔者认为还可增加敏捷治理原则<sup>②</sup>。以上原则在国内外具有普遍共识，承载着人工智能应当予以遵循的价值目标，如公平公正是各国宪法和联合国《公民权利和政治权利国际公约》普遍规定的基本原则。此外，良好的法律制度能引导人们作出法律所要求和期望的行为，促进生产力的发展。《示范法》中产业促进的内容仍比较粗略，《人工智能法》还可将《上海市促进人工智能产业发展条例》《北京市促进通用人工智能创新发展的若干措施》等地方立法中的成熟条款提炼为基本制度，具体措施再借助委任性或准用性规范由前者规定，融贯立法的主体与分支。

其次，可将《数据安全法》《中华人民共和国个人信息保护法》（以下简称《个人信息保护法》）中的数据分类分级制度引入立法，作为基本制度之一。数据分类分级是全球公认的治理制度，但中国现有规范在数据分类分级标准上尚未统一，导致数据孤岛、信息壁垒现象层出不穷，《人工智能法》应实现数据互联互通，这是《示范法》所未涉及的。举例而言，《个人信息保护法》将个人信息分为一般信息和敏感信息，《数据安全法》将数据分为一般数据、重要数据、核心数据。是故，《人工智能法》需规定由权威部门（如全国信息安全标准化技术委员会）统筹协调，整合多元数据分类分级标准，并借助转介条款进一步确定场景化清单。比如，援引国家标准《数据安全技术 数据分类分级规则（GB/T 43697—2024）》的规定，依据“行业领域—业务属性—专门要求”进行数据分类，再综合考虑“影响对象”（如国家安全、经济运行、社会秩序、公共利益、组织权益、个人权益）与“影响程度”（如特别严重、严重、一般）进行数据分级。

最后，可将《推荐管理规定》中的算法备案制度引入立法并进一步完善。因算法黑箱、算法歧视等缘故，人工智能风险一定程度上由算法决定，<sup>[26]</sup>依据风险级别设置不同程度的合规义务是通行做法。

《推荐管理规定》中的算法备案制度有类似功能且已纳入《合成管理规定》《暂行办法》，亦可引入《人工智能法》。但该制度有两点缺陷：一是针对的算法限于“具有舆论属性或者社会动员能力”；二是以告知作用为核心，唯有风险非常明显或损害发生后，才以事后溯源追责的方式发挥作用<sup>[27]</sup>。弥补第一个缺陷关键在于实施分类分级，将风险达到一定程度的算法均纳入审查。对此，有观点指出欧盟的“四分

① 徐钢.法律如何监管人工智能? [EB/OL].(2023-09-19)[2024-03-15].[https://mp.weixin.qq.com/s/qzPjhq\\_svPYuyuzgEDI-1g](https://mp.weixin.qq.com/s/qzPjhq_svPYuyuzgEDI-1g).

② 敏捷治理原则围绕透明度、适应性、协作以及对交付价值的关注，这些原则为有效的决策和治理奠定了基础。敏捷治理原则主张积极地管控风险，有助于实现风险的灵活治理和精准治理。

法”过于严厉，且“不可接受风险”的标准模糊，应进行“两分法”式的风险管理（如《示范法》）<sup>①</sup>。但也有学者主张采用高、中、低风险的“三分法”，并以《中华人民共和国生物安全法》《医疗器械监督管理条例》为例证<sup>②</sup>。笔者认可后者，并主张辅以英国情景关联的风险识别方法，各别类型的具体目录由国务院统一发布。弥补第二个缺陷的关键是充实备案类型。当下算法备案制度系告知型和后设型备案，约束力有限且易滞后，仅可用于中风险的算法；对于高风险的算法，应引入许可型备案，即由监管机构事先对备案算法的材料进行实质审查，唯有设置防范措施和保护机制后方可投放市场或投入使用，并建立“全生命周期监管”；低风险的算法无备案的必要，只需相关人工智能遵守数据标注规则，使用户有所认知即可。

### （三）制度重构：责任义务体系

人工智能监管不能仅靠“最佳化命令”，必须设置“强制性规则”和“阶梯式责任”，实施“软硬搭配”的融贯治理。目前，《推荐管理规定》《合成管理规定》引导下的算法监管主要围绕“服务提供者”展开，要求其承担主体责任并履行系列数据、算法的安全义务<sup>[26]</sup>。但该做法存在两个问题：一是弱化了技术支持者的义务要求，往往将其作为服务提供者的辅助，导致居于技术运行关键地位的技术支持者未承担应有义务和责任；二是忽略了服务使用者的应有义务，人工智能引发的侵权问题并不总是来自于技术支持者或服务提供者，也可能是用户恶意滥用导致的。<sup>[29]</sup>《暂行办法》虽然规定了服务使用者的善意使用，但仍未就技术支持者的责任义务作出规定。《示范法》则对人工智能研发者、人工智能提供者的责任义务进行了规定，而忽略了对服务使用的要求，亦有不足。

人工智能的责任义务需打破以“服务提供者”为中心的传统治理路径，重构为“研发者—提供者—使用者”三方担责。有学者指出，《人工智能法》作为基础性法律，应将人工智能用户、服务提供者、部署者等主体的权利、义务、责任理清楚，这是重中之重<sup>③</sup>。三方的责任义务划分如下：第一，研发者与提供者在处理对象、业务承担等方面存在区别，但两者均处于人工智能应用链的前端且存在身份重合，如企业既研发人工智能技术又将智能服务投入市场，故两者应遵守共通的、必要的义务，如公开透明义务、解释说明义务、风险管理义务等；第二，研发者主要从事算法设计、数据标注、模型训练等研究和开发活动，该过程中可能因数据污染、算法变量不适配等缘故，导致错误生成、非法生成或歧视性生成，故其应另外承担数据来源合法义务、数据质量保障义务、歧视性防范义务等；第三，提供者直接面向用户提供人工智能服务，应当确保用户知悉生成内容来自机器，并及时处理用户使用过程中遭遇的问题，故其应另外承担数据标识义务、通知和补救义务、及时受理投诉举报义务等；第四，使用者作为用户，位于人工智能应用链的末端，只需保证不恶意使用、不非法滥用，参照《暂行办法》中的善意使用义务即可。

除责任义务划分之外，还有必要就责任类型和责任认定作出安排。第一，从责任类型来看，人工智能应用全链条过程会涉及民事责任、行政责任甚至刑事责任。从发展与安全并重的原则出发，综合考虑三种法律责任，优化三种法律责任的衔接。第二，责任认定方面采用过错原则。当行为主体违反自身承担的义务时，即可认定其过错，除非其能够证明自身无过错。当研发者与提供者、此研发者与彼研发者、

---

① 吴阳.人工智能法纳入立法计划背后：专家建议划定红线[EB/OL].(2023-06-08)[2024-02-23].<https://static.cdsb.com/micropub/Articles/202306/ed5ea6859cdbcfd574bd87544b7b5c0fe.html>.

② 徐钢.法律如何监管人工智能？[EB/OL].(2023-09-19)[2024-03-15].[https://mp.weixin.qq.com/s/qzPjhq\\_svPYuyuzgEDI-1g](https://mp.weixin.qq.com/s/qzPjhq_svPYuyuzgEDI-1g).

③ 吴阳.人工智能法纳入立法计划背后：专家建议划定红线[EB/OL].(2023-06-08)[2024-02-23].<https://static.cdsb.com/micropub/Articles/202306/ed5ea6859cdbcfd574bd87544b7b5c0fe.html>.

此提供者与彼提供者等不同主体构成共同过错时，应当承担连带责任。第三，可借鉴欧盟《人工智能责任指令》中的“证据披露规则”和“因果关系推定”，前者是指受害方穷尽一切适当努力仍无法从被告方收集其掌握的、可能造成损害的特定高风险人工智能的相关证据时，法院可指令相关主体披露证据，如对方拒绝披露，则推定其未尽到注意义务。后者是指当一个相关过错和损害被证明，且可以合理地认为该过错影响人工智能的输出或导致无法输出时，则推定该过错与人工智能的输出或无法输出存在因果关系<sup>①</sup>。

## 五、结语

全球智能竞争逐渐从技术之争发展到规则之争、治理之争。人工智能伦理、原则等软性规范难以满足实践需求，国家强制力保障的硬性立法成为时代治理主题。为抢占国家人工智能治理的规范高地，西方发达国家不断加快人工智能立法进程。2024年3月13日，欧洲议会以523票赞成、46票反对、49票弃权通过了《AI法案》官方最终版本，全球即将迎来第一部人工智能监管立法。对于中国而言，人工智能立法草案已经纳入《国务院2023年度立法工作计划》，而且自2024年全国两会以来，有关立法的呼声日渐高涨。本文对国际人工智能立法作比较考察，归纳出统一垂直型立法和分散网格型立法两种模式，并分析了两种模式的特点和优劣。在此基础上，本文结合中国人工智能实践，提出西方两种人工智能立法模式并不适合中国，并尝试提出第三种立法模式——领域融贯型立法。立足于领域融贯型立法模式，本文对中国未来《人工智能法》的基本规定、基本原则、基本制度和责任义务等重点内容作了展望，尝试为下一步立法工作提供指引。因由《人工智能法》从立法设计到生效实施还有较长的过程，期间可能出现新的变化和问题，相关理论研究应当持续推进。值得注意的是，即便《人工智能法》公布生效，也不意味着伦理、原则等软法规范失去价值，“软硬结合”才能实现融贯治理。总之，如何让人工智能在发展与安全之间取得最佳平衡，并持续增进人类共同福祉是未来立法与学界研究的重大课题。

## 参考文献

- [1] SAG M. Copyright safety for generative AI[J]. *Houston Law Review*, 2023, 61(2): 295-344.
- [2] BAKINER O. The promises and challenges of addressing artificial intelligence with human rights[J]. *Big Data & Society*, 2023, 10(2): 1-13.
- [3] 胡蝶飞. 委员建议尽快推进人工智能立法[N]. *上海法治报*, 2024-03-05(2).
- [4] 郑志峰. 人工智能立法的一般范畴[J]. *数字法治*, 2023, 1(6): 55-75.
- ZHENG Z F. The general scope of artificial intelligence legislation[J]. *Digital Law*, 2023, 1(6): 55-75.
- [5] 江国华. 立法模式及其类型化研究[M]//刘茂林. *公法评论：四卷*. 北京：北京大学出版社，2007.
- [6] 曹海晶. *中外立法制度比较*[M]. 北京：商务印书馆，2016.
- [7] ROBERTS H, BABUTA A, MORLEY J, et al. Artificial intelligence regulation in the United Kingdom: a path to good governance and global leadership?[J]. *Internet Policy Review*, 2023, 12(2): 1-31.
- [8] SELBST A D, BAROCAS S. Unfair artificial intelligence: how FTC intervention can overcome the limitations of discrimination law[J]. *University of Pennsylvania Law Review*, 2022, 171(4): 1023-1093.
- [9] HELBERGER N, DIAKOPOULOS N. ChatGPT and the AI act[J]. *Internet Policy Review*, 2023, 12(1): 251-265.
- [10] 姚志伟, 李卓霖. 生成式人工智能内容风险的法律规制[J]. *西安交通大学学报(社会科学版)*, 2023, 43(5): 147-160.

<sup>①</sup> AI liability directive[EB/OL]. (2023-10-11)[2024-02-23]. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0496>.

YAO Z W, LI Z L. Legal Regulations on the risks of generative artificial intelligence content[J]. Journal of Xi'an Jiaotong University(Social Sciences),2023,43(5):147-160.

[11] 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[J].法律科学(西北政法学报),2024,42(3):3-17.

ZHANG L H. What kind of 'artificial intelligence law' does China need: the basic logic and institutional framework of China's artificial intelligence legislation[J]. Science of Law(Journal of Northwest University of Political Science and Law),2024,42(3):3-17.

[12] 邵长茂.人工智能立法的基本思路[J].数字法治,2023,1(5):1-7.

SHAO C M. The basic approach to artificial intelligence legislation[J]. Digital Law,2023,1(5):1-7.

[13] 赵毅宇.中国商事调解立法模式选择及其展开[J].法学杂志,2023,44(3):157-172.

ZHAO Y Y. The choice and implementation of legislative models for commercial mediation in China [J]. Law Science Magazine,2023,44(3):157-172.

[14] 张守文.经济法的立法路径选择[J].现代法学,2023,45(1):118-131.

ZHANG S W. Choices legislative path selection of the economic law[J]. Modern Law Science,2023,45(1):118-131.

[15] 布朗斯沃德.法律 3.0：规则、规制和技术[M].毛海栋,译.北京:北京大学出版社,2023.

[16] 郭春镇.生成式 AI 的融贯性法律治理：以生成式预训练模型(GPT)为例[J].现代法学,2023,45(3):88-107.

GUO C Z. The coherent legal governance of generative AI: taking the generative pre-training model (GPT) as an example[J]. Modern Law Science,2023,45(3):88-107.

[17] 杨铜铜.论立法目的类型划分与适用[J].东岳论丛,2023,44(2):177-190.

YANG T T. On the classification and application of legislative purposes[J]. Dong Yue Tribune,2023,44(2):177-190.

[18] 侯东德.人工智能法的基本问题及制度架构[J].政法论丛,2023,30(6):61-72.

HOU D D. Basic problems and institutional framework of artificial intelligence law[J]. Journal of Political Science and Law,2023,30(6):61-72.

[19] 龙卫球.科技法迭代视角下的人工智能立法[J].法商研究,2020,37(1):57-72.

LONG W Q. Artificial intelligence legislation from the perspective of iterative development of science and technology law[J]. Studies in Law and Business,2020,37(1):57-72.

[20] 龙卫球.人工智能立法的“技术—社会+经济”范式：基于引领法律与科技新型关系的视角[J].武汉大学学报(哲学社会科学版),2020,73(1):65-76.

LONG W Q. "Technology-Society+Economy" paradigm of AI legislation:from the perspective of leading the new relationship between law and technology[J]. Wuhan University Journal(Philosophy & Social Science),2020,73(1):65-76.

[21] 钱弘道,康兰平,申辉.数字法治的基本原理和实践进路[J].浙江大学学报(人文社会科学版),2022,52(9):5-24.

QIAN H D, KANG L P, SHEN H. The basic principles and practical approach of digital rule of law [J]. Journal of Zhejiang University(Humanities and Social Sciences),2022,52(9):5-24.

[22] 雷磊.融贯性与法律体系的建构：兼论当代中国法律体系的融贯化[J].法学家,2012,27(2):1-16.

LEI L. Coherence and the construction of legal system: how to improve the coherence of contemporary Chinese legal system[J]. The Jurist,2012,27(2): 1-16.

[23] 周安平.常识法理学[M].北京:北京大学出版社,2021.

[24] 陈亮,张翔.人工智能立法背景下人工智能的法律定义[J].云南社会科学,2023,43(5):162-170.

CHEN L, ZHANG X. Legal Definition of Artificial intelligence in the Context of Artificial Intelligence Legislation[J]. Social Sciences in Yunnan,2023,43(5): 162-170.

- [25] 刘金瑞.生成式人工智能大模型的新型风险与规制框架[J].行政法学研究,2024,32(2):17-32.  
LIU J R. Regulatory framework for new risks of large generative AI models[J]. Administrative Law Review,2024,32(2):17-32.
- [26] 高进,刘聪. 算法歧视的敏捷治理--以 D 短视频平台为例[J].西安交通大学学报(社会科学版),2024,44(1):121-132.  
GAO J, LIU C. Agile governance of algorithmic discrimination:take the D short video platform as an example[J]. Journal of Xi'an Jiaotong University(Social Sciences), 2024,44(1):121-132.
- [27] 张吉豫.论算法备案制度[J].东方法学,2023,16(2):86-98.  
ZHANG J Y. On record-filing system of algorithms[J]. Oriental Law,2023,16(2):86-98.
- [28] 张凌寒.生成式人工智能的法律定位与分层治理[J].现代法学,2023,45(4):126-141.  
ZHANG L H. The legal positioning and hierarchical governance of generative AI[J]. Modern Law Science,2023,45(4):126-141.
- [29] 陈建兵,王明.负责任的人工智能：技术伦理危机下 AIGC 的治理基点[J].西安交通大学学报(社会科学版),2024,44(1):111-120.  
CHEN Jianbing, WANG Ming. Responsible artificial intelligence: governance fundamentals for AIGC in the ethical crisis of technology[J]. Journal of Xi'an Jiaotong University(Social Sciences),2024,44(1):111-120.