

中国人民大学法学院 数字法学教研月报

2026年第3期（总第27期）

2026年5月26日



本期看点

【数字法治大事件】人工智能治理体系加速完善，《人工智能应用伦理安全指引1.0》正式发布，系统划定全链条伦理安全边界；三部门联合印发《智能体规范应用与创新发展的实施意见》，明确19个典型应用场景与分类分级治理框架。平台算法治理取得实效，《生活服务类平台算法负面清单（试行）》在订单分配、收入抽成等七大领域落地见效。理论建构与网络治理同步推进，张东刚发表《人工智能时代中国自主知识体系建构》重要文章；中央网信办全面部署短视频内容标注工作，多部门联动整治网络谣言与自媒体未规范标注乱象，《中国网络诚信发展报告（2026）》同步发布。

【研究动态】本期研究覆盖数字法学全领域核心议题：本期研究覆盖数字法学全领域前沿：人工智能治理聚焦智能体分层治理、大模型法律边界、生成式AI侵权责任与版权配置等核心议题；网络平台治理探讨超大型平台社会责任与数据污染治理；数字行政与智能司法研究智慧城市立法、自动化行政法律保留与数字法院审判权运行；财产权益保护围绕数据财产权、数据资产担保与商业秘密保护展开；个人信息保护深入分析

数据可携权、处理捆绑禁止与调取正当程序；公共数据开放研究共享法律责任、授权运营收益分配与国家所有的宪法检视，为数字法治建设提供多元理论支撑。

【教研活动】个人信息保护法治研讨会系统总结中国高水平保护实践；珞珈网络治理大讲堂反思数据三权分置制度；数字世界反不正当竞争法国际讲座、数字转型全球价值链重构国际研讨会搭建中外学术对话平台；华东政法大学研讨网络直播打赏法律问题，人大举办生成式人工智能治理专题研讨会；数字中国峰会大模型与涉外法治分论坛发布多项成果；《人工智能法》浙大版学者建议稿圆桌论坛共商立法大计，推动学科建设与实务深度融合。

【数字法评】

《建构中国自主的数字法学知识体系》，《民主与法制》周刊2026年第16/17期，作者：杨东。

《生成式人工智能服务提供者差异化注意义务的规范构造》，《数字法治》2026年第2期，作者：万勇、吴翔。

本期目录

数字法治大事件3	公共数据开放研究.....33
《人工智能应用伦理安全指引 1.0》发布.....3	
专家解读 推动人工智能创新发展行稳致远 筑牢伦理安全治理屏障.....5	
头条! 张东刚: 人工智能时代中国自主知识体系建构.....8	
《生活服务类平台算法负面清单(试行)》实施取得初步成效.....12	
国家网信办、国家发展改革委、工业和信息化部联合印发《智能体规范应用与创新发展实施意见》.....14	
《智能体规范应用与创新发展实施意见》答记者问.....18	
中国网络社会组织联合会正式发布《中国网络诚信发展报告(2026)》.....20	
中央网信办全面部署推进规范短视频内容标注工作.....20	
虚构政策、蹭炒热点、误导舆论 多部门联动整治网络谣言乱象.....21	
网信部门严管“自媒体”未规范标注信息来源行为.....22	
研究动态23	
人工智能治理.....23	
网络平台的治理.....26	
数字行政与智能司法.....26	
数字时代的部门法问题.....27	
数字法学基础理论研究.....30	
数字时代的财产权益保护.....30	
个人信息保护研究.....32	
	教研活动36
	“个人信息保护法治研讨会”在京举行.....36
	珞珈网络治理大讲堂第十八期——数据三权分置的反思和重构.....37
	Unfair Competition Law in the Digital World 讲座圆满举办.....38
	“数字转型: 互联世界中的全球价值链重构”国际学术研讨会在香港成功举办.....41
	华政顺利召开“网络直播打赏前沿法律问题”专题研讨会.....49
	中国人民大学人工智能治理研究院主办“生成式人工智能治理”专题学术研讨会.....50
	第九届数字中国建设峰会“大模型、人工智能体与涉外法治前沿分论坛”顺利召开.....54
	“人工智能外包的合同治理创新”讲座圆满举办.....56
	人大人工智能治理研究院主办大模型训练数据的使用规则研究.....57
	“AI 善治·新智新声”人工智能法治在线课堂暨硕博研习(春夏季)第三场圆满举办.....59
	第三届网络与信息法学学科建设论坛在福州成功举办.....61
	《人工智能法》浙大版学者建议稿圆桌论坛在杭州召开.....62
	数字法评66
	《建构中国自主的数字法学知识体系》.....66
	生成式人工智能服务提供者差异化注意义务的规范构造.....71

学术顾问: 王利明

编委会: 张新宝 丁晓东 王莹 张吉豫

编辑部: 阮神裕 毕坤阳 来唯希 吕昊然 邓语鑫 梁因格 王昊

联系方式: RUCdigitallaw@163.com

本期编辑: 梁因格

数字法治大事件

导言：

在人工智能加速融入经济社会运行、平台生态治理持续深化的背景下，我国数字法治建设正进一步迈向“智能化治理”与“平台秩序重塑”协同推进的新阶段。2026年5月，《人工智能应用伦理安全指引1.0》正式发布，并配套推出专家解读文章，对人工智能应用中的伦理风险、安全边界、责任体系与治理原则进行系统阐释，强调在推动人工智能创新发展的同时，建立覆盖研发、训练、部署、应用全过程的伦理安全治理机制。这标志着我国人工智能治理正在从以算法备案、生成式人工智能服务规范为核心的技术监管，逐步延伸至人工智能社会嵌入过程中的伦理秩序建构与风险预防体系完善。与此同时，国家网信办、国家发展改革委、工业和信息化部联合印发《智能体规范应用与创新发展的实施意见》，围绕数据安全、内容可信、行为边界和行业应用等问题提出系统性要求，体现出我国人工智能治理正加快形成“创新促进+风险治理”并重的路径，并推动智能体产业发展与治理规范同步完善。

值得关注的是，在制度治理持续深化的同时，人工智能时代的数字法学理论建构问题也受到高度重视。张东刚发表《人工智能时代中国自主知识体系建构》一文，强调应立足中国数字化发展实践与治理需求，加快构建具有中国特色、中国话语的自主知识体系，推动人工智能时代数字法治理论创新。这反映出我国数字法治建设正在从制度规范供给进一步延伸至数字治理基础理论与自主知识生产能力建设，数字法治领域的话语体系与理论体系建构正不断深化。

在平台算法治理领域，《生活服务类平台算法负面清单（试行）》实施取得初步成效，针对“大数据杀熟”、价格歧视、诱导消费等问题，通过负面清单机制推动平台在推荐逻辑、价格机制与用户权益保护之间实现更加透明和可审计的制度平衡，体现出算法治理正从原则倡导走向场景化约束，并进一步向数字消费和平台经营秩序领域延伸。

在网络生态治理方面，中央网信办全面部署推进规范短视频内容标注工作，重点强化生成合成内容、演绎摆拍、虚构情节等信息的显著标识要求。与此同时，多部门联动整治虚构政策、蹭炒热点、

误导舆论等网络谣言乱象，并加强对“自媒体”未规范标注信息来源行为的监管治理，推动信息来源可追溯、传播责任可识别、舆论生态可治理。相关治理举措也表明，网络内容治理正进一步向传播透明化与主体责任化方向深化。

此外，中国网络社会组织联合会正式发布《中国网络诚信发展报告（2026）》，围绕网络诚信建设、平台信用治理和数字社会信任机制建设进行系统总结，表明我国数字法治建设已从安全与风险治理进一步延伸至数字社会信任秩序塑造。

总体来看，从人工智能伦理安全治理到智能体规范发展，从自主知识体系建构到平台算法治理，从短视频内容标注到网络谣言整治与网络诚信建设，本月数字法治建设呈现出技术治理、平台治理、理论建构与网络生态治理协同推进的鲜明特征。这表明我国数字治理体系正在不断从技术合规与内容监管，迈向对智能社会运行秩序、平台责任体系与数字信任机制的整体性制度建构，在创新发展与风险防控之间不断寻求更加精细化的制度平衡。

《人工智能应用伦理安全指引1.0》发布

原载：“网信中国”微信公众号

5月19日，在2026年中国网络文明大会人工智能赋能网络文明建设分论坛上，全国网络安全标准化技术委员会（以下简称“网安标委”）发布了《人工智能应用伦理安全指引1.0》（以下简称《指引》）。

为进一步引导人工智能应用坚持以人为本、智能向善，推动人工智能应用相关方正确认识和妥善应对应用活动中的伦理安全影响，促进人工智能应用在规范有序、安全可控的轨道上健康发展，《指引》给出了人工智能应用伦理安全理念与原则，明确了人工智能应用开发、服务提供和应用使用等安全指引。

网安标委秘书处相关负责人表示，《指引》的发布，正是对引导人工智能应用尊重人的主体地位、维护公平正义、保障合法权益、促进社会信任，

推动人工智能技术始终朝着有益、安全、公平方向发展这一时代课题的积极回应，体现了人工智能治理坚持积极稳妥、开放包容、协同共治的实践导向，也体现了网络文明建设对人工智能时代技术向善、价值引领和秩序塑造的主动回应。

TC260

全国网络安全标准化技术委员会技术文件

TC260-005

人工智能应用伦理安全指引 1.0

Ethics-Safety Guidelines for Artificial Intelligence Applications 1.0

2026-05-19 发布

全国网络安全标准化技术委员会发布

摘要

近年来，人工智能不断加速发展，生成式人工智能、智能体等新技术及其应用层出不穷，人工智能对社会、生产、生活的重构性影响逐步显现。在人工智能积极、深刻地改变社会运行及个人生活方式并带来诸多便利的同时，也带来许多伦理安全挑战。

为进一步确保人工智能安全可控，统筹人工智能发展与安全，保障强化人工智能对国家经济、社会、生态等方面的持续推动作用，帮助人工智能应用相关方在各应用场景开展相关活动时，更好地兼顾发展、安全以及伦理各方面影响，本文件给出了人工智能应用伦理安全的理念与原则，提出了开展相关活动的基本要求，提供了各方实践的应用指引。本文件为原则性、参考性技术文件；涉及个人信息、自动化决策、内容标识、算法治理、知识产权等事项时，应与现行法律法规和部门规章协调适用。

目次

前言 1

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 人工智能应用伦理安全影响 1

5 伦理安全理念与原则 1

 5.1 伦理安全理念 1

 5.2 伦理安全原则 1

6 伦理安全指引 1

 6.1 通用指引 1

 6.2 应用开发指引 1

 6.3 服务提供指引 1

 6.4 应用使用指引 1

前言

本文件由全国网络安全标准化技术委员会（SAC/TC260）发布。

本文件起草单位：清华大学、中国电子技术标准化研究院、上海交通大学、四川大学、北京科技大学、阿里巴巴集团、华为技术有限公司、北京深度求索人工智能基础技术研究院有限公司等。

本文件主要起草人：薛澜、梁正、郝春亮、贾开、王姣、赵静、张妍婷、申卫星、朱旭峰、吕飞霄、庞祺敬、王净宇、曹雄、李芒、宋雨鑫、吴宗泽、傅宏宇、林秋诗、吴少卿、王晓薇、邵萌、薛晖、张荣、李妮莉、陈迪思等。

人工智能应用伦理安全指引

1 范围

本文件给出了人工智能应用伦理安全影响，提出了人工智能应用伦理安全理念与原则，规定了人工智能应用开发、服务提供和应用使用等安全指引。

本文件可为组织和个人开展人工智能应用活动提供指导，也可对相关主管部门、行业组织和有关机构推进人工智能伦理安全治理提供参考。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注明日期的引用文件，仅该日期对应的版本适用于本文件；不注明日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 25069 信息安全技术 术语

GB/T 41867 信息技术 人工智能 术语

3 术语和定义

GB/T 25069和GB/T 41867界定的以及下列术语和定义适用于本文件。

3.1

人工智能应用 artificial intelligence application

具有功能特性的人工智能的使用，该人工智能在利益相关方场景中运行以实现预期结果。

[来源：ISO/IEC 5339:2024(en), 3.1]

3.2

人工智能应用伦理安全（简称“伦理安全”） ethics-safety of artificial intelligence applications

开展人工智能应用活动，处理人工智能与人、社会、自然之间相互关系时，从安全影响角度出发，应保障的伦理价值或秩序的规范与准则。

3.3

应用开发者 developer

开展人工智能应用的理论研究、技术创新、数据归集、模型开发、算法迭代等相关活动，或利用人工智能理论或技术形成具有特定功能、满足特定需求的系统、产品或服务的组织或个人。

3.4

服务提供者 provider

在工作与生活场景中，利用人工智能技术向具体对象提供服务的组织或个人。

3.5

使用者 user

在工作与生活场景中使用人工智能系统、产品或服务的组织或个人。

4 人工智能应用伦理安全影响

主要影响包括：

- a) 人类主导权影响——人工智能行为超出人类所预设、理解和可控的范围，在对人工智能应用关键节点的监督、干预与纠偏不足时，可能冲击人类在社会运行和治理过程中的主导地位。
- b) 公共秩序影响——人工智能应用于社会运行和行为决策，在技术快速演化而社会回馈相对缓慢的情况下，可能影响行为规范、市场秩序与社会信任，最终对社会基础公共秩序带来冲击。
- c) 个体认知与社会价值影响——人工智能普遍应用形成的新社会环境可能对个体认知与社会价值体系带来冲击，产生人类过度依赖、个体与现实社会脱节等问题。
- d) 社会分化和歧视影响——人工智能应用对不同群体可能产生差异化影响，放大偏见与歧视，造成结构性不利后果，影响公平公正、加剧社会分化。
- e) 生命健康与基本权益影响——人工智能因其全新的应用方式，可能对自然人的生命健康、人身安全、人格尊严、隐私、财产、劳动等基本权益造成影响，也可能对组织的合法权益造成侵害或产生不利影响，甚至可能对国家安全、公共安全带来危害。
- f) 可持续生态影响——不合理的人工智能技术路线选择与应用模式可能会带来系统性生态压力，影响人与自然的可持续发展。

5 伦理安全理念与原则

5.1 伦理安全理念

在开展人工智能应用活动时，应将伦理安全要求贯穿全过程，从安全影响角度出发，妥善处理人工智能与人、社会、自然之间的相互关系，以造福人类、服务社会和可持续发展为导向，推动人工智能朝着有益、安全、公平方向健康有序发展，避免不合理的人工智能技术路线选择与盲目应用，保障人工智能始终服务于增进人类共同福祉。

5.2 伦理安全原则

主要原则包括：

- a) **增进人类福祉**：坚持正确价值观，坚持智能向善、以人为本，推动人工智能服务人的全面发展和社会共同利益，促进提高生产生活水平、改善公共服务、提升社会治理效能。鼓励人工智能发挥积极作用，拓展人类能力、增进社会福祉，保护个人劳动就业权利，事先评估并防范人工智能应用可能引发的就业替代或失业风险。
- b) **尊重生命权利**：坚持生命至上、尊严优先，尊重和保障人的生命健康、人身安全、人格尊严和基本权益。保障人的自主决策、维护个人主体性、避免过度依赖，涉及生命健康、人身安全和人格尊严等重要场景的人工智能应用，应以保障人的生命权利和基本权益为底线，避免对人的身心健康、人身安全和人格尊严造成不利影响。
- c) **坚持公平公正**：坚持公平包容和机会均等，避免人工智能应用造成不合理差别对待或加剧既有不公。坚持消除技术偏见和歧视，确保人工智能不对特定民族、信仰、国别、性别等群体以及特定组织或服务造成不公平影响。关注特殊群体与弱势群体权益，鼓励人工智能增进公共服务公平性与可及性。
- d) **合理控制风险**：坚持发展和安全并重，统筹人工智能应用的积极价值与潜在影响，强化风险意识和底线思维，推动人工智能在安全、可靠、可预期的范围内健康有序

2

应用。审慎开展可能对国家安全、公共安全、生命健康等产生重大影响的人工智能应用，提前防范人工智能应用被非法使用、恶意利用、滥用等。

- e) **保持公开透明**：提升人工智能应用的透明度，推动其以更加清晰、可感知、可理解、可追溯的方式运行，增进社会对人工智能的认知与信任。
- f) **保护隐私安全**：尊重人的隐私权益，避免人工智能应用侵犯合理预期下的隐私边界，保护个人和家庭隐私空间，增强社会公众对人工智能应用的信任。
- g) **确保可控可信**：确保人工智能应用的主导权归属人类，鼓励人工智能应用安全开发。在关键环节设置人类控制机制，建立应急处置与人工干预机制，确保人工智能的运行始终处于人类控制之下向善发展，防范人工智能脱离人类监督或威胁人类生存发展。
- h) **敏捷共治**：加强安全互认，推动治理规则、技术标准、安全基准等方面国际衔接，提升敏捷治理水平，推动各方协同共治。推动相关主体共同履行人工智能应用伦理安全治理责任，加强协作配合，凝聚治理合力，避免因技术壁垒、封闭排他等加剧数据鸿沟，提升社会人工智能应用素养，共同保障人工智能健康有序发展。
- i) **普惠共享**：构建合作共生、能力共建、收益共享的人工智能生态，不利用人工智能强迫他人、实施垄断等。培育开源创新生态，鼓励人工智能模型、工具组件、评测基准等全方位技术开源，同步提升开源生态安全能力。鼓励不同路径技术探索，共享人工智能知识成果、最佳实践与安全治理经验，推动人工智能应用成果更广泛惠及社会。

6 伦理安全指引

6.1 通用指引

主要指引如下：

- a) 在开展人工智能应用活动前，预先评估应用目的及影响，充分保障国家安全、公共利益、组织及个人权益。
- b) 以集体主义、人的全面发展等原则为基础，构建符合中国国情和文化传统的伦理规范体系。
- c) 正确认识人工智能的应用价值、能力边界及潜在影响，避免盲目依赖、片面夸大和过度宣传。
- d) 保留关键环节的人类判断、监督、干预和纠错，避免人工智能过度替代人类决策。
- e) 重视隐私和个人信息保护，防止过度收集、不当使用、泄露或者滥用相关信息。
- f) 关注人工智能应用可能带来的偏见、歧视和机会不均等影响，重视未成年人、老年人、残障人士等易受影响群体保护。
- g) 全面提升人工智能应用伦理素养水平，加强人工智能伦理应用影响的安全风险提示反馈和改进，及时发现并采取相应措施。
- h) 适时开展伦理安全影响评估，及时识别伦理安全风险并完善伦理安全防护措施。
- i) 涉及生命健康与人身安全的，坚持以人为本、安全优先、审慎适用，以保障生命健康与人身安全为底线，确立人工智能应用的适用边界与责任安排。
- j) 涉及社会治理与公共服务的，坚持公共利益优先、公平公正，在提质增效的同时，保障公共服务的普惠性、可及性与程序正当，保障公众知情、质疑、复核与救济渠道有效运转。
- k) 涉及信息资讯与传播的，坚持真实可信与清朗生态导向，推动人工智能提升优质内容供给与传播效率，同时维护传播秩序与受众权益，避免出现违法违规内容，强化对生成与推荐的风险治理，避免人工智能可能带来的信息茧房、认知误导、认知退化等问题。
- l) 涉及知识发现与生产的，坚持求真务实、可验证与学术诚信，鼓励人工智能在科研探索、工程研发与高等教育等活动中发挥辅助增益作用、更好释放人的创造性与主体性，确保关键结论与重要判断具备必要核验与责任承接，强化证据、引用与署名规范，尊重知识产权与成果归属边界。
- m) 涉及金融活动的，坚持稳健审慎，注重风险控制与消费者保护。对可能影响个人财产权益、交易机会与交易条件的关键环节，应设置清晰规则，便于责任追究与结果复核，关注对公平竞争与市场秩序的影响。

6.2 应用开发指引

主要指引如下：

- a) 统筹考虑人工智能应用伦理安全，识别满足伦理要求的安全开发需求，审慎关注人工智能的合法合规应用情况，避免单纯以性能或效率目标驱动迭代开发与优化改进。
- b) 满足我国人工智能科技伦理审查要求，审慎开发具备高度自主性人工智能应用，重点评估其失控风险以及其对产业和社会的影响。
- c) 理性认知人工智能应用的技术能力边界与潜力，正确看待“幻觉”等问题，对缺乏依据或者存在较大不确定性的输出作出必要管控以及提示说明。
- d) 将安全可控、公平公正、隐私保护、机会均等理念贯穿数据选择、目标设定、算法设计、技术开发、产品研发、测试评估、运行维护等全过程。
- e) 面向未成年人、老年人、残障人士等易受影响群体的特点，在应用开发过程中采取专门设计考量，其安全提示、透明度标识及隐私授权等采用符合年龄认知的语言，并符合适用的无障碍获取要求，确保易受影响群体能够清晰感知。
- f) 在技术可行、符合行业普遍技术水平的前提下，提高开发过程透明度，结合人工智能应用场景和影响程度等方面，制定并提供安全配置、安全部署的操作指南，及时、准确、完整对外说明人工智能应用的功能、局限、安全风险和可能的影响。
- g) 设置事故信息追溯机制，保留与风险事件相关的必要信息，确保在发生争议、损害或事故时能够复盘分析并支撑责任追究。
- h) 推动不同组织、个人间合作与互信，促进良性竞争与多元化技术路线并行发展，协同开展伦理安全风险评估。

6.3 服务提供指引

主要指引如下：

- a) 选择符合伦理影响规范的人工智能应用目标，理性引导社会认知服务能力与边界，及时提示可能风险，避免对使用者造成误导。
- b) 提供符合科技伦理审查要求的产品或应用，加强人工智能应用全生命周期的跟踪审查。
- c) 服务涉及国家安全、社会公共利益、组织和个人重大生命财产安全时，人工智能应仅承担辅助决策作用，不提供直接决策依据。
- d) 以清楚明确且便于操作的方式向使用者提供拒绝、干预及停止使用人工智能服务的机制。着力避免提供使用者无法便捷地强行通过断电、关停系统等方式停止的人工智能服务。
- e) 尽量减少对使用者隐私和敏感个人信息的处理，遵循合法、正当、必要、最小化原则，杜绝偷采、强采、滥采，防止信息泄露以及不当披露。
- f) 如需收集使用者数据用于训练模型、改进服务的，宜持续对使用者提示收集数据的状态，并显著告知关闭方式。
- g) 设置事故应急处置机制，持续监控服务运行过程，主动识别发现伦理安全风险，及时处置伦理安全问题，确保合法合规应用，必要时采取人工紧急干预、中止应用等手段。
- h) 面向未成年人、老年人、残障人士等易受影响群体服务时，着重提升服务便利性，避免造成使用障碍。
- i) 服务涉及个人权利义务、公共资源分配等场景时，重点提升公平性，并完善风险提示。

6.4 应用使用指引

主要指引如下：

- a) 正确认知人工智能，学习人工智能伦理安全基本知识，增强对虚假、错误、误导性内容的识别能力，增强对人工智能相关物理安全风险、决策安全风险的认识，形成理性、审慎的使用习惯。
- b) 适度使用人工智能，正确认知人工智能情感服务原理，以人工智能作为辅助真实生活的工具，避免过度依赖、过度沉迷，避免通过人工智能过多替代现实交往与真实活动。
- c) 妥善运用人工智能，通过人工智能帮助学习、生活、创作，促进个人主体性提升和创造性发挥。
- d) 合法使用人工智能，尊重他人尊严与合法权益，不借助人工智能对他人进行误导、骚扰、操纵，不利用人工智能伪造、冒用他人身份或仿冒权威主体，不使用人工智能实施攻击、破坏、盗窃等违法违规行。
- e) 妥善保护自身敏感信息，审慎向人工智能提供、上传或分享敏感个人信息、隐私信息，避免个人文件、商业秘密等意外泄露，重点防范涉及国家安全、公共利益的信息被不当传播。
- f) 提高人工智能素养，积极履行人工智能生成内容标识等人工智能相关义务，及时反馈伦理安全风险，帮助建设良性人工智能生态。

专家解读 | 推动人工智能创新发展 筑牢伦理安全治理屏障

原载：“网信中国”微信公众号

近年来，人工智能深度融入经济社会各领域，深刻改变人类生产生活方式。同时，算法偏见、隐私泄露、滥用风险等伦理安全问题愈发突出，侵害个人权益、扰乱社会秩序甚至危害国家安全风险不

断上升。习近平总书记高度重视人工智能治理，强调要把握人工智能发展趋势和规律，加紧制定完善相关法律法规、政策制度、应用规范、伦理准则。近日，全国网络安全标准化技术委员会发布《人工智能应用伦理安全指引 1.0》（以下简称《指引》），系统划定人工智能应用伦理安全边界与底线，为人工智能健康有序、安全可控、向善发展提供规范化依据与实操性遵循。

一、统筹发展与安全，健全中国特色人工智能伦理安全治理体系

我国立足国情与治理实践，防范化解人工智能伦理安全风险，逐步构建以顶层设计为引领、柔性指引与刚性规制协同的中国特色人工智能伦理安全治理体系。柔性指引层面，先后出台《新一代人工智能伦理规范》《关于加强科技伦理治理的意见》以及《人工智能安全治理框架》1.0版、2.0版，将伦理要求嵌入人工智能全生命周期，系统防范安全风险。刚性规制层面，通过《互联网信息服务算法推荐管理规定》《生成式人工智能服务管理暂行办法》《人工智能拟人化互动服务管理暂行办法》以及新修订的《中华人民共和国网络安全法》，完善伦理安全规范、强化风险监测评估与合规监管。在顶层统筹层面，我国依托科技伦理治理、网络信息治理、产业治理等多方面机制，推动形成多部门协同、社会各方参与的治理格局，通过《全球人工智能治理倡议》参与全球治理，通过《国务院关于深入实施“人工智能+”行动的意见》进一步探索形成智能向善理论体系。

我国人工智能伦理安全治理体系持续健全完善，伦理风险防控加速从制度层面逐步走向实践落地，现阶段仍面临诸多现实堵点与突出难题。人工智能加速赋能千行百业，出台《指引》既是推动人工智能规范合规、健康有序发展的内在要求，也是有效防范化解人工智能应用层面伦理安全风险的现实举措。

一是筑牢伦理安全防线，保障人工智能健康有序发展。《指引》的出台，立足人工智能发展新阶段，围绕人类主导权、公共秩序、个体认知与社

会价值、可持续生态影响等重点方向，系统提出伦理安全理念、原则与行为规范，有助于推动人工智能在可控、可靠、安全轨道上发展，防止技术异化和失控滥用，确保人工智能始终服务于人类共同福祉、服务经济社会发展。

二是规范产业发展秩序，推动人工智能负责任创新。《指引》围绕应用开发、服务提供、应用使用等全链条提出明确要求，强调将以人为本、生命至上、公平公正、公开透明、可控可信等伦理安全要求贯穿人工智能应用活动全过程，推动企业在技术创新和产业应用过程中更加注重伦理安全责任与治理。这既有助于营造规范有序、健康可持续的产业生态，也有利于推动人工智能产业在规范有序中实现更高质量发展，进一步提升产业创新能力、应用水平和社会信任度。

三是提升全民使用素养，防范技术滥用风险。

《指引》强调正确认识人工智能的能力边界和潜在影响，倡导增强自主判断、妥善运用人工智能，明确提出提升公众对虚假错误信息、误导性内容和伦理安全风险的识别能力。同时，《指引》鼓励建立风险提示、投诉举报、反馈纠偏等机制，推动形成政府、企业、社会公众多方参与的协同治理格局，有助于提升全民人工智能应用素养和风险防范意识，减少技术滥用对社会造成的不良影响。

二、强化技术指引，全方位筑牢伦理安全防线

一是为各应用主体提供明确行为准则。《指引》围绕应用开发者、服务提供者 and 应用使用者三类主体，分别提出系统性规范要求，明确人工智能研发、部署、运营、使用全过程中的责任边界和行为要求。在开发环节，强调提升可控性、公平性、透明度等目标贯穿应用开发与更新，要求开发者开展伦理安全影响评估、设置事故信息追溯机制；在服务提供环节，强调理性引导、隐私保护、风险提示和应急处置；在应用使用环节，倡导妥善运用、审慎提供敏感信息、尊重他人合法权益。通过建立全链条责任体系，《指引》有助于推动各主体形成规范意识和责任意识，减少人工智能应用活动中的

伦理安全失范行为。

二是防范化解人工智能伦理安全风险。人工智能风险具有复杂性、隐蔽性、扩散性和外溢性等特点，一旦失控，可能对个人、社会乃至国家层面造成较大影响。《指引》聚焦人工智能可能带来的人类主导权弱化、社会秩序影响、价值体系冲击、偏见歧视、合法权益侵害、可持续生态压力等问题，提出覆盖事前预防、事中监测、事后处置的系统性治理要求。例如，《指引》要求建立必要的应急处置与人工干预机制，防范人工智能脱离人类监督；要求持续监控服务及运行过程，识别发现伦理安全风险并及时处置；要求设置事故信息追溯机制，实现责任追溯。这些措施有助于提升人工智能风险治理能力，增强技术应用的安全性和可控性。

三是引导人工智能向善发展。《指引》坚持以人为本、智能向善的理念，强调人工智能应服务于经济社会发展和人民福祉提升，而非削弱人的主体地位或加剧社会不公。《指引》强调公平包容、机会均等和普惠共享，关注未成年人、老年人、残障人士等重点群体权益，倡导人工智能提升公共服务公平性与可及性。同时，鼓励开源创新、知识共享和多元技术路径探索，推动人工智能应用成果更广泛惠及社会公众。这有助于推动人工智能技术始终与社会公共利益保持一致，引导人工智能始终坚守正向价值、坚持向善发展。

四是完善伦理安全协同治理体系。人工智能治理是一项系统工程，既涉及法律法规、伦理规范、技术标准，也涉及社会监督、国际合作和公众参与。《指引》在理念层面提出敏捷共治、协同治理等原则，在实践层面提出伦理安全评估、风险监测、信息提示、投诉反馈等制度安排，为构建我国人工智能伦理安全治理体系提供了重要补充。同时，《指引》强调加强国际衔接与安全互认，推动治理规则、技术标准和安全基准兼容互洽、同步演进，有助于提升我国在全球人工智能治理中的活跃度和影响力，推动形成开放、合作、共赢的人工智能治理格局。

三、构建协同共进发展格局，护航人工智能

高质量发展

一是规范人工智能产业发展。《指引》的出台有助于进一步规范人工智能产业发展秩序，推动行业从重速度、重规模向重质量、重安全、重责任转变。通过明确伦理安全要求和风险防控机制，减少无序竞争、低水平重复建设以及因安全问题导致的产业风险，增强市场主体的合规意识和长期发展能力。同时，统一规范有助于增强社会公众和市场对人工智能产品与服务的信任度，推动人工智能实现规范、有序、可持续发展。

二是推进人工智能创新生态建设。《指引》并非限制技术创新，而是强调安全与发展并重，推动技术创新更加注重长期价值和社会责任。通过将可控可信、公平公正、隐私保护等要求前置到研发和设计阶段，有助于推动可信人工智能、隐私计算、可解释算法、安全评测等关键技术发展。同时，《指引》鼓励开源生态建设、多元技术探索和协同创新，有利于形成更加开放包容、健康可持续的人工智能创新生态，提升我国人工智能核心竞争力。

三是提升智能化社会治理水平。人工智能正在深刻改变社会治理模式。《指引》有助于推动人工智能在医疗健康、信息传播、科研教育等领域规范应用，在提升治理效率和治理能力的同时，防止自动决策失当、信息操纵、侵害知识产权等问题影响人身安全和社会公平。特别是在社会治理、公共服务等关键领域，《指引》强调普惠可及、程序正当、结果复核和公众救济机制，有助于维护社会公平正义，提升现代化治理能力和治理公信力。

四是健全公众权益保护机制。《指引》更加突出对个人合法权益的保护，围绕隐私保护、个人信息安全、知情权、选择权、申诉权等提出明确要求，有助于增强公众在人工智能时代的安全感和获得感。同时，《指引》强调妥善运用使用人工智能，防范过度依赖、沉迷使用、虚假信息误导等问题，并要求加强未成年人、老年人、残障人士等重点群体保护。这有助于推动形成安全、可信、健康的人工智能应用环境，更好保障人民群众合法权益。

头条！张东刚：人工智能时代中国自主知识体系建构

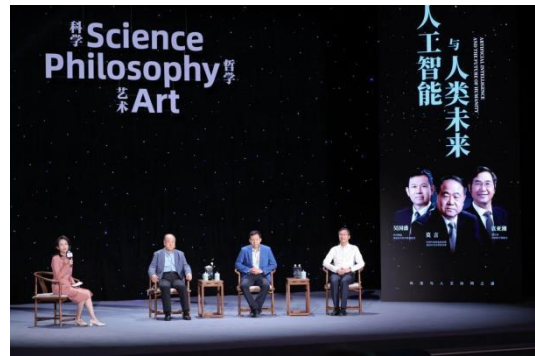
原载：“中国人民大学”微信公众号

今天《学习时报》头版头条刊发了中国人民大学党委书记张东刚署名文章《人工智能时代中国自主知识体系建构》现登载如下，以飨读者。



人工智能驱动哲学社会科学的深刻变革

纵观人类文明史，每一次划时代的技术革命，都不仅是以生产工具为代表的生产力飞跃，更是认识世界、改造世界的思维方式的深刻变革。从马克思主义的观点看，人工智能作为当代最活跃、最具渗透性的“一般智力”形态之一，代表着生产力的质变性突破。这种跃升必将要求并推动作为上层建筑重要组成部分的哲学社会科学，在研究预设、方法路径、评价标准等方面，发生系统性、结构性的重塑。



▲中国人民大学创办“科学、哲学与艺术：科技与人文协同之对话会”。2025年9月，莫言、袁亚湘、吴国盛参加首场对话会，与1000余名人大师生共话人工智能和人类未来。

拓展认知疆域，重塑研究范式。研究范式是特定学术共同体在某一历史时期所共同遵循的基本信念、理论框架与方法规范，构成知识生产与理论建构的深层基础。范式的转型往往先行于知识体系的重构，并为后者提供前提性条件。因此，剖析人工智能对研究范式的影响，是理解其赋能自主知识体系建构的逻辑起点。从哲学基础来看，人工智能重塑研究范式涵盖本体论、认识论与方法论三个层面。在本体论范式上，人工智能推动社会实在图景发生结构性变化，催生了新的、亟待认识的社会实在，研究对象不再局限于物理空间中的人类行为，而是扩展至人机交互、算法驱动、数据留痕等新型社会现象。在认识论范式上，人工智能通过引入数据驱动、模式识别、关联分析等新型认知方式，使知识可以从海量数据中“涌现”出来，挑战了“无理论不认知”的传统信条。在方法论范式上，人工

当前，以生成式人工智能为代表的智能技术浪潮，正以前所未有的广度与深度渗透至人类社会各领域，引发生产力与生产关系的深刻变革。习近平总书记深刻指出，“哲学社会科学是人们认识世界、改造世界的重要工具，是推动历史发展和社会进步的重要力量”。人工智能作为这一变革的核心驱动力，与哲学社会科学的结合，超越了简单的工具性应用，触及研究范式、知识生产逻辑与学术生态的深层结构。这要求我们深刻理解工具理性和价值理性的辩证统一，深刻把握二者“双向赋能、相互塑造”的内在逻辑：一方面，自觉运用人工智能这一革命性力量，为建构中国自主的知识体系提供全新动能；另一方面，以哲学社会科学引领人工智能的健康发展，确保其服务于中国式现代化与创造人类文明新形态的伟大实践。

智能极大拓展了哲学社会科学的研究工具与方法谱系，计算社会科学、生成式社会模拟、自然语言处理、图学习模型等新型研究方法迅速兴起，与传统定性、定量方法形成互补与融合。



▲“中国人民大学涉外法治大模型 2.0”全面服务于企业出海及法律实务部门，智能支撑全国法学科研究与教育，搭建全球数据合作平台和法律智能服务平台。



▲“中国人民大学涉外法治教学智能体”入选 2026 世界数字教育大会首批高等教育智能体。

革新认知方式，变革知识生产。知识生产是人类社会对客观世界进行认知加工、形成系统化知识成果的创造性活动，其规模、速度与质量决定着知识体系的完备程度与更新能力。从历史演进来看，传统哲学社会科学的知识生产主要采取个体化的“手工作坊”模式，学者凭借个人学养积累，通过文献研读、田野调查等方式提出问题、形成观点并著书立说，知识传播依赖于期刊、著作与课堂，知识获取主要发生在学术共同体内部，呈现出周期长、门槛高、受众窄等特征。人工智能的兴起正在推动知识生产向智能化、协同化、普惠性方向转型。**在知识的提出环节**，人工智能极大地增强了问题发现与假设生成的能力，通过大规模文献挖掘、学术趋势分析和社会舆情监测，为研究者提供数据驱动的选题建议。**在知识的传播环节**，人工智能利用智能推荐系统、学术知识图谱与多语言互译工具，实

现了知识的即时推送、精准匹配与无障碍传播。**在知识的接受环节**，人工智能通过提供智能摘要、知识问答、关联推荐、交互式学习等新型接受方式，使研究者能够快速把握知识脉络、精准定位核心观点，从而加速知识的内化、转化与新知识的再生。



▲中国人民大学发布的“亿心通”社区书记智能体，是国内高校自主研发的首个深度融合大模型与基层治理实践的生成式人工智能系统。

重构学术生态，重建评价体系。学术评价体系是对哲学社会科学研究成果、研究人才及学术机构进行价值判断的标准与方法集合，在知识体系建构中发挥着“指挥棒”与“风向标”的关键作用。学术评价体系设定何为“好研究”“真问题”“高水平成果”的判断标准，潜移默化地引导着研究者的选题方向、方法取向和价值追求，从而深刻影响着知识体系建构的路径选择与资源分配。在人工智能赋能下，学术评价体系呈现出三个鲜明特征。**一是评价主体多元化。**智能算法可以对研究成果进行自动化的规范性检测、创新性评估和影响力测算，形成人机协同的评价模式。**二是评价维度融合化。**通过对海量数据的自动采集与智能分析，综合考量研究成果的学术创新性、社会影响力、文化传承力以及国际能见度等多个维度，为克服“以刊评文”“以数论质”的弊端提供了技术可能。**三是评价过程动态化。**通过实时追踪研究进展及阶段性成果，为研究者提供持续的优化建议；同时，评价数据可以累积形成学者的“学术成长档案”，为人才评价提供更加立体、长时段的过程支撑。

在人工智能赋能中实现知识体系跃升

人工智能在赋能创新发展、激发范式变革的同时，也潜藏着不容忽视的风险与挑战。例如，算法依赖可能导致研究主体性的弱化，智能生成内容的

滥用可能侵蚀学术原创精神，技术门槛的差异可能加剧学术资源分配的“马太效应”。因此，人工智能的赋能作用，必须服务于自主知识体系建构的核心目标，确保赋能行稳致远、安全可控。从本质上来看，自主知识体系内在包含着体系化、学理化、标识化的建构标准。**体系化是前提**，没有体系化，知识便是零散杂陈的碎片，难以形成理论合力；**学理化是核心**，没有学理化，体系便是有形无神的空壳，难以具备思想深度；**标识化是表征**，没有标识化，学理化成果便难以在国内外学术场域中产生影响。在人工智能深度介入的条件下，体系化、学理化、标识化被赋予了新的时代要求，呈现出新的发展特征。

以智链接，实现更加系统协同的体系化，推动知识要素从零散分布走向有机融通。一个成熟的知识体系，必须在概念、范畴、命题、理论之间建立起层次分明、相互支撑的有机联系。在传统语境下，体系化建设主要依赖学者的个体耕耘与学术共同体的代际积累。在人工智能条件下，要实现更高层次的体系化，必须充分发挥人工智能技术在知识整合、关联与演进方面的独特优势。**在知识整合方面**，要利用多模态数据处理能力，对人文社科、自然科学以及中外学术资源进行大规模自动化聚合，消除信息孤岛，拓展体系化的广度。**在知识关联方**，要借助知识图谱、语义网络与自然语言处理技术，自动识别概念之间的逻辑联系，推动知识要素从机械拼凑走向有机融通，提升体系化的深度。**在知识演进方面**，要依托动态学习与自适应能力，实时追踪全球学术前沿，持续更新知识图谱，增强体系化的活力。通过以智链接，不仅能够有效弥补传统体系化的固有不足，更能推动自主知识体系实现从零散分布到协同集成的根本转变。



▲ 2026年5月22日，中国人民大学在杭州举行第12场中国自主的知识体系成果发布会，发布智能经济领域重磅成果。

以数挖掘，实现更加深刻透彻的学理化，推动理论建构从经验描述走向规律提炼。学理化是从纷繁复杂的经验现象中提炼出具有普遍解释力的规律、原理与范畴，这一过程就是所谓的“抽象力”。抽象力是人类特有的认知能力，它依赖于研究者的理论直觉、思辨智慧与创造性综合。人工智能的引入并非替代研究者的理论思维，而是作为“认知伙伴”，在人机协同中拓展抽象思维的广度与深度。

一是提升规律识别的敏锐性。人工智能以其强大的数据挖掘与模式识别能力从海量政策文本、社会数据与典型案例中自动提取高频语词、分析潜在关联，为研究者提供理论建构的经验基础，使规律发现更加敏锐、更加系统。

二是强化理论验证的严谨性。人工智能通过对比分析、异常检测与因果推断等技术，协助研究者验证理论假设、识别逻辑谬误、评估理论的适用范围与边界条件，提升学理化的科学性与严谨性，避免出现以点概面式的经验主义错误。

三是激发原创思维的创造性。人工智能可以基于已有知识进行类比推理与跨域迁移，启发研究者提出新的分析框架与理论命题，为理论创新提供原创洞见。

三是激发原创思维的创造性。人工智能可以基于已有知识进行类比推理与跨域迁移，启发研究者提出新的分析框架与理论命题，为理论创新提供原创洞见。



▲ 中国人民大学研发 TaoSight 智能分析平台，为社会科学提供数据清洗与分析工具。

以算识别，实现更加鲜明有力的标识化，推动话语传播从被动输出走向主动引领。标识化是自主知识体系建构的表征，核心任务是提炼和传播一批如“全过程人民民主”“新质生产力”“人类命运共同体”等具有原创性、识别度和影响力的标识性概念。人工智能的引入，将以其独特的运算优势超越传统的标识化模式，在识别、提炼、传播三个

环节实现全面革新。在**概念识别上实现主动发现**。人工智能依托海量文献数据库、典籍资源库、政策文本库、学术成果库，开展全域语义检索、关联聚类分析、脉络谱系梳理，高效挖掘“学术萌芽”，实现对标识性概念全覆盖、深层次、谱系化的智能识别。在**概念提炼上实现精准凝练**。人工智能通过逻辑推演、跨学科关联比对等技术，对识别出的思想元素进行归类整合、去芜存菁，让概念提炼更加学理化、规范化、体系化。在**概念传播上实现精准触达**。人工智能能够发挥智能转译、场景重构、精准推送等传播优势，对内实现学术概念向大众话语的通俗转化，对外进行多语种智能阐释和跨文化语境适配，提升中国学术思想的叙事能力与传播效能。

以中国自主知识体系引领人工智能向善发展

马克思曾深刻警示：“技术的胜利，似乎是以道德的败坏为代价换来的。”技术从来不是价值中立的。它诞生于特定的社会关系之中，其设计、应用与后果必然承载着设计者与使用者的价值取向、利益诉求与文化逻辑。如果缺乏正确的价值引导与制度约束，技术进步并不必然带来社会福祉，反而可能加剧不平等、固化偏见，甚至威胁人类主体地位。因此，人工智能的健康发展，绝不能陷入“技术决定论”的迷思，需要自主知识体系为其设定价值坐标、规划发展航道、筑牢安全屏障。

强化价值锚定，确保人工智能发展以人民为中心。人工智能向何处去？这不仅是技术问题，更是关涉发展为了谁、依靠谁的重大价值问题。习近平总书记强调：“为什么人的问题是哲学社会科学研究的根本性、原则性问题。”这一论断同样为审视和规划人工智能发展提供了根本遵循。人工智能作为社会生产的革命性力量，其研发决策、算法设计、数据应用乃至最终的社会影响，必然承载并传导着特定的价值取向与利益偏好。倘若缺乏主动、明确的价值引领，技术逻辑可能自发地滑向资本增殖、流量至上或效率唯一的单向维度，甚至可能固化偏见、加剧不公。中国哲学社会科学自主知识体系，以马克思主义为指导，扎根中华优秀传统文化，立足于中国式现代化的伟大实践，这套知识体系所

凝结的关于社会公平、人的全面发展、共同富裕、人与自然和谐共生等价值理念与规律性认识，构成了中国社会的价值共识。这一价值内核应当贯穿人工智能从算法设计、数据训练到场景应用的全过程。它要求人工智能的发展坚持以人民为中心，确保技术进步的红利为全体人民所共享；符合社会主义核心价值观，抵制技术滥用带来的伦理失范；服务于中华民族伟大复兴，成为建设中国式现代化的强大助力。



▲中国人民大学“AI+”创研课融合前沿AI技术与多元学科知识，构建起一套以学生自主探索为核心、教师精准指导为支撑、团队协作共进的创新教学模式。

深化伦理建构，为人工智能发展校准准则与边界。人工智能引发的治理难题，核心在于如何应对其模糊责任主体、挑战既有法律框架、冲击传统伦理观念的全新特性。从算法歧视的防范、数据隐私的保护，到人工智能生成物的权责界定、自主系统决策的追溯与问责，都已成为关涉社会正义、权利保障与全球治理秩序的深刻伦理与法律命题。习近平总书记指出，“要整合多学科力量，加强人工智能相关法律、伦理、社会问题研究”。这为构建适应智能时代的社会规则体系指明了方向。中国自主的知识体系必须在这一领域进行前瞻性、体系化的理论供给与规范建构。通过哲学、伦理学、法学等学科深度协同，立足中国国情与实践，致力于将社会主义核心价值观转化为可操作、可审查、可问责的具体伦理准则与技术标准。通过前瞻立法研究、参与标准制定，勾勒出清晰、坚实的治理“红线”，将“负责任的人工智能”从理念倡导落地为贯穿创

新全过程的制度约束，为技术进步构筑兼顾活力与秩序、创新与安全的规则航道。



▲中国人民大学发布“明新大模型 1.0”，推动人工智能在新闻生产与舆情治理领域的深度融合，为应对智能时代新闻传播领域的重大挑战提供解决方案。

融入人文精神，为人工智能发展注入灵魂与温度。人工智能的终极目标应是赋能于人、服务于人，促进人的全面发展，而非替代人、异化人。纯粹的技术理性无法自行回答关于价值、意义与美好生活的根本问题，甚至可能因过度追求效率、标准化与可控性而加剧社会的工具化倾向。因此，人工智能的健康发展，也需要人文精神的深度介入与价值校准。文学、历史学、艺术学等人文科学所蕴含的对生命意义的追问、对美好情感的珍视、对创造力的礼赞，是防止技术应用走向“冰冷”的温暖底色。将中华优秀传统文化中“天人合一”“仁者爱人”“和而不同”等智慧融入人工智能发展，能够为其提供不同于西方工具理性至上的另一种文化基因与发展路径。总之，通过知识体系的人文滋养，能够引导人工智能成为丰富人类体验、激发创造潜能、促进文化繁荣的赋能者，使其真正服务于人的全面发展与社会的文明进步。

（原文刊载于《学习时报》2026年5月25日1版，作者系中国人民大学党委书记。）

《生活服务类平台算法负面清单（试行）》实施取得初步成效

原载：“网信中国”微信公众号

贯彻落实党中央决策部署，为更好维护新就业群体和广大人民群众合法权益，2026年1月，中央网信委印发《生活服务类平台算法负面清单（试行）》。随后，中央网信办会同有关部门精心组织实施，北京、上海、江苏、浙江、广东等地网信部门迅速行动，督促外卖、网约车、货运、网购、在线旅游、票务领域重点平台对照《负面清单》，深入开展自查自纠。美团、淘宝闪购、淘宝天猫、京东、滴滴、高德、T3、百度、满帮、货拉拉、拼多多、抖音、携程、去哪儿网等平台积极响应，已实施优化改进措施63项，承诺遵守算法要求139项，限期推进125项，在提升算法透明度、公平性、合理性方面取得初步成效。

1. 优化订单分配算法。

订单优先就近分配。滴滴出行、T3出行将距离作为首要派单影响因素，绝大多数订单由距离最近的司机接单。货拉拉不指派订单，全部由司机自主抢单，按照“就近优先”原则确定接单司机的比例达93%。**抵制不合理的差异化派单。**多家重点平台承诺不会根据司机/骑手性别、年龄、价格承受能力、历史接单情况进行差异化派单。美团外卖派单主要参考顺路程度和超时概率，不涉及骑手个人画像。满帮承诺所有订单均由司机自主接单。**设定同时接单上限。**美团外卖、淘宝闪购、京东外卖，结合自身业务情况，均设定了骑手同时接单上限，并允许骑手在上限范围内自主调整。**提升接单自主权和选择权。**美团外卖、淘宝闪购、京东外卖为不同类型骑手，提供了每日至少4次无条件拒单或转单机会。

2. 优化时间预估算法。

设置宽容灵活送达时间。美团外卖、淘宝闪购、京东外卖已在消费者端，以时间段形式展示预估配送时间。美团外卖将探索在部分场景下，把骑手配送时间在合理预估时间点基础上向后预留不少于15分钟。**疏通“最后一公里”送餐难。**淘宝闪购试点“取送分离”配送模式，对于复杂内部环境的地点，通过固定人员固定地点集中取送，与外卖配送

人员分离，提升配送效率，解决复杂场景配送难问题。美团外卖构建“进入难”和“上楼难”场景自动识别能力，已覆盖40万个小区的通行类型、410万个楼栋的上楼方式。**基本取消超时扣款。**美团带头取消超时罚款，淘宝闪购实现超时免罚全覆盖，京东外卖正逐步减少直接扣款处罚措施。**优化防疫劳算法。**各重点平台相继上线防疫劳措施。美团外卖推出“4小时和8小时提醒，累计跑单时长12小时强制下线”的防疫劳举措。滴滴出行司机24小时内累计在线服务10小时，强制下线休息6小时。

3. 优化安全保障算法。

完善安全保障机制。美团外卖、淘宝闪购、京东外卖对骑手超速、闯红灯等危险行为进行语音提醒。美团外卖在全国投放超过70万顶智能安全头盔，淘宝闪购提供智能耳机、智能头盔等智能装备24万件，提升骑手骑行安全保障能力。滴滴出行实时更新路网数据，提升规划路线安全性。**激励安全行驶行为。**美团外卖、淘宝闪购上线“安全分”机制，通过现金奖励、权益奖励等措施，引导骑手自觉安全行驶。

4. 优化收入抽成算法。

保障劳动者报酬。滴滴出行将抽成上限从29%下调至27%，上线“返佣宝”自动返还月均抽成超过25%的部分。高德将打车信息服务费上限下调至9%。满帮、滴滴送货、货拉拉建立低价运费预警拦截机制。满帮等长途货运平台完善运费纠纷快速处置、法律追讨、权益补偿、赔偿先付等机制。滴滴送货建立货主信用评价规则，从源头预防运费拖欠问题，首创未支付运费全额垫付机制。货拉拉持续完善运费垫付机制，针对长时间拖欠运费、催费无果情况，对符合条件的司机予以垫付。**承诺不转嫁营销成本。**滴滴出行、T3出行、货拉拉、满帮等承诺，用户端使用的各种优惠活动不会降低司机端运费。**保持规则公平统一。**美团外卖、淘宝闪购、京东外卖承诺，不以骑手接单量、价格承受能力等因素对同类型骑手差异化设定配送费。滴滴出行、T3出行承诺，不依据在线时长、出车天数、优惠接单等情况进行差异化抽成。

5. 优化定价算法。

规范动态定价算法。各重点平台均承诺，不基于消费者浏览记录、支付意愿、支付能力、消费记录、消费偏好等信息，对同一商品或服务在同等交易条件下设置不同价格。**规范自动调价工具。**携程主动下架“酒店AI生意助手”（调价助手）。抖音承诺尊重平台入驻商家的自主定价权，不提供自动调价工具。**促进流量分配公平有序。**抖音承诺，不会将商家是否参与促销优惠活动与搜索排名、流量分配挂钩。淘宝天猫、京东商城等承诺，不利用算法实施不合理的限制交易、过度推荐等行为。

6. 优化申诉处理。

合理保障各方权益。美团外卖、淘宝闪购、滴滴出行等平台承诺，对无法判别是非或各有道理的情况，按照有利于司机/骑手、消费者原则分别处置。货拉拉平台坚持“证据不足一律无责”，对司机轻微过错免罚。**约束算法自动处置处罚。**高德将“停止派单”“列入黑名单”等影响司机重大权益的决定100%转由人工复审。滴滴出行、T3出行提前向司机下发预处理通知书，保障司机申诉权利。**强化正向引导示范。**各重点平台均将司机/骑手在工作过程中见义勇为列入正向考核。滴滴发起“正能量在路上”公益计划，满帮设立公益基金会，美团外卖、淘宝闪购、京东外卖等设置现金奖励，对司机/骑手善举善行予以表彰。

7. 提升算法透明度。

持续推进算法核心机理公开。各重点平台正统一调整算法公示专区入口，承诺进入算法公示专区点击不超过2次，采用图文结合等形式直观解释算法机理与逻辑，逐步实现核心算法100%公示。**强化算法变动协商。**各重点平台正逐步完善算法集体协商和事前公示制度。滴滴出行对涉及口碑值等算法核心逻辑的重大调整，提前至少30天进行线上公示。满帮将订单分配算法调整，由提前7天通过站内信通知变更为“7天征求意见+7天公示”。**优化订单信息展示。**各重点平台进一步推动订单信息公开透明。滴滴出行将订单费用明细的可查询时间从30天延长到90天。淘宝闪购综合评定订单重量、

品类、数量、路段等配送难度因素，展示对应补贴额度，便于骑手清晰感知每笔订单收入影响因素。

中央网信办相关负责人表示，生活服务类平台算法治理取得了初步成效，但个别平台仍存在“选择性整改”“别人不改我不改”等问题，与广大新就业群体和人民群众的期盼期待相比还有不少差距。下一步，中央网信办将会同有关部门，持续抓紧抓实《负面清单》实施工作，进一步压实平台主体责任，适时组织专项检查，评估效果、督促进度。发现存在严重问题的，依法依规处置处罚。

国家网信办、国家发展改革委、工业和信息化部联合印发《智能体规范应用与创新发展实施意见》

原载：“网信中国”微信公众号

近日，国家网信办、国家发展改革委、工业和信息化部联合印发《智能体规范应用与创新发展实施意见》（以下简称《实施意见》）。《实施意见》旨在落实国务院《关于深入实施“人工智能+”行动的意见》，促进智能体规范应用与创新发展。

《实施意见》提出，智能体是具备自主感知、记忆、决策、交互与执行能力的智能系统，是人工智能产品及服务的重要形态。随着大模型等新一代人工智能技术迅猛发展，智能体正加速与网络空间、物理世界深度融合，深刻改变人类生产生活方式和社会治理模式。

《实施意见》明确智能体发展要坚持安全可控、规范有序、创新驱动、应用牵引的基本原则，并提出了4个方面举措：**一是夯实发展基础**，完善技术底座，构建标准协议。**二是守牢安全底线**，明确产品准则，防范安全风险，完善治理体系，强化行业自律。**三是强化应用牵引**，围绕科学研究、产业发展、提振消费、民生福祉、社会治理等方向，提出

19个典型应用场景。**四是建设创新生态**，促进产业合作，强化应用推广。

《实施意见》提出，国家网信办、国家发展改革委、工业和信息化部会同有关方面加强统筹谋划，完善配套政策，形成工作合力，推动重点任务落实落地。

智能体规范应用与创新发展实施意见

智能体是具备自主感知、记忆、决策、交互与执行能力的智能系统，是人工智能产品及服务的重要形态。随着大模型等新一代人工智能技术迅猛发展，智能体正加速与网络空间、物理世界深度融合，深刻改变人类生产生活方式和社会治理模式。为落实国务院《关于深入实施“人工智能+”行动的意见》，促进智能体规范应用与创新发展，制定本实施意见。

一、基本原则

以推动科技创新、提升治理能力、构建产业生态、增进民生福祉为导向，坚持安全可控，将智能体安全、可靠、可信作为发展的底线要求，贯穿智能体技术研发、应用部署与推广的全过程，切实防范系统性风险。坚持规范有序，适应智能体技术演进规律，构建与现有政策法规衔接顺畅、行业自律自治、底线红线清晰的治理体系，有序推进智能体落地应用。坚持创新驱动，加强理论创新、技术创新、工程创新联动，体系化突破智能体关键技术，完善政产学研用协同机制，构建开放共享的智能体生态，提升产业创新活力。坚持应用牵引，重点围绕科学研究、产业发展、提振消费、民生福祉、社会治理等实际需求，发挥典型应用场景示范效应，先易后难、循序渐进，促进智能体技术验证、产品迭代、应用落地。

二、夯实发展基础

夯实技术底座，健全标准体系，降低智能体研发、适配、应用门槛，为丰富智能体产品及服务奠定基础。

（一）完善技术底座

1. 强化基础技术研发。

持续提升通用基础模型性能,支持行业发展细分领域专用模型,形成适应不同场景和设备的模型产品矩阵。面向智能体训练与运行,提升高质量数据集供给能力。加强智能体任务理解、任务规划、工具使用、长期记忆、互认互通、群体协同等技术攻关,提升智能体泛化能力。

2. 完善智能体工具链。

开展智能体底层框架研究,加快研发感知、记忆、决策、交互、执行等关键组件,完善智能体研发、测试、部署、运维等工具链。发展对抗样本检测、行为异常检测等安全与治理工具,提升对智能体非合规行为的发现、干预、阻断、恢复能力。

(二) 构建标准协议

3. 建立智能体标准体系。

制定智能体标准化工作指导文件,形成智能体标准框架,系统布局关键技术、重要产品、数据交换、应用场景、质量评测、安全保障、可信认证等标准体系,加快制定智能体与软件工具、应用服务、硬件外设接口等基础标准。加强智能体互连协议(AIP)等智能体互连关键国家标准、行业标准的推广应用。支持医疗、交通、媒体、公共安全等领域制定强制性标准。鼓励企业按照相关标准研发产品服务,提升智能体规范性。积极参与国际标准制定。

4. 布局发展智能互联网。

研究建立智能互联网体系架构,探索建立智能体注册平台,提供智能体数字身份管理、检索发现、能力声明等服务,支持开发者、部署方式、接口协议、合规认证等信息查询和管理。提升多智能体协同能力,研究智能体身份标识、可信互连、合规支付、安全防护、冲突解决等基础技术。发挥互联网协议第六版(IPv6)技术优势,提升智能体端到端通信能力。探索建立智能互联网监测指标体系。

三、守牢安全底线

坚持以人为本、智能向善、多元共治、安全稳妥,营造规范发展、鼓励创新的制度环境,促进智能体健康有序发展。

(一) 明确产品准则

5. 完善政策法规和伦理规范。

加快研究智能体相关政策法规及伦理规范,发挥专业机构内容资源和审核把关优势,确保智能体行为符合法律法规及主流价值观。防止智能体利用数据优势、人格化技术实施传播不良价值观、算法压榨等行为,防范未成年人、老年人沉迷成瘾、情感依赖等风险。做好与人工智能伦理审查等制度衔接。

6. 明确决策权限。

在遵守法律法规、尊重社会公德和伦理规范前提下,厘清仅限用户本人决策、需由用户授权决策和智能体自主决策等各种决策方式的合理边界及所需权限。确保用户对智能体自主决策享有知情权和最终决策权,智能体执行操作不得超出用户授权范围。

7. 加强行为管控。

发展规则内嵌、行为围栏等技术,确保智能体在公共场所、隐私场所、专门场所等的行为合法合规。探索利用区块链等技术,建立重要应用场景智能体行为可验证、可追溯机制,防范智能体不当行为引发重大风险。

(二) 防范安全风险

8. 提升内生安全能力。

研究智能体数据安全、个人信息保护、密码防护、攻击检测、权限管理、行为控制等安全技术,提升智能体系统安全保障能力,防范数据投毒、隐私泄露、算法篡改、系统漏洞、运行失控等安全风险。研究智能体安全检测技术,探索建立智能体安全评估体系。

9. 加强供应链安全。

制定智能体开发、部署、应用、维护等全周期安全规范,加强模型接入、应用程序接口调用、扩展工具使用等环节安全管理。探索建立智能体供应链安全信息共享和预警机制,及时发布风险提示,提升安全保障能力。

10. 化解应用衍生风险。

完善智能体常态化风险识别、预警及干预机制,强化人机协同审核、拦截阻断等风险处置能力,防

范系统性安全风险。强化智能体应用安全管理，避免智能体被用于自动化攻击、隐私侵犯、虚假信息生成传播、网络诈骗等违法犯罪行为。

（三）完善治理体系

11. 构建分类分级治理框架。

根据应用场景和潜在影响，审慎稳妥开展智能体分级治理。对于敏感领域及重点行业，由网信部门联合行业主管部门确定开放场景，根据相关法律法规、监管要求和安全防护标准，实行备案、检测、问题产品召回等管理措施。对于部分生活娱乐、日常办公等低风险领域，完善智能体评估测试工具，通过合规自测、信息报告、分发平台管理、行业自律等实现高效治理。

12. 健全合规服务体系。

强化智能体风险监测预警、检测评估、咨询、认证等专业服务供给，引导行业积极研发智能体监测工具。开展智能体功能、性能、质量、合规等第三方评测服务，推动认证与检测结果互通互认，为用户选择智能体提供参考。编制并发布智能体技术及应用成熟度报告，为产业研发应用提供参考。（四）强化行业自律

13. 引导行业加强自律。

鼓励行业组织、主要企业联合制定行业自律规则，明确智能体功能合规、算法治理、知识产权保护、公平竞争等规范细则。指导智能体开发平台、分发平台、服务提供者建立公平合理的平台规则、用户服务协议及隐私政策，明确供需双方权责，保障产业健康发展。加强智能体应用风险宣传教育，提升用户安全意识。

14. 探索信用评价机制。

指导行业组织建立智能体市场主体自愿参与的信用评价机制，对于技术滥用、诱导消费、虚假宣传、隐瞒缺陷信息等行为进行信用评价，依法依规开展失信惩戒。引导智能体开发者、开发平台、分发平台、服务提供者等参与信用评价，共同营造良好发展环境。

四、强化应用牵引

积极稳妥推动智能体典型场景应用，牵引技术

产品优化提升，探索形成可复制、可推广的智能体落地应用模式。

（一）科学研究

15. 科研探索。

研发理论推演、模拟仿真等智能体，挖掘潜在技术路径。强化智能体信息关联整合、知识体系构建等能力，提升自然科学、哲学社会科学研究发现能力。促进智能体与科学仪器、实验平台融合，实现方案设计、实验操作、数据处理、结果分析等全流程智能化。

16. 研发辅助。

发展软件开发智能体，提升需求分析、架构设计、代码生成与测试等全流程开发能力。促进智能体与计算机辅助设计（CAD）、计算机辅助工程（CAE）等软件结合，提供设计方案生成、仿真验证、参数调优等功能。

（二）产业发展

17. 智能制造。

研发生产管理智能体，动态优化生产排程、资源分配和工序衔接，推动智能体在工业互联网领域应用，提升企业精细化管理水平。提升智能体工艺参数优化、加工精度检测、产品缺陷识别等能力，促进智能体与数控机床、工业机器人、自动化产线等融合，促进提质增效降本。

18. 能源资源。

研发大气、水体、土壤、噪声等环境要素感知智能体，提升自然灾害、环境污染等风险预警能力。依托智能体强化对国土空间资源全周期管理能力。依托智能体实现能源、金属等矿产资源高效勘探。发展电力调度、用电监测、电网维护等智能体，提升电力资源使用效率。

19. 交通运输。

研发交通安全监管、应急指挥调度等智能体，提升违章违规行为识别、交通基础设施风险预警、重点车辆（船舶）监管、事故快速响应等能力。优化交通监测调度智能体性能，发展交通载运工具管控智能体，提升路网、水网、空域的通行效率。

20. 农业生产。

研发农业服务智能体,开展农技指导、病虫害诊断与防治等服务。推动智能体在种植养殖、高效育种等环节应用,推进农业智能化转型。推动智能体与智能农机具、智慧大棚、农业服务平台融合,提升农业生产效率。

21. 金融服务。

研发金融风控智能体,提升信贷审批、交易监控、账户安全等环节风险识别能力。完善智能体异常检测、合规审计功能,提升信贷违约预测、信用卡盗刷拦截、反洗钱监测等能力。

(三) 提振消费

22. 终端应用。

推动智能体赋能互联网应用及服务,优化在线购物、出行导航、生活缴费、日常办公等服务体验。推动智能体与手机、电脑、汽车、家居、可穿戴、消费级机器人等终端设备协同发展,提升跨应用、跨设备任务完成能力。

23. 文化旅游。

研发文学、音乐、绘画、视听、演艺等内容创作智能体,促进优秀文化传播推广。发展智能导览、多语种翻译、适老适残等旅游服务智能体,提升旅游服务水平。

24. 商业服务。

提升智能体客服能力,提供7×24小时咨询、预约、售后等服务。发展导引、清洁、仓储、配售等具身智能体,提升餐饮、零售、住宿、物流等商业场所的运营效率。探索通过具身智能体提供低成本家政、养老、托育、助残等服务。

(四) 民生福祉

25. 教育教学。探索课件生成、作业批改、学情分析等智能体,提高教师工作效率。依托智能体开展个性化学习方案制定,完善智能导学、答疑辅导、虚拟助教等功能。支持在线教育平台研发智能体,提供终身学习服务。

26. 医疗健康。

提升医学影像分析、疾病诊断推理、定制化诊疗方案生成等医疗辅助智能体性能,探索药品管理、手术排程、病历管理等智能体,提升医疗服务效率。

稳步发展预问诊、报告解析等智能体,提升患者体验。

27. 人力资源。

探索智能体在就业促进、技术技能人才培养评价、劳动关系公共服务等领域应用,提升就业服务能力。发展社会保险、劳动争议仲裁、欠薪治理等智能体,保障劳动者合法权益。

28. 信息服务。

探索智能体在网络内容建设管理中的应用,鼓励信息发布部门和内容传播平台研发用户分析、选题策划、采编加工、分发推荐、智能审核、舆论引导、情绪疏导、实时翻译等智能体,实现多模态信息、跨领域信息的高效整合。

(五) 社会治理

29. 政务服务。

探索事项辅助审批智能体,推动政务审批流程智能化。发展政策咨询智能体,提供全天在线的政务咨询、流程指引等服务。探索主动推送适配政策、服务提醒及办理指南,加快从“人找服务”向“服务找人”转型。

30. 司法服务。

探索全流程办案辅助智能体,提升案件材料梳理、案件信息录入、证据审查、辅助法律文书生成等能力。发展法律宣传、法律咨询、法律监督等智能体,为群众提供高效便捷的在线司法服务。

31. 公共安全。

探索监测预警、应急处置、救援调度、协同治理等智能体,提升安全生产监管和防灾减灾救灾等能力。提升智能体异常行为识别、潜在威胁预警、动态防控处理能力,维护公共安全。推动具身智能体在灾害救援、安防巡检、危险品处置等领域落地应用。

32. 城市治理。

探索智能体在城市规划、建设与治理环节应用,支撑智能建造、房屋管理、城市基础设施安全运行等工作,提升城市治理专业化水平,提升城市人居环境质量。

33. 招标投标。

探索招标投标智能体,实现招标投标活动全链路智慧管理,保障全过程规范高效。提升招标投标交易、服务和监管的智慧化水平,实现招标科学合理、评标公平公正和监管穿透高效。

五、建设创新生态

畅通供需渠道,促进研发侧、需求侧高水平互动,形成市场牵引、内驱发展的智能体产业生态。

(一) 促进产业合作

34. 培育开源创新力量。

引导国内人工智能开源社区加强智能体布局,开展智能体与开源芯片、开源操作系统、开源大模型兼容适配。引导企业、高校、科研机构积极参与智能体框架、交互接口、工具链等开源项目,推动技术体系融通发展,加快提升国际影响力。

35. 搭建产业协作平台。

发挥智能体相关生态联盟、技术验证实验室等产业协作平台的作用,协同产业链上下游开展智能体共性技术研发、标准制定、评估认证等工作,开展智能体技术与产业应用复合人才培养。引导互联网应用、智能终端等领域企业共建生态,探索建立互利共赢的合作模式。

(二) 强化应用推广

36. 构建应用推广渠道。

推动建立智能体软件商店、行业供需信息发布平台,引导智能体企业积极发布产品,形成集聚效应。开展智能体应用供需对接活动,采取公开招标、揭榜挂帅等方式吸引智能体企业定制化开发相应产品。引导整机、软件等企业基于智能体研发产品和服务,培育用户市场。

37. 推进重点场景开放。

推动重点领域开放智能体应用场景,在产业集聚区、重点行业、重点领域开展智能体应用试点,打造一批具有引领作用的示范项目。发展市场化、专业化的智能体技术转化服务机构,探索智能体应用场景,提升技术成果转化效率。促进行业数据共享开放,支撑重点场景智能体训练部署。

38. 积极培育全球生态。

依托世界人工智能大会、世界互联网大会等国

际平台,交流展示智能体技术创新成果。推动终端设备、软件企业适配智能体,引导相关企业做好海外合规建设,推动智能体适应当地法律法规和文化习俗。

六、保障措施

国家网信办、国家发展改革委、工业和信息化部会同有关方面加强统筹谋划,强化资源整合和力量协同,完善配套政策,形成工作合力,推动重点任务落实落地。建立并完善智能体发展评价指标体系,加强智能体规范应用与创新发展的监测评估、滚动实施和动态调整。

《智能体规范应用与创新发展实施意见》答记者问

原载:“网信中国”微信公众号

近日,国家网信办、国家发展改革委、工业和信息化部联合印发《智能体规范应用与创新发展实施意见》(以下简称《实施意见》)。国家网信办有关负责同志就《实施意见》回答了记者提问。

1. 问:请介绍一下《实施意见》的出台背景。

答:近年来,以手机助手、终端智能管家、云端智能体等为代表的智能体产品加速涌现,呈现规模化应用态势,极大便利人们工作生活。同时,智能体高自主性、高权限等特性也带来了隐私泄露、越权操作、行为失控等安全风险,需要统筹发展与安全,促进智能体规范应用和创新发展。

党中央、国务院高度重视人工智能发展。习近平总书记在中共中央政治局第二十次集体学习时强调,要把握人工智能发展趋势和规律,加紧制定完善相关法律法规、政策制度、应用规范、伦理准则,构建技术监测、风险预警、应急响应体系,确保人工智能安全、可靠、可控。2025年8月,国务院印发《关于深入实施“人工智能+”行动的意见》,以科学技术、产业发展、消费提质、民生福祉、治理能力等领域为重点,提出到2027年,率先实现人工智能与重点领域广泛深度融合,新一代智能终

端、智能体等应用普及率超70%的阶段性目标。

制定出台《实施意见》是贯彻落实《关于深入实施“人工智能+”行动的意见》的具体举措，以推动科技创新、提升治理能力、构建产业生态、提升民生福祉为导向，营造良好政策环境，发挥典型场景示范效应，统筹推动智能体高质量发展、高水平安全、高效能治理。

2. 问：《实施意见》的总体思路是什么？

答：在制定《实施意见》过程中，重点把握以下4个方面：

一是强化思想引领。坚持以习近平新时代中国特色社会主义思想为指导，贯彻落实习近平总书记关于网络强国的重要思想，深入实施“人工智能+”行动，推动智能体高质量发展。

二是统筹发展和安全。遵循技术发展客观规律，夯实产业技术底座，完善治理体系，积极稳妥推动智能体规范应用，构建发展与安全协同促进的格局。

三是坚持应用牵引。深入挖掘典型应用场景，推动形成以创新带应用、以应用促创新的良性循环，促进智能体赋能千行百业。

四是守牢安全底线。坚持以人为本、智能向善、多元共治、安全稳妥，明确智能体发展底线红线，将安全可控贯穿智能体技术研发、应用部署与产品推广全过程。

3. 问：《实施意见》在规范智能体应用方面有哪些具体要求？

答：《实施意见》将智能体安全、可靠、可信作为产业发展的底线要求，规范有序推动智能体落地应用。

一是明确产品准则。要进一步完善政策法规和伦理规范，做好智能体权限管理、行为管控等工作，为智能体相关产品研发提供指引。

二是防范安全风险。发挥技术手段作用，提升在智能体内生安全、供应链安全、应用衍生风险等方面的风险防范能力，实现智能体开发、部署、应用、维护全周期安全管理，有效防范系统性风险。

三是完善治理体系。适应智能体技术演进规律，根据应用场景和潜在影响，构建审慎稳妥的分类分

级治理框架，推动建设智能体合规服务体系，做到既“放得活”又“管得好”。

四是强化行业自律。鼓励智能体相关企业、行业组织、研究机构加强责任意识，共同制定行业自律规则。引导智能体开发者、开发平台、分发平台、服务提供者参与信用评价，共同营造良好发展环境。

4. 问：《实施意见》围绕推动智能体创新发展部署了哪些重点工作？

答：《实施意见》围绕智能体技术突破、场景应用和生态建设等重点方向，体系化推动智能体创新发展。

一是夯实发展基础。通过强化基础技术研发、完善智能体工具链，为产业提供高水平技术底座。建立智能体标准体系，降低智能体研发、适配、应用门槛。前瞻布局多智能体协同、智能互联网等前沿领域，为产业持续演进打好基础。

二是强化应用牵引。围绕科学研究、产业发展、提振消费、民生福祉、社会治理等方向，提出19个典型应用场景，牵引智能体技术产品优化提升，探索可复制、可推广的落地应用模式。

三是建设创新生态。加强培育开源创新力量、搭建产业协作平台，推动产业高效协同，提升创新活力。通过构建应用推广渠道、推进重点场景开放，畅通供需渠道，形成市场牵引、内驱发展的产业生态。积极培育全球生态，推动国内外技术融通发展，营造开放共享的国际合作环境。

5. 问：下一步就落实《实施意见》有哪些工作举措？

答：国家网信办、国家发展改革委、工业和信息化部将会同有关方面切实推进《实施意见》落地见效。聚焦智能体技术研发、场景开放、安全治理等关键环节，完善配套政策，形成工作合力，推动重点任务落实落地。同时，加强智能体规范应用与创新发展的监测评估、滚动实施和动态调整。

中国网络社会组织联合会正式发 布《中国网络诚信发展报告 (2026)》

原载：“网信中国”微信公众号

5月20日，2026年中国网络文明大会互联网企业社会责任分论坛在广西南宁举行，中国网络社会组织联合会正式发布《中国网络诚信发展报告(2026)》(以下简称《报告》)。《报告》以习近平总书记关于网络强国的重要思想为根本遵循，系统总结网络诚信建设的内涵演变、基本特征、治理理念，全面梳理2025年我国网络诚信建设的丰硕成果与突破进展。这是中网联连续编写发布的第六份年度《中国网络诚信发展报告》。

今年《报告》分为正文和附录两大部分，正文部分包括政策理论、主体实践、重点专题、发展建议共4大篇章。附录部分包括2025年中国网络诚信十件大事、2025年网络诚信建设工作案例。

报告研究认为，2025年，国家从法规制度层面大力推进网络诚信建设，中央网信办、国家市场监督管理总局等有关部门，印发互联网平台价格行为规则、直播电商监督管理、人脸识别技术应用安全等制度文件，深入开展整治AI技术滥用、自媒体乱象等“清朗”系列专项行动，督促指导网站平台压实主体责任，推动网络诚信建设稳步推进、成效显著，呈现多元协同与责任明晰相统一、虚拟治理与现实赋能相统一、法治约束与文化浸润相统一、动态治理与前瞻防控相统一的鲜明特征。

据了解，《报告》与往年相比，既有传承，也有创新，主要呈现三个特点：

一是守正创新，进一步丰富网络诚信建设的理论体系。立足我国互联网发展治理实践，《报告》研究提出，从基础诚信向协同诚信、全域诚信的发展过程中，网络诚信从技术约束、制度规范向全面治理、高质量发展演进，形成坚持党的领导、坚持以人为本、坚持良法善治、坚持系统观念、坚持创

新驱动的治理理念。

二是立足实践，全面呈现多元主体参与网络诚信建设的经验成效。《报告》总结展示各地各部门、行业组织、各类企业、广大网民等多元主体参与诚信文化传播、城市文明建设、科技向善赋能、行业自律规范等方面的有益探索与成熟经验，具体体现了各类主体加强网络诚信建设的生动实践。

《报告》记录2025年中国网络诚信建设各领域的十件大事，标注年度工作的重要成果。

三是紧扣热点，深入剖析重点领域风险挑战及对策建议。结合新技术发展动向、社会大众关心热点，《报告》以专题形式深度剖析网络谣言治理、人工智能+医疗服务、互联网金融、数字文旅、网络公益等重点领域工作经验与共性难题，分析指出网络诚信建设面临新技术衍生失信问题、网上网下协同治理难度加大、信用评价标准体系不统一等风险挑战，提出规范新技术新应用管理、深化跨域联动合作、凝聚多方合力等对策建议。《报告》的发布，着力于打造网络诚信建设成果交流展示平台，力求为社会各界了解我国网络诚信建设发展状况提供参考借鉴。

中央网信办全面部署推进规范短 视频内容标注工作

原载：“网信中国”微信公众号

针对部分短视频内容来源不清、真假难辨、混淆视听等突出问题，今年以来，中央网信办全面部署推进规范短视频内容标注工作。1月以来，指导网站平台深入清理虚假摆拍等违规短视频52万余个，严惩违规账号6.8万余个，发布治理公告54期，并集中曝光典型案例。3月，指导抖音、快手、腾讯、小红书、哔哩哔哩、微博、淘宝、京东、拼多多、支付宝、美团、百度等12家平台先行先试，完善内容标注标签，对内容标注功能进行优化和测试。

近日，中央网信办总结12家平台试行经验，

指导各地各网站平台全面推进落实三项工作：

一是规范短视频内容标注标签，明确网站平台必须为用户提供6类“必选标签”，并可根据自身实际提供其他“可选标签”。其中，“必选标签”包括“含有虚构演绎内容”“含有AI生成内容”“含有营销信息”“内容为转载”“内容为个人观点”和“无需标注”。真实生活记录类短视频可选择“无需标注”标签，该标签不在短视频页面呈现。

二是将内容标注设为短视频发布必经环节，发布者必须从“必选标签”中选择一项，才能发布短视频。

三是对新增短视频标注情况加强审核，对存量短视频进行分批回溯，对未规范标注的，进行补标或纠正，并对相关发布者进行教育警示，推动实现短视频内容应标尽标。规范短视频内容标注是维护短视频内容生态、保障网民合法权益的重要举措，也是一项长期工作。下一步，中央网信办将加大对网站平台的指导督促和监督检查力度，对未按要求进行标注的账号和主体责任落实不力的网站平台，依法严惩并公开曝光。希望各网站平台、短视频发布者和广大用户自觉遵守标注规范，共同打造健康、有序、可信的短视频环境。

虚构政策、蹭炒热点、误导舆论 多部门联动整治网络谣言乱象

原载：“网信中国”微信公众号

2026年4月，网络谣言主要集中在公共安全、民生政策、农业生产、旅游消费等领域。造谣者或假冒官方、炮制虚假骗局，或歪曲事实、编造离奇故事，误导公众认知、扰乱社会秩序、影响市场预期。网信、公安等部门联动核查，及时发布权威辟谣信息，打击曝光造谣传谣行为，持续净化网络环境。

捏造发布虚假政策信息，借机引流、涉嫌诈骗。

所谓“点链接领‘2026年度综合补贴’”，

系不法分子套取公民个人信息的“钓鱼”链接。还有传言称4月1日起，“70岁以上老人就医新规落地”“出台新规，全国严查后排安全带”，前者很快被有关主管部门证实为无中生有，后者则被公安交管部门指出是对既有法规的过度解读。“5月1日起医生不24小时值班了”则是对现行政策的张冠李戴。少数自媒体为谋取私利，利用信息差编造“吸睛”谣言，制造政策焦虑，暗藏诈骗陷阱，损害群众切身利益。

炮制涉农涉产不实信息，违背规律、干扰生产。

4月正值春耕关键期，“割青麦作饲料”谣言卷土重来、持续炒作，干扰正常农业生产秩序。“四川德阳千亩农业基地被查”“江苏南通大葱种植地免费拿”等均为凭空捏造的不实传言，“湖北已建成多条标准化人造大米生产线”是利用AI恶意生成的虚假内容。清明节后，黄鳝大量上市，一条流传近30年的陈年旧谣“黄鳝养殖被喂避孕药”再次被翻炒，影响相关消费。该谣言无视黄鳝“性逆转”特性，刻意制造食品安全恐慌，借机博取流量、非法牟利。“库布其沙漠投放10万头黑猪治沙”的说法，嫁接真实的治沙图片、视频，并无科学依据，误导公众认知。

炒作社会民生话题，渲染焦虑、扰乱秩序。

网传“环保新规导致鸡蛋涨价”纯属造谣，相关部门并未出台所谓“新规”，该谣言是对若干法规的牵强拼凑，刻意混淆公众认知。涉旅游类谣言同样多发，“青岛景区清明免票通知”“男子九寨沟一夜被宰10万”等虚假信息，蹭炒假日流量，误导群众出行，有损地方文旅形象。此外，“深圳超大医院遇冷、八万床位空置”的传言严重失实，误导公众对公共医疗事业的客观认知。

网信、公安、广电等部门高效联动、重拳整治。中央网信办发布《关于加强网络直播打赏规范管理的通知》，明确要求打赏互动不得含有违法和不良信息。国家广播电视总局整治“AI魔改”违规视频乱象，清理相关违规视频23000余条、处置违规账号100余个。公安机关依法打击造谣传谣行为，对

“湖北人造大米生产线”“男子九寨沟一夜被宰 10 万”“江苏南通大葱种植地免费拿”等造谣传谣相关责任人依法予以行政处罚。

网信部门严管“自媒体”未规范标注信息来源行为

原载：“网信中国”微信公众号

近期，一些“自媒体”账号在发布涉时政等领域信息时，未规范标注信息来源，误导公众认知，破坏网络生态。主要情形包括未标注国内外时事、公共政策、社会事件的信息来源，未标注 AI 生成标识，未标注虚构演绎标签。网信部门督促网站平台深入自查自纠，依法依规处置违规账号 9.8 万余个。现将其中典型案例通报如下：

1. 抖音“青青国际”“极客科普馆”、快手“名妍”、哔哩哔哩“晋说”“军武-史记”等“自媒体”账号，集纳涉美伊等国际时事信息，未标注信息来源，公众难以溯源信息原始出处，无法辨别信息真伪。涉及的账号已被依法依规处置。

2. 抖音“小胖农业”“乡村发展”、快手“农民大姐”“清华鱼妈”、哔哩哔哩“商业策略”等“自媒体”账号，发布涉农业农村、教育、养老等领域公共政策相关信息时，未标注信息来源，公众

无法获得准确、完整的权威信息，可能会基于碎片化内容对政策作出错误理解。涉及的账号已被依法依规处置。

3. 抖音“萌萌哒”“用户赵先生”、快手“农村户外小张 1”“玉婷”、微博“金毛治郁系”、哔哩哔哩“哎呦哎呦小然子”等“自媒体”账号，通过人工智能技术，制作并发布金毛抱小孩、大猩猩护崽对峙鳄鱼、老虎和橘猫开车回家探亲、小猫玩游戏眼花后戴眼镜、柴犬厨神做菜等视频，未添加 AI 生成标识，易误导不了解 AI 的网民，难以区分虚拟与现实边界。涉及的账号已被依法依规处置。

4. 抖音“刘百川商业”“马瑞（在农村）”、快手“玉雕师风云”、微信视频号“遇见翡翠的夏天”、哔哩哔哩“加油嘉宜”等“自媒体”账号，以剧情摆拍、虚构情节等方式，发布外卖员送餐遭歧视、婆媳冲突、代际冲突、未成年人无人看护、丑化农村等内容，未标注虚构演绎标签，借机博取流量，以负面叙事渲染消极情绪，挑动群体对立。涉及的账号已被依法依规处置。

网信部门将指导网站平台全面规范短视频内容标注工作，优化标注功能，将标注设为短视频发布必经环节。提醒广大“自媒体”创作者，在信息内容制作和发布过程中，切实增强信息来源标注意识，按要求主动、规范标注，确保发布信息真实、完整，共建共享真实可信的网络环境。

（技术编辑：来唯希）

研究动态



人工智能治理

1、AI 智能体跨平台代理执行的竞争法评价（鲍奕含）

来源：《法律科学(西北政法大学学报)》2026年第3期

随着人工智能技术的发展，具备跨平台代理执行用户指令能力的 AI 智能体正逐步从辅助工具转变为执行主体。在跨平台代理执行行为迅速涌现的背景下，平台企业与智能体提供者就该行为的正当性产生了争议。争议焦点在于跨平台代理执行仅获得用户授权但未经平台授权，是否破坏市场竞争秩序，构成不正当竞争行为。本文基于该行为给平台带来的竞争损害限度不构成实质性替代、平台负有互联互通义务及用户享有相关权益三方面原因判断行为具有正当性，但主张应划定行为的正当性边界。跨平台代理执行不应超出用户授予的最小必要权限，需满足合理限度并兼顾多元利益，从而促进智能经济的高质量发展。

2、OpenClaw 类自主智能体的分层治理（傅宏宇）

来源：《法律科学(西北政法大学学报)》2026年第3期

人工智能正经历从“生成式对话助手”向“自主执行智能体”的范式转型，新的风险特征对安全治理带来新的挑战。面对智能体能力的非线性扩张，单纯依赖事前预判和准入限制，或采取事后修补的被动式监管路径，都无法准确识别和及时管控风险。基于实现能力与规制同步演进的敏捷治理思路，应穿透风险的外观，根据智能体能力与风险的伴生性原理，通过“本体—交互—生态”的风险分层治理方案，对智能体本体层、交互层、生态层三个层面的核心能力与伴生风险进行解析，并在每一层面提出对应的治理策略，为智能体发展提供可操作、可迭代的安全保障体系。

3、论大模型的法律能力边界（王禄生）

来源：《法律科学(西北政法大学学报)》2026年第3期

大模型需要借助法律知识习得、法律事实建构、法律规范发现与法律论证推理能力完成法律任务。受内生技术逻辑决定，大模型的法律能力边界呈现出不均衡与非线性分布的样态。在法律知识习得维度，大模型长于显性、隐性与非具身默会知识的习

得,却失于动态知识的习得;在法律事实建构维度,大模型长于描述性建构、弱于规范性建构;在法律规范发现维度,大模型长于规范解释、弱于规范识别与规范自洽、失于规范检索与规范续造;在法律论证推理维度,大模型长于内部证成而弱于外部证成。因此,应基于大模型的法律能力边界,划分法律任务的风险等级,形成明确的大模型适用指南,在此基础上建立兼具启发性、反思性与监督性的递进式法律领域人机协作框架,并从认知素养、运用素养与评估素养三个维度塑造法律职业群体的人机协作能力。

4、生成式人工智能服务提供者的阶段性注意义务与侵权责任(李丹)

来源:《现代法学》2026年第2期

生成式人工智能技术的迅猛发展在推动社会创新的同时,亦因技术特性引发多阶段侵权风险,亟须构建与技术生命周期适配的生成式人工智能服务提供者阶段性注意义务与侵权责任规则。针对不同阶段的侵权风险,生成式人工智能服务提供者在模型训练阶段需履行数据合规与算法公正义务,违反义务者应适用过错责任原则,独立承担侵权责任;在运行应用阶段侧重内容管理与用户监督义务,其中,技术原生型侵权场景适用过错责任原则并辅以证据开示与因果关系推定规则,用户诱发型侵权场景则承担过错补充责任;在优化迭代阶段强化模型改进与“通知—处置”义务,生成式人工智能服务提供者未履行改进义务的,应承担过错责任,未及时处置侵权内容的,依据用户参与度差异承担自己责任或与用户承担连带责任。

5、论著作权法上的创作行为——以人工智能生成内容可版权性争议为切入点(涂藤)

来源:《清华法学》2026年第3期

人工智能生成内容的可版权性争议反映了司法实践、理论研究对著作权法上的创作行为的忽视与误解。对此,须探索创作行为的一般理论与构成要件。比较司法史表明,创作行为的价值基础兼具主观与

客观要素。现行法下,创作行为的主观要件应解释为主体创作作品的概括性自然意思,客观要件则为对作品中单独不受保护的最小单元的“选择和安排”。二者共同决定了创作行为与作者身份的构成,体现了创作行为在作品可版权性认定中的独立地位,修正了要求作者对作品具备具体控制力的理论。侵权判定场合,创作行为亦影响作品保护范围的划定,可解释特定情形下何以仅禁止原样使用行为,而不延伸至实质性相似表达。

6、合成数据的“合成”风险与法律规制(霍敬裕)

来源:《东方法学》2026年第2期

作为人工智能领域的“新燃料”,合成数据依托数据脱敏和多维特征重组的技术路径,既解决了传统数据应用的隐私合规困境,亦突破了跨境协作的信息壁垒,实现了数字治理从数据资源化向数据工程化的范式跃迁。然而,既有的数据治理范式尚无法有效识别和规制合成数据全生命周期中呈现的溯源性风险、过程性风险和衍生性风险,数据仿真度的提升与特征维度的扩展客观上导致主体资格认定、客体权属界定以及应用场景规制等法律关系的复合性异变。治理的目标不再是单纯“管理”既有的数据,而是“塑造”合规的数据集。为此,需重构以结果为导向治理中的“数据控制者—处理者”二元架构,转向原始数据控制者在初始训练集构建阶段、算法开发者在模型优化过程中和运营方在应用场景拓展时的三元责任分配模态,实现合成数据从生成到传播再到价值实现的全链条治理。

7、医疗人工智能产品责任中的发展风险抗辩规则(方乐坤)

来源:《东方法学》2026年第2期

从医疗技术进步需求、防止责任绝对化的制度效应等角度衡量,应将发展风险抗辩规则继续适用于医疗AI产品责任之中。应秉持优先保障安全利益、兼容发展利益、以技术进步推动实现高水平医疗安全的理念,对医疗AI场景下制造商的理性或注意义务作适度扩张设定,构建一套合理安全的发展风

险抗辩规则体系。鉴于医疗行业的高敏感性，应将医疗 AI 产品责任中的发展风险抗辩事由设定为相对免责事由，令制造商对算法运行中的不可控环节履行警示告知义务，对缺陷产品负担召回等义务。应根据医疗 AI 产品的危险程度弹性把握合理发展风险的判定标准。“黑箱”医疗的不可解释性不应该当然解释为合理发展风险；应倡导与促进社会安全价值观相一致的负责任的医疗技术创新，尽可能使医疗 AI 产品有效满足可解释性要求。

8、生成式人工智能侵害肖像权的侵权责任（程啸）

来源：《法学论坛》2026年第3期

生成式人工智能侵害肖像权和声音权益的侵权责任中，如何认定肖像与声音的“可识别性”以及怎样妥当处理侵害肖像权与侵害个人信息权益的侵权责任的关系是两大疑难问题。肖像的可识别性应以社会一般人能否以视觉或听觉的方式通过外部形象识别出特定自然人作为判断标准。人脸替换引发的侵害肖像权案件中，关键在于去除面部特征后的视频还能否使社会一般人仅凭肉眼而识别出该权利人。如果虚拟人脸只是与某个或某些自然人的脸有一定相似性的，但只要社会一般人不能凭此识别出特定的自然人，也不能认定为肖像。利用人工智能技术生成的新的声音有无可识别性，应以社会一般人能否依靠听觉识别出特定自然人为标准。肖像或声音既受到肖像权的保护，作为个人信息，也受到个人信息权益保护。权利人可以选择主张侵害肖像权或个人信息权益的侵权责任。个人信息权益与肖像权的合理使用规则能够相互适用。

9、生成式人工智能服务提供者的注意义务（姚佳）

来源：《法学论坛》2026年第3期

生成式人工智能服务提供者注意义务的界定，是认定侵权责任的核心前提。服务提供者因部署具有固有风险（如“幻觉”现象）的人工智能系统而持续开启风险，且具备事实上的风险控制能力而应尽相应注意义务，履行该义务应符合行为效益与风险防范负担之权衡原则。现行法相关规定及平台用户协议

已设定了覆盖内容标识、内容安全管控、算法透明、数据质量保障、个人信息保护等多维度义务，侵权法中的交往安全义务在此难觅独立的适用空间。过失认定仍应坚持以“理性人”标准为基石，但需将注意义务的适用范围从单纯的人类行为合理性拓展至“人机交互系统的整体安全性”，以契合技术特性与风险结构。据此，服务提供者的注意义务可类型化为技术特性维度的注意义务、提供服务过程中非技术维度的注意义务以及程序性的注意义务。这一分析框架在侵权法体系内为过失认定提供了类型化的判断基准，有助于合理界定服务提供者的民事责任边界，实现权益保障与行为自由的协调。

10、人工智能风险的财税法协同治理（胡元聪）

来源：《中国法学》2026年第2期

财税法协同治理人工智能风险具有独特优势。但当前的财税法协同治理在原则协同、规范协同和功能协同层面上存在不足，不利于实现对该风险的有效治理。财税法协同治理的创新需要实现理论跃升、制度跃迁和功能跃阶。在理论跃升层面，应坚持财税法治理的原则协同、规范协同和功能协同；在制度跃迁层面，应推动财税法激励与约束的内部协同和跨界协同；在功能跃阶层面，应通过原则协同、规范协同和功能协同，实现风险治理的效能跃进。最终形塑构架精巧、脉络清晰的人工智能风险财税法协同治理框架，从而体现中国财税法的法治价值，彰显中国财税法的法治自信。

11、论人工智能安全管理义务及其承担者的刑事责任（皮勇）

来源：《中国法学》2026年第2期

对于人工智能系统自主决策和控制造成的危害后果，现行法律规制存有空白。若仅采取民事、行政法律手段，不足以遏制此类严重危害，需要引入刑法治理。关于刑法治理的路径，人工智能系统犯罪主体说难以实行，而妨害人工智能安全管理秩序犯罪立法的路径能够兼顾人工智能安全与发展，其中，采取违反人工智能安全管理义务犯罪的立法方

案能够有效遏制前述危害。人工智能安全管理义务是为防控人工智能安全风险而设定的特殊、有限的规范责任。为有效惩治严重违反该义务的行为，有必要在刑法中增设违反人工智能安全管理义务罪。在罪行构造上，应依据义务犯、风险分配等刑法学理论，以有效降低人工智能安全风险为标准，合理规定该罪的人工智能安全管理刑法义务、主体和罪过等要件。

12、生成式人工智能版权侵权责任分层配置研究 (刘友华)

来源：《法学杂志》2026年第3期

生成式人工智能技术的市场化应用，蕴含开发者、服务提供者等多主体贡献。各主体行为在提升技术效能的同时加剧了版权侵权内容的传播风险。国内外相关案件频发表明风险已成为现实并日趋加大。我国司法实践多以违反《生成式人工智能服务管理暂行办法》设置的公法注意义务，判令服务提供者承担版权侵权责任，忽略了其他主体对于版权侵权行为的控制能力与获益现实，不仅不利于侵权行为预防，还可能造成寒蝉效应，阻碍技术创新与应用。生成式人工智能版权侵权责任应分层配置，在技术维度、法理维度以及经济学维度均具有正当性。开发者的版权责任应结合过失侵权规则认定，赋予其法定注意义务，基于邻近性、可预见性规则科以“理性机器人”标准，将是否违反一般注意义务交由司法裁判。对服务提供者适用“通知一必要措施”规则。开发者和提供者的责任关系应结合共同故意侵权、共同过失侵权以及无意思联络的共同侵权规则具体认定。

网络平台的治理

1、论超大型数字平台对从业者的社会责任(喻术红、李梦云)

来源：《法学评论》2026年第3期

平台从业者的权益保障目前得到了一定程度的改善，但依然面临困境。超大型数字平台用工行为的负外部性缺陷，是从从业者权益受损的重要原因。当

下有必要打破传统劳动法保护的路径依赖，立足超大型数字平台的社会责任，探讨其对平台从业者的劳动权益保障义务。根据利益相关者理论，平台从业者可能具备员工和非员工两类利益相关者地位，超大型数字平台据此应履行不同范围的劳动权益保障义务。为了确保超大型数字平台履行对从业者的社会责任，应明确超大型数字平台具体的劳动保障法律义务，敦促其在用工链条中劳动尽责，并构建有效的平台企业社会责任评价机制推动其履行社会责任。

2、数据污染的平台治理责任(张喆锐)

来源：《东方法学》2026年第2期

生成式人工智能和社交媒体的发展加剧了数据污染的现代风险。泛滥的污染数据信息可能会导致个体认知偏差、民主协商失序以及社会信任脆弱等风险。公法应当对数据污染的平台治理加以规制。首先，在义务主体上，对污染数据信息的主要审查义务不应由个体而应由平台承担。其次，在义务内涵上，一方面，应在区分事实逻辑类、价值偏见类以及低价值类污染数据的基础上，对数据的可靠性程度进行类型划分。另一方面，应区分信息智能生成阶段与信息传播阶段，结合模型学习、模型应用、社交媒体平台传播以及具体应用场景的差异，确定不同类型平台的可靠数据保障义务。最后，在不法后果上，应构建合比例的行刑规制体系，以《刑法》第286条之一作为数据污染平台治理的最后法保障。

数字行政与智能司法

1、人民城市理念下智慧城市治理的立法保障(王怡)

来源：《法律科学(西北政法大学学报)》2026年第3期

智慧城市数据驱动的治理范式在重塑城市运行逻辑的同时，催生了数据效能最大化与权利保障刚性化之间的内在张力。这一张力并非智慧城市的内生缺陷，而是技术赋能、治理体系与价值目标尚未深

度耦合的阶段性表现。智慧城市建设应实现治理价值、制度设计与技术应用的有机统一。人民城市理念为智慧治理体系的构建提供了根本价值遵循。为推动智慧城市治理与权利保障的精准适配，应构建技术赋能与权利保障互促共生的城市治理生态，在国家立法引领下发挥地方立法的细化与创新功能，结合比例原则、正当程序原则、数据权利束理论，针对数据采集与开放共享、算法应用、特殊群体倾斜保护、应急场景权利平衡等关键环节，构建精细化、可操作、前瞻性的权利保障规则体系，为实现数据效能与权利保障从张力并存到协同整合的转换提供系统性的法学方案。

2、自动化行政的法律保留（周维栋）

来源：《法制与社会发展》2026年第3期

在数字技术驱动行政行为自动化运行的过程中，传统的依法律行政容易异化为依“算法”行政，不仅导致行政权力逸脱法律的控制，还使得公民权益受到技术的过度规训。在技术的裹挟下，自动化行政的法律风险呈现出“技术驱动性、形态隐蔽性、影响广泛性”的鲜明特征。这既源于数据与算法的技术缺陷，也来自行政行为形式创新对行政相对人权益保护的挑战。法律保留作为依法行政原则的核心支柱，通过“事先授权”与“边界划定”的规制逻辑，可以为行政权的技术化行使设定明确的法治底线与“规范密度”。在保留范围上，法律的必要性对于行政的“自动化”与自动化的“行政”显然不同，传统行政行为自动化实施无需适用法律保留，而新型自动化行政行为应仅保留“干预性”行为。在保留强度上，从“立法层级”与“立法内容”两个维度，应通过“法律—法规—规章”三级授权适配不同干预程度，并紧扣技术特性明确自动化系统的技术标准与行政相对人的程序权利，以实现技术风险精准调控与法律介入强度的合理平衡。

3、数字法院背景下审判权运行逻辑研究（洪冬英）

来源：《政治与法律》2026年第5期

数字法院建设是司法领域的重大变革，通过数据整

合、算法建模与流程节点控制，将以算法模型为代表的数字技术深度嵌入司法体系，并与司法责任制中适法统一、质效监管以及责任追究等机制相耦合，对审判权运行产生影响。在法官办案层面，数字技术既促进了信息共享与适法统一，也可能压缩法官的自由裁量空间。在院长、庭长监管层面，数字技术推动监管逻辑从静态审批转向动态预警，从集中行权转向节点控制，形成更具渗透性的数字监管格局。这一过程在提升审判规范化程度与效率的同时，也带来了技术理性与司法价值之间的张力。未来，应坚守数字技术的辅助定位，明晰其应用边界，通过构建高质量且可控的技术实施框架，完善制度保障与闭环监管机制，推动数字法院建设提质增效，在赋能审判与保障法官主体地位之间寻求平衡。

数字时代的部门法问题

1、数字时代的公共失信联合惩戒及其法律规制（冉克平）

来源：《政法论坛》2026年第3期

在数字时代，公共失信联合惩戒是社会信用体系建设的重要环节，系基于信用信息以及算法评估失信状态而实施的风险管控措施。公共失信联合惩戒是公共管理机构行使数字权力的具体表现，呈现“违法违规—严重失信—联合惩戒”的逻辑。公共管理机构既要保障公民权利不受数字权力的恣意侵害，也要增强违法行为人失信后果的可预测性。“严重失信”的适用范围通常限于故意实施的严重违法违规行为，原则上不包括违约行为和违反道德的行为。公共失信联合惩戒属于新型社会治理方式，但是过于偏重解决现实问题的“有效性”而忽视权力运行的“正当性”。公共失信联合惩戒以分配正义为价值基础，其本质是对“行为人违法失信人格”的制裁而非对“行为人违法行为”的制裁，在法律属性与运行机制上与经典行政行为存在差异。公共失信联合惩戒措施应当摒弃形式主义解释论而采取实质主义解释论，符合比例原则和不当联结禁止原则的要求。

2、算法控制下签约主播劳动权益保障困境的制度回应（艾琳）

来源：《法学评论》2026年第3期

随着人工智能实践和算法技术的飞速发展，作为新就业形态的职业主播规模日益壮大。在算法和平台的双重隐蔽控制下，签约主播的劳动发生异化并引发了更深层次的劳资关系冲突，面临着劳动身份界定不清、超时劳动以及劳动报酬获得难等问题。这些困境的背后，是平台通过数据垄断、流量分配与情感劳动要求，对签约主播实施严格且隐蔽的劳动管理。面对从属性标准在新就业形态语境下解释力不足的现状，可以从两个层面重构签约主播的劳动权益保障逻辑：在理论层面，可将技术从属性、类雇佣特征等作为辨析劳动关系的新思路，结合现有规范进一步将劳动关系认定和基本劳动权益保障解绑；在实践层面，以落实签约主播最关心的劳动报酬权和休息权为切入点，构建算法解释机制并强化平台的用工责任，以期实现直播平台和签约主播之间权利义务关系的平衡。

3、从价格到数据：数字时代反垄断分析框架之转型（朱战威）

来源：《清华法学》2026年第3期

既有“价格中心型”反垄断分析框架仅仅聚焦于可量化的价格因素，不仅导致反垄断法在实施中不断遭遇挑战，而且难以契合数字时代的市场竞争规律。该分析框架面临的时代困境主要表现为零价模式改变了价格因素的作用机理，跨界竞争造成价格分析的技术难题，数据隐私超出了价格理论的解释范畴。为更契合数字时代竞争实践，反垄断分析的核心要素应当从价格转换至数据，由此确立“数据导向型”分析框架。相应地，反垄断法亦需通过制度变革进一步体现数据的核心分析功能，包括将数据作为市场地位认定的核心要素、重构行为违法性判定标准以及将数据隐私明确纳入调整范围，从而提升在数字时代的解释力与回应性。

4、云服务定价模式的垄断风险与梯次监管方案（李鑫）

来源：《东方法学》2026年第2期

在云服务市场中，超大型云服务提供商惯用云信用额度与数据出站费用的定价模式吸引客户。然而，高面值、大范围的云信用额度筑高了市场壁垒，高昂且具有不确定性的数据出站费用加强了客户锁定效应，两者的结合不仅削弱了竞争对手的竞争能力，也损害了消费者利益，造成了竞争损害。由于定价模式的垄断风险难以被认定为典型垄断行为，且反垄断需在产业发展与维护竞争秩序之间保持平衡，因此，应针对云服务行业属性、风险阶段适用梯次监管而非事后规制路径。在梯次监管路径下，算力资源因无法满足要件不能作为行业必需设施，临时措施与经营者承诺则能够达到监管云服务的及时性、预防性、救济性目标。实践中，临时措施与经营者承诺所包含的救济措施应重点关注以下几种：提升云服务的价格透明度、推动云服务提供商之间的互操作性、实现云服务之间的可携带性。

5、数字纪检监察的逻辑起点与法治化治理的体系化展开（李智伟）

来源：《法学论坛》2026年第3期

数字纪检监察是在党和国家监督体系中嵌入大数据与算法技术，对权力运行实施全周期、穿透式监督的数字化治理形态。其以数据驱动、平台整合与模型识别为基本特征，在提升监督实效的同时，也带来数据驱动下的合法性制约、算法赋能下的可解释性不足、体系监督下的权责边界模糊以及数字形式主义对监督实效的消解等法治化挑战。立足其生成逻辑，应以合法性原则、正当程序原则和人的主体地位原则确立数字纪检监察法治化的治理理念。在此基础上，从形式与实质两个层面展开其法治化路径：形式层面以保障数据处理的合法性与程序正当性为目标，通过规范数据流动、约束算法过程并强化监督体系的依法协同，实现技术运行的可控与可审；实质层面以提升监督结论的实体正当性与腐败认定的精准性为目标，借助数字技术强化事实识别，并通过矫正形式主义倾向，使监督回归功能本位，进而形成形式与实质相统一、内在贯通的数字

纪检监察法治化治理体系。

6、中国平台用工的特点和新就业形态劳动者权益保障——基于不完全劳动关系的法理学思考（娄宇）

来源：《法制与社会发展》2026年第3期

中国的平台用工合规和新就业形态劳动者权益保障问题需要基于中国国情来解决。劳动者与平台企业及相关机构之间可能存在多种法律关系。劳动关系的本质系支配性管理。不完全劳动关系是具备较弱的人格从属性和组织从属性、一定的经济从属性的独立法律关系，而非一种居于民事劳务关系和劳动法关系之间的过渡类型。类推适用其他法律无法为新就业形态劳动者提供系统性保障。现行劳动者权益保障制度的法理依据是从属性及其内部各类要素。可以将不完全劳动关系这一标识性概念作为出发点，将从属性要素与劳动者权益保障之间的法理关联性作为原创性理论，构建中国平台用工的自主劳动法知识体系。在该理论的指导下，可以进而发掘不完全劳动关系中的平台算法、数据信息、劳动者自主决定工作时间和工作量三大要素，解释三者与各类劳动者权益的关联性，为新就业形态劳动者权益保障立法构筑坚实的理论基础。

7、金析为证：大数据分析报告证据化的二元模式（陈如超）

来源：《中国法学》2026年第2期

公安机关近年大力推动金析为证，即将侦查人员的资金分析报告全面改造成鉴定人的司法鉴定意见。然而，在侦查阶段，公安机关实际上存在侦查型资金分析（侦查人员的资金分析研判）与鉴定型资金分析（鉴定人的资金分析鉴定）两种类型。根据侦查人员与鉴定人的诉讼角色及诉讼功能差异，金析为证应当采取二元模式：侦查人员的资金分析报告作为资金数据检查笔录；鉴定人的资金分析报告作为法定的鉴定意见。基于金析为证的实践探索与理论分析，刑事司法中各类大数据分析报告的证据化路径可以采取“大数据检查笔录+司法鉴定意见”

的双轨制模式。证据类型的二元化既有助于规范大数据分析报告的生成机制与评价机制，又能推动大数据类司法鉴定工作的规范化与制度化。

8、数智时代税收征管技术辅助权的法律规制（李蕊）

来源：《法学研究》2026年第3期

税收征管中数智技术的深度嵌入，引致“技术权力化”“权力技术化”的双向融合发展。征管科技企业基于行政委托和自我赋权，实质获得了具有一定公权属性的功能意义上的税收征管技术辅助权。因耦合了技术、资本、权力三重运行逻辑，税收征管技术辅助权在行使时可能越界侵蚀税收征管权，压缩侵害纳税人权利。由于技术逻辑与税收征管法治逻辑之间的巨大张力，传统上限制公权、保障私权的税法规范难以发挥有效的规制作用。为确保税收征管技术辅助权的规范运行，应在行政委托合同中明确征管科技企业的技术服务标准和质量，并设定征管机关的单方变更权、解除权等行政优益条款，通过健全算法审查机制、严格全过程测试方案验证、完善审计跟踪机制等强化征管机关的税收数字监管，基于自动化决策的影响和复杂程度确定人工介入的时点和方式，通过风险分类实现对税收征管技术辅助权的差异化规制。

9、数智时代犯罪化立法的优化路径（李海良）

来源：《法学》2026年第5期

数智技术的快速演进加剧了刑事立法的滞后问题。我国刑事立法对技术风险的总体回应，非但未能充分解决危害规模、规制门槛与风险形态三重理论难题，反而存在实害评价标准模糊、前置法与刑法衔接不畅、法律拟制尚无专门规范的立法症结。法益侵害的实质性迭代、前置法体系的规制失灵及“补丁式解释”的困局，强化了刑事立法优化的需求。积极刑法观与消极刑法观在数智风险治理上均存在自身难以克服的不足，应当确立以担当刑法观为核心理念的立法理论范式，实现立法理念从“是否要立法”的二元争议向“如何负责任地立法”的转

变。担当刑法观以社会治理主动性与权利保障审慎性为辩证内核，并对数智领域犯罪化立法的动因、对象与程度，分别进行目的正当性、行为类型化与损益相当性审查。未来，新型数智犯罪立法应当重塑适应规模化与累积性犯罪后果的评价基准，调整衔接前置法的罪状与罪名，以及构建共犯行为正犯化、预备行为实行化的罪名体系。

数字法学基础理论研究

1、数据安全法益的体系定位与规范适用（阎二鹏）

来源：《政法论坛》2026年第3期

当前理论面向与实践面向对数据安全法益存在双向认知困惑，前者面临数据安全法益独立性的证成困境，后者则导致司法实践对狭义数据犯罪的无序扩张与罪名选择等困境。数据与信息在形式与规范层面均有差别，对两者的混淆解读会导致数据犯罪构成要件的解释困境，数据安全无法等置于信息安全，在前置法的形塑指引下，数据安全应着眼于确保其处于有效保护与合法利用状态的动态安全，从而与关注数据自身安全结果的静态安全形成区别。动态数据安全法益的独立内涵需从数字生态系统本身所建构的技术法则中寻找，通过限缩数据类型可实现对获取型数据犯罪的认定困境化解以及对破坏型数据犯罪的认定纠偏，从而为验证动态数据安全法益的机能提供实践路径。广义数据犯罪的多元属性决定了数据安全法益保护的立法安排，仍应维持从动态数据安全到数据权益的多元保护路径，并通过聚焦较“获取”更为前端的“非法侵入”行为、界分作为工具的算法犯罪与作为对象的数据犯罪在立法逻辑上的差异等划定其保护范畴。

2、人机价值对齐的法哲学路径（郑玉双）

来源：《法制与社会发展》2026年第3期

人机价值对齐是应对人工智能技术伦理挑战的积极姿态，但其在概念内涵上存在争议。对人机价值对齐的内涵目前存在技术路径和伦理路径两种理解。技术路径强调保障人工智能技术应用的安全，伦理路径强调人与机器的价值互动，这两种路径都未能充分展示人机价值对齐作为创造性和制度性

实践的属性。既有研究针对人机价值对齐的对象、方式和方法论框架已经进行了一些有益的探索。在此基础上，应当从伦理学视角转向法哲学建构。在人工智能技术的冲击下，社会规范发生了物理学转型，亟需规范空间的再造。从法哲学视角出发，可以建构人机价值对齐的可能规范框架，即人机价值对齐的主要实践机制在于构建计算规范。计算规范是价值表达，具有建构性和解释性。计算规范表达了社会基本善，可以参照法律作为价值实践的丰富经验，通过制度性累积而不断完善自身的实践机制，实现实质的人机价值对齐。

3、数字法律预防的理论基础与制度构建（郑智航）

来源：《中国法学》2026年第2期

数字法律预防是在网络空间、数据应用、人工智能系统等数字环境中，运用系统性、前瞻性的数字技术手段预先识别、评估、规避潜在法律风险，从而减少违法行为发生、防范权益侵害的综合性预防机制。数字法律预防具有鲜明的架构化特点。数据驱动是数字法律预防的认识论基础，抽象风险下的法秩序和公共安全是数字法律预防的主要保护客体，行为阻断是数字法律预防的主要效果机制，“全息统御”是数字法律预防的时序位置，人机协同是数字法律预防执行机制的鲜明特征。数字法律预防整体上契合现代法律价值，但受数字法律预防底层结构的影响，不可避免地与现代法律追求的部分价值存在一定的张力或冲突。沿着“法律目的一法律关系一预防行为一技术规制”这一脉络构建数字法律预防制度，对于减轻或化解数字法律预防与现代法律价值之间的张力或冲突具有重要意义。

数字时代的财产权益保护

1、公法形成数据财产权的机理（刘连泰）

来源：《现代法学》2026年第2期

《民法典》第127条用了留白式的表述，将数据财产的保护转介到该法典之外。用民法上的物权、债权或者知识产权保护数据财产面临体系性难题，对数据财产使用规制模式有诸多短板。公法形成数据

财产权更灵活,有制度优势,应当通过公法形成数据财产权。“数据二十条”有大量有关数据权的修辞,形成了数据资源持有权、数据加工使用权和数据产品经营权等权利类型。经由公法形成的数据财产权,整体上与民法典融贯。公法形成数据财产权,对财产权体系有较大影响,改变了公法和私法在财产权领域的分工。

2、网络虚拟财产的二元划分与结构性权利配置(赵自轩)

来源:《法学评论》2026年第3期

网络虚拟财产包括映射类财产与建构类财产两种类型,二者在产生机制与权利内容上显著不同。映射类财产是用户人格利益或财产利益在网络空间中的映射,用户对其享有控制权、使用权与流转权。平台提供者对映射类财产的管理应当符合程序正义与实质正义,涉及用户重大利益的虚拟财产管理性条款应取得用户的“单独同意”。平台提供者应通过规则完善与技术创新充分保障用户依法行使虚拟财产流转权。建构类财产权的内容直接取决于集体共识,但受到国家监管政策的严格约束。为此,建构类财产权的内容不得危害公序良俗,其流转应当依法依规有序进行。

3、企业数据资产担保的制度构造及其实现(李鸣捷)

来源:《现代法学》2026年第2期

企业数据资产的法律定位是企业数据控制者利用权,将来数据资产可以作为担保客体。企业数据资产担保可采抵押或权利质押构造,并应重视担保财产保全规则的适用,加强对担保权人的权利保护。就企业数据资产担保的设立而言,数据的无体性、强流动性无碍于客体特定要件的满足,且宜将“特定”缓和为“可特定”。就企业数据资产担保的公示而言,其公示方式为登记。在效力模式层面,可暂且采登记对抗模式,待未来时机成熟时改采登记生效模式。企业数据资产担保的实现方式包括变价与就收益主张权利,前者应以市场法为主要评估方

法,同时就评估人员资质设置数据领域的准入条件,后者应紧扣“原始担保物—收益”的价值演变路径以确定收益范围,并借助区块链等技术支持实现对收益的有效监控。

4、论数据持有事实及其变动的规范意义(宁园)

来源:《法学评论》2026年第3期

数据持有事实及其变动乃数据产权制度构建的客观土壤,有必要回应因此产生的规范难题。数据持有事实本身将产生三个规范难题,即不受持有行为控制之数据是否受到法律保护,是否有必要为所有持有事实提供一般性保护,持有事实的法律效力如何设置。对此,应确立数据持有保护的一般性规则,明确持有事实仅具有财产特定化功能,无推定效力和公信力。数据持有的变动将引发自主持有与他持有并存、有权持有与无权持有并存等复杂数据持有关系,并产生新的数据产权分配难题。对此,可尝试确立间接持有保护规则,明确持有者无权处分数据或进行非合意数据生产时的数据产权分配规则。

5、从支配权到规制权:知识产权法律属性的再审视——兼论数据权能否成为知识产权新类型(唐艳)

来源:《法制与社会发展》2026年第3期

知识产权是支配权乃学界通说,而著作权和专利权不具有支配权自用自足的特征,故知识产权支配权论并不恰当。通过创设一个与支配权同位阶的“规制权”概念,即权利人得直接禁止或许可他人法定的特定行为的权利,知识产权的“法律之力”得以昭示。商标权兼具支配权和规制权混合属性。从权利结构和制度化进程视角观之,规制权的客体为法定的特定行为。从法经济学和分析法学维度论之,规制权的行使方式为禁止或许可的规制自由。知识产权规制权论的提出,纾解了后现代反本质主义哲学思潮引发的知识产权客体探寻困境,还使知识产权在民事权利体系中因彻底摆脱了“类物权”“支配权”等传统理论枷锁而更加独立,并为新型财产权的创设提供了新的法律技术工具。例如,数

据权可被赋予规制权的“法律之力”而成为知识产权新类型。

6、论数据无主物（许恋天）

来源：《法学杂志》2026年第3期

数据确权贯穿数据要素从无主到有主的历史进程，数据产权建构难以忽视事实上无主的部分数据或数据资源。有必要针对数据这种特殊的无体物创造性适用传统旨在确权的无主物制度，通过法律拟制构建促进数据流通的数据无主物规则。数据无主物是数据财产权体系的建构起点和转化中介，可以类型化为自然无主数据与拟制无主数据，前者开启数据财产权从无到有的首次配置，后者推动数据财产权从有到无再到有的迭代配置，进而摆脱数据确权面临的初始权属空白、价值循环阻滞两大困境。为充分发挥数据无主物激活数据流通、减少数据闲置的功能，规避先占衍生的数据垄断等失序风险，需明确数据公有物向数据无主物的释放边界，厘清数据无主物转化为数据私产的正当前提。目前，可以对数据无主物进行适当的类型化并明确一般处置规则，同时在公共数据开放、授权运营及商业化平台采集等特殊场景下形成协调数据公有物、无主物与私有财产的数据可采集制度，进而构建多方平行持有与共同合理利用的处理原则，为数据产权的多层次初始分配提供必要制度路径。

7、数据集合商业秘密保护的力度与限度（李安）

来源：《政治与法律》2026年第5期

商业秘密是数据财产权益保护的新近实践动向。对于数据集合的商业秘密诉求，不能简单主张保护或不保护，应在保护力度与限度的互动关系中对其进行辩证分析。数据集合尽管由公开信息组成甚至面向公众开放，但仍有商业秘密属性，且数据集合的商业秘密保护在认知经济性、创新与投资激励水平、数据流动与利用效率等方面具有制度优势，可与数据不正当竞争保护并行，因此不宜完全拒绝数据集合商业秘密保护。权益保护的承认与权益限制的确认是同在的，权益限制是证成权益保护的“后

果性说理”，应在给予商业秘密保护的同时强调限制：一是在事前加强数据集合的商业秘密客体资格审查，二是在事后限定数据集合商业秘密的权能行使范围。数据集合的商业秘密保护是有边界的、可控制的，因此是可欲的。

个人信息保护研究

1、赋权与规制：数据可携权的二元保护构造（徐则林）

来源：《现代法学》2026年第2期

数据可携权关乎个人信息自决与数据要素流通，其法律保护存在以个人信息保护法为基础的“赋权保护”与以反垄断法为基础的“行为规制”两种路径。然而，实践中前者常因承载过高的功能期待而陷入空置或异化的困境，后者则面临过度干预而抑制创新的风险。摆脱困境的关键在于探寻能够整合二者的系统性制度架构。赋权范式侧重于保障个体对基础性个人信息的自主控制，规制范式则侧重于维护市场竞争秩序，二者在数据可携权所涉义务主体的范围、所涉数据类型的边界、所涉转移时限的要求、所涉转移方式的条件等诸多方面均存在客观差异。在系统论法学视角下，为促进数据要素市场的健康发展，应确立一种层次分明、先后有序、功能互补的数据可携权二元保护构造：在赋权范式下保护最低限度的数据可携权，在规制范式下保障进阶数据的可携性，二者各司其职。此外，还需考虑纳入以数据中介为代表的补充机制，以解决复杂数据的技术性难题，增强该二元构造应对现实复杂性的韧性与实效。

2、中国个人数据处理捆绑禁止的路径再造：从绝对禁止到相对允许（傅雪婷）

来源：《清华法学》2026年第3期

我国数据立法正在紧锣密鼓地制定中，《个人信息保护法》作为我国第一部专门规范个人信息处理的法律，深刻地影响着数据流通利用。该法第16条规定了捆绑禁止，即处理者不得以个人拒绝或撤回同意为由，拒绝提供产品或服务。上述禁令同时适

用于存在合同关系、无合同关系及合同磋商情形，仅在处理为提供产品或服务所必需时构成例外。该必需性判断与个保法第13条第1款存在交叉，表现为无合同关系时依客观直接相关性判断，有合同关系时引入主观标准辅助的构造。违反捆绑禁止将导致同意无效，合同效力亦受影响。从立法者解释及学说立场看，我国倾向于对捆绑采绝对禁止立场，部分司法判决则表现出松动的迹象。上述“大适用范围+强后果+绝对禁止”的规制模式无法适应数字社会的发展，亦不利于信息自决、营业自由、数字经济的良性互动。对此，应构建相对允许捆绑的规制模式，基于个人与处理者权力关系、产品服务提供状况、合理替代给付、利益平衡等因素做实质化衡量，扩展不应被禁止的例外情形，实现数据保护与利用的平衡。

3、生成式人工智能中的个人信息保护（李西岭）

来源：《法学论坛》2026年第3期

生成式人工智能以大规模数据驱动为基础，其运行机制贯穿数据收集、标注、模型训练、内容输出以及反馈与再训练等多个阶段。在以生成式人工智能的生命周期为分析框架的基础上，应聚焦于其中个人信息保护问题最为集中的数据收集、数据标注以及内容输出与反馈阶段，通过梳理各环节中个人信息来源复杂化、“已公开信息”合理预期失衡、委托处理责任模糊以及生成内容再识别风险等核心难题，进而构建来源限制、安全评估、退出与纠错机制等制度设计以调整个人信息规模化利用的正当边界。

4、个人信息调取的正当程序（张迪）

来源：《法学研究》2026年第3期

实践中，法律规范相对滞后和粗疏，个人信息调取的授权规范不明确、调取程序失范、程序性保障缺失等多重问题亟待化解。侦查规范体系改革的功利化和被动化、技术治理下个人信息共享的政策推动、个人信息权益侵犯的隐蔽性和无形性等因素叠加，进一步加剧了个人信息调取的法律控制难题。

个人信息调取的规范体系以基本权干预之判定为起点，由调取之法律授权、调取之令状程序、调取之程序保障等构成正当程序三元框架。我国个人信息调取正当程序的三元结构应包括：以隐私权为中心的比例控制；以检察机关为中心的准司法令状制度；以个人信息权益为中心的程序保障。据此，应在比例原则之下运用“五要素衡量法”和“五要素干预法”完善授权规范，明确令状审批主体、申请主体和具体内容，构建四元保障程序。

5、个人信息分类赋权与规则展开（巩姗姗）

来源：《法学杂志》2026年第3期

个人信息权利属性界定与内容构造，直接决定个人信息保护强度与个人信息利用限度。在数字技术和智能技术深度融合趋势下，个人信息处理底层逻辑已发生深层变革，个人信息控制论的制度功能日渐式微。诞生于计算机单机时代的抽象化个人信息控制权，既无法涵摄数字时代涌现的新型个人信息形态，亦构成个人信息交易流通的制度性障碍。随着个人信息类型持续扩张、处理方式迭代升级与安全风险动态演化，个人信息权利体系更迭成为必然。遵循个人信息分类确权的核心原则，在唯一标识类个人信息之上，配置具有支配权属性的控制权；在非唯一标识类个人信息之上，配置具有请求权属性的防御权。个人信息收集与交易规范构造如下：个人信息处理者处理唯一标识类信息，必须事前取得个人明确同意；处理非唯一标识类信息，可采用事后退出的授权模式。立法应当明确禁止唯一标识类信息的交易与流通，而非唯一标识类信息可依法自由流通。个人信息分类确权理论的规范价值体现在双重维度：其一，有效降低个人信息处理者合规和运营成本；其二，为个人信息合法交易与有序流通奠定坚实的理论支撑，构筑严密的制度框架。

公共数据开放研究

1、公共数据共享法律责任的理论逻辑和实践价值（郭文涛）

来源：《东方法学》2026年第2期

针对公共数据共享存在的“共享不充分、权责不清、安全无保障”等典型问题以及执法司法“认定难、适用乱”等实践困境，厘清其法律责任的理论逻辑，并探索其实践价值、执法司法指引及精准运用路径，对助力数字法治政府建设具有重要意义。我国数据地方立法实践经历了“政府数据共享—政务数据共享—公共数据共享”的演变过程，范围逐渐扩大，但央地两种法规均鲜明地将数据共享的法律责任归纳为责令改正、处分、刑事责任三种。在公共数据共享中，责令改正反映了同级或层级监督关系，处分反映了公共管理和服务机构与其工作人员之间的惩戒关系，刑事责任反映了犯罪后的刑事法律关系。这三对法律关系背后都遵循着坚实的理论基础，分别是内部行政法理论、特别权力关系理论、罪刑法定原则。公共数据共享法律责任以责令改正、处分、刑事责任为核心，以民事责任与行政处罚款为补充，核心责任是因公共数据公权属性而聚焦公法主体权力的规制。与上位法同类责任相比，其在规制对象、价值目标与适用场景上具有特殊性，可以为同类责任形态的区分与适用提供理论参照。三类核心责任层级递进、功能互补，形成“预防—惩戒—制裁”闭环，其逻辑对于数字法治政府制度设计、实施与拓展，具有重要借鉴意义，并有助于构建统一、协调的数字法治责任框架。

2、公共数据资源化的理论构造与制度展开（冯洋）

来源：《东方法学》2026年第2期

公共数据资源化是指公共数据集合经加工处理并通过法定机制确定其权属，最终转化为可供市场流通利用的数据要素资源的过程。关于资源开发利用和发展社会生产力的宪法条款为公共数据资源化奠定了根本法依据。价值创造是公共数据资源化的核心功能，其制度空间在与公共数据的安全保障、公共服务和社会治理功能的协调中得以厘清。公共数据资源化应当明确为国家义务。同时，应当确立公共数据资源国家所有权，并构建所有权与用益权分离的二元权利结构，突出用益权的价值增益功能。通过法律的调整以及《公共数据条例》的制定，

围绕数智基础设施、授权与确权、功能协调与互补、组织与激励等关键事项开展制度建设，为公共数据资源化及其价值创造提供全面的制度保障。

3、论公共数据授权运营收益分配的政府参与（褚睿刚）

来源：《东方法学》2026年第2期

收益分配是公共数据授权运营的基础制度，政府在其中扮演三重重要角色。政府参与公共数据授权运营收益分配遵循“初次分配+再分配+第三次分配”三阶构造方策，并在各分配阶段通过不同的制度与方式参与其中。政府通过授权运营机构收取运营服务费的方式间接参与收益初次分配，再将所分享的收费收益与课税收入汇入数据财政，并通过财政支出直接实施再分配，最后通过助推慈善捐赠与分享经济等方式对第三次分配进行嵌入式治理。实践中，上述各分配制度均存在不同程度的设计不足乃至规则空白问题，须进行体系性优化。具体而言，公共数据运营服务费定价应突破成本补偿的桎梏，采取更加市场化的方式，明确政府与运营主体“成本补偿收益+其余收益”的对应分配结构，并合理确定政府内部各数据提供部门间的收益再分配比例。公共数据课税宜采用传统课税模式，可在授权运营的多个环节课征多类税种，并以税收优惠方式促进、引导公共数据要素市场发展。此外，政府可通过主导公共数据产品共享平台建设，加强对慈善捐赠与分享经济的法律激励，并辅之以柔性的合作监管方式，助推收益第三次分配。

4、“公共数据归国家所有”的宪法检视（胡萧力）

来源：《法学研究》2026年第3期

围绕公共数据的权属界定，当前地方立法实践中存在数据汇集型、数据维护型、数据经营型三种确权逻辑。相关立法多笼统借用经济学产权概念与政策性表述，欠缺严格的法学维度的正当性证成与上位法依据支撑。公共数据深度嵌入行政过程、数字基础设施运转架构和社会协同创新网络，其运行遵循信息媒介逻辑，与自然资源属性迥异。公共数

据利用秩序的建构,不宜直接套用宪法上关于自然资源国家所有的法理,而应在宪法第12条所确立的社会主义公共财产规范框架下展开。公共数据归国家所有,应理解为由国家承担防止侵占、保障合理利用、促进公平利用之义务的规范要求。未来应通过中央立法明确公共数据治理的基本原则、数据汇

集的合法性要件、授权运营的目的与基本程序、私法规则的适用边界等,并配套财政监督、程序保障与数据流通的统筹机制,推动形成体系完整、逻辑融贯、普惠包容的公共数据利用生态。

(技术编辑:毕坤阳)

教研活动

“个人信息保护法治研讨会”在京举行

2026年5月13日,作为纪念习近平总书记“5·17”重要讲话发表十周年系列学术活动之一,由中国社会科学院法学研究所主办,中国社会科学院文化法制研究中心、中国社会科学院法学研究所宪法与行政法研究室、中国社会科学院法学研究所网络与信息法研究室承办的“个人信息保护法治研讨会”在北京成功召开。来自中国社会科学院、中央财经大学、北京理工大学、对外经济贸易大学、中国信息通信研究院、同济大学等高校与科研机构的专家学者齐聚一堂。本次会议聚焦我国个人信息保护法治建设的最新成就与理论发展,围绕我国个人信息保护法治经验展开研讨,为构建更加公平、公正的全球数字治理体系提供坚实的学理支撑与智力支持。

在致辞环节,中国法学会网络与信息法学研究会会长**姜伟**指出,随着我国《网络安全法》《数据安全法》《个人信息保护法》等法律的深入实施,个人信息保护的法治屏障作用愈发凸显。他强调,我国政府获取数据有着明确的法定权限、严格的程序约束和有效的监督救济机制;应当系统性地讲好中国法治故事,以高质量的合规实践展现我国负责任的大国法治形象。中国社会科学院学部委员、法学研究所所长**莫纪宏**表示,我国在个人信息保护领域所形成的丰富实践与显著成效,已成为提升国家数字竞争力与国际规则话语权的战略基石。他强调,面对我国个人信息法治建设的高水平成果,法学研究应当坚持问题导向,融贯理论与实践,系统梳理并提炼我国个人信息保护的成功经验,为数字经济规范健康有序发展提供更稳定的理论支撑和制度预期。

会上,中国社会科学院法学研究所网络与信息法室副主任(主持工作)**周辉**代表研究团队,围绕关于比较法视野下中国高水平个人信息保护的研究报告作了专题介绍。研究报告系统论证了我国个

人信息保护是体系完备、标准严格、执行有力、救济有效的高水平法治实践。报告指出,从《宪法》《民法典》到《个人信息保护法》等专门立法,我国构建了极其严格的保护规则,在赋予个人信息相关权利、确立合法性基础逻辑等方面的标准严于部分发达国家和地区。特别是在国家机关调取数据维度,我国法律设置了非常严格的权限限制和程序约束。此外,我国还建立了涵盖行政监督、检察监督、审计监督在内的立体多层次全面监督机制,以及完善的行政和司法救济途径,向国际社会清晰展现了全球数字治理的中国主张与实践定力。

在专家研讨环节,七位法学学者结合各自研究领域,从多维度深入论证并高度评价了我国个人信息保护的法治水准。中国社会科学院法学研究所研究员**董坤**指出,我国司法机关在办案中对公民个人信息的保护已形成“事前审批、事中监督、事后救济”的全链条体系化高标准保障机制,彰显了法治温度。中央财经大学法学院教授**刘权**强调,我国法律在政府获取数据时严格贯彻“最小必要原则”与比例原则,受正当程序与安全保障机制的强力双重约束,在公共利益与个人信息权益之间实现了精细的法治均衡。北京理工大学法学院教授**洪延青**提出,我国法院处理了大量涉及个人信息的案件,有力的司法实践切实印证了我国公民行权的便利度;我国庞大且专业的数据保护官(DPO)群体的合规实践,更是中国高水平数据保护的鲜活写照。对外经济贸易大学法学院教授**许可**表示,中国已经通过一系列立法和全球倡议构建了坚实的数据安全法治基础。他结合具体涉外法律实务指出,中国企业在个人数据跨境传输方面展现出了极高的合规水准,充分彰显了中国个人信息保护法治的体系完备。中国信息通信研究院互联网法律中心主任**何波**总结道,我国确立了以人为本的保护理念,构建了严密的规范体系,在规则制定与法律实施层面取得了举世瞩目的历史性成就,为企业营造了良好的法治环境。同济大学法学院助理教授**朱悦**指出,我国个人信息保护具备丰富的制度语境和恰到好处的法治颗粒度,整体治理体系已充分彰显出扎实的实质效能。同济大

法学院副教授**王伟**通过分析巴西与欧盟实现数据保护充分性互认的经验，系统比较了中国、欧盟和巴西在数据保护法方面的制度等效性与差异，进而提出我国应借助国际和区域间数据执法机构（如 Privacy Commissioner/Authority）及数据保护执业（如 DPO）等专业社群的对话与互动，稳步推进跨国数字合作与区域共识的达成。

本次研讨会的成功举行，不仅全方位、多层次地展示了我国个人信息保护法治建设的丰硕成果，也为深化涉外法治研究凝聚了重要共识。与会学者一致表示，将持续深耕中国个人信息保护法治经验的理论阐释，积极推动扎实的国内法治成果转化为清晰的国际叙事，为助力我国企业高质量发展、推动构建更加公平公正的全球数字治理体系贡献中国智慧。

珞珈网络治理大讲堂第十八期——数据三权分置的反思和重构

武汉大学网络治理研究院珞珈网络治理大讲堂第十八期于法学院 217B 会议室顺利举行。来自武汉大学网络治理研究院、武汉大学法学院、中南财经政法大学法学院等单位的专家学者及师生代表齐聚一堂，围绕“数据三权分置”的制度逻辑、现实困境与未来重构展开深入研讨，现场学术氛围浓厚，讨论热烈。

本次会议由武汉大学网络治理研究院副院长、武汉大学法学院教授**袁康**主持，中南财经政法大学法学院教授、数字法学学科负责人**李昊**作题为“数据三权分置的反思与重构”的学术报告，武汉大学法学院**王德夫**副教授、**宁园**副教授担任与谈人。



会议伊始，**袁康**教授结合我国数据基础制度建设背景，指出“数据三权分置”作为中国数据治理

领域的重要制度创新，在理论界引发了广泛讨论。在《数据二十条》及近期国家数据局相关文件出台的背景下，重新思考其制度逻辑具有重要意义。随后，他对主讲人及到场师生表示欢迎。

中南财经政法大学法学院教授、数字法学学科负责人**李昊**教授首先围绕“数据三权分置”的理论脉络、制度争议与未来重构路径展开讨论。他指出，在数字经济背景下，数据已逐渐成为新的经济增长点。近年来，无论是数据资产入表，还是数据要素市场建设，都在推动数据资源加速进入经济运行体系。然而，国家数据局成立之后，数据要素市场化配置仍面临诸多现实障碍，其中一个核心问题便是“数据究竟如何被使用、如何真正转化为经济增长动能”，其背后涉及数据权利结构与制度设计的根本逻辑。

李昊教授进一步指出，“数据三权分置”试图通过对数据资源持有权、数据加工使用权和数据产品经营权进行区分，回应数据流通、利用与收益分配中的复杂关系，但在制度实践中面临诸多争议。例如，数据权利边界是否清晰、不同主体之间的权利配置是否协调、数据权利能否真正实现市场化流通等，均有待进一步厘清。

在此基础上，李昊教授结合我国数据要素市场建设的现实情况，对“数据三权分置”的制度逻辑进行了反思。他指出，数据不同于传统物权客体，具有可复制性、非排他性与高度流动性，不能简单套用传统财产权结构。当前制度讨论中，部分观点过于强调“确权”逻辑，却忽视了数据利用过程中公共利益、个人信息权益与产业发展的动态平衡。未来的数据制度建设，应更加注重数据流通机制、收益分配机制与治理机制之间的协同，而非仅停留于静态的权利划分。



在主讲人发言结束后的与谈环节，武汉大学法学院**王德夫**副教授围绕数据法律治理相关话题展开交流分享。他指出，当前我国数据领域政策文件密集出台、法条修订频繁，《反不正当竞争法》等法律相继增补数据相关条款，政策体系繁杂且处于动态调整阶段。结合行业实践，他强调数据复制存储的边际成本极低，纠正了大众对于数据复制必然高成本的认知误区；企业依托防火墙、合规合同、刑法规制等方式已构建完备的数据保护体系，并不欠缺数据安全防护手段，核心痛点是缺少合法的数据权属名分。王德夫老师重点剖析了数据确权落地难的核心症结，认为最大阻碍是社会伦理矛盾是公众难以接受个人数据归平台所有。《数据二十条》则是确权暂缓落地下的过渡方案，为中大型数据平台提供合规依据。他预判，未来数据资源将高度集中于国家算力基础设施与头部平台，依托API接口实现时效化利用，当前优化现有合规机制是务实选择。



武汉大学法学院**宁园**副教授结合自身研究体会，分享了对数据产权制度的反思与见解。她认同主讲人对数据领域相关问题的反思，重点指出数据产权制度构建中存在的核心问题：一是概念选取不合理，将经济学、日常生活中的“数据资源”“数据产品”直接作为法律概念并赋予新定义，增加了论证与实践难度；二是数据担保缺乏可操作性，与传统动产抵押担保差异显著，目前无实际实现案例，其市场可行性存疑；三是相关政策推进较为谨慎，且规则本身存在未充分讨论、与既有制度沿革不符等问题。

宁园老师进一步提出，当前数据产权研究陷入困境，核心原因是市场实践不够活跃，缺乏足够参考样本，难以发现需制度解决的具体问题。同时，研究存在客体思维依赖和权利本位导向的误区，忽视了数据的动态性和共享性，且个人信息保护与数据产权规则存在衔接难题，易因个人信息违法导致数据产权无效。她也肯定了数据产权工作的积极进展，认为未来产权制度应朝着开源化、兼顾多方权益的方向发展，既要扩大产权享有范围，也要探索数据合理使用，才能充分发挥激励作用，推动数据要素市场化运用。



Unfair Competition Law in the Digital World 讲座圆满举办

2026年4月17日上午，由中国人民大学法学院主办，中国人民大学知识产权学院、未来法治研究院、人工智能治理研究院联合协办的“Unfair Competition Law in the Digital World: AI Challenges and Perspectives”讲座在中国人民大学举办。本次讲座聚焦人工智能背景下反不正当竞争法面临的新问题，由德国慕尼黑大学法学院**Ansgar Ohly**教授主讲。中国人民大学法学院院长**杨东**教授、知识产权学院副院长**郭禾**教授出席活动并致辞。中国人民大学法学院**孙靖洲**助理教授主持讲座。

杨东院长代表中国人民大学法学院对**Ohly**教授的到访表示热烈欢迎。他指出，人工智能技术的发展正在深刻影响数字市场竞争秩序，AI生成内容、模型模仿、数据抓取、算法操纵等问题，对传统反不正当竞争法规则提出了新的挑战。本次讲座围绕数字环境中的反不正当竞争法展开讨论，对于

推动中德学术交流、深化人工智能法治研究具有积极意义。杨东院长预祝讲座圆满成功，并期待未来中国人民大学法学院与慕尼黑大学、德国马克斯·普朗克研究所继续深化交流合作。



中国人民大学法学院院长、教授 杨东

郭禾教授介绍了 Ohly 教授的学术背景，并回顾了中国人民大学法学院与慕尼黑大学法学院长期以来的学术交流。郭禾教授指出，Ohly 教授长期从事知识产权法、反不正当竞争法、商业秘密法等领域研究，此次到访并开展专题讲座，有助于促进中德知识产权领域的学术对话，也为人工智能背景下相关法律问题的研究提供了有益参考。



中国人民大学法学院教授、知识产权学院副院长 郭禾

Ohly 教授以“登山观雾”为生动比喻展开演讲：将“知识产权法”比作视野清晰、边界明确的山峰，规则体系成熟、保护范围稳定；而“反不正当竞争法”则如同山谷中的浓雾，概念弹性强、判断标准模糊，在人工智能技术冲击下面临巨大挑战。他指出，当前全球学界对 AI 与版权法的讨论已较为充分，但针对“AI 与反不正当竞争法”的交叉研究仍明显不足，这正是本次讲座的核心关切。



慕尼黑大学法学院教授 Ansgar Ohly

Ohly 教授结合《巴黎公约》第 10 条之二、TRIPS 协定等国际规则，围绕人工智能场景下的不正当竞争风险，重点讨论了三个方面的问题。

第一，AI 生成内容可能引发混淆行为及商标类不正当竞争问题。Ohly 教授以 Getty Images 诉 Stability AI 案、AI 生成商标图像、AI 广告中使用奢侈品品牌形象等案例为例指出，AI 生成内容可能再现他人商标、商业标识或水印，但是否构成侵权，关键取决于其是否被投入商业使用。单纯生成图片通常不构成商标意义上的使用；但如果将生成内容用于商品、服务或商业宣传，则可能构成商标侵权或假冒行为。至于 AI 平台或模型开发者是否承担间接责任，目前仍有待进一步讨论。

第二，AI 训练数据、模型权重和模型蒸馏等问题可能涉及模仿行为的法律规制。Ohly 教授提出，应当首先考察著作权、专利、外观设计、商业秘密等专门规则是否能够适用；只有在专门规则无法解决，且存在明显市场失灵的情况下，才应谨慎适用反不正当竞争法一般条款。他认为，AI 训练数据和模型权重在符合秘密性、价值性和保密措施要求时，可以纳入商业秘密保护范围。但对于 AI 生成内容，如果不能获得著作权保护，原则上不宜通过反不正当竞争法一般条款扩大保护，以免压缩公共领域和竞争自由。

第三，数字经济中的“黑暗模式”也可能构成不正当竞争。Ohly 教授将“黑暗模式”概括为通过界面设计、默认选项、弹窗提示、时间压力等方式影响消费者选择的行为。典型情形包括虚假“最后席位”提示、Cookie 授权诱导、隐藏拒绝按钮、反复弹窗要求同意等。他强调，并非所有营销引导都

违法，关键在于是否存在虚假、误导或明显违背诚实信用原则的操纵行为。相关行为可能落入虚假宣传、误导性遗漏或一般条款规制范围，但法律适用仍应保持审慎。

在与谈环节，北京大学法学院**刘银良**教授、清华大学法学院**蒋舸**副教授、中央财经大学法学院**李陶**副教授、中央民族大学**熊文聪**教授、中国人民大学法学院**金海军**教授、中国人民大学法学院**张吉豫**教授分别围绕讲座内容进行评论。

刘银良教授指出，数字世界与现实世界相互交织，使不正当竞争行为的认定更加复杂。反不正当竞争法应维护公平竞争秩序，但在AI时代仍需保持谦抑，不能仅因某种利益具有商业价值就直接赋予法律保护。



北京大学法学院教授 刘银良

蒋舸副教授围绕反不正当竞争法一般条款与知识产权法的关系展开评论。她指出，德国和美国司法实践均较为重视知识产权法的优先适用，而中国部分案件中存在直接诉诸一般条款的倾向。她认为，应进一步明确一般条款的适用边界，避免反不正当竞争法过度扩张。



清华大学法学院副教授 蒋舸

张吉豫教授围绕AI模型的可版权性、开源模型许可和AI代理等问题进行讨论。她指出，AI模型结构、权重组合与软件之间存在一定的关联，但仍需区分功能性要素与表达性要素。面对AI代理、自动交易、跨平台交互等新型行为，法律应在保护合法权益与保留创新空间之间取得平衡。



中国人民大学法学院教授 张吉豫

李陶副教授从技术与法律衔接的角度，对模型蒸馏、训练数据保护、商业秘密认定和反向工程规则进行了分析。他指出，AI相关纠纷的处理应当建立在准确理解技术事实的基础之上，特别是要区分模型复制、模型模仿、API访问和反向工程等不同技术行为。



中央财经大学法学院副教授 李陶

熊文聪教授围绕中德反不正当竞争法一般条款的适用差异展开评论。熊文聪教授认为，在人工智能与数字经济场景下，应进一步明确一般条款的适用条件，避免仅以“投入较大”或“具有商业价值”为由扩大保护范围，从而在维护竞争秩序、保护创新投入与保障竞争自由之间取得平衡。



中央民族大学法学院教授 熊文聪

金海军教授结合中国司法实践指出，近年来数据、平台和AI相关不正当竞争案件持续增加，一般条款的适用边界值得关注。他强调，反不正当竞争法的功能在于制止不公平竞争，而不是一般性地禁止模仿。司法裁判应尊重知识产权法的制度边界和市场竞争规律。



中国人民大学法学院教授 金海军

在提问交流环节，现场师生围绕中德一般条款适用差异、AIGC专门权利、AI训练中的利益分配、屏幕爬取、AI代理、商业秘密与反向工程等问题与Ohly教授进行了交流。

针对是否应为AIGC设立专门权利的问题，Ohly教授表示应保持谨慎。他认为，欧盟数据库特殊权利的经验表明，创设新的专有权可能带来保护过强、边界不清和退出困难等问题。对于包含人类创造性投入的AI生成内容，现有著作权法已有一定回应空间；对于缺乏人类创造性投入的内容，则不宜轻易赋予专有权。

针对AI代理和屏幕爬取问题，Ohly教授指出，反不正当竞争法不应被用来保护既有商业模式本身。只要相关行为未侵犯商业秘密、未违反有效合同约定，也未实施混淆或误导，就应为技术进步和商业模式创新保留空间。

针对API开放与商业秘密保护的关系，Ohly教授认为，开放API并不当然意味着模型权重、训练方法或整体结构已经公开。只要相关信息仍处于秘密状态，并采取了合理保密措施，仍可能构成商业秘密。但通过侵入系统、破坏技术保护措施等方式获取信息，则可能超出合法反向工程的范围。

本次讲座围绕人工智能背景下反不正当竞争法的适用边界展开深入讨论，系统回应了AI生成

内容、模型模仿、训练数据保护、商业秘密、黑暗模式等数字经济中的前沿问题。



“数字转型：互联世界中的全球价值链重构”国际学术研讨会在香港成功举办

2026年4月20日至21日，由中国人民大学法学院、香港城市大学法律学院、荷兰鹿特丹伊拉斯姆斯大学法学院主办的“数字转型：互联世界中的全球价值链重构”(The Digital Shift: Redefining Global Value Chains in a Connected World)国际学术研讨会在香港城市大学顺利举行。来自多所知名高校及研究机构的学者、专家齐聚一堂，围绕人工智能、平台治理、全球价值链重构及数字法治等前沿议题展开了为期两天的深入研讨。

4月20日上午，会议正式开幕。香港城市大学法律学院院长、基本法讲座教授林峰教授，香港城市大学香港商务及海事法研究中心主任陆飞鸿教授，鹿特丹伊拉斯姆斯大学Koen Swinnen教授，以及中国人民大学法学院副院长丁晓东教授先后致欢迎辞。



香港城市大学法律学院院长、基本法讲座教授
林峰

林峰教授在致辞中对与会嘉宾表示诚挚欢迎。他指出，人工智能与大数据已成为影响社会运行的关键力量，也给法学教育带来了如学术诚信管理、考核机制调整等现实挑战。林峰院长重点介绍了城大法律学院的发展远景，强调学院正致力于构建比较研究高地，通过丰富的学术资源支持，推动国际法律体系的对话。他鼓励国内外学术机构依托城大四大研究中心开展深度合作，共同探索新兴技术下的法律规制。



香港城市大学香港商务及海事法研究中心主任
陆飞鸿教授

陆飞鸿教授从私法与实务视角出发，分享了数字化对公共空间治理的启示。他通过香港城大行人隧道数字化改造以及电梯图像数据处理等案例，论述了公私合作、数据保护与公共安全之间的平衡，并强调了数字化时代的法律创新对人才培养的重要性。他诚邀在座专家通过开设短期强化课程等形式，为城大弟子带来前沿的国际法学视野。



鹿特丹伊拉斯姆斯大学 Koen Swinnen 教授

Koen Swinnen 教授介绍了由荷兰政府资助成立的伊拉斯姆斯法律与数字化中心（ECLD）的运行情况。该中心旨在整合法学研究力量，聚焦数字化带来的法律议题。Swinnen 教授表达了将此美国

际研讨会常态化、机制化的愿景，期待通过建立长期的国际伙伴关系，共同攻克数字化转型中的全球性法律难题。



中国人民大学法学院副院长 丁晓东教授

丁晓东教授在致辞中指出，当前全球价值链的塑造力正从资本、劳动力向数据、平台算法转移，数字化转型不仅是经济和技术转型，更是一场深刻的法律转型。丁教授表示，本次会议聚焦平台责任、人工智能主体地位及跨境数据流动等前沿议题，具有鲜明的时代意义。在全球法律治理体系中，构建“共同语境”与“共享词汇”是各方协作的基础，人大法学院将始终致力于推动中、欧、美法学界的深度对话，为数字化法治贡献力量。

会议第一专题“AI, Platform and GVCs”聚焦人工智能与平台对全球价值链的结构性重塑，由鹿特丹伊拉斯姆斯大学的 Alberto Quintavalla 教授主持。



莱顿大学 Gianclaudio Malgieri 教授

莱顿大学 Gianclaudio Malgieri 教授发表了题为《数字化依赖、“不健康状态”与AI驱动平台中的脆弱性》的主旨演讲。Malgieri 教授从法律视

角对“脆弱性”进行了重构，指出脆弱性并非个体的静态属性，而是一种由权力失衡、非对称性依赖以及社会政治环境共同塑造的动态关系。他认为欧盟《人工智能法案》中对脆弱性定义过度倾向于个人主义和消费者保护视角，而忽视了权力关系、行政控制及系统韧性的影响。他认为应建立一种跨学科的综合治理框架，将脆弱性界定为权力失衡导致的基本权利高风险状态，以此作为AI时代维护个人尊严的法治基石。



清华大学法学院院长 崔国斌教授

清华大学法学院院长**崔国斌**教授围绕《限制机器人访问的用户协议条款之合法性》进行发言。崔教授结合国内外AI智能体引发的最新司法争议，深入探讨了平台通过用户协议禁止自动化工具访问的法律效力。他在承认平台在数据控制、资源保护及安全维护方面的合法权益的基础上，引入著作权法中的“合理使用”规则，审视机器访问是否应享有法定豁免。崔教授指出，治理核心在于界定平台能否通过格式条款“合同化地剥夺”技术创新所必需的合理访问权。他认为，法律应在私法自治与数字公共利益之间寻求平衡，防止平台利用生态优势通过不合理的协议条款阻碍技术的进步与公平竞争。



香港城市大学法律学院 何天翔副教授

香港城市大学法律学院**何天翔**副教授以《受版权保护的“产品”、数字转型与全球价值链》为主

题进行分享，深入剖析了数字化转型对版权产业价值链的重构作用。他指出，生成式AI的崛起不仅带来了替代人类原创作品的风险，更可能通过黑箱化训练造成对“全球南方”国家的文化剥削，使价值链陷入缺乏创新的停滞循环。为此，他呼吁建立以人本主义为核心的法治框架，通过落实AI训练数据强制透明化、引入技术检测与收益共享机制，以及完善国际的强制许可制度，确保版权收益能够公正回流至人类创作者，维护全球价值链的公平与活力。

会议第二专题“Emerging Challenges in GVCs”聚焦全球价值链面临的前沿挑战，由机器人与人工智能法律协会**Martin Ebers**教授主持。



北京大学法学院副院长 戴昕副教授

北京大学法学院副院长**戴昕**副教授围绕AI能动体行为的规范预期，深入探讨了AI能动体行为准则的构建范式。他提出，社会对AI的预期存在“模范”（要求其超越人类，更公平、更安全）与“镜像”（仅要求其反映人类水平以实现服务普惠）两种核心逻辑，并指出不同的范式选择将直接影响法律对准入门槛与注意义务的设置。结合中国监管实践，戴教授强调，AI治理的本质是回归人类中心主义，在技术介入社会生活前，应首先厘清人类对不同功能定位下AI系统的真实规范诉求。



乌特勒支大学 Cristiana Santos 助理教授

乌特勒支大学 **Cristiana Santos** 助理教授聚焦欧盟“暗黑模式”的规制与执法，梳理了“暗黑模式”如何通过操纵性界面设计扭曲数字化价值链。她指出，此类行为利用人类直觉偏见，在隐私保护、账号管理及注意力分配等领域对用户造成广泛的心理与经济损害。欧盟《数字服务法》（DSA）已针对暗黑模式确立了欺骗、操纵与损害三大规范维度，但目前仍面临非物质损害认定难、执法碎片化等挑战。为此，她主张构建全球化、结构化的暗黑模式执法数据库，通过统一的法律语言与技术检测手段，强化司法威慑力，从底层机制上遏制数字经济中的“操纵型价值链”。



阿姆斯特丹大学 Marija Bartl 教授

阿姆斯特丹大学 **Marija Bartl** 教授发表了题为《算法之前：全球价值链中的私法、预分配与人工智能》的演讲。她通过引入“预分配”概念，指出私法制度决定了主体在进入市场前的法律禀赋，而当前私法正从福利国家的权利保护转向新自由主义的“信息披露”范式，导致权力失衡加剧。Bartl 教授强调，AI 算法与全球价值链的结合制造了海量的信息不对称与价值提取机制，使领先企业得以逃避社会责任。她主张通过重构公平的预分配机制来保障所有参与者的劳动尊严与机构代理能力，防止人类在 AI 驱动的经济体系中沦为无用之人。

第三专题题目为 **Data Protection, Data Governance and GVCs**，聚焦数据保护、数据治理与全球价值链的交叉领域，探讨了数字经济核心生产要素的制度构建与价值分配问题。



纽卡斯尔大学 Darren McCauley 教授

纽卡斯尔大学 **Darren McCauley** 教授以《欧盟数据中心的公正数字循环》为题展开分享。他指出，AI 产业爆发式增长推动数据中心大规模扩张，已引发欧盟层面的重点监管关注。其研究结合政策分析与实地访谈，从修订后的能效指令、统一数据中心评级体系、跨境资源流动三个维度切入，揭示了当前数字循环模式中存在的领土正义缺失与发展负担分配不均问题，呼吁构建兼顾生计、技术与社会正义的可持续数据中心发展体系。



鹿特丹伊拉斯姆斯大学 Koen Swinnen 教授

鹿特丹伊拉斯姆斯大学 **Koen Swinnen** 教授围绕数据商品化核心议题，就数据作为抵押品的可行性问题进行发言。他从荷兰财产法基础理论出发，论证了当前法律框架下数据无法直接作为抵押品的制度根源，并介绍了英国、苏格兰在数字资产财产权立法上的最新突破。他指出，数据抵押落地的核心障碍在于银行对资产估值、价值波动、处置渠道及数据安全的顾虑，其团队正通过对荷兰主流金融机构的实证调研，为后续立法与行业实践提供依据。



浙江大学 李汶龙研究员



香港大学 罗佳坤博士

浙江大学李汶龙研究员与香港大学罗佳坤博士以《谁分得收益？AI 驱动游戏开发中的价值创造、攫取与分配》为题发表联合分享。他们基于 2025 年全球游戏开发者调研数据与 Roblox 平台实证案例，揭示了一个核心悖论：AI 极大提升了游戏开发效率、拓展了价值创造的边界，却并未改变价值向平台与头部 IP 所有者高度集中的格局，普通创作者与生产环节从业者的收益份额反而被进一步压缩。他们呼吁建立与技术变革相适应的全链条治理体系，推动价值分配与实际贡献相匹配。



中国人民大学 孙靖洲助理教授

中国人民大学孙靖洲助理教授聚焦具身智能时代跨境数据流动的全新治理命题展开分享。她指出，具身 AI 将全球竞争焦点从算力转向高保真物理行为数据，中国在此领域拥有独特的资源优势，

但同时面临价值分配失衡、产业安全风险、监管法律空白与政企激励错配四大核心挑战。她提出，传统以防范即时危害为核心的治理模式已不再适用，应转向以技术主权与产业竞争力为核心的主动治理范式，建立多维度的战略数据跨境流动评估体系。

第四专题题目为 Accountability and GVCs，围绕全球价值链中的问责机制展开，探讨数字时代企业责任、平台治理与法律责任体系的重构路径。



布里斯托尔大学 Paula Giliker 教授

布里斯托尔大学 Paula Giliker 教授以《母公司是否应为其海外子公司的侵权行为承担替代责任？》为题展开分享。她分析了受害者起诉母公司的现实考量，系统梳理了英国普通法下替代责任与直接责任两种救济路径的司法实践与核心判例，深入解读了欧盟最新修订的企业可持续发展尽职调查指令的重大调整与制度局限。她指出，当前跨国企业责任治理仍面临侵权法与公司法价值冲突等核心难题，为构建更公平的全球价值链责任体系提供了重要的比较法视角。



中国人民大学法学院副院长 丁晓东教授

中国人民大学法学院丁晓东教授围绕平台责任核心制度，就“通知-删除”机制的理论重构问

题进行深入讨论。他对比分析了美欧中三大法域通知-删除机制的演变路径与制度差异，指出传统“归因”与“豁免”二元理解的局限性。他创新性地提出将通知-删除定位为“报告-治理机制”，主张区分个案侵权与大规模系统性侵权进行差异化规制。他进一步指出，数字经济的发展正推动侵权法从传统的个案责任追究向系统性治理转型，为全球平台责任制度改革提供了全新的理论视角。



鹿特丹伊拉斯姆斯大学 李舒助理教授

鹿特丹伊拉斯姆斯大学李舒助理教授以《产品安全、产品责任与全球价值链》为题发表主题发言。他指出，传统以物理产品为核心的监管与责任制度，正面临跨境电商碎片化交易与AI产品持续迭代的双重冲击。他对比了欧盟与中国应对挑战的差异化改革路径，揭示了数字时代产品治理从“实体监管”向“信息治理”转型的深层趋势，呼吁建立全球性合作框架，共同应对跨境产品安全外部性问题。



香港城市大学 陈扬助理教授

香港城市大学陈扬助理教授以《人工智能秘密的登记制度》为题展开分享。他指出，企业对AI算法、训练数据等信息的过度商业秘密主张，已形成“秘密治理”困境，严重阻碍了对AI决策的有效监督。他深入剖析了商业秘密法的结构性缺陷，提出建立带可信第三方审查机制的AI秘密登记制

度，通过事前范围界定与事后公共监督相结合的方式，在保护企业合法权益的同时，破解AI透明度难题。

2026年4月21日上午，“数字转型：互联世界中的全球价值链重构”国际学术会议在香港城市大学继续开展第五专题和第六专题的研讨。机器人与人工智能法学会Martin Ebers教授、香港中文大学李治安教授、香港城市大学洪凯晋助理教授、德国杜塞尔多夫大学郑柯娜博士后、香港大学孙皓琛教授、香港中文大学柯丽娜副教授、鹿特丹伊拉斯姆斯大学Alberto Quintavalla副教授、中国人民大学李子硕讲师等海内外多位权威学者参会并发表主题演讲，共同探讨数字时代全球价值链的法律与治理前沿问题。

第五专题题目为Global Perspectives and GVCs，聚焦人工智能与全球价值链的治理规则、权力结构与竞争秩序，四位学者从不同维度展开深度剖析。



机器人与人工智能法学会 Martin Ebers 教授

Martin Ebers教授以《如何治理AI价值链：来自欧盟与美国的洞察》为题进行了发言。他深入分析了通用人工智能(GPAI)价值链的流动性特征，指出上下游角色重叠使安全保障并非单一主体的责任，下游参与者可以增强或削弱上游设定的安全措施。他对比了欧盟《人工智能法案》与美国加州、纽约州的最新立法尝试，强调在复杂的技术生态中，加强跨链条的协同监管与责任分配是确保技术合规与社会安全的关键，并呼吁构建更加平衡的责任分担框架，以应对公共采购中存在严重的信息与技术不对称。



香港中文大学法律学院 李治安教授

香港中文大学法律学院教授**李治安**围绕人工智能治理的“北京共识”，系统解析了中国在开源大模型和芯片自主研发等领域的强劲势头，在产业落地方面，中国在工业机器人领域占据绝对优势，并在自动驾驶、大型语言模型（LLMs）和无人机等关键赛道表现突出。中国构建了包含政策指导、行政法规和基础立法的监管体系，立法和监管呈现出“小快灵”和“小切口”的特征。李教授借用“北京共识”这一概念，总结了中国在 AI 治理中展现出的三大核心逻辑：强调创新、实验与容错；使用新工具控制混乱；追求技术自决。他认为这一“北京共识”正为全球数字主权提供新范式，同时也为重塑全球价值链中的其他发展中国家提供了宝贵经验。



香港城市大学法律学院 洪凯晋助理教授

香港城市大学法律学院助理教授**洪凯晋**围绕全球价值链中的算法权威指出，人工智能技术的广泛应用正在深刻改变商业合同中自由裁量权的行使方式，合同中原本由人行使的裁量权，逐渐转化为由算法驱动的自动化决策，导致传统的法律控制机制在算法决策面前逐渐失灵。他强调，这种转变正在全球价值链内部催生一种由算法主导的新型私人治理模式：商业合同是全球价值链的制度基

础设施，其跨国触达力往往比国家层面的法律规制更为高效，跨国主体通过合同构建起一个去中心化的网络化法律秩序，将生产协调、风险分配以及合规标准内化其中。在此趋势下，算法权威具有官僚化冷漠特征，往往边缘化了非经济维度的社会文化考量，导致主导企业的权力进一步硬化和去人格化，尽管这种系统设计对拥有技术主导权的企业而言高度可预测，但对于价值链中的其他弱势方而言，则显著增加了决策的不可预测性。



德国杜塞尔多夫大学 郑柯娜博士后研究员

德国杜塞尔多夫大学博士后**郑柯娜**认为科技巨头的商业模式正在从平台锚点向扩张性生态系统，再向人工智能驱动的生态系统转型。核心资源如算力、数据、人才正日益集中在少数几家大企业手中，导致了对赢者通吃的市场担忧。通过对比欧盟《数字市场法案》与英国《数字市场、竞争与消费者法案》在监管灵活性上的差异，指出当前竞争法规在应对 AI 供应链各层级的复杂挑战时仍显不足，并进一步建议监管机构必须在“加强竞争干预”与保护 AI 创新之间寻找微妙平衡，并寻找更具实验性和演化性的方案。

第六专题题目为 Innovation, Competition and GVCs，主要聚焦数据治理、跨境监管、算力环保、平台内卷等创新与竞争核心问题展开研讨，为全球价值链规制提供新思路。



香港大学法律学院 孙皓琛教授

香港大学法律学院教授**孙皓琛**以《第二次隐私革命：中美两国的功利主义数据治理》为题发表演讲。他指出全球正经历从基于人权尊严的传统隐私观向以市场效率、消费者福利和公共安全为核心的功利主义治理模式的第二次隐私革命转型，其底层驱动力源于数据作为商品在交换、创新和预测中所展现出的巨大经济价值。通过比较中国的数据产权模式与美国的合同规制路径可以看出，数据商品化过程中可能导致的身份异化及算法权力对个人自由的侵蚀。主张通过确立算法知情权并引入算法尊严与集体尊严的概念来重塑数字时代的透明度与问责机制，以应对 AI 权力不对称带来的社会治理难题。



香港中文大学法律学院 柯丽娜副教授

香港中文大学法律学院副教授**柯丽娜**指出，全球价值链的理想状态是单一审查、全球统一救济以及可预测的数字规则；但现实却是多重审查、冲突的补救措施、严重的双重追诉风险，以及欧美等不同司法管辖区在数字监管上的严重分歧。竞争法的属地性干预干扰了跨境创新与生产，增加了企业的合规成本与投资不确定性，迫使跨国公司建立监管缓冲区，甚至导致小型价值链参与者退出跨境贸易。她进一步强调竞争法的功能在于保护市场竞争而非特定的竞争者，其在解决价值链内部的分配公平问题上存在天然局限，竞争法正日益成为贸易政策的延伸，表现出明显的地域和政治偏见。在此基础上，她建议企业需从知识产权、数据和关键投入的维度重新评估自身的竞争法风险，承认各司法辖区的不同并接受竞争法的限制。



鹿特丹伊拉斯姆斯大学 Alberto Quintavalla 副教授

鹿特丹伊拉斯姆斯大学副教授 **Alberto Quintavalla** 深入探讨了欧盟算力扩张与环境责任之间的尖锐矛盾，并指出当前欧盟针对数字化转型与环境保护的监管响应孤立化、数据中心可持续性的法律框架缺乏统一性等问题可能导致不同管辖区在吸引投资时陷入牺牲环境标准的逐底竞争。为此，他提倡构建一种综合运用强制性指令、私法救济、市场手段及软法引导在内的“明智组合”（Smart Mix）监管框架来破解治理难题。



中国人民大学经济学院讲师、未来法治研究院研究员 李子硕

中国人民大学经济学院讲师、未来法治研究院研究员**李子硕**发表了题为《解构平台算法驱动的内卷化竞争：典型场景与监管路径》的报告，剖析了平台经济中普遍存在的“内卷化”现象，指出不同于传统的恶性竞争，平台算法驱动的内卷化竞争本质上是一种“三方博弈”。平台为了获取长期利润，往往倾向于通过算法剥削和挤压中小商家，而非直接损害消费者。他认为这种模式具有深层且客观的经济学根源，内卷化竞争更侧重于审视价值链内部

的分配公平性与可持续性预期。在探讨监管路径时，他提出了直接规制与间接规制两种进路，并指出直接的价格干预仅在特定条件下能同时提升福利与公平，而在大多数场景下，监管者必须在最大化整体消费者福利与实现更公平的价值分配这一根本性冲突中进行权衡。



本次会议由中国香港城市大学法律学院香港商务及海事法研究中心、鹿特丹伊拉斯姆斯大学法律数字化中心组织，并得到了中国人民大学未来法治研究院、全球法学教育合作联盟的大力支持。会议聚焦数字转型背景下全球价值链重构的重大理论与实践议题，汇聚全球法学、科技治理、数字经济领域顶尖专家，搭建跨国家、跨学科、跨机构的学术对话平台，为人工智能治理、跨境数据监管、平台竞争规制、算力可持续发展等关键领域提供多元视角与治理思路，助力全球数字治理体系的协同创新与完善发展。

华政顺利召开“网络直播打赏前沿法律问题”专题研讨会

2026年4月25日，“网络直播打赏前沿法律问题”专题研讨会在上海吉臣维景酒店顺利召开。本次研讨会由华东政法大学数字法治研究院、上海市法学会数字法学研究会及《数字法学评论》编辑部联合主办，汇聚众多法学领域专家学者与司法实务工作者，共同聚焦网络直播打赏领域前沿法律问题，开展深度交流。

本次会议邀请了众多学界、司法界及行业相关领域的专家参与，其中包括中国社会科学院大学、中国传媒大学、对外经济贸易大学、华东政法大学、广东财经大学等高校的学者；同时，上海市高级人民法院、上海市第一中级人民法院、上海市第二中级人民法院、上海市普陀区人民法院、上海市嘉定

区人民法院、上海市浦东新区人民法院、杭州市中级人民法院、合肥市中级人民法院、苏州市中级人民法院、无锡市中级人民法院、宿迁市中级人民法院、北京互联网法院、广州互联网法院等单位的专家也出席了本次会议，各方齐聚一堂，旨在凝聚行业共识，助力直播行业规范健康发展。



本次会议共分为三个单元，第一单元的主题是“直播打赏的基础法律关系与平台治理”。在本单元中，各位专家重点探讨了用户、主播、平台三方的法律关系定性，打赏行为的法律效力与退款争议，未成年人与夫妻共同财产打赏等问题。厘清主播、用户、平台三方的法律关系，是司法审理此类纠纷的核心前提。其中，多数观点认为用户与平台、主播与平台之间构成网络服务合同关系；而用户与主播之间的法律关系则需分类界定，需区分平台内打赏与平台外私下转账、馈赠等行为，结合行为目的与实际往来，判定为有偿服务关系或无偿赠与关系。针对未成年人打赏，需要考虑未成年人权益保护和家长的监护责任约束；针对配偶一方使用夫妻共同财产打赏，需要考虑配偶利益和公序良俗原则。在退款追责环节，应结合平台与主播的收益分配模式，建立梯度化责任机制，依据实际权责与过错程度，合理认定主播和平台责任。

第二单元围绕“夫妻共同财产直播打赏的司法认定与无效返还”这一核心主题展开深入研讨。该单元重点探讨了夫妻一方直播打赏行为的法律性质、违背公序良俗的认定标准，以及打赏款项返还的主体和比例等问题。多数观点认为打赏行为需结合具体场景定性，区分正常消费型打赏与违背公序良俗的打赏；公序良俗的认定应秉持审慎原则，主要针对存在婚外不正当关系的打赏行为，不宜将单

纯线上暧昧言辞轻易认定为背俗。打赏款项返还需区分平台、主播、MCN机构的实际收益情况，平台在尽到合规监管义务、无过错的情况下，一般不承担返还责任，主播返还范围通常结合其实际到手收益及过错程度确定，避免过度扩大背俗认定范围或过度加重一方责任。

第三单元以圆桌讨论的形式展开，议题进一步延伸至数字经济背景下平台治理、司法价值导向等多方面内容。与会人员结合各地司法实践与典型案例，指出当前直播打赏纠纷的裁判标准存在地域差异，理论研究与司法实践之间亦存在偏差，例如“配偶单方大额打赏是否超出家事代理权”虽为学界广泛探讨的理论议题，但在实际司法审判中鲜有涉及。直播打赏法律关系需区分正常娱乐消费与违背公序良俗的行为，合理区分责任、酌情判定返还范围与比例。

本次研讨会围绕直播打赏法律问题展开深入研讨，与会人员一致认为司法层面需秉持规范、保障、促进数字经济发展的原则，同时兼顾法律效果和社会效果的统一。当前，直播打赏纠纷仍存在诸多法律问题与现实难题，包括司法裁判标准尚不统一、多方主体法律关系界定模糊、公序良俗界定标准不一等。未来，期待学界与业界持续深化交流，进一步完善相关治理机制，为直播行业健康发展提供坚实的理论支撑与实践指引。

中国人民大学人工智能治理研究院主办“生成式人工智能治理”专题学术研讨会

2026年4月25—26日，由中国人民大学人工智能治理研究院主办的第六次“人工智能治理学术研讨会”在中国人民大学立德楼成功举行，本次研讨会由中国人民大学交叉科学研究院、新时代智慧治理学科交叉中心、信息学院、新闻学院、书院建设与管理中心协办。本次研讨会以“生成式人工智能治理”为主题，汇聚了来自中国人民大学、中国科学院数学与系统科学研究院、广州大学、对外经

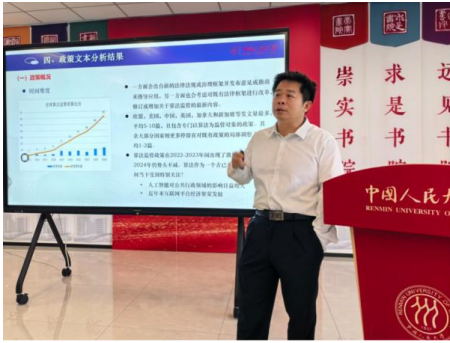
济贸易大学、北京科技大学、北京知链科技有限公司等高校、科研机构 and 单位的专家学者、业界工作者，跨计算机科学、管理学、数学、经济学、法学、新闻学等多个学科领域，共同围绕生成式人工智能治理的核心议题、实践难点与发展路径展开深度研讨与交流，为完善生成式人工智能治理体系、推动生成式人工智能规范健康发展凝聚跨学科智慧，贡献力量。

中国人民大学哲学院教授、人工智能治理研究院院长刘玮向所有与会的专家、学者、学生表示热烈欢迎。他介绍了人工智能治理研究院的建设目标和宗旨，提出生成式人工智能已深度融入日常生活，机遇与治理挑战相伴而生，希冀各位师生在两天的研讨中深入交流，碰撞思想，为当下面临的问题提供真知灼见。



25日上午第一场讨论由刘玮教授主持，讨论围绕全球算法监管、人工智能生成文本识别等议题展开。

中国人民大学公共管理学院教授、人工智能治理研究院研究员刘鹏以《全球算法监管的趋势特征、要素识别与国别分析研究——基于政策文本量化分析》为题进行分享。他聚焦全球算法监管政策的特点与趋势，从公共管理视角出发，探讨算法在社会各领域应用引发的问题及监管需求，提出了包括算法影响评估、数据质量管理、算法审计及信息披露在内的风险管理策略，强调了主体赋权、算法备案和记录追溯等措施在提高算法透明度和合规性方面的作用。



中国人民大学信息资源管理学院教授、人工智能治理研究院副院长**任明**以运用信息学定律设计人工智能生成文本的识别方法为核心问题，发表《人工智能生成文本识别方法研究：信息学定律的视角》主题分享。她探讨了AI生成内容带来的真实性危机，阐述了信息学定律在揭示人类和AI生成文本差异中的应用，利用齐普夫定律等规律，提出了一种基于统计指标的AI文本识别方法，为人工智能治理领域提供了有效的识别视角和治理策略。



25日上午第二场讨论由**任明**教授主持，讨论围绕多模态网络暴力检测、人工智能与健康工作预期寿命、人智交互下的用户行为等议题展开。

中国人民大学信息学院讲师、人工智能治理研究院研究员**傅哲**围绕《基于情境感知的多模态网络暴力检测》分享研究成果。他指出，现有网暴检测多依赖事后互动数据，缺乏事前干预能力。在此背景下，他提出“网络情景感知”视角，构建四维前置风险指标：攻击回升指数、攻击焦点匹配度、传播极化指数和互动毒性系数，通过多模态融合、多

专家系统等方法实现仅凭借视频内容即可预测网暴发生概率，突破“无互动不预警”的瓶颈。



中国科学院数学与系统科学研究院博士后**彭慧**以《人工智能与健康工作预期寿命》为题展开分享。她探讨了AI对健康和工作预期寿命的双重影响，指出AI显著延长了健康工作预期寿命和工作预期寿命，但增加了不健康工作的年限，对健康构成一定的威胁。她提出，行业和地区分析显示，AI在劳动密集型和知识密集型产业带来收益，但在东部地区对健康构成更大威胁。适度AI渗透率有益于人机协作，高水平则可能损害健康。



中国人民大学信息学院副教授、人工智能治理研究院研究员**余艳**以《人智交互下的用户行为》为题，分享了其最新研究成果。她首先讨论了AI在优化型任务与战略性任务中的不同表现，指出AI在优化型任务中可显著提升质量，而在战略性任务中效果较差。她探讨了情感人工智能在提升服务失败容忍度中的作用，通过实验方法和经济计量分析，提出借助情感人工智能提升服务失败容忍度的策略。基于视频和文本实验，她讨论了拟人化特征对人工智能失败容忍度的非线性影响，验证了喜爱情感对服务失败容忍的正向影响，以及心理所有权在其中的调节作用。



25日下午,论坛第三场研讨由**余艳**副教授主持,与会专家融合新闻学、社会学等跨学科视角展开讨论。

中国人民大学新闻学院教授、人工智能治理研究院研究员**卢家银**以《全球人工智能治理的政策图景》为题展开分享。他围绕人工智能政策在全球范围内的目标设定与工具选择展开,通过分析46份全球人工智能政策文件,探讨了不同发展水平国家的政策差异及其背后的逻辑,分析了政策目标与工具选择的适配性问题。他发现,政策目标上呈现强国家、重市场、弱社会的特点,工具应用上以环境侧为主,存在工具与目标失衡问题。他指出,未来政策需平衡国家与市场作用,加强社会权利与公平利益的治理。

广州大学新闻与传播学院讲师**黎樟浩**以《人工智能信任的全球差异:文化与媒体的结构性影响》为题,探讨了不同国家对AI的信任差异。他构建FAT框架探讨了公众、教育水平、国家文化、媒体态度等AI信任影响因素,发现其中公平性最受公众关注,透明性和问责性则因离公众较远而不显著。他的研究首次运用国家文化维度解析AI信任问题,揭示了权力距离、个人主义与长期导向对AI信任的影响。他建议,AI开发者应强调公平性,政府监管不必担忧过度影响信任,而媒体应正向引导AI认知。

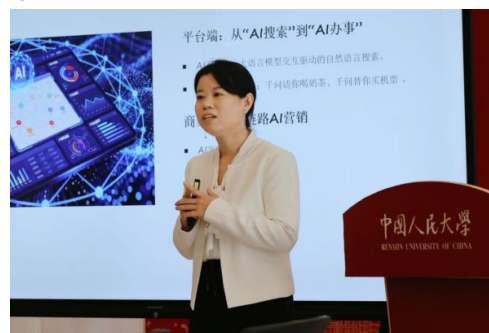


中国人民大学新闻学院讲师、人工智能治理研究院研究员**王裕平**以《人工智能与虚假信息的治理》为题,强调虚假信息治理的重要性。他从虚假信息的定义入手,重点探讨了人工智能在判断信息真伪中的作用。他分享了利用智能体进行自动化实时核查图片真实性的方法,通过反向检索、证据分析和大模型推理生成核查报告。



第四场讨论由**卢家银**教授主持。三位发言人围绕AI披露标签与消费者意愿、生成式人工智能模型评测、共情生成式AI对中学生写作表现的影响等议题展开。

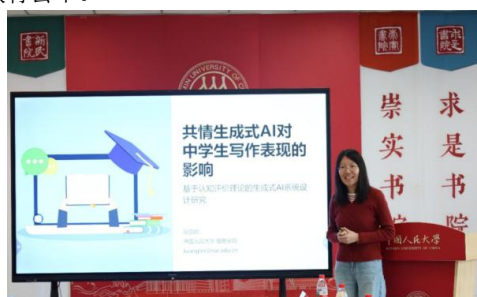
对外经济贸易大学国际经贸学院副教授**金悦**以《AI披露标签与消费者UGC补偿搜寻行为》为主题展开分享。她指出消费者在线购物时对商家生成内容(MGC)的不信任,特别是AI生成图片的透明度缺失导致的购买意愿下降。对此,消费者可能更多依赖用户生成内容(UGC),如评论,来补偿信息不足。



中国人民大学信息学院讲师、人工智能治理研究院研究员**王文轩**带来《社会科学启发的生成式人工智能模型评测》的主题分享。他讨论了国产大模型在文化价值观表达上的优势,分析了当前评测模型在信度和效度上的局限性,包括数据泄露、角色扮演和理解等距性等问题。他指出,大模型安全对

齐是一个复杂且长期的探索过程，要面临目标冲突、不匹配泛化、浅对齐等挑战。

中国人民大学信息学院讲师、人工智能治理研究院研究员**匡丽妮**以《共情生成式AI对中学生写作表现的影响》为题，基于认知评价理论，设计共情AI系统在学生评估写作建议时提供情绪调节，以促进更有效的作文修改。她介绍了一个基于高考英语作文评分标准的写作辅助系统，系统通过AI提供修改建议并进行共情解释，旨在提升学生写作表现。通过实验，她发现，共情AI系统显著提升学生写作表现，尤其对低能力学生帮助更大，促进教育公平。



论坛第五场讨论于4月26日上午举行，由中国人民大学信息学院教授、人工智能治理研究院研究员**程霖森**担任主持人。两位发言人围绕生成式人工智能治理问题展开深度解读。



中国人民大学信息学院教授、人工智能治理研究院研究员**杨波**以《学术科研中的生成式AI治理》为题展开分享。他讨论了生成式AI给科学研究带来的信任危机，提出问题的解决可能不在于AI本身，而在于人的使用方式和学术环境、评价机制的调整。他提出，在AI时代，科研评价需关注成果的创新性而非研究手段，同时探讨了如何规范使用AI以鼓励期望行为，确保科研目的的实现。



中国人民大学商学院教授、人工智能治理研究院研究员**王刊良**带来题为《生成式人工智能前沿问题》的分享。他结合了技术与社会相互塑造理论、IT生产率悖论，提出技术是替代还是增强现有工作决定了社会接受度，呼吁深入研究技术对社会、行业和企业的影响。基于媒体丰富性理论和同步性理论，他强调不同媒介在复杂度与任务匹配上的重要性。他还探讨了人类知识的显性与隐性部分，指出编码知识仅是冰山一角，隐性知识在人类行为中占据主导地位，且从未知到已知、从意识到表达的转化过程充满挑战。



论坛的最后一场讨论由中国人民大学法学院副教授、人工智能治理研究院研究员**黄尹旭**主持。三位报告人围绕冲突价值观评测、人工智能赋能文旅服务、对生成内容的安全探索等话题展开讨论。

中国人民大学法学院副教授、人工智能治理研究院研究员**彭雅丽**以《AI冲突价值观评测》为题进行了分享。她探讨了主流大语言模型在跨文化和跨境背景下将抽象伦理原则转化为具体法律伦理判断的能力，发现模型在伦理基线上存在一定程度的一致性，但文化起源和提问语言可能影响输出。她通过构建关怀范围、决策主体和变革取向三个正交维度，揭示了AI生态中普遍存在的自由主义价值

观倾向，指出英文语料的主导地位及技术路线的一致性可能是导致这一倾向的原因。



北京科技大学经济管理学院副教授**傅诗轩**以《人工智能与沉浸式虚拟技术驱动下的文旅服务交互设计》为题展开分享。她探讨了在文旅行业，人工智能与虚拟现实技术在文化遗产数字化表达中的应用及问题，包括表达偏差、文化偏见、历史厚重感流失等风险。她提出了通过技术手段识别错误表达、增强人机协作以优化文化遗产数字化表达的策略。此外，她还关注了博物馆场景下的文化内容开发与数字博物馆建设，强调多学科融合在提升文化价值表达与传播中的作用。



北京知链科技有限公司总裁**刘全宝**以《基于云端对生成内容的安全探索》为题展开分享。他探讨了 AI 和区块链技术在发展过程中面临的安全隐患和数据隐私问题，指出数据全生命周期中存在安全漏洞，呼吁关注技术应用的伦理和安全。他指出，由于大模型可能携带有害指令，且使用过程中存在权限滥用和数据安全风险，因此需要重构 AI 应用流程，引入 MCP 协议，确保上下文可控、数据有水印、责任可追溯。

在为期一天半的研讨会上，学者和学生们围绕相关议题展开了热烈而有深度的讨论。

人工智能治理研究院将围绕数智社会治理的核心议题持续开展学术研究和成果交流，推动多学科的融合、治理创新和科技向善。

第九届数字中国建设峰会“大模型、人工智能体与涉外法治前沿分论坛”顺利召开

4月29日，第九届数字中国建设峰会在福建省福州市开幕。本届峰会由国家发展改革委、国家数据局、国家网信办、工业和信息化部、福建省人民政府共同主办，主题为“加快数智技术创新发展，深入推进数字中国建设”。中共中央政治局常委、国务院副总理**丁薛祥**出席开幕式并发表主旨讲话。

作为本届峰会唯一的法治论坛，由中国人民大学主办，中国人民大学法学院、中国人民大学福州研究院、福建省电子信息集团承办的“大模型、人工智能体与涉外法治前沿分论坛”于当日下午在福州市长乐区数字中国会展中心 203 会议室顺利举行。本次分论坛得到中国人民大学涉外法治学院、中国人民大学未来法治研究院、教育部哲学社会科学创新团队“新科技革命与未来法治创新团队”、福州大学法学院、福建省法学会数字法学研究会、闽江大学法学院等单位大力支持，并由《数字法治》编辑部提供学术支持。来自政府机关、高校、企业等单位的 200 余位领导嘉宾、专家学者和业界代表参加了本次分论坛。

中国人民大学党委副书记**郑水泉**，福建省委常委、政法委书记**石好勇**，中国法学会网络与信息法学研究会会长、最高人民法院咨询委员会副主任委员**姜伟**，中国人民大学法学院院长**杨东**出席开幕式并依次致辞。福州市委常委、政法委书记**朱训志**，福州市教育局副总督学**苏夏铃**及承办单位中国人民大学福州研究院院长**郭海鹰**、福建省电子信息集团党委副书记**郑志峰**出席开幕式。

郑水泉表示，人工智能正深刻重塑生产生活方式，但也在安全治理、规则供给和涉外法治等领域带来了新的挑战。近年来，中国人民大学发挥法学、人文社会科学与人工智能等多学科优势，推进交叉

融合，研发涉外法治大模型等标志性成果；坚持问题导向，深化与司法机关、企业和数据机构合作，服务数字法治建设；拓展国际合作，搭建人工智能法治研究与教育交流平台。期待以本次论坛为契机，进一步凝聚学界、实务界与产业界共识，共同推动人工智能在安全可信、法治化轨道上行稳致远。



中国人民大学党委副书记 郑水泉

姜伟表示，随着模型算法的迭代加速，智能体应用场景持续拓展，人工智能已成为推动数字经济高质量发展的关键力量。面对技术创新与风险治理并存的新形势，人工智能法治建设应当以充分发挥人工智能引擎作用为目标，加强法治保障，切实防范人工智能的现实风险，以法治之力促进科技创新发展。同时，应当积极参与全球治理，统筹推进国内法治与涉外法治。



中国法学会网络与信息法学研究会会长、最高人民法院咨询委员会副主任委员 姜伟

杨东表示，中国人民大学法学院始终关注技术变革、产业创新与数字文明时代的法治回应，积极服务数字中国建设。中国人民大学法学院依托未来法治研究院，深度参与数字法治与人工智能立法、数字中国法治建设工作。在习近平总书记考察调研

中国人民大学并发表重要讲话四周年之际，学校正式成立涉外法治学院，着力培育法科复合型涉外法治人才，为全国首创。未来，人大法学院将继续依托数字中国峰会、世界互联网大会等平台，深化政产学研合作，推动法治建设与法治人才培养协同发展。



中国人民大学法学院院长 杨东

中国人民大学法学院副院长、未来法治研究院副院长丁晓东主持开幕式。



中国人民大学法学院副院长、未来法治研究院副院长 丁晓东

论坛现场举行中国人民大学福州研究院涉外法律咨询服务中心（福州）揭牌仪式。在郑水泉副书记、姜伟会长的见证下，中国人民大学法学院院长杨东，中国人民大学福州研究院院长郭海鹰共同为“中国人民大学福州研究院涉外法律咨询服务中心（福州）”揭牌。该中心依托中国人民大学法学院、涉外法治研究院建设，立足福州、辐射福建、服务东南，旨在积极支持福建省涉外法治工作布局，为海丝中央法务区和中国（福建）自由贸易试验区建设提供高质量的涉外法治智力支持与人才支撑，为高质量共建“一带一路”提供有力法治保障。



揭牌仪式

中国人民大学法学院分别与上海市知识产权局、浙江省绍兴市新昌县人民政府签署合作协议。



中国人民大学法学院与上海市知识产权局签约

自2022年以来，中国人民大学法学院杨东教授团队和上海知识产权局围绕数据产品知识产权领域，有力推动上海市数据产品知识产权登记试点落地并在全国范围内进行推广。把涉外法治大模型的涉外法律英国、越南问答指令数据作为全国范围内首个高校数据产品知识产权进行登记，这是我国首例经过实质性审核、被官方机构登记确权的涉外法治数据集。后续两团队之间将继续保持紧密的战略合作和共同研究，包括数据产品知识产权国际规则研究、数据产品知识产权实质性审查标准研究等，并携手建立数据产品知识产权创新人才培养基地和探索培养数据产品知识产权交叉复合人才。



中国人民大学法学院与浙江省绍兴市新昌县人民政府签约

中国人民大学法学院与新昌县人民政府将围绕法治政府理论研究、涉外法治大模型建设、区块链安全与反洗钱实验室建设等方面开展战略合作，进一步深化基础理论研究，加强技术领域研究，打通成果转化通道，切实推动各项合作开花结果。

论坛发布中国人民大学法学院人工智能立法主题报告和人工智能法治知识平台软件系统，福建省电子信息集团所属企业锐捷网络发布了行业先进智能体“阿在”，杭州汉资信息科技有限公司发布了“数字中国”应用场景之“司法信用修复”，执行领域分级治理的智能化应用模型。

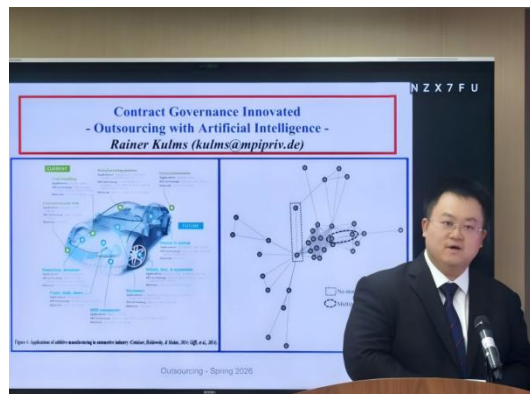
来自政府机关、高校、企业等单位的与会嘉宾围绕“人工智能发展的法治回应与制度创新”和“人工智能全球治理变革中的涉外法治应对”两大议题展开讨论，并针对“全球视野下人工智能治理的机遇与挑战”进行圆桌对话。

杨东在总结发言中表示，中国人民大学法学院高度重视涉外法治学科建设与人才培养，作为中央依法治国办、教育部联合确定的全国涉外法治人才协同培养创新基地秘书处单位，在全国高校率先开展涉外法治战略人才班与创新人才班招生培养工作，发布全国高校首个涉外法治大模型并入选教育部首批高等教育智能体。本次分论坛集中展示了中国人民大学法学院在人工智能法治、涉外法治和数字法治领域的研究成果与实践布局，也进一步凝聚了政产学研各界共识。希望以本次论坛为契机，与各方持续深化协同合作，推动形成更多具有理论深度、实践价值和国际影响力的数字法治成果，为深入推进数字中国建设贡献法治智慧。

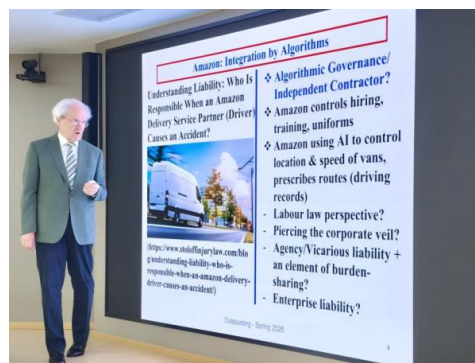
“人工智能外包的合同治理创新”讲座圆满举办

2026年5月6日，由中国人民大学未来法治研究院、民商事法律科学研究中心、涉外法治研究院、普通法中心联合主办的“未来法治与数字法学前沿讲坛”在明德法学楼顺利举办。德国马克斯普朗克比较法与国际私法研究所高级研究员 Rainer

Kulms 教授应邀主讲“人工智能外包的合同治理创新”，中国人民大学法学院张吉豫教授、包丁裕睿助理教授、陈哲立讲师与谈，吴至诚副教授主持，校内外师生共同参与研讨。



Kulms 教授系统回顾人工智能发展历程，指出生成式 AI 已深度融入服务外包、供应链、金融、医疗、智能制造等领域，传统合同治理与私法规则面临前所未有的挑战。他以亚马逊配送、银行数据外包、智能投顾、AI 医疗手术、按需制造五大场景为切入点，深入剖析现实法律困境：平台通过算法对配送员实施全流程管控，劳动关系与责任认定陷入争议；银行外包数据处理引发消费者隐私保护与责任划分难题；智能投顾自动化决策带来受托人义务与损害赔偿争议；医疗 AI 与手术机器人则涉及知情同意、医疗过失及多方责任分配。他强调，AI 技术突破合同相对性，算法治理替代传统人工管理，现有代理、雇主责任、法人格否认等制度难以适配高度整合的数字化生态。为此，应转向功能性企业责任，以实质控制与风险收益匹配为标准重构责任体系。



评议环节，与会嘉宾结合中外实践展开对话，提出普通用户使用 AI 代理进一步加剧责任复杂性，透明化难以充分保障权益，保险机制与实质控制标准或为破题关键。中国司法在平台用工、金融外包中坚持“不得以外包转嫁责任”的经验，为全球治理提供借鉴。互动中，现场师生就全自动 AI 应用边界、数据利用与隐私保护冲突、监管路径选择等问题积极提问，主讲人逐一回应。

本次讲座紧扣数字法治前沿，为 AI 时代外包合同治理创新、责任规则完善提供了重要理论参考，有力推动了人工智能法律治理的学术交流。



人大人工智能治理研究院主办大模型训练数据的使用规则研究

2026年5月6日，由中国人民大学人工智能治理研究院主办的第二十五期人工智能治理沙龙举行。本次沙龙围绕大模型训练数据的使用规则展开，由中国人民大学交叉科学研究院、高瓴人工智能学院博士后李铭轩进行主题分享，北京理工大学法学院助理教授、工信部智能科技风险法律防控重

点实验室研究员**包晓丽**担任对谈嘉宾，沙龙由中国人民大学哲学院副教授，人工智能治理研究院研究员**王福玲**主持。



李铭轩首先指出，在当前的法律环境下，获取海量训练数据面临着著作权、个人信息权益及数据财产权等排他性权利的法律限制，以及 Robots 协议、API 服务协议等合同约定的约束。这些规则在保障安全与维护公平方面具有正当性，但过强的排他性也会大幅推高许可成本，进而阻碍大模型的创新进程。

基于法律经济学视角，李铭轩引入了“市场失灵”理论来透视数据获取的结构性难题。在大模型数据市场中，外部性与交易成本是导致资源配置低效的主因。例如，非营利性科研训练虽能产生巨大的社会价值，但开发者往往缺乏支付高昂对价的能力；同时，语料库涉及的权利主体动辄以千万计，海量的搜寻、谈判及事后监管成本使得自发的市场合意极难达成。这种情形下，法律必须主动介入，通过制度设计来替代失灵的市场机制。

李铭轩随后分析了现有的各种数据利用模式的局限性。开放共享模式虽能提供 PB 级别的公共语料，但高质量私有数据开放度不足且存在来源合法性争议；授权使用模式无论是一对一授权还是集体管理，都面临管理成本过高、收益分配微薄及易筑起竞争壁垒的风险；法定许可与合理使用模式虽能降低门槛，但前者定价难度大，后者则可能实质损害权利人的经济利益。

最后，他提出了规则完善的双重路径：市场环境规则与市场补充规则。前者旨在通过明确开源数据定性、完善数据确权及引入“延伸性集体管理”等机制来优化市场运行环境；后者则主张在市场失

灵时，通过法律解释适度拓宽“合理使用”的边界。例如，在人工智能训练场景下，应综合考量识别成本与个人意愿，允许在特定范围内合理处理公开信息，为技术创新预留必要的公共领域空间。

包晓丽对主讲人的报告做了背景补充，介绍了两个实践案例，案例一是开源的 ImageNet 标注数据虽然已声明仅限非营利性科研使用，但市场上大量商用模型都在违规使用。案例二是美国大模型厂商 Anthropic 使用受版权保护的电子书进行训练而被起诉，被告辩称模型仅吸收了逻辑与参数，并未复制原作者的“表达”，最终法院认定不构成违法，但需支付一定的法定许可金。

随后包晓丽就大模型训练规则的法理与实务提出了四个维度的追问：首先她质疑了以“社会整体福利”为由牺牲数据产权人利益的正当性；其次是探讨现有的四种数据利用模式是否本质上只是“已公开数据”与“未公开数据”在不同场景下的交叉应用；第三是营利性的界定，她追问是否应以“营利/非营利”作为划分集体协商收费与合理使用免费的界限；最后是社区规则与法律的关系，她关注在开源生态中，技术社区自发形成的治理规则能否在一定程度上排除或替代国家法律的强制性规则。



针对包晓丽指出的法律在判定效率时往往以社会整体福利为首要标准，但这在伦理上存在困境：如果数据对个体的价值远超其对大模型开发者的价值，为何要强制个体为了社会福利而让渡不开放的自由？王福玲从伦理学视角作出回应，个体在拥有数据所有权之前，其成长与创作已经受益于社会长期积累形成的肥沃土壤。因此，数据拥有者有

责任在特定情形下回馈社会，这种义务类似于企业的社会责任，通过短期的利益让渡来优化整体生存环境，最终为个体带来更广阔的长远收益。

李铭轩回应授权使用不仅针对非公开数据，同样适用于已公开数据，数据利用模式的维度不能按照“是否公开”来简单划分。针对以营利性作为合理使用划分界限的观点，他明确了合理使用不止局限于非盈利使用，营利性使用只要构成不替代原作品市场的“转换性使用”仍有被认定为合理的空间。在社区规则与法律效力的权衡上，他强调应视其是否促进效率及是否侵害合法权益而定。若自发形成的社区规则能极大促进效率，法律理应尊重；但若是数据寡头利用强制性的点击合同排除用户合法合理使用权以扩张私权，此类社区规则即为无效。

有同学提出，高质量数据持有者的贡献与获益之间存在巨大的“时间差”，且数据一旦共享，个人对数据的专有权和安全性将面临损失。特别是在插画等艺术领域，AI虽未直接复制作品，但通过学习画风、配色等层面替代了原作者的市场，这种剥夺感在现有仅保护“表达”的版权框架下难以得到合理补偿。李铭轩回应到，在AI能穷尽一切表达的时代，法律关于“保护表达、不保护思想”的底层逻辑正面临革命性挑战。对于涉及尊严、隐私等核心利益的数据，应坚持不可让步的刚性规则；而对于纯粹的经济利益，则可以尝试“先使用后付费”的二次分配机制，通过税收或事后补偿来消解创作者的公平性顾虑。

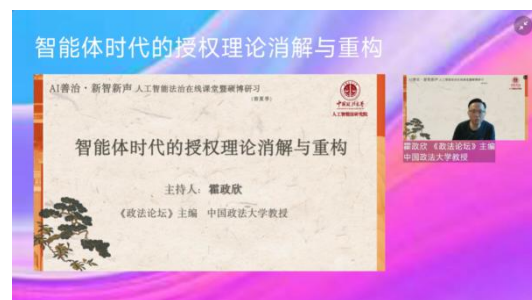
有同学建议参考欧盟的“访问权”制度，同时追问大模型嵌套智能体时，个人开发者与基座平台之间的归责逻辑。对此包晓丽老师明确了“数据空间”与“访问权”的功能差异。她强调，国内推行的行业“可信数据空间”主要功能是解决信任问题，而非无条件的数据共享，欧盟访问法和这个不太相关。李铭轩认为，AI责任规制应秉持“风险最小化”原则。在因果链极长的“黑箱”模型中，法律应将责任施加给能以最低成本防范风险的主体。对于因技术局限产生的不可预见风险，除了侵权归责外，

未来还应引入保险、赔偿基金等多元化的社会分配制度来共同承担。

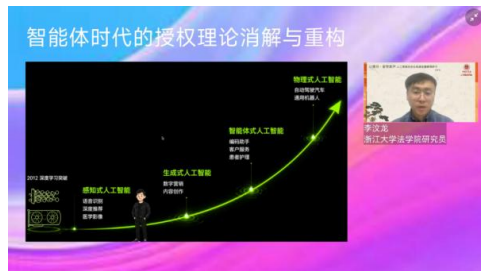
本次沙龙从宏观的市场失灵理论探讨到微观的具体法律规则解释，呈现出大模型训练数据的使用不仅在挑战传统权益的控制边界，也在重构数字时代的资源配置机制。法律系统面临的核心挑战不再是如何通过严格的排他性来静态守护数据财产，而是如何在呼唤创新的技术浪潮中，通过灵活适用与解释“合理使用”等规则，构建兼顾公平与效率的数据治理生态。

“AI善治·新智新声”人工智能法治在线课堂暨硕博研习(春夏季)第三场圆满举办

5月7日，“AI善治·新智新声”人工智能法治在线课堂暨硕博研习(春夏季)第三场顺利举办。本场活动由中国政法大学人工智能法研究院牵头，联合多所知名政法院校与科研机构共同打造，以线上研学形式，为全国各高校硕博学子搭建了一个前沿、鲜活且务实的学术交流与思想对话平台，助力硕博学子深耕人工智能法治领域、拓宽学术视野。



活动由《政法论坛》主编、中国政法大学霍政欣教授主持。霍教授指出，当前人工智能智能体技术迅猛兴起，自主决策能力的跃升正在重塑人机协作模式与法律关系格局，传统授权理论面临前所未有的挑战与重构压力。本场讲座聚焦智能体时代的授权逻辑消解与重构这一前沿命题，期待与各位嘉宾共同深入探讨，凝聚学术共识，为智能体法律规制研究提供新思路与理论支撑。



主讲环节中，浙江大学光华法学院**李汶龙**研究员以“智能体时代的授权逻辑消解与模型责任建构”为题作主题分享。李老师指出，个人信息保护、知识产权许可和民法委托代理等传统数字领域中的授权机制，本已面临诸多争议；而智能体技术进一步聚合并放大了这些问题，正在系统性动摇既有授权逻辑。李老师重点分析了智能体场景下授权面临的三重挑战：一是自然语言交互和提示词表达具有高度开放性，容易引发无权代理、越权代理和表见代理等问题，但传统民法代理规则又难以直接适用于智能体；二是用户授权意志往往在与模型持续互动中逐渐形成和变化，而非预先明确存在，导致授权边界不断被稀释；三是多智能体系统具有复杂的跨模型、跨平台调用结构，容易使原本单向的授权关系在层层转译和外包中发生断裂。治理路径上，李老师提出，首先应从形式化授权转向结果控制和能力边界设定。其次应建立结构化的权限表达体系，发展机器可读、可验证、可追踪的授权凭证，减少对开放性提示词的过度依赖。再者，应构建系统性的模型责任机制，将责任配置前移至模型设计、能力供给、权限管理和身份识别等环节，通过对被授权模型的识别、行为链条的记录、责任与损害的溯源，降低无权代理、越权代理等风险。智能体治理不能仅仅依赖事后侵权救济，而应在事前通过能力设计、权限架构和责任分配，回应智能体时代授权逻辑所面临的深层理论与制度挑战。



与谈环节，三位专家从多维度、全视角针对智能体时代的法律问题展开交流，为线上研修的硕博研究生开阔学术眼界、带来思想启发。对外经济贸易大学副校长**梅夏英**教授认为，智能体的本质仍是高级自动化工具，赋予其独立人格或代理地位既无必要也不可行。他指出，智能体带来的核心问题属于公共安全领域的人机关系调整范畴，应依靠公法规范与技术标准加以规制，而非将其简单套用私法代理制度框架，否则将导致理论逻辑的错乱与治理实效的双重失落。



中国人民大学法学院**张吉豫**教授指出，可通过构建“三重结构”理论来弥合智能体授权的效力鸿沟，即搭建人类理解层、机器理解层与法律表达层，通过分层设计确保授权的有效性与其可执行性。她同时提醒，过于严格的事前审查机制可能抑制技术创新，建议以推荐性标准或指南为主要形式，结合避风港原则，在安全保障与产业发展之间寻求动态平衡。



蚂蚁集团研究院**李振华**院长就智能体的授权问题从多个角度进行分析。李院长认为，智能体授权分为人对智能体的授权、app对智能体的授权，且上述两层授权在逻辑上有所不同。他提出可以通过规范化的意图转换机制，将自然语言转换为结构化的表达，形成可约束、可验证、可追溯的授权基础。此外，还应重视授权凭证生成、存储和管理，

可采用加密凭证和区块链技术来确保授权的可追溯性和不可否认性。



硕博研习环节，来自东南大学的博士研究生**肖远航**、中国政法大学的硕士研究生**郭玉洁**就社会学理论在法学中的应用、跨部门法授权统一性以及智能体越权判断标准的细化等问题，与主讲嘉宾、与谈专家展开探讨，充分展现了新生代法学研究者饱满的学术热忱与扎实的思辨探索能力。主讲人与与谈专家针对学生的提问进行了细致答疑、精准点拨，交流富有启发、研讨成效显著。

第三届网络与信息法学学科建设论坛在福州成功举办

2026年5月9日，第三届网络与信息法学学科建设论坛在福建福州成功举办。本届论坛以“‘人工智能+’背景下网络与信息法学的学科回应与实践探索”为主题，由中国法学会网络与信息法学研究会主办，福建师范大学法学院承办，《中国法学》《法学研究》《中外法学》《法学家》《法商研究》《法学》《法律科学》《法学评论》《政法论坛》《法制与社会发展》《现代法学》《比较法研究》《环球法律评论》《清华法学》《政治与法律》《当代法学》《法学论坛》《法学杂志》《华东政法大学学报》《中国刑事法杂志》《东方法学》《行政法学研究》《中国法律评论》《法律适用》《国家

检察官学院学报》《财经法学》《国际法研究》等期刊提供学术支持，《数字法治》编辑部参与学术合作。论坛汇聚了高校、司法机关、科研院所、期刊出版单位、企业和实务部门的专家学者，共同围绕“人工智能+”背景下网络与信息法学学科建设的理论与实践问题展开交流研讨。本届论坛由主论坛和四个专题分论坛组成。论坛开幕式由福建师范大学法学院院长、福建省法学会副会长**杨垠红**主持。福建师范大学校长**郑家建**，福建省委政法委副书记**林良灌**，中国法学会网络与信息法学研究会常务副会长兼秘书长、同济大学法学院特聘教授**周汉华**，福建省人民检察院二级巡视员**张时贵**，福建省高级人民法院审判委员会委员、研究室主任**范宁宇**分别致辞。各位领导嘉宾从政法协同、学科前沿、检察实务、审判实践等多元视角切入，深入阐述了网络与信息法学学科服务国家战略、回应时代需求的重要意义与实践路径。开幕式后，中国法学会网络与信息法学研究会常务副秘书长，中国人民大学法学院教授、未来法治研究院执行院长**张吉豫**正式发布《人工智能法课程（示范方案）》。主旨演讲环节由中国社会科学院学部委员、《法学研究》主编**陈甦**主持。中国政法大学比较法学院院长、《比较法研究》主编**解志勇**，中国人民大学法学院教授、《法学家》副主编**韩立余**，上海社会科学院法学研究所研究员、《政治与法律》常务副主编、编辑部主任**陈庆安**，吉林大学法学院教授、《法制与社会发展》副主编**侯学宾**，公安部十一局局长**张钢强**，中国法学会网络与信息法学研究会副会长、腾讯集团法务副总裁**刁芸芸**，福建师范大学法学院院长、福建省法学会副会长**杨垠红**等专家学者围绕数字法学体系、数据进出口管制、网络信息犯罪的法律适用、数据产权登记制度、网络与信息法学人才培养、AI时代的产业法治、区域数智法学建设等议题作主旨演讲，紧扣“人工智能+”背景下的学科前沿理论与实践展开深入研讨。当日下午，会议设置“学习贯彻习近平法治思想与习近平总书记关于网络强国的重要思想”“网络空间法治实践与司法实务前沿”“网络与信息法学学科建设与人才培养”

“网络与信息法学学科理论体系构建”等四个分论坛。与会代表围绕人工智能生成内容的著作权问题、网络违法惩处机制、网络与信息法学学科人才培养、跨学科理论体系构建、人工智能赋能法学教育等前沿热点议题展开了深入研讨，与会专家交流热烈，充分体现了论坛鲜明的问题导向、实践导向与前沿导向。闭幕式于当日17时40分举行，由中国法学会网络与信息法学研究会副秘书长，《数字法治》执行主编、编辑部主任、编审兰丽专主持。福建师范大学法学院党委书记盛子同、中国法学会网络与信息法学研究会常务副会长兼秘书长、同济大学法学院特聘教授周汉华分别作总结发言，系统梳理总结本届论坛的研讨成果，并就后续工作部署安排。

《人工智能法》浙大版学者建议稿圆桌论坛在杭召开

在促进人工智能创新与安全发展，捍卫中国在全球人工智能治理新秩序话语权的时代关口，我国人工智能立法者肩负现实而紧迫的责任。浙大法律人愿以《人工智能法》浙大版学者建议稿之“砖”，引全国乃至全球法律人之“玉”。

2026年5月9日下午，由浙江大学数字法治研究院、浙江立法研究院、浙江大学立法研究院联合主办的“《人工智能法》浙大版学者建议稿圆桌论坛”在浙江省杭州市浙江宾馆花港厅顺利召开，邀请全国专家学者共商立法大计。

开幕式



开幕式由浙江大学智库高端人才、光华法学院研究员孙军工主持。浙江大学光华法学院党委书记、

浙江立法研究院、浙江大学立法研究院院长郑春燕代表主办方致欢迎辞，浙江大学数字法治研究院院长孙笑侠阐述了建议稿草案的基本原则。



郑春燕书记指出，浙大推出《人工智能法》学者建议稿具有显著优势：浙江大学是我国学科门类最齐全的高校之一，具有法学与计算机等学科交叉研究的深厚传统。自2015年起，浙江大学光华法学院就成立互联网法治研究中心，长期深耕数字法治领域，具有现行优势和深厚积淀；地处数字经济蓬勃发展、正成为世界级人工智能产业集聚中心的杭州，浙大团队深度参与《杭州市创新活力之城条例》《人工智能产业促进条例》等地方立法起草，直面人工智能产业发展的实践需求；她表示本次圆桌论坛幸得全国各地的学者、实务专家莅临，期待听取各方宝贵意见，进一步完善建议稿，以备后续的正式发布。



孙笑侠院长就《人工智能法》浙大版学者建议稿的“立法导向”阐述了“能—智—事适配”原则。他表示，人工智能立法首先要解决的根本矛盾是发展与安全的平衡问题，需对人工智能的全生命周期进行风险拆解。针对当前风险评估中的过度归因以及国际上某些地区的“扼杀式监管”，他提出人工智能的风险根源在于“能、智、事”三者错配。该原则包括场景特异性、安全与发展平衡、责任网络

适配三个维度，可细化为六项制度体现：阶段化准入与分级监管；按 AI 能力、认知水平、场景复杂度分级准入；场景特异性使用条款；采用“以能测智、以能代智”的评估方法；动态责任分配与链条化治理；技术标准嵌入法律规范。随后，孙院长接受专访，就“科际法学”的异质型、“人工智能法”中国特色、立法边界、技术创新与法律治理等问题进行深度回答。

建议稿发布

本环节由浙江大学光华法学院特聘副研究员**陆家豪**主持并发布《人工智能法》浙大版学者建议稿，浙江大学光华法学院“百人计划”研究员**李汶龙**介绍了建议稿的研究背景、方法与整体框架。



据李汶龙介绍，《人工智能法》浙大版学者建议稿广泛吸收近年人工智能立法成果，变革了传统的全维度风险或全产业促进的立法思路，以“能一智一事适配”为基础，统筹发展与安全，坚持全生命周期管理与风险适配的敏捷治理，贯穿“以人为本、智能向善、公开透明、责任追溯、公平公正与多元共治”的价值导向。建议稿构建了总则、原则层与执行层三大结构，涵盖了基础奠定、监管体制、执行机制与风险救济等规则设计。



本次圆桌论坛分为两个专题，专题一为人工智能立法的结构和基本范式研讨，由浙江大学光华法学院研究员**魏斌**主持。



复旦大学国际问题研究院教授、全球人工智能创新治理中心副主任**蔡翠红**指出，人工智能立法应更加注重治理链条的完整性和运行通畅性，在制度设计中明确分类分级规则、责任主体划分、风险评估程序、监管工具配置、申诉救济渠道和法律责任机制，以提升法律的可操作性和执行力。同时，面对人工智能技术的快速迭代，立法应通过监管沙盒、试点示范和动态调整机制保持必要的开放性与适应性。她还强调，中国人工智能法应具有更强的国际视野，不仅服务于国内治理，也应成为中国参与全球人工智能治理的重要制度接口，增强中国规则的国际可解释性、可对话性和可合作性。



清华大学智库中心助理研究员**刘云**认为，人工智能治理目前有五种可能的立法定位，欧盟的风险法，美国的权利法，职权法，日韩的产业促进法，关注“智能经济新业态”的商业法。他提出应进一步明确《人工智能法》浙大版学者建议稿的基本定位，并思考如何落实国家“十五五规划纲要”对人工智能相关问题的设想。



浙江理工大学数据法治研究院执行院长**郭兵**建议，从关键技术、典型应用等多重维度确定立法的调整对象，避免将调整对象限定为大模型或者生成式人工智能等有限的范围；立法框架应体现出发展与安全并重的原则理念，合理分配发展导向和安全导向的规则比重；条文设计应体现出提取公因式的特点，协调衔接好相关条文与现有的部门规章、国家标准以及伦理规范等制度规范的关系。



香港大学法律学院教授**孙皓琛**认为，条文设定应考虑监管成本，并注重与《民法典》《个人信息保护法》等现有法律相关法条的衔接，明确现有法律的调整范围，避免重复立法。



专题二为关键制度设计与治理工具的落地研讨。本环节由浙江大学光华法学院“百人计划”研究员**吴亦全**主持，他指出，本专题围绕专家建议稿

具体制度落地的四个细分问题进行讨论：一是治理对象与监管主体的区分，二是治理思路的变化，三是治理工具的选择，四是治理尺度的把控。



西南政法大学高等研究院院长**郑志峰**指出，需厘清全文研发、提供、部署、使用等阶段的主体统一性，人工智能分级分类管理中风险定义和人工智能应用场景的关系等问题。



广东财经大学法学院教授、人工智能法研究中心主任**姚志伟**认为，在涉及责任设定的条款上，归责原则与人工智能的分级分类需要进一步讨论，法定连带责任的设定应考察是否存在共同故意的基础。



北京理工大学网络空间国际治理研究基地主任、法学院教授**洪延青**提出，数据质量直接决定人工智能模型的能力边界。因此，在人工智能立法中，数据治理早已超越企业内部合规与技术性安全的狭隘范畴，它同时深刻影响着产业创新活力、国家

数据主权以及全球数字治理格局。这意味着，人工智能立法不能停留在简单的安全底线思维上，而需要对数据治理更加丰富的设计：既要规范数据采集与使用的边界，也要构建激励高质量数据供给的机制；既要保障数据安全与个人权益，也要为产业创新和国际协作留出空间。这是一场从“管住”到“管好”的立法理念跃升。



中国信通院政策与经济研究所主任工程师程莹提出，新形势下人工智能治理存在三个维度的重要转向，一是智能体发展背景下从“注意力经济”向“执行力经济”的治理转向；二是AI情感能力提升背景下从安全治理向人机伦理治理的深化；三是AI感知能力提升下从虚拟空间到现实物理空间的治理转向。本次专家建议稿有效平衡了发展和安全，积极回应了技术跃迁带来的制度诉求。

自由讨论

在专题讨论后，浙江大学数字法治研究院副院长高艳东主持自由讨论环节。广东省人大法制委原

主任委员王波、安徽省人大法制委员会主任委员吴斌、湖北省人大常委会法工委副主任郑文金、浙江省人大常委会法工委副主任吴恩玉对活力和秩序的平衡、全生命周期管理、人工智能产业发展保障、立法定位、立法权限等问题提出了宝贵建议。

闭幕式



浙江大学光华法学院副院长、教授陆青对本次圆桌论坛作总结致谢，并表示，人工智能治理是浙江大学法学院的重要研究领域，本次圆桌论坛对《人工智能法》浙大版学者建议稿的进一步完善指明了三个方向，一是强化问题意识，回应实践需求；二是推进规范体系化；三是增强章节逻辑严密性，为后续完善提供了宝贵经验。本次建议稿的提出恰逢其时，浙大法律人将精益求精，推动《人工智能法》浙大版学者建议稿的公开发表，为我国人工智能立法贡献智慧。

（技术编辑：吕昊然）

数字法评

《建构中国自主的数字法学知识体系》

此处删除了原文脚注，全文请参见《民主与法制》周刊 2026 年第 16/17 期，转载或引用请注明出处。
作者：杨东

自主知识体系是支撑国家发展、保障意识形态安全，筑牢道路自信、理论自信、制度自信、文化自信的根本依托，其核心使命在于立足中国实践、回答时代命题，以原创性理论阐释中国道路、总结中国经验。在此背景下，建构中国自主的数字法学知识体系，是新时代法学创新发展的战略举措，是完善中国哲学社会科学自主知识体系的内在要求，为数字法治建设提供了学理支撑，具有重大的理论、实践与时代价值。

一、建构中国自主的数字法学知识体系的时代意义

建构中国自主的数字法学知识体系，是深入学习贯彻习近平法治思想，推进法治中国建设、服务数字中国战略的重大理论与实践任务。在中国式现代化加速推进、国际竞争日趋激烈、思想文化形态多元的背景下，加快构建中国哲学社会科学自主知识体系，被确立为“十五五”时期的重要任务。

中国自主的数字法学知识体系是中国哲学社会科学自主知识体系的关键板块。作为数字文明时代的法学新形态，中国自主的数字法学知识体系立足中国数字治理实践，突破传统法学研究边界，构建独立的理论范式、学科框架与话语体系，有效填补哲学社会科学在数字治理领域的理论空白。中国自主的数字法学知识体系以学术自主、学科自立、话语自信为核心，与经济学、社会学等学科的自主知识体系形成协同联动，丰富了中国哲学社会科学的整体内涵，全面提升中国哲学社会科学自主知识体系的完整性、前沿性与原创性。

建构中国自主的数字法学知识体系，是推动马克思主义法学理论中国化时代化在数字场域落地

生根的重要路径。习近平总书记在党的二十大报告中提出“在法治轨道上全面建设社会主义现代化国家”，这是对既有法治命题的重大突破、重大创新、重大发展，要求全面推进国家各方面工作法治化，将法治建设贯穿经济、政治、文化、社会、生态文明建设等各领域各方面。在此背景下，中国自主的数字法学知识体系以辩证唯物主义与历史唯物主义为指导，紧扣我国数字经济、数字社会、数字政府建设实践，将马克思主义立场、观点、方法融入数据治理、算法规制、数字金融等理论研究，坚守人民立场与实践品格，摒弃西方资本本位、技术至上的理论局限，实现马克思主义法学理论与数字法治实践的深度融合。同时，中国自主的数字法学知识体系以原创理论、核心范畴与多元方法论为支撑，统筹学科体系、学术体系、话语体系一体化建设，整合各类法学研究方法，规范研究方向、明晰学科边界，有效破解数字法学研究碎片化、跟风化问题，引领学科实现体系化、规范化、自主化发展。

中国自主的数字法学知识体系是培养德才兼备高素质数字法治人才的重要依托。坚持立德树人根本任务，融合价值引领、知识传授与能力培养，构建适配交叉学科的课程与实践体系，强化数字法治理论、算法伦理等专业教育，坚守法治信仰与德法兼修标准，破解传统法学人才与数字时代需求脱节的结构性难题，为国家战略实施输送复合型法治人才。实践赋能层面，中国自主的数字法学知识体系立足超大规模社会治理需求，聚焦数据安全、平台治理、智慧司法、涉外数字法治等重大命题，提供理论方案与制度供给，推动法治与数字技术深度融合，统筹发展和安全、国内法治和涉外法治，全面提升国家治理现代化水平，为中国式法治现代化高质量发展注入强劲数字动能。

二、中国自主的数字法学知识体系的历史逻辑

从历史发展看，中国自主的法学知识体系正在经历由依附走向自主的深刻转型。其演进可从学术史、学科史和话语史三个方面加以把握，完整呈现出从外部移植到本土自觉、从单一范式到多元融合、从传统法理到数字转型的发展轨迹。数字法学

作为法学发展历史进程的最新阶段，承接了百年法学发展的历史积淀，实现了自主知识体系的整体性升华。

在学术史维度，中国法学先后完成四次范式更迭：新中国成立初期形成以继受苏联法律观为主的法学范式，确立政治优先的研究逻辑；改革开放后重构权利义务范畴，实现向规范本位的范式转型；20世纪90年代中后期形成社科法学、实证法学、法教义学三足鼎立的方法论格局；21世纪以来，数字法学新范式应运而生，重塑法律推演逻辑，提出“平台—数据—算法”（PDA）三维治理框架，实现技术实证与法理批判的有机统一。在学科史维度，法学建制同步历经四次变革：新中国成立初期通过院系调整植入苏联学科模式，形成工具理性导向的教育体系；改革开放后推进学科标准化，固化行政化二级学科壁垒；新时代催生社会治理法学、监察法学等本土原创学科；数智时代打破学科边界，推动法学向跨学科、场景化治理转型，形成计算理性主导的交叉学科新格局。在话语史维度，法学话语完成五轮迭代：从新中国成立初期移植苏联法制话语，到改革开放后构建权利本位现代法治话语，加入世界贸易组织后转向国际规则参与型话语，再到人类命运共同体理念引领下的全球治理话语，最终依托全球法学教育合作联盟，实现数字法治话语的全球引领与历史性跨越。

数字法学作为中国自主的法学知识体系最新成果，在学术、学科、话语三大维度实现了全方位突破。学术层面，中国自主的数字法学知识体系推动法学研究方式由传统规范分析向数据驱动、技术融合转型，突破传统因果关系理论桎梏，形成复合型研究范式，以本土原创理论回应全球数字治理难题，坚守法治主体性与规范性。学科层面，中国自主的数字法学知识体系消解部门法壁垒，拓展自主学科版图，创新人才培养与评价体系，完成从制度移植到自主建制的历史跨越。话语层面，中国自主的数字法学知识体系凝练网络空间主权等标识性概念和原创性理论，构建中国特色数字法治话语体系，主动创制全球数字治理规则，全面提升中国法

学的国际话语权和影响力。

由此可见，数字法学并非传统法学的简单外延，而是中国法学自主知识体系的全新标杆。数字法学以学术原创筑牢理论根基，以学科革新完善知识体系，以话语自主彰显理论自信，三者协同发力，推动中国法学迈入自主创新、自主建构、自主引领的全新历史阶段，为中国式法治现代化与人类数字法治文明建设提供坚实的学理支撑。

三、中国自主的数字法学知识体系的理论基础

（一）党的领导是建构中国自主的数字法学知识体系的根本政治保证

建构中国自主的数字法学知识体系必须坚持以习近平新时代中国特色社会主义思想为指导，坚持正确政治方向，牢牢掌握意识形态主导权，深化“两个结合”的学理阐释，以历史唯物主义与辩证唯物主义揭示数字法治发展规律，推动党的创新理论体系化、学理化融入数字法学研究全过程。

建构中国自主的数字法学知识体系是一项跨领域、系统工程，需强化顶层设计与统筹推进，打破学科壁垒与部门分割，搭建跨学科研究平台，实现数字法学理论研究与数字法治实践的深度对接。同时，以思想政治工作凝聚学术共识，引导广大法学研究者坚守学术初心，将个人学术研究与国家战略需求紧密结合，树立“为人民做学问”的价值导向，将党的领导优势转化为推动学术创新的内生动力，为中国自主的数字法学知识体系建设筑牢政治根基与组织保障。

（二）习近平法治思想为建构中国自主的数字法学知识体系提供根本遵循

马克思主义是中国哲学社会科学之魂与旗帜。习近平法治思想作为马克思主义法治理论中国化时代化的最新成果，是建构中国自主的数字法学知识体系的指导思想与理论基石。必须以习近平新时代中国特色社会主义思想为指导，将马克思主义立场、观点、方法贯穿数字法学研究、教学与实践全过程，坚持把马克思主义基本原理同中国数字法治具体实际相结合、同中华优秀传统文化法律文化相结合，立足数字中国实践提炼规律、阐释理论，摒弃

照搬域外理论的依附性思维,提炼具有本土特质的标识性概念和原创性理论,彰显中国法学的主体性与原创性,确保数字法学发展不偏向、不走样,始终服务于党和国家事业发展大局、服务于国家战略、服务于人民根本利益。

2025年11月,中央全面依法治国工作会议在京召开,习近平总书记对全面依法治国工作作出重要指示,强调要“更加注重法治与改革、发展、稳定相协同”。建构中国自主的数字法学知识体系,必须坚守马克思主义法治理论的“魂脉”,传承中华优秀传统文化法律文化的“根脉”,统筹理论传承与前沿拓展,推动知识体系创新、研究方法更新与研究领域延伸;秉持开放包容原则,强调批判吸收世界法治文明有益成果,做到甄别借鉴、有机融合,避免理论水土不服;坚持系统思维,统筹推进学科体系、学术体系、话语体系建设,协同推进教材体系、教学体系、课程体系建设,全面提升中国自主的数字法学知识体系的整体性与完备性;立足时代视野,直面全球数字治理变革,主动回应人类法治共同命题,提出具有世界意义的数字法治理论。

四、建构中国自主的数字法学知识体系的原则方法

建构中国自主的数字法学知识体系,必须严格遵循习近平总书记提出的“以中国为观照、以时代为观照”根本要求,立足中国实际,解决中国问题,形成建构中国自主的数字法学知识体系实践方法。

(一) 坚持走中国特色自主创新道路,增强理论自信与制度自信

扎根中国实践是建构中国自主的数字法学知识体系的根本立足点,需从中华优秀传统文化、红色法治基因与新时代数字经济、数字治理实践中汲取养分,以马克思主义中国化时代化最新成果为指引,阐释中国数字法治的内在逻辑与制度优势。中华优秀传统文化是数字法学的文化根脉,红色法治基因是数字法学应有的精神底色。建构中国自主的数字法学知识体系,必须破除对西方理论与评价标准的路径依赖,清醒认知西方资本主导、防守型治理等模式的内生缺陷。推动中华优秀

传统法律文化创造性转化、创新性发展,将民为邦本、德法共治、无讼和谐等理念融入数字治理。赓续红色法治血脉,坚守人民立场与使命担当,使自主知识体系兼具民族性、时代性与实践性。

中国超大规模数字市场、完备的数字基础设施、世界领先的数字经济成就与独特的数字治理实践,为自主理论创新提供了丰厚土壤。数据要素市场化、智慧司法、监管科技等本土实践,是西方理论无法涵盖的原创性素材。增强理论自信与制度自信,核心在于将中国数字治理经验进行学术化提炼与理论化升华,打造具有中国标识的概念范畴,构建适配本土国情的理论框架,确立中国数字法学的学术主体性。

(二) 紧跟数字革命时代大势,把握数字法学的本质特征与发展规律

人类社会正经历从工业文明向数字文明的历史性跃迁,大数据、人工智能、区块链等技术深刻重塑社会结构与法律关系,传统法学范式面临“代差危机”。数字法学必须主动拥抱技术革命,突破传统法学认知框架,推动人工智能、大数据等前沿技术与法学研究深度融合,实现法学研究范式的革命性升级。建构中国自主的数字法学知识体系,需创新法学研究方法 with 组织模式,聚焦算法伦理、数据安全、平台责任、人工智能伦理等前沿议题,开展前瞻性研究,构建适配数字文明的新理论、新范式、新方法,抢占数字时代知识生产制高点。

要加快建设高水平跨学科协同创新平台,重构人才培养体系,培育兼具法治素养与数字技术能力的复合型人才,破解数字法治人才供给短板。国内高校应主动布局数字法学科建设,深耕智慧司法、法律人工智能等领域研究,以技术创新赋能理论创新,精准把握数字法学的时代内涵与发展规律,推动中国数字法学研究抢占时代先机、引领学术前沿。

例如,当前虚拟货币相关跨境违法犯罪已成为数字时代法治建设面临的新型严峻挑战。部分案件中,不法分子将诈骗犯罪所得转换为虚拟货币后转移至境外,部分虚拟货币被外国政府部门查扣没

收,凸显虚拟货币跨境流通带来的资产追回、司法协作、法律规制等现实难题。对此,需要积极研究跨境资产追回与风险防控路径,加快虚拟货币监管、跨境数字资产治理等领域立法进程,出台更具实操性的政策举措,健全风险防范、执法协作、资产追缴机制,切实维护人民群众财产安全与国家经济金融安全。

(三) 服务人类命运共同体与全球治理,提升国际传播力与知识贡献度

自主知识体系不是封闭自守,而是兼具本土性与开放性的开放包容体系。中国自主的数字法学知识体系既要立足中国实践,批判吸收世界各国法治文明有益成果,避免照搬照抄;又要创新中国话语与叙事体系,面向世界讲好中国数字法治故事,提出回应人类共同难题的中国方案,形成融通中外、引领未来的知识格局。应当立足全球视野,践行全球治理倡议,主动参与全球数字治理体系变革。要以人类命运共同体理念为指引,围绕数字经济、人工智能治理等全球性议题设置学术议程,参与国际规则制定,将中国数字法治方案转化为国际共识。

创新对外话语体系,以国际化叙事讲好中国数字法治故事,提升中国数字法学理论的国际传播力与认可度。强化涉外法治战略布局,构建跨国数字风险防控机制,增强我国在数字金融、数据跨境流动等领域的规则话语权。通过搭建全球学术合作网络、主办高端国际会议、推动学术成果海外传播等方式,将自主知识体系转化为全球制度型公共产品,弥合全球数字治理赤字,彰显中国法治的世界价值,全面提升中国法学的国际话语权和影响力。

五、建构中国自主的数字法学知识体系的实践路径

建构中国自主的数字法学知识体系是涵盖学理创新、人才培养、方法变革、国际传播的全链条系统工程,必须坚持以习近平新时代中国特色社会主义思想为统领,以科学方法为路径,围绕自主理论体系、自主人才培养、自主方法范式、自主话语传播协同突破,形成可落地、可推进、可跃升的完整实现机制。

(一) 自主理论体系: 打造具有标识性的学术

高地

理论体系是自主知识体系的骨架。必须立足中国数字经济与数字治理实践,突破西方单一“权利范式”局限,原创性提出并发展“利益—权利”二元共生理论。以数据要素为核心,兼顾静态权利保障与动态价值分配,构建数据要素四权分置、平台—数据—算法三维治理等原创性理论,形成能够阐释中国实践、回应全球问题的自主学术高地。

(二) 自主人才培养: 塑造新时代复合型数字法治人才队伍

人才是自主知识体系的第一资源。必须打破传统法学单科培养模式,面向数字时代培养德才兼备、懂法懂数,具有家国情怀和国际视野的复合型法治人才。加快发展数字法学、计算法学、涉外数字法治等新兴学科,改革学科评价体系,激励广大学者立足本土、面向前沿,培养一支能够引领数字法学发展、支撑国家战略的高水平人才队伍。

(三) 自主方法范式: 以数字技术赋能法学研究变革

方法论革命是知识体系现代化的关键标志。构建以“平台—数据—算法”为核心的第五科学范式,推动法学研究从逻辑推理向数据驱动转型、从法律条文解释向系统治理转型、从单一学科向交叉融合转型。第五科学范式超越实验科学、理论科学、计算科学、数据发现这四代传统科学范式,形成以人工智能为驱动、以算法为主导、兼具跨域综合性与知识原创性的新型研究范式,为破解数字时代复杂法律议题、实现法学研究范式代际升级提供了全新方法论路径。运用大数据、人工智能等技术开展计算法学、实证法学研究,构建法律专用大语言模型,同时以法治规范技术向善,防范算法垄断与技术异化,形成技术理性与法治理性有机统一的自主方法论体系。

(四) 自主话语传播: 提升国际规则塑造力与全球影响力

自主知识体系最终要走向世界、贡献世界。必须强化国际传播能力,主动设置全球数字治理议题,参与国际规则制定。将中国数字法治实践转化

为国际通用学术话语与制度型公共产品，有效应对跨境数据执法、长臂管辖、数字贸易壁垒等外部挑战，提升我国在数字经济、数据安全、人工智能治理等领域规则中的话语权和影响力，为全球数字治理贡献中国智慧。2025年10月，第四届21世纪世界百所著名大学法学院院长论坛暨人工智能时代全球法治大会在中国人民大学召开，会上中国人民大学法学院、牛津大学法学院等联合发起成立“全球法学教育合作联盟”，共同推动构建跨文明对话新范式、孵化未来法治创新者、探索人类共同法治

框架。这是中国自主的数字法学知识体系话语传播的重要实践。

面向未来，建构中国自主的数字法学知识体系，必须坚持以习近平新时代中国特色社会主义思想为指导，牢牢把握数字文明发展机遇，持续推进理论创新与制度创新，加快形成具有中国特色、中国风格、中国气派的数字法学体系，为中国式现代化和全球数字治理贡献中国智慧、中国方案。

生成式人工智能服务提供者差异化注意义务的规范构造

此处删除了原文脚注，全文请参见《数字法治》2026年第2期，转载或引用请注明出处。

作者：万勇、吴翔

内容提要：生成式人工智能重构了信息生产传播秩序，传统的网络版权注意义务体系因风险生成机制变化、平台角色转型陷入适用失配困境。基于网络服务提供者侵权责任理论，结合技术与风险机理，应构建以风险分层和参与程度为核心的差异化注意义务规范体系。该体系以服务提供者的参与度与控制力为轴设置阶梯式注意义务，依据作品类型的可识别性调节注意义务强度，同时将技术可行性、多元价值衡量作为适用边界，通过分层、分型的注意义务配置，在权利保护、治理效率与表达自由间实现平衡，为生成式人工智能服务提供者注意义务的司法认定与制度建构提供理论支撑。

生成式人工智能正以“数据输入—模型训练—人机交互—内容输出”的链式过程重塑知识生产与信息传播秩序，使信息生成不再遵循传统的“创作者—平台—受众”单向线性结构，呈现出由算法系统与用户指令共同驱动、由平台规则与模型能力共同塑形的复合样态。在此情形下，生成式人工智能服务提供者在内容生成与传播链条中的功能与角色不断发生变化，使传统的网络版权治理结构面临重新阐明的迫切需求。我国围绕生成式人工智能服务提供者的注意义务已出现具有标志性的裁判探索，即区分不同类型的注意义务并提出相应的认定标准，形成区分阶段与场景进行责任认定的裁判思路，为生成式人工智能服务提供者的侵权责任规则的建构提供了重要样本。

现有的生成式人工智能服务提供者注意义务议题讨论，整体上呈现以下面向。其一，规范渊源层面，《生成式人工智能服务管理暂行办法》对生成式人工智能服务提供者提出了合规义务，但侵权责任的成立与边界仍应回到我国《民法典》侵权责任编、《著作权法》及有关司法解释的制度框架中

加以类型化涵摄与评价。其二，规则演化层面，注意到“通知—删除”规则并非仅是版权人与平台利益博弈的偶然产物，而是一种依据信息搜集筛选能力对侵权责任作结构性分配的制度安排，从而要求在新技术语境下重新界定其适用方式，并制定责任再分配方案。其三，规范构造层面，尝试将被动的“通知—删除”规则转向主动的“语料监管与模型审查”，并据此发展出事前语料与模型风险防范、事中用户行为干预与高危提示、事后标识与投诉处置等可操作的分层路径。总体来看，现有研究的共识在于拒绝以结果导向的严格责任简单替代精细化规则治理，转而以注意义务为中心进行责任分配，并倾向在既有的规范中寻找可移植的框架构造。

不同人工智能产品的可解释性与可控制性存在结构差异，单一化的归责原则与构成要件标准无法适配不同风险等级的责任体系构建需求。基于此，本文拟在网络服务提供者侵权责任理论框架基础上，结合生成式人工智能的技术原理与风险机理，以《民法典》过错责任的一般条款与网络侵权的“通知—必要措施”结构为基本骨架，提出以风险分层与参与程度为中心的差异化注意义务的规范构想。在不否定既有责任限制机制之制度价值的前提下，期待回答生成式人工智能服务提供者的注意义务应如何根据参与程度与风险等级分配进行差异化设计的问题，亦即何种条件下应当更积极地预防与管理，何种条件下应当以程序性义务为限并保留创新空间。

一、生成式人工智能服务提供者注意义务的失配显现

（一）生成式人工智能对既有注意义务结构的冲击

1. 生成式人工智能服务形态及其风险生成机制的变化

传统网络版权治理以“用户上传—平台存储/链接—权利人通知—平台响应”为基本链条展开，这种结构的隐含前提是侵权内容通常能够被权利人稳定识别并具体定位为某个文件、某条链接或某

个页面位置，平台据此快速采取移除或断链等措施，即可以较低成本达到治理效果。在这一链条中，侵权的生成与扩散主要由用户触发，网络服务提供者被设定为对内容不作实质选择的外部性载体，其角色特点是外部性、工具性与被动性。由此，注意义务围绕“是否知道或应当知道某一具体侵权事实”展开，其规范重心在于对既成风险的反应，而非对风险生成过程的干预。“通知—删除”规则、“红旗”规则以及“重复侵权人”制度均已内嵌于这一风险生成假设之中。

然而，生成式人工智能的运行模式并非以存储、展示某一特定作品为常态，而是“训练数据—模型参数—输出内容”的连续过程，由此，侵权风险不再只发生在某条具体链接或某次具体上传行为中，而是可能以记忆化输出、风格化复现、诱导式生成等方式在模型输出中反复出现，因而呈现出结构性的规模化风险外溢。作品作为训练材料被统计性摄取，模型通过模式识别与相似性学习获得模仿与再生产能力，从而在推理阶段以极低边际成本生成大量内容。当输出具有可复制性并可被二次输出、再加工时，侵权不再是点状偶发事件，而更接近一种可被持续放大的风险产量，使“个案识别—个案处置”的治理逻辑面临结构性失配。

此外，生成式人工智能的风险并不止于输出端的偶发相似，更在于其商业模式将风险嵌入“规模—速度—迭代”的扩张逻辑之中。生成式人工智能服务提供者通过开放 API、插件与 workflow，把模型能力外溢给下游应用，并借由更频繁的更新、更大范围的功能开放与更高的吞吐换取市场份额与数据回流。一旦训练集、对齐策略或安全过滤存在系统性缺口，侵权便不再呈点状、偶发分布，而更可能在高并发生成、二次输出与再训练回路中被放大为面状扩散。于是，传统注意义务构造所依赖的“权利人可定位、平台可响应、风险可止损”的治理节奏被显著打乱。网络服务提供者删除某一侵权内容，可能落入移动靶困境：输出并非稳定存储于可定位的位置，而是按需生成、对外不可见或仅对单一用户可见。正因为风险不再稳定地附着于可定位

的对象，而更以可复制产量的方式出现，网络服务提供者在治理链条中的地位与功能亦随之发生变化，为后续注意义务的规范重构明确了前提。

2. 平台角色从技术工具向内容生成与分发主体的转变

在传统网络语境中，法律之所以愿意在相当范围内承认网络服务商具有“中立性”，是因为其对内容的选择、编辑与生产能力有限，且对信息的处理多为自动化、非选择性的技术操作。而注意义务的相对克制，正是建立在传统网络服务商缺乏内容控制能力与风险预见能力的认识基础之上。也就是说，传统网络服务商既难以事前识别不法性，亦难以可接受成本对海量信息作语境化判断，故对其责任的判断多围绕“知情—响应”的构造展开。

与此不同，生成式人工智能服务提供者通过模型架构与训练策略直接塑造输出空间，又通过如 API 开放、插件生态、批量生成等产品功能决定输出规模与传播速度，其角色定位已经从“承载传播”转向“组织生产”。技术及其制度化应用并非天然导向公平，反而可能在既有权力结构中进一步累积与放大其优势地位。生成式人工智能服务提供者往往通过提升模型能力、加快迭代更新与拓展应用场景来获取市场竞争优势，相应地，其对风险结构的塑造也就愈加显著。

生成式人工智能服务提供者应被视为区别于传统网络服务商的“新型信息服务主体”，其责任的讨论不能机械地沿用既有的网络平台规则，而应结合其服务形态做结构性调适。正因如此，生成式人工智能服务提供者角色的变化促使注意义务评价从“是否足够谨慎”转向“如何控制风险、控制到何种程度”，使责任判断更接近制度设计问题，而非单次侵权归责问题——因为生成式人工智能服务提供者在事实上既是风险生成的共同塑造者，也是风险流动与扩散的关键控制者。从“Safe harbour”走向“AI harbour”的关键，不在于给生成式人工智能服务提供者贴上服务提供者或内容生产者的标签，而在于重新评估其在风险生成机制中究竟具有何种主动性与控制力，并据此重构

责条件与合规门槛。需要注意的是,生成式人工智能服务提供者的角色从技术工具走向组织者,并不意味着注意义务无限上升,而是强调注意义务的评价不能仅停留于是否履行了合理注意义务,还应进一步考察服务提供者对技术活动强度及其风险扩散后果的控制与分配责任,并将服务规模、功能配置与风险控制能力纳入整体考量之中。

(二)传统注意义务在生成式人工智能服务提供者责任配置中的局限

1.以合理注意为中心的注意义务构造及其适用前提

以合理注意为核心的注意义务构造之所以能够在互联网平台责任中运转,依赖于“可识别的行为、可定位的风险、可归因的损害”这样一个隐含前提。因而,在网络侵权治理结构中,合理注意义务的制度设计通常体现为网络服务提供者在接到合格通知后采取反应措施以实现侵权风险的事后遏制,而不被期待对海量内容作普遍、前置的实质审查。正是在这种风险可对象化、可定位的技术环境中,“通知—必要措施”规则所确立的合理注意义务标准能够在保障网络服务提供者自身高效运转的前提下实现制止知识产权侵权行为与免除相应侵权责任的双重效果,但该规则作为事后介入的被动应对机制,在一定程度上仍不可避免地构成对侵权行为的阶段性放任。

进一步来说,这一构造形成于网络服务提供者主要作为信息传输或存储中介的技术环境,其规范前提在于侵权风险主要源于用户行为,而网络服务提供者通常不对内容作实质选择,仅在明知或应知侵权时负有采取合理措施的义务。与此相对应,“通知—删除”规则将“知道或者应当知道”收束为“红旗式”的具体认知,从而在网络服务提供者缺乏有效控制能力的情形下,避免其被置于过度风险之中。需要强调的是,“知道或者应当知道”并非纯粹主观心理事实,而是一个可被外部因素与规则结构支撑的规范评价,即便《民法典》未就第1197条中的“知道或者应当知道”给出具体的判断公式,学理上仍可通过若干可考量因素来界定网络服务

提供者在特定场景下是否已进入可归责的风险知悉状态。由此,传统合理注意义务被理解为一种以个案侵权识别为中心的事后控制机制,其规范重心并不在于改变风险生成结构本身,而在于对既成侵权的反应是否及时、是否必要且是否适度。

传统过错责任中的合理注意预设风险源于可被识别的具体行为,并可在事后由裁判者参照当时可期待的注意水平进行回溯性评判。概言之,注意义务在结构上依赖“行为可分解、风险可追踪、因果可解释”的条件。以注意水平为中心的注意义务构造,更像是一种面向单次行为的谨慎义务。这要求网络服务提供者在既定业务边界内“行动更小心”,而不是要求其对自身业务形态与技术结构进行持续性的自我约束。因此,在一般情形下,网络服务提供者并不当然负有算法披露与全面过滤的主动审核义务。即便存在事前过滤,通常也被定位为补充性、有限性与自治性的前置安排。当然,传统合理注意义务仍可通过扩大“应知”范围,或强化个案审查来应对新型风险。这一框架在从Web1.0到Web2.0的“存储—传播”链条中尚能自治,是因为风险的放大节点主要在于用户上传与外部传播,此时网络服务提供者被设定为被动反应者。

生成式人工智能服务提供者的风险生成机制恰恰打破了上述适用前提,因为其服务机理把风险从内容搬运推向内容生成。生成式人工智能的输出建立在规模训练、参数微调以及面向海量用户的即时生成之上,侵权风险往往并非由单次输出行为独立触发,而是由平台在模型训练、调用接口、分发机制与商业化优化的组合活动中被结构性放大。正因如此,合理注意义务难以继续以单点行为为抓手实现可归责的边界划定。传统“中立免责—明知担责”二分责任模型难以解释生成式人工智能产业链中多主体、分阶段的风险生成结构。生成式人工智能服务提供者在定位上更像是内容共创者与风险把关者,需要在数据集构建、模型训练、部署分发等环节配置分层化注意义务。换言之,风险并非只在“具体侵权内容出现之后”才可识别,相当一部

分风险在训练数据来源、输出约束、产品化入口设计时就已被制度性决定。若仍沿用以“通知后反应”为中心的合理注意义务构造，既会造成风险治理的结构性滞后，也会使“知道或者应当知道”的规范评价失去可操作性。

2. 在规模化、商业化与快速迭代场景下的规范失灵

以合理注意为核心的注意义务构造在规模化、商业化与快速迭代的互联网平台场景中容易出现规范失灵问题。生成式人工智能在规模化参数记忆与检索增强等机制下，特定条件下的可复现性与可预测性会被放大。就其技术机理而言，模型训练对海量语料的统计性摄取与输出阶段的概率性拼接共同构成风险的两端：前者决定了潜在侵权素材的进入阈值，后者则决定了既有表达被再现、近似或记忆化复现的外显方式。二者的叠加使得侵权不再主要表现为可被识别定位的单一作品副本，而更像是由系统性活动持续产出的风险流。纯粹的事后救济在现代风险情境下具有结构性局限：一方面，损害恢复到未发生状态往往只是理论假设；另一方面，救济的延迟与高成本会导致救济不足与低效率。因此，风险生成机制之变已足以动摇以往以可定位对象为中心的责任配置前提。

从风险社会的视角看，传统侵权制度的重心偏向个体正义与事后救济，在控制集体性风险方面存在天然不足；相应地，集体性风险治理更强调通过行为标准与强制信息披露来实现预防性控制。正是在这一结构下，如果仍把合理注意理解为一种静态、单点、事后的标准，将在制度上诱发两种看似对立但同样失灵的效果。具体来说，一方面，若按静态标准从宽理解生成式人工智能服务提供者的注意义务，其能够凭借“技术中立”对抗过错认定，导致系统性风险外溢而无法被有效吸收；另一方面，若按结果导向从严理解生成式人工智能服务提供者的注意义务，裁判将倾向把“能做”推导为“应做”，进而以事后归责逻辑反向迫使其承担事实上的全面审查义务。这两种路径表面上分属规范的宽严两端，本质上却具有同一缺陷——二者均未把生

成式人工智能服务提供者的风险生成活动拆解为可治理的节点，也就无法在“可行性、误伤率、外部性”之间形成可重复适用的裁判结构。

由此可见，规范失灵并非仅指传统注意义务构造不重要，而在于这一构造的功能实现倚重于事后反应与行为谨慎的治理结构，已不足以回应生成式人工智能服务提供者通过产品设计与商业策略持续生成风险的现实。因而，生成式人工智能语境下的注意义务的重构任务，应当将其还原为侵权责任体系中的归责要素，用以回答“风险应当由谁承担、在何种条件下承担、以何种强度承担”的问题。换句话说，生成式人工智能服务提供者责任的核心冲突，正在从“是否中立”转向“如何把风险产量纳入可被评价的义务变量”，据此为后续的阶梯化义务结构奠定规范基础。

二、生成式人工智能服务提供者注意义务的法理基础

（一）注意义务作为侵权责任配置的规范基础

1. 注意义务作为侵权责任成立要素的规范意涵

从侵权制度体系演进看，风险社会为该制度提出的核心任务已不止于事后填补，而更强调预防损害与风险分配。过错责任原则是侵权制度中的基本归责原则，其核心要义为：若一方未采取经济上合理的预防措施而导致损害发生，则需承担法律责任。在过错责任的法经济学表述中，所谓“注意义务”可理解为法律所要求的预防性行为强度，目标是在给定约束下实现社会福利最大化。其功能不在于评价行为人的主观善恶，而在于判断行为是否构成法律上可归责的风险生成。由此，注意义务的核心规范意义在于通过设定可预期的行为标准，将损害结果与行为人的风险控制建立规范上的归属关系。

从制度功能上看，注意义务不仅是过错责任的构成要件，更承担着控制责任边界的制度功能。即便损害具有可预见性，责任也不当然成立，而是要通过“是否存在对特定受害人群体的义务”来避免过错责任的无限扩张。就此而言，注意义务要件在

现代侵权制度中越来越呈现出“关闸效应”：通过限定可保护利益与可被期待的注意范围，把可预见风险转化为可归责风险。

注意义务之所以被置于侵权责任构成的门槛位置，并不在于其具有抽象的道德劝诫性，而在于它以可操作的外在标准，将过错从主观心理事实转译为可评价的行为偏离。这种转译的功能意义在于，它允许法官在不必穷尽行为人的内心动机与真实认识的情况下，仍能依循是否可预见、可避免、可负担的逻辑路径，对行为作出规范评价，并据此将损害风险在社会成员之间进行可解释的分配。进一步而言，注意义务同时具有事前行为期待与事后责任判断的双重维度：在行为发生之前，注意义务表现为法律对风险控制行为的规范期待；在损害发生之后，注意义务则转化为判断行为人是否违反该期待的责任标准。

应当强调的是，注意义务虽具有风险预防功能，但并不意味着实现风险的绝对消除，而是在预防成本与预期损害之间进行权衡，追求社会效益最大化意义上的有效预防，使侵权责任得以在矫正正义与制度效率之间维持最低限度的可接受性。一方面，矫正正义强调侵权制度所处理的是加害人与受害人之间的双边关系。责任成立需要将损害与特定行为人的不当风险施加相连接，而注意义务正是完成这种连接的规范纽带——它要求被告对其制造并外溢的风险作出与其行为性质相称的回应。另一方面，制度效率强调注意义务通过“成本—效益”校准机制避免过度预防。当预防成本远超侵权损失的可避免部分时，将风险强行内化于行为人并不当然正当。换言之，注意义务之合理是责任可归责的条件之一。没有这一条件，侵权责任要么滑向严格责任，要么陷入不可操作的主观探询。

基于上述认识，注意义务的规范构造必须围绕风险控制能力与行为期待之间的对应关系展开，即行为人仅在其具有合理控制能力的范围内承担风险防范义务。正是这种能力与义务之间的对应关系，使注意义务成为侵权责任配置中的核心中介机制。当生成式人工智能以持续输出、即时分发、规

模扩张的方式运转时，传统过错责任的优势在于仍可沿用风险、预防、成本的结构解释。注意义务判断本身就是在划定何种风险可被纳入过错责任的归责范围，生成式人工智能服务提供者并非因制造了内容就当然担责，而是因其在算法设计、数据训练、部署方式与技术迭代中系统性地产生并放大了可预见的侵权风险，从而进入注意义务评价的射程。因此，注意义务的规范内涵需要在保持过错责任基本结构的前提下进行结构性调整，而该调整的关键即在于重新界定风险控制能力与责任配置之间的对应关系。

2. 知识产权法中注意义务功能扩展的体系依据

知识产权法语境中的注意义务并不止于一般侵权制度上的谨慎要求，而更承担着对信息“生产—传播”结构进行责任性塑形的制度功能。在传统侵权制度语境中，注意义务常以避免对他人法益造成不合理风险为中心，用以衡量行为人是否未尽合理谨慎义务。在网络服务提供者版权侵权场景中，风险是否合理取决于损害大小、权利识别成本、许可交易成本、合理使用的不确定性，以及平台措施可能造成的过度删除或过滤。因此，版权法中的注意义务天然带有强烈的制度协调色彩，制度的设计既要回应权利保护与信息流通的双重目标，亦要对网络服务提供者技术措施的误伤后果保持敏感。若仍将注意义务仅定位为个案层面的谨慎标准，便会出现侵权风险结构已变而注意义务结构未变的失配。这正是知识产权法将注意义务推向更制度化的原因，它不再只关注是否足够谨慎，而开始关注网络服务提供者应当如何组织其业务形态以抑制系统性侵权风险。

知识产权法之所以能够承载这种扩展，关键在于其责任体系本就包含以治理义务换取条件性责任限制的结构安排。“避风港”规则的核心是以事先履行必要措施为条件，换取责任限制的确定后果。具言之，权利人承担初步识别与指向性通知成本，网络服务提供者承担接到合格通知后的反应义务，并通过是否采取必要措施来决定其责任边界。

当网络服务提供者的地位从中立工具转向内容组织者或分发者，注意义务不应停留在事后响应，而应当进入以技术能力为界定尺度的再分配逻辑。一方面，网络服务提供者对作品传播的参与度与控制力改变了其在侵权链条中的角色定位，从而改变了注意义务的强度与形态；另一方面，不同作品类型的可识别性与合理使用的不确定性改变了预防措施的误伤概率与社会成本，从而反向限定注意义务的上限。这也解释了为什么知识产权法中的注意义务更像一种结构性责任分配规则——它不仅是评价有没有尽到谨慎义务，更是在配置谁更适合承担哪一类预防义务，并把公共利益作为义务边界的内在变量。

反证地看，如果将注意义务一般化为普遍前置审查，其在知识产权法内部反而缺乏可持续的正当性。一方面，版权侵权的识别并非总是“红旗式”的显而易见，诸多情形需要进入合理使用、权利归属与许可范围等领域判断。过度依赖自动化通知并缺乏语境化审查，会导致覆盖过宽而抑制合法表达与创新活动，从而在制度上引发“寒蝉效应”，这意味着以保护版权为名反而侵蚀版权法本应维护的鼓励表达与创新生态。另一方面，强制性过滤往往以系统性预防为代价，把难以算法化的法律判断外包给技术系统，这会显著牺牲网络运行效率并增加社会成本，可能对初创型和中小型网络服务提供者造成沉重负担，尤其是在过滤技术的许可费用和系统维护成本上，进而对市场竞争与创新形成结构性挤压。总的来说，把网络服务提供者的责任完全推向严格或推向豁免都可能失衡，较可行的做法是通过制度设计将其风险控制能力与所应承担的注意义务类型对应起来，并在涉及表达与创新的场景下谨慎配置阻断或过滤的边界。

（二）生成式人工智能服务提供者控制能力作为风险责任配置的正当依据

1. 最小成本规避者理论与风险内化的责任逻辑

从制度效率视角看，法律责任在原则上应配置给最有能力以最低成本预防、限制或消除损害的一

方。在风险治理的归责选择上，更具解释力的路径是遵循“最小成本避免者”思路，其判断核心在于某一主体是否具备对相关风险的认知以及对该风险的控制程度，且此种认知与控制足以预防由此产生的损害。简言之，由具备信息优势与治理工具、能够以更低成本识别并降低风险的一方承担更高阶的注意义务，才可能实现“成本—效益”的最优匹配。在版权侵权的间接责任结构中，责任配置是与预防成本、损害规模、信息条件相绑定的，在直接侵权人分散、权利识别困难、执行成本高的情形下，若中间主体能以较低成本改变侵权发生概率或损害扩散路径，则使其承担一定责任的规则，更可能实现更优的预防性投入效果。将责任置于“最小成本避免者”的制度意义，不在于道德谴责，而在于通过风险内化机制对未来的“最小成本避免者”形成可预期的谨慎激励，从而将风险治理转化为可执行的行为规范。这一路径的逻辑是谁能以更低成本减少更多风险，而非谁更靠近内容就当然应审查内容。

侵权制度并非只评价一次性的谨慎程度，而是评价持续性活动所制造的可预见风险是否通过制度化方式压降到合理区间。生成式人工智能的动态学习能力决定了风险会随运行与交互而变化，因此生成式人工智能服务提供者的注意义务应转向以“行为合理预测性”为导向的持续监控与动态纠偏要求。在网络服务提供者的版权侵权行为中，适当的分析单位并非某一次的具体侵权，而是将其作为整体的治理结构。具体来说，网络服务提供者在每一次侵权中的参与样态往往高度同质，在这种情况下，责任认定的关键问题转化为其应当被要求以何种强度、何种方式去区分内容并抑制侵权。正是在此意义上，生成式人工智能服务提供者的控制能力成为责任配置强化的规范连接点——其能够通过训练数据治理、输出机制约束与分发链路设计，在总体上显著降低侵权风险的产量与扩散速度，从而以更低的社会成本实现更有效的侵权风险预防。相较之下，如果将风险治理完全交由权利人通过逐案发现、逐案通知的模式实现，这会把制度运行成本

推向最不具规模经济的一端，从而导致救济失灵与过度诉讼的双重外部性。尤其是在生成式人工智能服务场景中，侵权风险往往并非来自单次输出的主观恶意，而源于模型在统计结构上对表达特征的可迁移复现能力。因此，注意义务随活动增强而呈阶梯递增，实质上是侵权制度的规范回应——当风险由规模化活动持续生产时，合理注意的对象应当从“个案谨慎”转向“活动强度”，以确保预防成本与风险产量之间存在可核验的比例关系。

然而，责任强化并不能以生成式人工智能服务提供者有能力采取措施作为充分理由，而必须围绕可治理风险点展开，并通过“可行性、误伤率、外部性”的三重检验，方能转化为合理且可操作的注意义务。首先，即便生成式人工智能服务提供者在技术层面上更有能力采取措施，也不能据此推导出其应当承担最大化义务，因为注意义务的社会成本不仅包括合规投入，还包括由过度删除或过滤引发的表达损失与创新抑制等负外部性。其次，权利人通常更加充分地掌握权属信息、比对素材或授权链条，在此情况下要求权利人提供结构化通知、指纹库或可机器读取的权利声明，往往比要求生成式人工智能服务提供者在不确定性极高的场景中预测侵权风险更节省总成本。最后，即便在生成式人工智能服务提供者是“最小成本避免者”的场景中，注意义务仍应被限定为可行且不过度误伤的“技术一程序”组合，而不应一般化为对全部生成内容的前置过滤。由此可见，风险内化真正要实现的是把生成式人工智能服务提供者“选择风险—获得收益”的商业逻辑，与其承担相称治理成本的责任逻辑相扣合。因而，更合乎法理的路径不是全面审查，而是分层、分型、分场景的注意义务阶梯。

2. 控制能力优势与类型化归责视野下的责任逻辑

在生成式人工智能责任结构中，生成式人工智能服务提供者控制力构成责任配置的规范连接点，根本原因在于侵权风险在其技术结构中具有可归责的生成与外溢机制。从侵权制度理论看，风险归责的正当性取决于风险是否可被归属于特定主体

的行为活动。生成式人工智能服务提供者凭借其其对数据、算法与模型的实质性控制，成为风险治理链条中最具能力预防侵权的主体，因而需承担相应的注意义务。是否要求生成式人工智能服务提供者采取更强的预防措施，关键在于侵权风险大小、潜在损害与预防成本之间的比较：只有当预防成本在整体上低于预期损害时，提高注意义务与引入更强的技术措施才具有合理性。进一步而言，生成式人工智能服务提供者之所以被视为责任强化的优先对象，还在于其具备结构性的信息优势与规则优势。生成式人工智能服务提供者通过集中化的信息与规则设定对内容的生成、传播与访问进行深度控制，这使其能够有效识别、筛选和调整用户生成的内容，还能够通过算法规则、用户协议和内容管理规范直接影响侵权内容的传播速度和触达范围。正因如此，侵权风险与平台所建构的技术环境并非仅存在事实关联，而更可能形成规范意义上的因果联系。当生成式人工智能服务提供者处于信息与规则网络的关键节点，其通过设计与治理决策能够重构信息流、塑造社会理解，并进一步影响人们据以行动的理由时，即应被纳入与其“节点性影响”相匹配的规范责任框架。因此，生成式人工智能服务提供者的责任并非限于消极的监控义务，而是更应体现为对可治理风险点的主动管理、预防与控制。

然而，控制力与信息优势只能证成应当在生成式人工智能服务提供者侧配置更高的注意义务，这并不当然推导出其应承担全面的审查义务。原因在于，著作权法中的合理使用与实质性相似判断具有高度语境性与不确定性。若将注意义务直接等同于绝对阻断侵权的结果导向要求，极易诱发过度过滤并压缩合理使用规则的适用空间，从而对创新激励与表达自由造成不利影响。《民法典》侵权责任篇所要求的并非最大限度风险消除，而是在合理成本内防止可避免的风险。因此，法律不应把生成式人工智能服务提供者的注意义务简化为过滤或不过滤的二分命题，而应当根据其控制能力、信息识别技术的成熟度以及合理使用的法律界限来履行相应的责任。生成式人工智能服务提供者可以在识别

度高、重复性强、误伤风险低的对象上采取更强的预防措施，但同时必须通过理由说明、申诉复核、反通知与恢复等机制，把风险预防从结果主义拉回过程治理，从而确保在社会层面实现更加公平的版权保护。

侵权制度本身也并不要求在所有情境中适用同一套“风险一效用”的通用注意标准，而是允许注意义务标准随情境与风险结构而分化。注意义务在本质上属于抽象性规范范畴，需依托填充规则将该抽象义务转化为对具体行为的类型化认定标准，使其发挥应有的规范与指引功能。网络服务提供者并非同质主体，内容生成类中介与基础设施类中介存在本质区别，类型化的框架可揭示不同网络服务提供者面对同一规范标准时的结构性差异。网络服务提供者控制能力在侵权制度中的最终意义，在于推动责任配置由抽象能力事实转化为具体的类型化注意义务结构。责任配置并非直接依据能力大小作出，而应通过制度化规则将能力差异转化为不同层级的注意义务。从网络版权治理的制度建构逻辑出发，类型化注意义务具有三重显著优势：其一，类型化规则可增强责任认定的可预见性，为行为主体明确风险防控的具体要求；其二，类型化结构能够约束个案裁量权的过度扩张，保障责任边界的稳定性；其三，类型化工具可针对不同风险场景实现差异化规制，进而提升制度运行效率。需要明确的是，类型化注意义务并非意味着责任的绝对固化，而是需考虑具体技术条件与风险结构的动态变化。基于此，生成式人工智能服务提供者控制能力的规范转化方向在于构建以风险可治理性为核心的阶梯式注意义务体系。

（三）作品类型差异作为注意义务内容的调节依据

1. 不同作品类型的可识别性与侵权风险结构

作品类型差异是决定注意义务内容能否被合理设定与有效履行的结构性因素，注意义务的高低会随作品类型与风险显著性而变化。

一方面，不同作品类型依托于不同的作品语言

与符号系统，这使得同一治理手段在识别精度、误伤概率与社会成本上呈现结构性差异，从而决定了生成式人工智能服务提供者需要采取的注意义务的强度与措施。传统上，知识产权法对于不同作品的保护要求并非一成不变，而是考虑到了作品的性质、用途和受众范围有所差异。作品类型之所以具有重要意义，在于不同类型往往对应不同的特别规则与利益结构。生成式人工智能服务覆盖从大语言模型到图像与音频生成器等多种模态，而不同模态的作品在可被机器识别、可被规则化治理的程度上差异显著。例如，图像、音频等作品往往具有较为稳定的感知特征，易于进入指纹比对、相似度阈值、拦截过滤的治理链条；与之相比，文本类作品尤其是叙事性、学术性或功能性文本，其生成结果高度依赖语境、用途与表达方式，侵权与合理使用之间的界限本身即具有高度开放性。正因如此，生成式人工智能服务提供者的责任不能仅依赖简单的“通知一删除”规则，而应以作品类型为轴对注意义务进行分层，把注意义务锚定在可实施的风险控制之上。如此一来，在能以低误伤成本识别侵权内容的领域，要求服务提供者采取更积极的预防措施，也就更具正当性。

另一方面，作品类型差异还会决定侵权发生的方式，从而改变注意义务应瞄准的风险点。以视觉类作品为例，侵权常以“高度近似再现”或“可识别角色/元素的再生成”呈现，属于相对典型的输出端风险。在此类情形中，侵权风险通常集中表现于生成结果的输出阶段，当生成内容对既有作品的独创性表达形成可识别的再现或高度近似时，侵权结果即可能发生。这种风险结构意味着，生成式人工智能服务提供者的注意义务不必一般化为对一切输出承担保证，但可以具体化为对高相似度、重复触发、明显指向特定作品或角色的生成请求与结果，建立更强的拦截与处置机制。

2. 不同作品类型的合理使用空间与注意义务边界

合理使用的不确定性并非制度缺陷，而是立法者为兼顾创新、表达与社会公共利益所保留的弹性

空间。生成式人工智能技术既引发训练阶段对作品使用的合法性争议，也引发输出阶段对保护范围与合理使用边界的再争论。在这种情况下，生成式人工智能服务提供者不仅要防范侵权行为，还要根据合理使用的标准，确保其在技术可行的情况下界定使用范围，从而避免对内容创作自由的过度抑制。生成式人工智能服务提供者在履行注意义务时应保持审慎与克制，避免以过度前置审查方式压缩合理使用空间、对信息流通与表达自由造成不当抑制。合理使用不确定性还要求其“输出替代效应”保持敏感，通过结合作品类型与合理使用的不确定性，生成式人工智能服务提供者的注意义务不仅是一个技术问题，更涉及法律上的平衡和公共政策的考量。

即便生成式人工智能服务提供者试图通过更有针对性的过滤来降低侵权风险，规范上也会随之提出更高要求：过滤规则会被期待把“合理使用”一并纳入考量，而这恰恰是生成式人工智能场景最不具可行性的部分。由于人工智能的算法决策受既定规则与条件触发结构所限，往往只能进行“条件式计算”，难以像人类一样在具体情境中作出平衡判断或例外处理。在文本生成、风格模仿与信息重组等场景中，侵权判断往往依赖个案化的价值衡量，而非技术可自动完成的事实识别。由于生成式人工智能服务无法完全区分有害内容与无害内容，且无法内部化无害内容所带来的全部社会收益，责任追究的威慑可能导致生成式人工智能服务提供者采取过度审查措施，即在移除有害内容的同时，也删除过多本属无害的内容，这种结果并不符合侵权制度所追求的最小成本规避目标。因此，生成式人工智能服务提供者的注意义务，应当围绕“风险控制能力”而非“结果正确性”展开，否则将不可避免地滑向事实上的严格责任。

三、生成式人工智能服务提供者注意义务的规范构造

（一）基于参与度与控制力的注意义务分层

1. 平台参与内容生成与传播程度的类型化区

分

“通知—删除”规则的确立，源于立法者对网络技术架构及功能边界的认知，是对网络侵权责任作出的特殊制度安排。该规则以技术工具性认知为规范基础，以推动互联网产业发展为价值导向，并在责任平衡与私权自治理念的共同塑造下，构建起以网络服务提供者技术能力与内容参与度为核心的责任认定体系。生成式人工智能语境下的注意义务分层，应以服务提供者在风险生成链条中的参与度与控制力为规范核心。常见的规范反证是，既然输出结果可能侵权，生成式人工智能服务提供者制造了风险，就应被一律视为内容提供者并承担同一强度的前置审查义务。然而，该结论的成立依赖于一个隐藏前提——生成式人工智能服务提供者在所有的服务形态中，都对输出内容具有同等的预见性与可控制性。

这一结论不可行的原因在于，生成式人工智能输出的作品高度依赖上下文语境，尤其在文字、视听作品的二次创作场景中，侵权与合理使用的区分往往需要结合目的、比例、替代性市场影响等因素进行综合判断。系统缺乏类似人类的注意力机制与记忆提取机制，即便储存了大量知识，也未必能在特定情境中根据当下语境信息，快速、灵活地调取并组织相关知识来完成判断。生成式人工智能服务提供者的控制力并非在所有场景都相同，因此对其施加的注意义务不宜采取“一刀切”的方式。当生成式人工智能服务提供者对信息传播的实际控制能力较弱时，责任配置更应保持克制，并主要通过竞争法等外部规范来约束其市场行为；当生成式人工智能服务提供者能够通过产品架构与规则设计对信息的可见性、传播路径与受众触达施加决定性影响，甚至在一定程度上“塑造”内容的呈现方式与传播结果时，要求其承担更高水平的注意义务具有规范正当性。

更稳妥的制度路径是以生成式人工智能服务提供者是否具有对风险链条的可塑性与可阻断性为标准，分别识别其在训练数据治理、输出控制与分发扩散中的位置，再据此确定注意义务强度。生

成式人工智能服务提供者参与度的类型化区分，首先应在“训练—生成—分发”的链条上识别其是否从工具性节点转化为内容生态的组织者，能否对风险的形成与扩散具有可塑性与可阻断性。从治理视角看，生成式人工智能服务提供者的规模、渗透性及社会影响力将直接影响政策制定者的规制方式。生成式人工智能服务提供者至少可能处于以下三类位置：其一，仅提供通用模型或 API，并将内容决策留给调用者。其更接近工具性技术服务提供者，对下游应用与终端用户行为的控制力相对较弱，因此在内容审核上宜配置较弱的注意义务，并可在一定范围内主张技术中立的免责抗辩。其二，在产品层面对提示词、模板、插件、检索增强与安全策略进行组合，并导向特定用途。此时，生成式人工智能服务提供者已在生成过程中形成实质介入。其三，生成式人工智能服务提供者不仅生成内容，而且以推荐、榜单、话题、商业投放等方式推动传播与变现，对内容分发具有结构性支配力。对三类平台若适用同一注意义务强度，实际上是忽略了“参与度差异→控制力差异→最小成本预防位置差异”的责任配置链条。

参与度区分的关键，不在于生成式人工智能服务提供者是否接近内容生产者的角色，而在于侵权风险究竟是在何处被生产、被放大、被商业化。生成式人工智能服务提供者的运营活动通常包含不同程度的人为干预要素，而干预程度的差异恰是界定其知识产权注意义务高低的关键依据。当生成式人工智能服务提供者在产品机制上引导生成，例如，以预设 workflow、风格化按钮、热门提示词推荐、检索增强等方式降低侵权生成的门槛，其角色已不再等同于中立技术通道；当生成式人工智能服务提供者在分发机制上放大输出，例如，将高互动内容纳入推荐与变现体系，其控制力就表现为对风险外溢的结构性影响。此时，要求生成式人工智能服务提供者承担较高层级的注意义务，并非将其等同于内容生产者，而是将负担配置给最易于承担并能有效控制风险的一方。

因此，类型化判断的出发点，不应在于生成结

果最终是否构成侵权，而应在于生成式人工智能服务提供者在内容生成与传播链条中的参与程度是否达到可归责的强度。这也解释了为何不应将注意义务一般化为前置审查。前置审查假定平台能在单次输出层面作出高质量的侵权或合理使用的判断，但在混合模式场景中，这一假定往往并不成立。过滤技术无法在语境层面稳定理解表达，因此其治理代价往往体现为“网中海豚”，大量本应受保护的合法表达在过滤中被连带屏蔽。与其把生成式人工智能服务提供者推向前置审查者的不可能角色，不如以参与度为基准引入差异化、可执行的类型化工具。

2. 注意义务随风险生成活动强度递增的阶梯结构

在确立“参与度—控制力”的分层轴之后，进一步的问题在于：不同参与强度应当对应何种注意义务结构。梯度型、递增式的注意义务正是对此问题的制度回应，其核心在于将注意义务从单次谨慎要求转化为与风险生成活动强度相耦合的递进性责任结构。阶梯义务的关键不在于更严而在于更准，其配置必须满足比例原则与法律确定性：注意义务不应通过含糊不清或不切实际的要求压缩网络服务提供者的经营自由，尤其应避免对中小网络服务提供者形成进入壁垒。因而，注意义务强度需要与服务提供者的技术与运营能力相适配，从而维持侵权制度的可归责边界。这种注意义务随风险活动递增的阶梯结构，使风险控制成本与风险产量之间形成合理稳妥的比例关系，从而避免将注意义务异化为事实上的严格责任。从法律技术角度看，“避风港”规则正是为了避免一般审查义务所导致的过度责任扩张，才通过设定有限且附条件的合规路径，在风险治理与创新激励之间实现制度平衡。

在具体结构上，阶梯式注意义务可以依据风险生成活动强度递增配置为三个层级：第一阶梯聚焦于数据训练与模型治理的最低合规标准，例如，训练数据来源与权利保留机制的透明化、基本的内容安全策略与重复侵权处置流程。其制度逻辑在于将源头治理转化为最低限度的注意义务门槛，使生成

式人工智能服务提供者对风险生成链条的基础性可控点承担责任。第二阶梯应当引入对可预见高风险用途的结构性约束,例如,对特定功能开放、插件生态、二次输出接口设定风控门槛。其制度逻辑在于通过对高风险使用路径的结构性干预,提高侵权性生成的进入成本与规模扩张门槛,从而在风险形成阶段实现前置控制。第三阶梯则应当要求生成式人工智能服务提供者对风险放大机制承担更强的控制义务,例如,对高传播通道设置降权、延迟发布、可追踪标识与反复侵权账户治理。其制度逻辑在于通过控制风险外溢与传播扩散的关键节点,降低侵权行为的社会外部性,并强化生成式人工智能服务提供者对风险放大效应的可归责性约束。总而言之,阶梯递增的依据在于生成式人工智能服务提供者在风险生成结构中的作用强度,而非抽象的身份标签。

(二) 基于生成式作品类型的注意义务调节

1. 高可识别性作品对应较高注意义务的配置

生成式人工智能所涉及的作品类型并非完全同质,必须区分“技术上可稳定识别的侵权风险特征”与“高度依赖具体语境的侵权判断”,生成式人工智能服务提供者之注意义务应当随作品类型的可识别性与风险显著性而分层调节。部分作品存在稳定、可重复提取的表达特征,侵权风险特征可被技术手段较为精确地捕捉,因而此类风险呈现出更强的可预期性与可重复性,这为强化生成式人工智能服务提供者注意义务提供了规范正当性基础。

对于高度标准化且具有强烈辨识特征的作品类型,例如,音乐作品、图像作品,生成式人工智能服务提供者对其生成与输出过程中的侵权风险具有较高的控制力,对这些作品实施的侵权行为已经可以通过现有的图像匹配等自动化技术手段进行精准识别。

进一步可以发现,高可识别性作品的风险并不总是由个别用户的偶发性行为触发,而常与生成式人工智能服务提供者在训练数据选择、模型调校、功能入口与输出机制上的安排直接相关。当特定提示语与侵权输出之间呈现出可重复、稳定的对应关

系时,生成式人工智能服务提供者的“必要措施”不应仅限于外部过滤或屏蔽,还应延伸到提示语规则优化和相关素材检索隔离等系统层面的治理措施,以降低同类侵权结果再次生成的概率。实践中,“广州奥特曼案”即以生成内容与权利作品之间的高度指认关系为背景,要求服务提供者采取“关键词过滤”等措施,且过滤强度应达到即使用户正常输入与该IP相关的提示词,也无法生成与权利作品实质性相似的图像。

基于高可识别性作品侵权风险呈现出可识别性强与可治理性明确的特征,更为合理且合比例的制度选择,应是相应提高生成式人工智能服务提供者的注意义务强度,并通过阶段化治理机制实现风险控制。其一,在训练语料管理环节,生成式人工智能服务提供者应当强化训练数据来源管理、权利保留与重复侵权数据的处置,降低高风险素材进入模型的概率;其二,在生成与交付环节,可对明显指向特定作品或角色的提示语与高相似度输出设置更严格的拦截阈值,并建立“重复触发一再上传一再生成”的联动治理;其三,在事后处置环节,应当完善投诉响应、可核验处置记录与针对重复侵权人的梯次处置。在必要时,及时更新安全规则、优化提示语与输出拦截策略,以降低同类侵权结果的再现概率。由此,强化注意义务的关键是将其锚定在风险高、可识别、可治理的节点上,实现对高可识别性作品侵权风险的系统性治理。

2. 高不确定性作品对应限缩注意义务的适用

相较于图像、音频等高可识别性作品,文字作品往往有更灵活的解释空间与更复杂的权利限制规则,其侵权风险难以抽象为可由固定规则稳定识别的模式。因此,在此类高不确定性作品场景中,生成式人工智能服务提供者注意义务的规范构造应当以对外部风险的可预见性与可控性为中心,而不宜将合理使用、表达自由等高度价值负载的判断,转化为平台的前置实体审查任务。相应的注意义务亦需受到限缩,将注意义务的内容转向过程治理,以维持过错责任框架的可运行性。

可能出现的规范反证是,既然生成式人工智能

已具备语义理解与相似度计算能力，是否应当同样强化其对文字作品的注意义务？实际上，这一推论未能考虑著作权法内部关于合理使用与表达自由的基本结构。就语境依赖型的表达而言，侵权判断的关键通常并非“是否存在某种相似”，而在于相似部分是否构成受保护表达的实质部分，以及该使用是否因目的、方式、必要性与市场替代效应而落入合理使用范围。侵权判断并非仅针对内容本身作抽象定性，而是更依赖使用语境、关系结构与具体呈现方式；若无视具体情境的自动化机制，可能将本不必然构成侵权的内容识别为侵权，进而在网络内容生态系统中引发“寒蝉效应”，从而反向压缩合法创作与信息流通空间。因此，在高不确定性作品场景中，要求生成式人工智能服务提供者承担实质性审查义务，不仅在技术上难以准确实现，更会在规范上造成责任配置失衡。

从侵权制度与知识产权法的交叉协调角度看，生成式人工智能服务提供者注意义务的扩张必须以不破坏著作权法既有的限制与例外结构为边界，否则将导致著作权法功能的异化。若对高不确定性作品同样施加较高的注意义务，常常并不能带来更有效的权利保护，反而会诱发服务提供者基于风险规避动机采取“一刀切”治理，进而损害用户的正当创作空间。换言之，当法律以高度不确定的实体判断作为触发注意义务的条件时，生成式人工智能服务提供者理性的策略选择往往不是更精准而是更保守。这一结果是将本应由权利人主张、由司法衡量的争议前置为生成式人工智能服务提供者的私力审查，并使著作权法的利益平衡结构被生成式人工智能服务提供者的免责激励所“俘获”。

基于上述理由，更可行的路径是，在高不确定性作品类型中将注意义务限缩为程序性与结构性义务，并以“通知—必要措施”规则为核心进行再构造：其一，生成式人工智能服务提供者应建立清晰、可达的权利主张入口与证据提交机制，并以合理期限响应通知请求；其二，在收到通知后，应采取限制再生成、优化提示与过滤策略等必要措施，且其强度应与侵权指向的明确性、重复性与可纠偏

性相匹配；其三，生成式人工智能服务提供者应建立配套的理由说明、申诉复核、反通知与恢复等纠偏机制，确保自动化治理的误伤成本可被制度化吸收。简言之，限缩生成式人工智能服务提供者对高不确定性作品的注意义务，旨在制度上保留合理使用与表达自由的空间，并维持生成式技术发展的必要弹性。

（三）差异化注意义务的适用边界与规范协调

1. 技术可行性约束下的注意义务适用边界

差异化注意义务是“技术—法律”耦合下的可行性约束，其设定必须能够被行为人以合理成本履行，并通过“现有技术水平”基准与“可指向、可验证、可纠偏”的结构实现外部证成。换言之，只有当注意措施能够在合理成本内被实施，并能被司法确认时，注意义务强度的配置才具备制度正当性。在此意义上，生成式人工智能服务提供者所负担的注意义务，应理解为以专业服务提供者一般能力标准为参照的较高注意义务，其强度应以同业平均水平为基准，而非以行业最优者为标准。基于此，在设定生成式人工智能服务提供者注意义务的强度时需参照以下四项核心因素：其一，生成式人工智能服务提供者的活动所制造的风险规模与概率；其二，生成式人工智能服务提供者的活动带来的社会效益；其三，生成式人工智能服务提供者采取注意措施的成本与负担；其四，生成式人工智能服务提供者对风险生成与扩散是否具有可归责的结构性责任。同时，还应结合服务性质、技术发展水平、替代设计可行性与成本、可采取措施及其效果，以及对行业的影响等因素对注意义务进行动态调适。

技术可行性并非抽象地限缩注意义务强度，而是要求把注意义务嵌入可操作的治理单元。首先，注意义务必须具有可指向性，即法律要求的治理措施能对应到明确的对象与触发条件，而不是泛化为对所有输出的事前审查。例如，在对象可识别且技术可行的重复侵权模式中，必要措施不应只停留在事后删除侵权链接，而应当是在制止已发生具体侵权的基础上，一并预防同一作品、同一模式、时间上紧密相连的“相同侵权”再次发生。其次，注

意义必须具有可验证性，即是否履行注意义务应能通过流程、文档与日志实现外部验证，而不是依赖抽象的主观判断。可验证性的要求意味着生成式人工智能服务提供者需把训练治理、输出约束、投诉处置与复核纠错等环节制度化，并形成可追溯的证据链，从而避免以“结果未避免”倒推违反注意义务。最后，注意义务必须具有可纠偏性，即治理体系必须允许对错误处置进行修正与恢复，避免误伤被制度化。技术可行性边界并不只回答能不能过滤，更要通过程序性保障机制对冲自动化治理的系统性误伤风险。例如，以理由说明、申诉复核、反通知与恢复、透明度报告为代表的程序性纠偏机制应被纳入技术可行性边界的组成部分。

生成式人工智能服务提供者是否违反注意义务的判断，应以“损害发生时的现有技术水平”为核心参照。只要服务提供者已采取符合当时技术能力的合理措施而损害仍难以避免，即不宜直接推定其违反注意义务。否则，注意义务一旦被设定为“超出现有技术”的强度要求，过错责任就可能在司法裁判中滑入严格责任，并对技术发展产生抑制效应。在此基础上，注意义务强度应当遵循比例原则与成本效益衡量。具体而言，对于可能造成重大损害且技术上具有可防控性的风险，应相应提高防控标准；对于技术上难以避免且发生概率较低的风险，则应适当降低义务强度，以避免因治理成本过高而导致制度整体效率受损。由此，注意义务的规范功能并非实现风险的绝对消除，而在于在合理成本范围内实现风险的可控制与可接受。从“成本—能力”结构看，让服务提供者对海量输出结果承担全面的有害内容筛查，既存在技术障碍，也不符合经济理性，反而会削弱其规范的正当性。

2. 多元价值化衡量下的注意义务体系平衡

在完成技术可行性边界的论证之后，差异化注意义务的规范构造必须转向规范体系中的价值维度。注意义务是以侵权制度的目的为依归的责任配置方案。侵权制度固然具有社会成本控制的维度，但效率并非其唯一目标，功利评价本身可以包含健康、自由、尊严等非经济性价值。因此，生成式人

工智能服务提供者注意义务的设定，必须在多元价值之间进行可说明的衡量，而非将其他价值排除出裁判与规范构造的视野。据此，生成式人工智能服务提供者注意义务的规范目标应当以“安全”为底线价值，即对显著、可预见且可避免的侵权风险，要求其提供最低限度的风险控制与程序保障。在此基础上，进一步与治理效率、产业发展、表达自由等价值作体系化协调，而不能被单一的风险规避逻辑所吞没。换言之，注意义务并不旨在实现权利保护最大化的结果导向，而是要在侵权制度的可归责性框架之中形成一套能够兼顾权利救济与制度可持续的责任配置方案。

注意义务作为一种风险预防机制，必须在保证治理效率与控制误伤风险之间建立可持续的平衡机制。面对人工智能的独特风险，治理逻辑应通过预防性行为与因应性制度的组合来构造风险应对框架。这意味着生成式人工智能服务提供者的治理措施既要避免预防不足导致权利保护落空，亦须防止过度预防压缩合理使用空间并造成系统性的“寒蝉效应”。生成式人工智能服务提供者的注意义务必须同时处理两类错误：一是漏拦，即侵权风险因处置不足而外溢；二是误伤，即合法内容因自动化治理被错误限制。因此，任何以零侵权为导向的单向度注意义务设计，都会在实质性相似、合理使用等高不确定性领域迅速放大误伤成本，并以效率换秩序的方式透支制度正当性。正因为误伤风险具有不可完全消除性，生成式人工智能服务提供者应倾向于以预防性努力与程序性保障的制度组合来实现风险预防的功能平衡。

此外，创新激励与表达自由应当构成注意义务强度的上限约束，生成式人工智能治理不应以牺牲创新与表达为代价换取表面合规。在动态过程中，更应依赖程序性救济与制度性纠偏机制维持均衡，避免片面追求权利保护导致公共利益受损。就此而言，生成式人工智能服务提供者在收到预警函或通知后，采取符合合理标准的技术过滤与程序措施即可，而不应被要求对指定内容实现绝对性的传播禁止。相反，如果以完全阻断侵权为唯一目标，生成

式人工智能服务提供者会在合规激励下以风险规避逻辑牺牲表达自由与文化创新，最终也会反噬版权制度对促进文化传播与文明进步的社会功能。因此，注意义务的合理强度应以激励充分预防为目

标，并承认技术发展阶段下适度容忍合理风险的必要性，以实现权利保护、创新发展与表达自由之间的可持续平衡。

（技术编辑：邓语鑫）