工信安全智库

瞭望系列报告

全球人工智能立法的主要模式、各国实践及发展趋势研究

国家工业信息安全发展研究中心

2024 年 12 月

“工信安全智库”系列报告编委会

编写组

撰　稿：刘芷君

审　稿：李　强　冯　媛　王花蕾　孙倩文

序

国家工业信息安全发展研究中心经过60余年的发展与积淀，在智库研究方面形成了丰硕的积累。2018年9月，中心推出“工信安全智库”品牌，立足深化供给侧结构性改革和加快建设创新型国家战略需求，围绕制造强国和网络强国建设任务，聚焦网络安全、数字经济、软件产业、产融合作等重点领域，开展基础性、战略性、先导性智库研究，为工业和信息化部、中央网信办、国家发展改革委等提供智力支持。

“工信安全智库”自2019年开始陆续推出“研判”“洞察”“瞭望”“指数”“案例”“编译”等系列研究报告，围绕党和政府决策急需的相关重大课题和关键问题，开展形势研判、专题调研、国际跟踪、景气测度、案例分析、报告翻译等方面的持续研究，为主管部门预见走势、把握机遇、应对挑战、谋划战略提供参考。

本次推出的瞭望报告总结了全球人工智能立法的两大逻辑导向和四大主流模式，列举了欧盟、美国、新加坡、中国等主要经济体的人工智能立法实践，研判了全球人工智能立法的发展趋势，并为我国完善人工智能法律体系提出了相关建议。

由于成稿仓促，加之水平有限，报告中难免有疏漏和错误之处，恳请批评指正。

编写组

2024年12月

摘 要

随着人工智能技术的不断发展和广泛应用，人工智能逐渐成为大国博弈的重要焦点，各国和地区都在积极推进人工智能立法，推动人工智能规范化应用，全球人工智能竞争逐渐从技术之争发展到规则之争、治理之争。纵观全球人工智能立法实践，可以发现不同国家和地区在人工智能立法方面体现出不同的逻辑导向和立法模式路径。欧盟以安全规制为导向，制定出台了全球首部统一的人工智能综合立法，体现出强监管倾向；美国以技术发展为导向，对人工智能的监管力度相对宽松，相关规定融合在现有立法中或对重点应用场景制定专门立法；新加坡致力于占据人工智能技术创新主动权，以技术发展为导向，通过制定标准、指南等非约束性的软法指导人工智能规范发展；中国则贯彻统筹发展和安全的理念，现阶段体现出融合立法和场景立法的特征，同时持续推进人工智能综合立法进程，不断完善人工智能治理法律法规体系。

随着全球人工智能立法和监管实践的持续推进，人工智能在各国的战略地位进一步加强，各国将更加关注安全和发展的平衡，柔性规则和硬性规定将同步发展，统一立法将成为国际主流趋势，人工智能国际治理规则体系将持续构建。我国应秉持统筹发展和安全的原则，顺应人工智能治理的全球趋势，强化立法顶层设计，制定人工智能治理制度框架，加强国际交流对话，提出人工智能治理的中国方案。

目 录

一、人工智能立法的逻辑导向和主要模式 1
（一）人工智能立法的两大逻辑导向 1
（二）人工智能立法的四大主流模式 3
二、主要经济体人工智能立法实践 6
（一）欧盟：以安全规制为导向，实行统一立法模式 6
（二）美国：以技术发展为导向，实行融合立法和场景立法模式 ... 9
（三）新加坡：以技术发展为导向，实行软法先行模式 12
（四）中国：统筹发展和安全，实行融合立法和场景立法模式 14
三、全球人工智能立法特点和趋势 20
（一）立法逻辑导向日趋融合，安全和发展并重成为全球共识 20
（二）软法先行成为普遍选择，软法硬法协同治理是重要方向 22
（三）统一立法成为大势所趋，多元立法模式呈融合互补趋势 23
（四）国际治理规则持续构建，人工智能立法全球化趋势显著 25
四、启示建议 27
（一）持续完善顶层设计，形成多元共治的治理格局 27
（二）划定安全治理边界，确立风险分级的管理机制 29
（三）强化国际交流对话，提出全球治理的中国方案 30
参考文献 33

人工智能是引领未来的战略性技术，是新一轮科技革命和产业变革的重要驱动力量，是必须抢占的科技制高点。党的二十届三中全会通过的《中共中央关于进一步全面深化改革 推进中国式现代化的决定》提出“建立人工智能安全监管制度”，对推动人工智能安全发展义重大。深入贯彻落实党的二十届三中全会精神，必须坚持全面依法治国，加快人工智能立法进程，推进人工智能治理法治化，更好发挥法治对于人工智能健康发展的引领、规范和保障作用。

一、人工智能立法的逻辑导向和主要模式

由于各国在人工智能技术发展、社会文化背景、法律传统等方面存在差异，各国人工智能立法体现出了不同的逻辑导向和立法模式。

（一）人工智能立法的两大逻辑导向

人工智能作为引领未来的战略性技术，是新一轮科技革命和产业变革的核心驱动力，因此对人工智能的立法规制需要衡量人工智能技术和产业的发展及其安全治理需求，然而这两者不可避免地存在一定的冲突与矛盾。一方面，人工智能技术的发展会带来更大的安全风险。例如，人工智能系统需要大量的数据进行训练和优化，将带来更大的数据泄露和个人隐私风险；人工智能技术将暴露更大的系统攻击面，增加网络攻击风险；自动驾驶汽车、智能医疗诊断等人工智能自主决策场景在系统决策算法存在缺陷或被恶意篡改的情况下，可能导致严重的安全事故；人工智能技术还可能被用于制造恶意软件、进行网络钓鱼等非法活动；如果人工智能技术被用于

监视和控制，将侵犯个人隐私和自由，引发严重的社会问题。另一方面，安全治理在一定程度上会制约人工智能技术的快速发展。一是导致合规成本增加。为了符合安全治理和合规要求，企业需要在技术研发、产品设计、数据收集与处理等各个环节增加投入，从而增加企业运营成本，或导致研发投入减少。二是限制技术创新。安全治理往往伴随一系列限制性规定，例如限制某些数据的使用方式或算法的应用范围以保护个人隐私、要求采用更为保守的技术方案以保障系统安全等，这些限制在一定程度上制约了人工智能技术的创新空间。三是提高市场准入门槛。安全治理往往会提高企业的市场准入门槛，可能导致一些中小企业或初创企业因难以承担高昂的合规成本而难以进入市场，从而削弱市场竞争活力。四是政策滞后性带来监管不确定风险。法律法规的制定和修订速度往往滞后于技术发展速度，导致企业在技术创新过程中面临不确定性和风险，或使企业在投资决策时更加谨慎，从而减缓技术发展速度。五是加大国际合作壁垒。不同国家和地区在人工智能、数据跨境流动、技术出口管制等方面的法律法规、政策标准和行业规范存在差异，可能导致国际合作面临障碍，从而制约人工智能技术的跨国界快速发展。

各经济体基于自身发展需求对人工智能展开法律规制，平衡人工智能安全与人工智能技术创新，体现出不同的立法导向。以美国为代表的国家以推动技术发展为导向进行人工智能立法。美国的法律、政治与文化一直信任市场力量，支持企业通过市场竞争为消费者与市场提供优质产品，这一导向在人工智能领域也得以体现。美

国对于国内人工智能发展强调市场主导与创新驱动，在风险防范方面强调行业自律而不是强有力的政府监管。在这种理念的指导下，美国的人工智能立法主要针对政府机构，对企业的限制仅限于消费者知情权、反歧视保护等方面。以欧盟为代表的经济体则以安全规制为导向推进人工智能立法。目前全球的产业与技术主要集中在美国与中国尤其是美国，而欧盟则几乎没有特别具有影响力的数字企业与领先技术，欧盟在人工智能等数字经济领域的技术硬实力逐渐降低，其影响力日益集中在规则制定的软实力领域。欧盟在价值取向上更注重人工智能本身的安全，对人工智能进行统一立法，既可以对数字产业进行有效规范，又可以利用立法对域外大型互联网与科技企业进行有效监管。

（二）人工智能立法的四大主流模式

通过立法规制人工智能已成为各国共识，各国不断探索采用符合本国国情的立法模式，以实现技术向善、规范运行的目的。立法模式是指一定时期内，一个国家或地区在某一领域制定法律的表现形式。纵观各国家和地区的人工智能立法实践，目前主要立法模式可以划分为统一立法模式、场景立法模式、融合立法模式、软法先行模式四种。值得注意的是，立法是一项复杂的国家行为，各国的人工智能立法并非只体现为单一的立法模式，往往会融合不同立法模式。

一是统一立法模式。统一立法模式是指国家制定一部专门的人工智能法案，对人工智能的定义、发展、安全治理、监督管理等作

统一规定，对人工智能进行垂直性和体系性治理，欧盟是统一立法模式的代表。整体上看，统一立法模式通常会对人工智能进行统一界定以确保法律规范的同一性和一致性，由专门机构统一负责监管执法以确保执法的专业性，还会明确违反法案所承担的法律责任。欧盟是统一立法模式的典型代表。

二是场景立法模式。场景立法模式是一国针对特定场景或特定领域制定不同的法案规范人工智能应用。与统一立法模式不同，场景立法模式下人工智能相关立法并不会对人工智能进行全面规制，而是对消费、金融、教育等不同场景下的人工智能应用进行规范，从而在推动人工智能创新发展的同时，确保在重要领域和场景下的规范应用和安全治理。美国是场景立法模式的典型代表。

三是融合立法模式。融合立法模式是针对人工智能的法律规定分散融合在其他部门法中，用以规范不同行业领域的人工智能安全应用。在融合立法模式中，对于人工智能的规范通常作为消费者保护、产品责任、个人信息保护等立法的重要内容之一，而不是作为专门规制的对象进行立法。美国也是融合立法模式的典型代表。

四是软法先行模式。软法是指不能运用国家强制力保证实施的法律规范。与软法相对应的概念是“硬法”，是指具有强制约束力的法律规范。硬法通常会经过严格的立法程序和审查机制，法律条文具有高度的确定性，难以较好地应对人工智能技术的复杂性和不确定性。因此，软法作为一种更加灵活、更具“实验性”的治理工具，被越来越多地应用于人工智能的治理中。在人工智能立法过程

中，优先采用不具有法律约束力但会产生实际规范效果的行为准则、政策指导等软法手段，能够增强法律规范的灵活性和对人工智能技术的适应性，降低监管成本，在一定时期实现对人工智能技术的有效治理。很多国家在人工智能立法最初阶段都是从软法开始的，新加坡是目前软法先行立法模式的典型代表。

比较全球主流的人工智能立法模式，会发现各个模式各有优缺点。统一立法模式能够最大程度集中国家力量，一方面能够协调一国内部法律规范，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。但统一立法模式明确界定人工智能相关概念范围，在面对技术更迭时易因灵活性和针对性不足而减损规范效果，还面临标准设定困难及风险动态变化的问题。场景立法模式和融合立法模式针对特定行业领域的人工智能应用进行立法规范，主要通过融入传统部门法或推出针对性法案进行规制，可最大限度地维护原有规范体系的稳定性，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”，又可留出试错空间推动人工智能技术迭代创新。但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。同时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。软法先行模式在人工智能处于高速发展的阶段能够为技术创新留出空间，通过指导性规范引导技术健康向善发展，

但软法模式下人工智能监管主要依靠柔性的伦理、原则或指南，缺乏强制性的法律责任，实际上很难发挥法律威慑作用。

二、主要经济体人工智能立法实践

人工智能属于新生事物，在人工智能立法方面各国家和地区均处于探索阶段，基于具体国情、发展实际和价值取向，采取了不同的立法路径。

（一）欧盟：以安全规制为导向，实行统一立法模式

欧盟在人工智能安全和技术创新的平衡中更加偏向安全治理，体现出安全规制导向。除了防范人工智能技术发展带来的安全风险外，欧盟高度重视人工智能安全规制还出于几方面考虑。一是保护公民基本权利。欧盟高度重视个人的权利和自由，无论是数据保护还是人工智能保护均体现出高度重视个人权利保护的立法导向。人工智能技术的快速发展和应用，对公民隐私权等基本权利构成了潜在威胁。欧盟通过制定严格的人工智能监管法案规范人工智能的应用，要求人工智能系统的开发者和部署者遵守透明度要求，保护用户数据的安全和隐私。二是提升欧洲数字市场的凝聚力和竞争力。在数字时代，欧洲面临着来自全球范围内的竞争压力。为了提升欧洲数字市场的凝聚力和竞争力，欧盟制定统一的法律框架来监管人工智能技术，为所有成员国提供了统一的法律指引，有助于消除监管碎片化，提升欧洲在数字领域的整体实力。三是维护欧盟的“数字主权”。在人工智能等前沿技术上，欧盟目前落后于美国和中国。为了维护自身的“数字主权”，欧盟希望通过强监管来规范外国技

术公司的行为，同时促进本地技术创新，有助于欧盟在数字领域特别是监管规则方面形成全球范式，掌握全球人工智能规则制定的主导权。

欧盟作为人工智能统一立法模式的典型代表，较早开始人工智能立法实践，经过多年的探索实践和立法程序，最终形成了全球首部专门的人工智能法案。欧盟自 2016 年起就不断探索推进对人工智能技术应用的监管体系建构，相关部门陆续发布《欧盟机器人民事法律规则》《算法的可问责和透明的治理框架》《可信人工智能伦理指南》《试行“可信赖人工智能的指引评估清单”》等指导性文件。在此基础上欧盟不断推进人工智能统一立法进程。2021 年 4 月，欧盟委员会提出了关于欧盟人工智能监管框架的建议，同时提出了《人工智能法案》提案的谈判授权草案，建议在立法中制定一个技术中立的人工智能系统定义，并根据“基于风险的路径”为不同风险的人工智能系统进行分类，并制定相应的监管举措。此后，欧洲议会和欧盟理事委员会针对该草案进行了多轮修订和讨论，于 2022 年 12 月形成了《人工智能法案》折中草案最终版。2023 年 6 月，欧洲议会以 499 票赞成、28 票反对和 93 票弃权，高票通过了《人工智能法案》谈判授权草案。此次谈判全面禁止了人工智能用于生物识别监控、情绪识别与预测性警务，同时也规定了生成式 AI 系统必须披露人工智能生成的内容，在选举中用于影响选民的人工智能系统被认为具有高风险。这份谈判授权草案推动了该法案进入立法程序的最后阶段，2023 年 12 月，欧洲议会、欧洲理事会和欧盟委员会三

方就《人工智能法案》达成协议。2024 年 2 月 2 日，欧盟 27 国代表在《人工智能法案》文本上达成临时政治共识。3 月 13 日，欧盟议会以 523 票赞成、46 票反对和 49 票弃权审议通过《人工智能法案》，标志着全球人工智能领域监管迈入新时代。欧盟《人工智能法案》对人工智能系统进行了统一界定和分类，并对风险进行分级管理，还对通用人工智能模型进行了规定，为各国提供了人工智能监管规则范例。

表 1 欧盟《人工智能法案》重点内容

重点事项	主要内容
法案适用主体及除外情形	法案旨在规范整个人工智能产业链的主体，适用主体包括与欧盟市场有连接点的人工智能系统提供商、使用商、进口商、分销商和产品制造商。法案也规定了一些适用的例外情形，包括但不限于（1）仅为科学研究和开发目的而专门开发和投入使用的人工智能系统或人工智能模型及其输出；（2）在纯粹个人非专业活动中使用人工智能系统的自然人；（3）专门为军事、国防或国家安全目的而投放市场、投入使用的人工智能系统。
人工智能系统的定义	法案明确人工智能系统是一种基于机器的系统，设计为以不同程度的自主性运行，在部署后可能表现出适应性，并且为了明确或隐含的目标，从其接收的输入中推断如何生成可影响物理或虚拟环境的输出，如预测、内容、建议或决定。
人工智能系统的分类	法案将人工智能系统分为两类：一是产品构成类，例如医疗器械、自动驾驶、轮船、玩具等产品内部包含的人工智能系统；二是独立于产品的辅助决策类，例如用于招聘雇佣、招生升学、移民筛查、执法检查等各类应用场景的人工智能系统。法案将这些不同类别的人工智能系统进行统一规制。
人工智能系统的风险分级管理	法案将风险分为不可接受风险、高风险、有限风险、最小风险四类，每个类别适用不同程度的监管要求。禁止类的风险主要包括利用潜意识技术扭曲个人或群体的行

重点事项	主要内容
	为，造成重大伤害、利用人工智能危害特定弱势群体、进行社会信用评分、利用人工智能进行特定侵入性的执法等类型。高风险主要包括应用于现有欧盟产品安全法范围的人工智能系统，以及在生物识别、关键基础设施、教育和职业培训、就业等特定领域应用人工智能系统。对于高风险的人工智能，人工智能系统的提供者需要承担一系列责任。
通用人工智能模型管理	法案明确通用人工智能模型是一个包括使用大量数据进行大规模自我监督训练的人工智能模型。该模型具有显著的通用性，无论以何种方式投放市场，都能胜任各种不同的任务，并可集成到各种下游系统或应用中，但在投放市场前用于研究、开发或原型设计活动的人工智能模型除外。法案将通用人工智能模型分为不具有系统性风险的通用人工智能模型和具有系统性风险的通用人工智能模型。如果通用人工智能模型具有高度影响能力或被欧盟委员会认定为具有高度影响能力，则被视为构成系统性风险，需要承担特殊责任，例如进行模型评价，评估和减轻系统性风险，追踪、记录和报告严重事件的相关信息与可能的纠正措施，确保足够水平的网络安全保护等责任。

（二）美国：以技术发展为导向，实行融合立法和场景立法模式

美国人工智能立法秉持着以技术发展为导向的传统，未制定联邦层面的综合性立法对人工智能实施统一监管。美国的法律体系对科技创新具有很高的包容性，对综合性的科技规制立法历来持审慎立场。在人工智能领域，美国秉持着技术发展导向，致力于强化其在人工智能领域的全球领导地位。在联邦层面，主要通过发布总统行政令指引人工智能规范化发展。总统颁布行政令虽然具有宪法或国会的授权，但其本身并不算严格意义上的法律，主要在特定事项上影响和引导联邦机构、企业等主体的行为。特朗普政府时期关于人工智能的总统行政令以促进人工智能发展为导向，旨在保持美国

在该领域的领导地位。2019 年 2 月，特朗普政府发布第 13859 号行政命令《保持美国在人工智能领域的领导地位》，要求联邦政府机构协调一致，推进“美国人工智能行动计划”，就美国联邦政府层面人工智能的研究、推广和培训做出全面部署，以确保美国在 AI 研发及相关领域的全球领先优势。2020 年 12 月 3 日，特朗普政府发布第 13960 号行政命令《促进联邦政府使用值得信赖的人工智能》，规定了除国家安全和国防领域以外的联邦政府机构在设计、开发、部署、购买或使用人工智能技术时需要遵循的 9 项基本原则。拜登政府发布的人工智能相关总统行政令则开始关注重点领域人工智能的安全治理和规范化应用。2023 年 10 月 30 日，拜登政府颁布第 14110 号总统行政命令《安全、可靠和值得信赖地开发和使用人工智能》，要求美国联邦政府机构在刑事司法、教育、医疗保健、住房和劳工等领域制定标准与规范。针对人工智能企业，该行政令援引《1950 年国防生产法》，要求企业在对国家安全或公共健康安全构成严重风险的人工智能模型训练方面与联邦政府共享信息。

美国在各行业领域立法中融入人工智能相关规定，或由各行业监管部门适用或解释相关规定以开展人工智能治理，体现出“融合立法”特征。联邦贸易委员会（FTC）等联邦机构以及各个州的立法通过适用或解释现有法律监管各个行业领域的人工智能应用。例如，在消费者保护领域，《联邦贸易委员会法》第 5 条“禁止不公平或欺诈行为”被监管部门应用于人工智能监管，FTC 基于该法对 OpenAI 公司进行了正式调查。在金融领域，《公平信用报告法》和《平等

信用机会法》均涉及对自动化决策的监管，美国金融机构在应用基于机器学习的信贷承销模型时均需符合相关规定。在航空领域，美国国会在通过《2018 年美国联邦航空局再授权法案》时增加了“建议联邦航空管理局定期审查航空领域人工智能的状况，并采取必要措施应对新发展”的内容。

为顺应人工智能的广泛应用，美国各州针对人工智能技术应用程度较深的特定场景进行了广泛的立法实践。为规制用工场景下的人工智能应用，伊利诺伊州 2019 年颁布《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。纽约市 2021 年通过第 144 号地方法律，禁止雇主和职业中介在没有进行偏见审计的情况下使用人工智能和基于算法的技术进行招聘、雇佣或晋升。为规范选举场景下的人工智能应用，密西西比州 2024 年 4 月通过了 SB 2577 号法案，将选举前 90 天内以深度伪造等方式蓄意传播意图伤害选举人以及影响选举结果的“数字化”行为定为犯罪。加利福尼亚州 2024 年 9 月颁布了三项新法案来遏制虚假信息和欺骗性选举内容的传播。纽约州 2023 年 5 月发布《政治人工智能免责声明（PAID）法案》，要求使用合成媒体发布政治信息的主体声明其通过人工智能生成并保存此类使用记录。佛罗里达州、密歇根州等州对在政治广告中使用人工智能行为进行了立法规范。此外，《科罗拉多州人工智能法》《犹他州人工智能政策法案》等州层面的人工智能立法主要规范人工智能应用过程中的消费者保护，加利福尼亚州的《前沿人工智能模型

安全创新法案》则聚焦于促进人工智能发展和监管人工智能带来的重大公共风险。

（三）新加坡：以技术发展为导向，实行软法先行模式

新加坡位于马来半岛南端，其陆地面积较小，自然资源有限，希望抓住人工智能发展机遇，建立强大的数字基础设施，营造支持创新的商业环境，吸引高质量人才和投资，从而实现发展和繁荣。英国媒体 Tortoise Media2023 年发布的《全球人工智能指数》显示，新加坡人工智能投资、创新和应用水平位居世界第三，仅次于美国和中国。新加坡以推动技术发展为导向，采取“软法先行”的模式，未出台强制性立法，而是采取“原则+指南+工具”的治理路径，通过发布非约束性的指南和建议实施人工智能治理，在最大程度推动人工智能技术创新和产业发展的同时引导人工智能安全规范发展。

在国家层面，通过出台人工智能战略分阶段实现人工智能发展和治理目标。2019 年 11 月，新加坡政府发布了首个国家人工智能战略（NAIS），提出了推动人工智能创新和应用的计划。2023 年 12 月，新加坡推出了国家人工智能战略 2.0 版本，承诺在基础设施建设、政府监管、人才培养、营商环境培育、数据安全和隐私保护等方面采取行动，计划在未来三到五年内，推进人工智能发展和价值创造，实现人工智能规范化应用，从而提升新加坡的经济和社会潜力。

在信息通信领域，研究发布人工智能治理相关模型框架、应用指南和实用工具。2019 年 1 月，新加坡信息通信媒体发展局和个人数据保护委员会联合发布了亚洲首个人工智能监管模式框架，为人

工智能的应用提供道德指导原则。2020 年 1 月，更新发布第二版框架，同时发布了配套的《组织实施和自我评估指南》，帮助各类组织机构评估其人工智能治理实践与模型框架的一致性，还发布了《人工智能治理案例汇编》，为行业实践提供有效参考。2022 年 5 月，信息通信媒体发展局推出名为“AI Verify”的人工智能治理测试框架和工具包，通过标准化测试验证人工智能系统的性能是否符合国际公认的原则。2023 年中，新加坡通讯及新闻部长宣布成立 AI Verify 基金会，以支持 AI Verify 的发展和使用。新加坡还计划将其提升为国际范围内使用的测试框架和工具包，并为国际标准的制定作出贡献。2024 年 5 月，AI Verify 基金会正式发布《生成式人工智能治理模型框架》，涵盖问责、数据、可信研发和部署、事件报告、测试和保证、安全、内容来源、安全与对齐研发、人工智能促进公益 9 个治理维度。

在金融领域，强化政企合作推进人工智能治理。新加坡金融管理局作为新加坡的中央银行和综合金融监管机构，率先在人工智能治理方面采取行动。2018 年，金融管理局与金融行业共同创建了一套“FEAT 原则”，即公平（Fairness）、伦理（Ethics）、问责（Accountability）和透明（Transparency），从而指导人工智能的负责任使用。2019 年，金融管理局宣布与金融行业合作创建“Veritas 框架”，指导金融机构将 FEAT 原则纳入其人工智能和数据分析驱动的解决方案中，旨在为金融部门的人工智能采用建立一个可信赖的环境。

在医疗卫生领域，发布指南提升患者对人工智能技术的信任。

作为亚太地区最大的医疗设备出口国之一，新加坡拥有成为领先医疗技术中心所需的强大制造业基础。近年来，新加坡日益面临医疗保健成本上升、慢性病负担加重、医疗人员短缺等问题，人工智能等创新医疗技术的发展为应对挑战提供了重要机遇。新加坡 2019 年发布的国家人工智能战略将医疗卫生领域作为人工智能发展的五大重点领域之一。通过促进人工智能研究投资、培养人才、建立配套数字基础设施等措施，加快人工智能创新和运用，有利于提高疾病检测效率，同时加强对慢性病分析预测和干预。为了强化医疗卫生领域的人工智能治理，2021 年 10 月，新加坡卫生部发布了《医疗人工智能指南》，促进患者提高在医疗过程中使用的人工智能信任并保障患者信息安全，强化了对人工智能医疗设备的监管和指引。

为强化人工智能网络安全水平，发布系列指南应对人工智能系统风险。2024 年 10 月 15 日，新加坡网络安全局（CSA）制定并发布了《人工智能系统安全指南》及其配套的《人工智能系统安全配套指南》，指导系统所有者在人工智能的整个生命周期内确保安全。这两个指南将有助于保护人工智能系统免受供应链攻击等传统网络安全风险和对抗性机器学习等新风险的影响。指南主要面向人工智能系统开发人员、人工智能系统运营人员、网络安全从业人员三类主体，针对规划与设计、开发、部署、运营与维护、数据和模型的终止处理等人工智能系统生命周期的五个重要环节提供安全指引。

（四）中国：统筹发展和安全，实行融合立法和场景立法模式

我国在人工智能立法的价值导向方面体现出统筹发展和安全的

理念。我国在不断完善现有法律法规的同时，不断推进人工智能治理的立法进程，2017 年，国务院发布的《新一代人工智能发展规划》提出了“三步走”的立法规划，即到 2020 年实现“部分领域的人工智能伦理规范和政策法规初步建立”，到 2025 年“初步建立人工智能法律法规、伦理规范和政策体系”，到 2030 年“建成更加完善的人工智能法律法规、伦理规范和政策体系”。在立法模式的选择方面，我国的科技立法一般体现出“小快灵”的特点，“小”即调整对象精准，“快”即针对突出问题及时立法，“灵”即增强立法的针对性、实效性。虽然目前统一的人工智能立法尚未出台，但我国将人工智能安全发展相关原则融入《个人信息保护法》等现有立法，还针对生成式人工智能、深度合成等场景出台了专门的部门规章，为人工智能的安全发展提供了及时的法律回应和依据，也体现出我国科技立法的“小快灵”特点。

在现有立法中融入人工智能安全发展的原则性规定以顺应人工智能技术的发展趋势。随着人工智能技术的不断发展，人工智能与各领域融合应用的程度不断加深，我国陆续在相关法律法规中融入人工智能安全发展的相关原则。一方面鼓励人工智能在相关领域的融合创新和技术赋能，例如《中华人民共和国突发事件应对法》《中华人民共和国基本医疗卫生与健康促进法》《中华人民共和国中小企业促进法》《国际邮轮在中华人民共和国港口靠港补给的规定》《未成年人网络保护条例》《无人驾驶航空器飞行管理暂行条例》《中华人民共和国海关风险管理办法》《网络暴力信息治理规定》

等，这些法律法规主要是在法条中加入鼓励采用人工智能技术的相关内容。另一方面强化对人工智能的安全管理，例如《中华人民共和国个人信息保护法》《网络数据安全管理条例》等，主要是强化对人工智能技术的规制以及施加提供人工智能服务的数据处理者义务。

表 2 我国现有法律法规中人工智能安全发展相关内容

名称	公布时间	人工智能安全发展相关规定
法律		
中华人民共和国突发事件应对法(2024 修订)	2024 年 6 月 28 日	第五十六条 国家加强应急管理基础科学、重点行业领域关键核心技术的研究，加强互联网、云计算、大数据、人工智能等现代技术手段在突发事件应对工作中的应用，鼓励、扶持有条件的教学科研机构、企业培养应急管理人才和科技人才，研发、推广新技术、新材料、新设备和新工具，提高突发事件应对能力。
中华人民共和国个人信息保护法	2021 年 8 月 20 日	第六十二条 国家网信部门统筹协调有关部门依据本法推进下列个人信息保护工作: …… （三）针对小型个人信息处理者、处理敏感个人信息以及人脸识别、人工智能等新技术、新应用，制定专门的个人信息保护规则、标准; ……
中华人民共和国基本医疗卫生与健康促进法	2019 年 12 月 28 日	第四十九条 国家推进全民健康信息化，推动健康医疗大数据、人工智能等的应用发展，加快医疗卫生信息基础设施建设，制定健康医疗数据采集、存储、分析和应用的技术标准，运用信息技术促进优质医疗卫生资源的普及与共享。 ……
中华人民共和国中小企业促进法	2017 年 9 月 1 日	第三十三条 国家支持中小企业在研发设计、生产制造、运营管理等环节应用互联网、云计算、大数据、人工智能等现

名称	公布时间	人工智能安全发展相关规定
		代技术手段，创新生产方式，提高生产经营效率。
行政法规		
网络数据安全管理条例	2024 年 9 月 24 日	第十九条　提供生成式人工智能服务的网络数据处理者应当加强对训练数据和训练数据处理活动的安全管理，采取有效措施防范和处置网络数据安全风险。
国际邮轮在中华人民共和国港口靠港补给的规定	2024 年 4 月 22 日	第五条　国务院有关部门和有关县级以上地方人民政府结合实际情况，综合运用大数据、人工智能等技术手段对靠港补给物资实行分类管理，完善不同种类物资的通关、仓储等管理措施，推动国际邮轮物资供应保障中心建设。
未成年人网络保护条例	2023 年 10 月 16 日	第二十六条 …… 网络产品和服务提供者应当建立健全网络欺凌信息特征库，优化相关算法模型，采用人工智能、大数据等技术手段和人工审核相结合的方式加强对网络欺凌信息的识别监测。
无人驾驶航空器飞行管理暂行条例	2023 年 5 月 31 日	第五条　国家鼓励无人驾驶航空器科研创新及其成果的推广应用，促进无人驾驶航空器与大数据、人工智能等新技术融合创新。县级以上人民政府及其有关部门应当为无人驾驶航空器科研创新及其成果的推广应用提供支持。 ……
部门规章		
中华人民共和国海关风险管理办法	2024 年 7 月 30 日	第四条　海关加强大数据、人工智能等现代科学技术在风险管理中的应用，推进风险信息共享，依托信息化系统开展风险处置，统一下达指令，提升风险管理智能化水平。
网络暴力信息治理规定	2024 年 6 月 12 日	第十二条　网络信息服务提供者应当在国家网信部门和国务院有关部门指导下细化网络暴力信息分类标准规则，建立健全网络暴力信息特征库和典型案例样本库，采用人工智能、大数据等技术手段和人工审核相结合的方式加强对网络暴力信

名称	公布时间	人工智能安全发展相关规定
		息的识别监测。
电子档案管理办法	2024 年	第三十五条 组织机构应当积极利用人工智能等信息技术，开展编研、展览和建设专题数据库等工作，不断开发档案数字资源。
应急管理行政裁量权基准暂行规定	2023 年 11 月 1 日	第三十四条 推进应急管理行政执法裁量规范化、标准化、信息化建设，充分运用人工智能、大数据、云计算、区块链等技术手段，将行政裁量权基准内容嵌入行政执法信息系统，为行政执法人员提供精准指引，有效规范行政裁量权行使。
农业综合行政执法管理办法	2022 年 11 月 22 日	第三十一条 县级以上人民政府农业农村主管部门应当依托大数据、云计算、人工智能等信息技术手段，加强农业行政执法信息化建设，推进执法数据归集整合、互联互通。 ……
企业环境信息依法披露管理办法	2021 年 12 月 11 日	第二十四条 生态环境主管部门应当会同有关部门加强对企业环境信息依法披露活动的监督检查，及时受理社会公众举报，依法查处企业未按规定披露环境信息的行为。鼓励生态环境主管部门运用大数据分析、人工智能等技术手段开展监督检查。
医疗器械临床使用管理办法	2021 年 1 月 12 日	第七条 卫生健康主管部门应当逐步完善人工智能医疗器械临床使用规范，鼓励医疗机构加强人工智能医疗器械临床使用培训。

针对重点领域的人工智能应用制定场景化规定和指导文件以推进人工智能安全治理。统一的人工智能立法的制定出台往往需要经历较长的周期，而在缺乏统一法律规范的阶段，我国采取“急用先行”的思路，采用“小切口”的暂行法，针对不同场景下的人工智能应用制定管理规定和指导性文件，以弥补监管空白。针对生成式人工智能应用场景，国家互联网信息办公室联合多部门于 2023 年 7

月 10 日发布《生成式人工智能服务管理暂行办法》，该部门规章以《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》《中华人民共和国科学技术进步法》等法律为上位法依据，聚焦违法信息、算法歧视、知识产权、人格权与个人信息保护、未成年人保护等问题，强调了生成式人工智能服务提供者的法律责任和义务。在信息服务与舆论传播场景下，国家互联网信息办公室、工业和信息化部、公安部 2022 年 11 月 25 日联合发布了《互联网信息服务深度合成管理规定》，规定深度合成服务提供者需要落实信息安全主体责任，规范相关技术发展，对“具有舆论属性或者社会动员能力”的深度合成技术进行“备案和变更、注销备案手续”“开展安全评估”等事前风险监管。国家互联网信息办公室、工业和信息化部、公安部、国家市场监督管理总局于 2021 年 12 月 31 日联合发布《互联网信息服务算法推荐管理规定》，明确了算法推荐服务提供者的信息服务规范和用户权益保护要求，要求具有舆论属性或者社会动员能力的算法推荐服务提供者履行备案义务。在司法应用场景下，最高人民法院于 2022 年 12 月 9 日发布了《关于规范和加强人工智能司法应用的意见》，着重从人工智能为司法工作提供全方位智能辅助支持、显著减轻法官事务性工作负担、有效保障廉洁司法等角度，明确了人工智能司法应用的主要场景，强调“无论技术发展到何种水平，人工智能都不得代替法官裁判，人工智能辅助结果仅可作为审判工作或审判监督管理的参考，确保司法裁判始终由审判人员作出，裁判职权始终由审判组织行使，

司法责任最终由裁判者承担”。在劳动者保护场景下，国家市场监督管理总局、人力资源社会保障部等多部委在《关于落实网络餐饮平台责任切实维护外卖送餐员权益的指导意见》中提出，对外卖骑手的考核应以“算法取中”等方式替代“最严算法”，合理确定订单数量、准时率、在线率等考核要素，适当放宽配送时限。

三、全球人工智能立法特点和趋势

随着人工智能技术的不断发展和各国人工智能治理实践的广泛推进，人工智能监管需求日益明确，各国人工智能立法过程中安全和发展的权重日益平衡，软法硬法协同治理、多元立法模式融合互补、人工智能立法全球化等趋势日益凸显。

（一）立法逻辑导向日趋融合，安全和发展并重成为全球共识

人工智能已经成为推动国际体系转变、影响国际关系走向的关键技术力量。一方面，人工智能发展广泛影响国家治理和社会生活，成为国家间实力竞争的关键指标。人工智能发展对生产力的提升、人类生活方式的改变具有巨大的推动作用。人工智能技术正在被国家、企业、个人等不同层次的行为体广泛使用，大模型广泛地应用于国家治理与社会生活的各个领域，由此带来的全局性影响也将进一步凸显。技术是各国提升硬实力的重要基础，也是改变国际格局力量对比的动力源之一，人工智能作为引领全球科技革命和产业变革的战略性技术，已经成为大国竞争博弈的重要靶标。另一方面，人工智能伴生安全风险，已成为各国面临的共同挑战。一是技术风险。人工智能系统的复杂性可能导致难以预测和控制的错误，进而

引发安全事故，同时也可能出现新的攻击手段，对人工智能系统的安全性构成威胁。二是伦理风险。人工智能的决策可能受到偏见和歧视的影响，导致不公平的结果。人工智能的自主性和智能水平不断提高，可能引发关于责任归属和道德决策的争议。三是社会风险。人工智能技术的广泛应用可能导致就业岗位的减少和社会结构的变革，人工智能的军事应用还可能加剧国际社会的冲突和战争风险。

发展和安全在各国人工智能立法价值选择中呈现融合发展趋势，各国均致力于探索安全和发展的平衡。发展和安全的天平偏向任何一方都可能引发一连串的不良后果，各国在人工智能的法律规制方面将很难再从发展和安全中进行具有倾向性的价值选择。高度重视技术创新、注重市场竞争的美国现阶段虽然在人工智能立法方面体现出技术发展导向，但是从其出台的总统行政令和备忘录、各州立法以及国会议员提出的法案来看，其对人工智能安全治理的重视程度正在不断提升。拜登签署的《安全、可靠和值得信赖地开发和使用人工智能》行政令将安全放在突出位置，其提出的八项重点行动前五项都旨在提升人工智能应用的安全性。2024 年 10 月 24 日，美国白宫发布了美国史上第一份关于人工智能的国家安全备忘录，同步发布“促进国家安全中人工智能治理和风险管理的框架”，指导建立风险管理、评估、问责和透明度机制。美国国会议员也多次提出人工智能监管相关法案，例如《人工智能和生物安全风险评估法案》《人工智能国家安全法案》《人工智能标签法案》等，体现出美国国会对于人工智能安全治理的关注。欧盟虽然偏向于强监管的

思路，出台统一的人工智能立法对人工智能应用进行全面规制，但其《人工智能法案》也将促进人工智能技术创新、支持初创企业发展作为重要内容，体现出在确保安全的同时促进发展的思路。而中国始终秉持统筹发展和安全的战略思维，在推进人工智能产业发展的同时注重通过融合立法和场景立法强化人工智能安全治理。可见，各国在人工智能立法规制的道路上正在不断强化发展和安全的平衡，最终引导人工智能走向实现发展和安全双轮驱动的可持续发展道路。

（二）软法先行成为普遍选择，软法硬法协同治理是重要方向

从全球来看，软法先行已经成为各国人工智能治理的普遍选择。美国亚利桑那州立大学 2021 年发布的一份关于软法治理报告的数据显示，在全球范围内确定了 634 个人工智能软法项目，其中 90%是 2017 年至 2019 年提出的。虽然各大经济体都在推进人工智能立法进程，但在人工智能治理的初期阶段，大多数国家都会通过制定原则框架、标准指南等软法指导和规范人工智能的安全应用，软法规范仍然是时下治理人工智能的主要法治工具。例如，欧盟在制定统一的《人工智能法案》之前，就已出台了《算法的可问责和透明的治理框架》《可信人工智能伦理指南》等一系列指导性文件；美国国防部、劳工部等部门为贯彻相关总统行政令要求，陆续发布了《人工智能原则：美国国防部关于人工智能道德使用的建议》《雇主使用人工智能的原则和最佳实践》等一系列人工智能相关的原则建议和指南；我国在出台人工智能相关硬性规定的同时也出台了诸如《关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导

意见》《国家人工智能产业综合标准化体系建设指南》等指导性文件。

从趋势上看，软硬兼施和多元主体协同共治将促进治理效能最大化。前文提到软法和硬法具有各自的优势和局限，随着人工智能技术更新迭代加快，监管和治理需求日益提升，探索软硬兼施的治理工具从而在灵活响应技术快速变革的同时促进人工智能的安全应用，已成为各国人工智能立法的重要趋势。各国在研究出台人工智能相关硬性法律规定的同时，政府机构、行业协会、企业主体等均在总结行业实践经验的基础上，出台一系列的指南、框架、协议等，为动态发展变化的人工智能技术提供指引。实现人工智能领域的软法硬法协同治理需要政府、行业、学术界和公众等多方面的共同努力。从我国来看，我国软法硬法协同治理还处于探索与实践过程中，还存在软法硬法衔接不畅、软法制定缺乏监管、软法执行缺乏保障等问题。未来，如何进一步探索软法硬法协同和多元共治机制，提升软法制定的科学性和合理性，真正发挥软法灵活治理、柔性监管的效力，同时优化软法与硬法的衔接和融合机制，将成为值得探讨的重点问题。

（三）统一立法成为大势所趋，多元立法模式呈融合互补趋势

从长远看，制定统一、综合性人工智能立法或将成为国际主流。虽然目前只有欧盟正式出台了首部人工智能综合法案，但从各国立法实践和趋势上看，制定综合性立法已是大势所趋。美国在短期内或不会进行联邦层面的综合性人工智能立法，但是其州层面的立法

一开始借鉴欧盟的立法实践，采取风险分级的思路对高风险的人工智能进行立法规制，在未来人工智能风险持续显露的情况下，不排除美国为有效应对风险，顺应全球的立法趋势，从州层面开始推进人工智能综合立法。从我国来看，近年来我国积极推进人工智能立法进程，国务院 2023 年、2024 年连续两年将人工智能法草案列入年度立法计划，人工智能领域的专家学者也集智发布了《中华人民共和国人工智能法（学者建议稿）》，体现出我国顺应人工智能技术发展潮流，推进人工智能综合统一立法的立场。

同时，场景立法、融合立法等多元立法模式作为重要补充也呈现同步发展的趋势。例如，欧盟在推出统一的人工智能法案的同时，还在产品责任和侵权责任等方面进行了立法实践，体现出开展融合立法和场景立法的趋势。在产品责任方面，欧盟委员会在 2022 年提出了新的产品责任指令草案，旨在对已经实施了 40 年的《产品责任指令》进行更新，以适应数字时代的新变化。根据这一草案，软件和独立人工智系统将首次被纳入“产品”的范畴，这些产品的制造者将对具有缺陷的产品承担严格责任。而非独立存在的人工智能系统如果被整合到产品中，或者对于产品发挥功能不可或缺，其制造者也需要承担严格产品责任。在侵权责任方面，欧盟委员会在 2022 年提出了《人工智能责任指令（草案）》，对基于“过错原则”的人工智能侵权进行了规定。这一草案可以视为《人工智能法》的侵权法配套立法，聚焦高风险人工智能的侵权责任，特别是高风险人工智能侵权中的举证责任。此外，前文提到美国、新加坡、中国等

尚未出台统一人工智能立法的国家，融合立法、场景立法等是规制人工智能安全应用的重要法律规范，未来也将不断发展。

（四）国际治理规则持续构建，人工智能立法全球化趋势显著

随着人工智能技术的全球传播和应用，其带来的风险和挑战已跨越国界成为全球性问题。一是人工智能技术的跨国应用使得风险具有全球性。一个国家的人工智能系统出现问题或漏洞，其影响可能会迅速扩展到全球范围。例如，一个重要人工智能系统遭受黑客攻击可能会泄露大量敏感数据，这些数据可能涉及多个国家的企业和个人。二是人工智能技术的全球供应链也增加了风险的跨国界性。人工智能技术的研发、生产、应用等环节往往涉及多个国家和地区。这意味着，如果一个环节出现问题，其影响可能会迅速传递到整个供应链，进而影响全球市场的稳定和消费者的权益。三是人工智能技术的全球竞争加剧风险的跨国界性。各国都在竞相发展人工智能技术以期望在全球市场中占据领先地位。这种竞争或导致一些国家采取过于激进或冒险的策略，从而增加人工智能技术的风险。

为有效应对人工智能带来的全球化风险，国际社会正逐步强化人工智能全球规则构建。一方面，联合国积极主导人工智能国际治理规则建设。2024 年 3 月 21 日，联合国大会通过首个关于人工智能的全球决议，呼吁推动开发“安全、可靠和值得信赖的”人工智能系统，以促进可持续发展。决议强调需要制定人工智能系统标准，鼓励联合国会员国和其他利益攸关方制定和支持有利于开发安全可靠的人工智能系统的监管和治理办法及框架，敦促会员国和其他利

益攸关方采取行动，与发展中国家合作并向其提供援助。另一方面，美欧等经济体积极推动国内规则国际化。美欧等发达经济体普遍将掌握人工智能规则制定主动权作为保持其在技术与产业领域优势地位的重要一环，尝试从国内、国际两个层面入手进行机制完善和规则构建。美国将“推动与盟友的合作”“建设新兴技术的世界领导者地位”作为基本立场，依托其主导的北约、美英澳三边安全伙伴关系、美日印澳四方安全对话等国际多边联盟，联合印太、欧盟等关键盟友，进行技术合作并推行共同的人工智能技术标准。欧盟《人工智能法案》确立了广泛的适用范围和域外效力，所有在欧盟境内投入市场或投入使用的人工智能系统均适用该法案，人工智能系统的供应商、分销商、运营商等相关主体，均为本法案的规制对象，即使系统供应商并未在欧盟设址，欧盟也可以通过该法案将其他国家的大型科技公司纳入监管范畴。

国际法将成为人工智能重要的法律规制手段。利用国际法规制人工智能是应对技术发展挑战、保障人权与基本自由、促进国际合作与交流以及推动技术创新与可持续发展的必要手段。当前，人工智能的国际法规制总体不成熟，尚处于初级阶段，内容上以伦理规则为主，形式上以非政府主体制定的软法为主直接规制规则缺位。未来，国际法层面的人工智能监管规则将持续完善，一方面将有效调整人工智能带来的跨国界争议，另一方面将有效约束各国依托人工智能技术开展的各类违法犯罪行为。目前，在国际公约方面已经取得了积极进展。2024 年 9 月 5 日，美国、英国和欧盟等签署了欧

洲委员会制定的《人工智能、人权、民主和法治框架公约》（以下简称“公约”），该公约是全球首个具有法律约束力的人工智能国际公约，旨在确保人工智能系统生命周期内的活动完全符合人权、民主和法治，同时有利于技术进步和创新。未来，直接规制人工智能的国际法或将陆续出台，人工智能国际规则体系将持续完善，不断推动人工智能的良法善治。

四、启示建议

参考各国家和地区人工智能立法实践，顺应全球人工智能立法趋势，应从我国国情出发，开展人工智能治理制度设计，建立多层次人工智能法律体系，向世界展现人工智能治理和立法的中国力量。

（一）持续完善顶层设计，形成多元共治的治理格局

一是加快推进人工智能综合立法进程。深入剖析全球人工智能的发展格局与我国人工智能的发展现状和价值体系，贯彻统筹发展和安全的立法导向，深入推进科学立法、民主立法、依法立法，借鉴全球人工智能立法和治理的实践经验，汇聚各方面力量和智慧加快人工智能立法进程，出台符合我国国情的人工智能综合立法，既满足当前人工智能治理的紧迫需要，又要为人工智能技术发展留足空间。有专家学者建议将我国人工智能法定位为人工智能治理的提纲挈领制度，参考《民法典》的立法经验，采用“总则式”的立法体例，明确人工智能技术、产品与服务必须遵循的基本原则，基本法律制度应遵循的一般性规则，同时为各行业领域开展人工智能治理和未来人工智能立法完善预留“接口”。

二是建立软硬协同的多层次人工智能法律体系。强化人工智能立法与《中华人民共和国国家安全法》《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等相关现行法律的立法衔接，避免立法重复和监管空白。未来我国人工智能综合立法出台后，应加快制定行政法规、部门规章等配套规则，并与生成式人工智能、算法、深度合成等现有管理规定做好衔接，必要时针对医疗、金融等重点应用场景出台专项规定。在健全硬法规则的同时完善软法治理，将软法作为促进硬法实施的重要补充和指引，规范软法制定流程，在相关国家标准和指南文件制定过程中援引法律法规明确的术语和指标，强化软法实施效果监督，实现人工智能领域软法与硬法治理的贯穿融合。

三是确立多元人工智能治理格局。治理主体多元化是实现科技立法与决策民主化的必然要求，要充分发挥政产学研用等多方主体作用，加强行业自律，制定行业规范和准则，发布行业最佳实践案例与合规指引，有效指导企业合规建设和使用者权益保护，鼓励人工智能服务提供者、使用者等市场参与者共享人工智能安全监测技术工具和手段，构建协同良好的人工智能治理生态。在治理对象方面，学界普遍认可生成式人工智能产业呈现“基础模型-专业模型-服务应用”的分层业态，建议构建契合产业发展形态的分层治理机制，在基础模型层以发展为导向促进数据要素供给、算力基础设施和模型建设，在专业模型层构建细分领域数据流通交易机制，在服务应用层建立人工智能应用问责机制。

（二）划定安全治理边界，确立风险分级的管理机制

一是建立动静结合的人工智能风险监测识别体系。有效识别人工智能风险是开展风险分级治理的基本前提。一方面，建议识别出人工智能风险较大的行业领域、应用广泛的通用人工智能系统和对生命、自由、人格尊严等重要人身权益影响重大的人工智能系统，适时推出重点行业领域和人工智能系统清单，对这些重点行业领域及重要人工智能系统展开重点治理。另一方面，建议建立人工智能风险的动态监测机制，从国家层面部署全国人工智能安全监测体系，建立风险监测平台，人工智能服务提供者及时报告服务提供过程中的安全事件，定期开展人工智能风险分析，研判风险走向和未来趋势，为长远开展人工智能治理和监管提供支撑。

二是合理确定人工智能风险分级原则和框架。依据风险进行分类分级管理是技术治理规则设定中普遍采用的原则。欧盟《人工智能法案》确立了风险分级监管框架，对各国人工智能立法产生了深远影响，人工智能风险分级管控已成为各国普遍采取的治理工具。由于各国在技术发展特点、监管环境、立法话语体系等方面存在差异，在具体的风险分级框架上应当结合本国国情进行差异化的制度设计。目前在人工智能领域确立风险管理原则已成为行业共识，但如何确定风险分级方式和级别划分，学界仍在探讨。例如有学者认为不宜在现阶段技术产业发展走向尚不十分明朗的情况下划定不同级别，而应划定“红线底线”以满足当下监管需求；还有学者建议将人工智能应用分为高风险、中风险、低风险三种类型。我国在《生

成式人工智能服务管理暂行办法》等文件中已体现了一定的风险分级的思路，建议在更高层级立法中确立风险分级的法律原则，并充分研判我国人工智能技术发展程度，在风险识别监测和分析研判的基础上，合理设计风险分级治理框架。

三是明确人工智能应用责任归属。在民事责任领域风险与侵权归责关联紧密，风险关系着过错责任和无过错责任[1]的判断。欧洲议会在《关于人工智能民生责任制度的决议》中建议对高风险人工智能系统的运营商施加无过错责任；对于未列入高风险人工智能附件中的人工智能系统，运营商承担过错责任。目前，我国现有的侵权责任规则无法很好地适应人工智能的技术特征，还需要根据人工智能的应用实际在立法中明确责任归属原则，以对人工智能的安全应用和责任划分提供法律指引。多位学者建议对不同风险级别的人工智能适用不同的归责原则，即对于高风险人工智能，适用无过错责任；对于有限风险人工智能，适用过错推定责任；对于低风险人工智能，适用过错责任。

（三）强化国际交流对话，提出全球治理的中国方案

一方面，积极参与国际规则制定。人工智能领域的国际规则同样可分为硬性规则和软性规则，硬性国际规则包括国际技术标准、接口协议、合同范本等，主要用于设置基础性技术底线和标准、确

[1] “过错责任”“无过错责任”“过错推定”是民法中侵权行为的主要归责原则，“过错责任”是指以行为人的过错作为承担民事责任的基本条件。按过错责任原则，行为人仅在有过错的情况下，才承担民事责任。“无过错责任”是指当事人实施了加害行为，虽然其行为无过错，但根据法律规定仍应承担民事责任。“过错推定”是指受害人若能证明其受损害是由行为人所造成的，而行为人不能证明自己对造成损害没有过错，则法律就推定其有过错并就此损害承担侵权责任。

保技术发展可控和向善；软性国际规则包括围绕特定议题形成的发展指南、原则共识、倡议协议等，既包括政府层面达成的共识，也包括由企业、非政府组织、科学家等民间力量推动形成的国际合作与规则。参与国际规则制定是我国承担大国责任，彰显良好国际形象的契机。要积极参与人工智能国际规则制定，支持联合国在国际合作中发挥中心作用，尊重国际法对人工智能的规范调整作用，明确我国安全发展的基本立场与态度，对于硬性国际规则，要强化国内规则与国际规则的制度衔接，积极参与国际标准制定，推动我国人工智能标准与国际标准体系接轨；对于软性国际规则，要主动发声，输出中国智慧，提升话语权，争取更多的国际合作资源，适时牵头和主导形成更多人工智能治理的相关指南、倡议等国际规则。

另一方面，强化人工智能治理国际合作。要探索与不同国家或国际组织开展人工智能国际合作的路径，对于美国等与我国竞争博弈较为激烈的国家，建议通过一轨或二轨外交等手段，积极开展对话，探索人工智能规则制定方面的共同利益和合作空间，共同推动人工智能健康有序发展。对于东盟等具有国际影响力且与我国具有合作基础的区域性大国或国际组织，建议依托“一带一路”等国际多边合作机制在人工智能治理领域深入合作，广泛凝聚共识，充分关注其他国家的人工智能治理诉求和治理逻辑，牵头发起或参与人工智能治理相关规则和倡议。除了政府层面的对话合作，还要推动和鼓励科学家群体、科技企业、智库、高校、行业协会等民间主体与全球范围内的同行开展深入交流合作，提升我国人工智能技术人

才参与国际规则制定的能力，为发出中国声音提供良好的渠道和平台，扩大影响力。

参考文献

[1]申卫星，张凌寒，周辉，等.推进有未来感的人工智能立法[J].探索与争鸣，2024，(10):4+177.

[2]张凌寒.中国人工智能立法需凝聚“总则式”立法共识[J].探索与争鸣，2024，(10):9-13+177.

[3]张凌寒.人工智能立法的理论阐释[J].华东政法大学学报，2024，27(05):5.

[4]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(西北政法大学学报)，2024，42(03):3-17.DOI:10.16290/j.cnki.1674-5205.2024.03.009.

[5]陈吉栋.以风险为基础的人工智能治理[J].法治研究，2023，(05):52-60.DOI:10.16224/j.cnki.cn33-1343/d.2023.05.001.

[6]薛澜.新兴科技领域国际规则制定：路径选择与参与策略[J].人民论坛·学术前沿，2023，(19):15-21.DOI:10.16619/j.cnki.rmltxsqy.2023.19.002.

[7]郑志峰.人工智能应用责任的主体识别与归责设计[J].法学评论，2024，42(04):123-137.DOI:10.13415/j.cnki.fxpl.2024.04.011.

[8]鲁传颖.人工智能：一项战略性技术的应用及治理[J].人民论坛，2024，(01):72-75.

[9]曾雄，梁正，张辉.人工智能软法治理的优化进路：由软法先行到软法与硬法协同[J].电子政务，2024，(06):96-107.DOI:10.16582/j.cnki.dzzw.2024.06.008.

[10]薛澜，贾开，赵静.人工智能敏捷治理实践：分类监管思路与政策工具箱构建[J].中国行政管理，2024，40(03):99-110.DOI:10.19735/j.issn.1006-0863.2024.03.10.

[11]周汉华.论我国人工智能立法的定位[J].现代法学，2024，46(05):17-34+217.

[12]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].西安交通大学学报(社会科学版)，2024，44(04):132-143.DOI:10.15896/j.xjtuskxb.202404013.

[13]宋华琳.人工智能立法中的规制结构设计[J].华东政法大学学报，2024，27(05):6-20.

[14]丁晓东.全球比较下的我国人工智能立法[J].比较法研究，2024，(04):51-66.

[15]郑志峰.人工智能产品责任的立法更新[J].法律科学(西北政法大学学报)，2024，42(04):3-17.DOI:10.16290/j.cnki.1674-5205.2024.04.014.

[16]黄静怡.“分级分类”与“契约”风险治理并行的人工智能监管制度构建——以欧盟《人工智能法案》为分析对象[J].海南金融，2024，(02):76-85.

[17]司伟攀.欧盟和美国人工智能立法实践分析与镜鉴[J].全球科技经济瞭望，2023，38(07):6-14.

[18]申卫星.全球人工智能法治发展与中国立场.清华大学人工智能国际治理研究院，https://mp.weixin.qq.com/s?__biz=MzU4MzYxOT

IwOQ==&mid=2247511339&idx=2&sn=fb0a3d03935222686180d9c4f6e7677b&chksm=fc62cfeb8bc4565eaf1a60d755db1a102cc966102dea895b35ac0e4cd03ec1ceefa367874641&scene=0&xtrack=1

[19]鲁传颖，田丽，封帅，周亦奇，张璐瑶，王玉柱，王天禅.建构人工智能发展的国际规则——趋势、领域与中国角色.上海国际问题研究院，清华大学战略与安全研究中心，https://ciss.tsinghua.edu.cn/upload_files/atta/1698891632222_19.pdf

[20]新加坡人工智能治理政策与法律体系解读，https://www.secrss.com/articles/67928

2024 年“工信安全智库”系列研究报告

报告编号	报告名称	发布时间
2024-lw-01	拜登政府以来美国网络安全政策举措走向及对我影响	2024 年 1 月
2024-dc-01	数实融合发展实施路径探索研究——基于深圳融合发展实践	2024 年 1 月
2024-dc-02	2023-2024 年我国工业互联网产融合作发展报告	2024 年 1 月
2024-yp-01	2023-2024 年度数字经济形势分析	2024 年 3 月
2024-yp-02	2023-2024 年度工业软件发展形势报告	2024 年 3 月
2024-yp-03	2023-2024 年度网络安全发展形势分析	2024 年 3 月
2024-dc-03	2023 年国外未来产业态势分析	2024 年 4 月
2024-dc-04	抢抓人工智能发展机遇，推进国产工业软件发展	2024 年 6 月
2024-dc-05	2023-2024 年我国百强互联网企业发展态势研究	2024 年 6 月
2024-dc-06	以开源开放促进大模型产业健康发展	2024 年 9 月
2024-dc-07	大模型背景下人工智能安全治理路径研究	2024 年 9 月
2024-yp-04	“十五五”时期我国数字经济高质量发展思路建议	2024 年 10 月
2024-dc-08	供应链金融数字化转型创新研究	2024 年 10 月
2024-dc-09	亚太地区数字营商环境发展报告	2024 年 11 月
2024-dc-10	2024 年前三季度我国电子信息产业投融资情况报告	2024 年 12 月
2024-dc-11	我国工业软件产业现状与地方典型做法研究	2024 年 12 月
2024-lw-02	全球人工智能立法的主要模式、各国实践及发展趋势研究	2024 年 12 月

联系人：王　蕊　　18612983082