数字科技与人工智能

Digital Technology and Artificial Intelligence

每月资讯 2025 年 7 月

上海市律师协会数字科技与人工智能专业委员会

上海市律师协会

数字科技与人工智能专业委员会

每月资讯

（2025 年 7 月）

目录

新规概览

1. 工业和信息化领域人工智能安全治理标准体系正式发布

内容摘要：2025 年 7 月 25 日，工业和信息化部技术委员会第一次成员大会在北京召开。会上，标委会发布了《工业和信息化领域人工智能安全治理标准体系建设指南（2025 版）》。这份指南首次提出了一个系统化的安全治理框架，旨在覆盖数据、算法、系统、应用四个层级的安全要求。它为未来的人工智能安全标准研制工作指明了重点方向，包括治理能力、基础安全、网络安全、数据安全等多个关键领域。

指南以人工智能赋能新型工业化为主线，遵循统筹发展和安全的基本原则，旨在构建工信领域人工智能安全治理体系。指南全面贯彻落实党的二十大和二十届二中、三中全会精神，坚持统筹人工智能发展和安全的基本原则，旨在完善人工智能标准工作的顶层设计，强化全产业链标准工作的协同性，统筹推进标准的研究、制定、实施以及国际化进程，为推动我国人工智能产业高质量发展提供坚实技术支撑。

下一步，标委会将深入推进人工智能安全治理标准体系建设，加大在治理能力、基础安全、网络安全、数据安全、算法模型安全、应用安全、赋能安全等重点领域的标准研制力度，积极参与国际标准法规协调制定，推进关键标准的宣贯实施。充分发挥产业主体在人工智能安全治理的重要作用，加速培育壮大人工智能产业，构建安全、可靠、繁荣的人工智能产业发展环境。

来源：https://mp.weixin.qq.com/s/h5ISBADQKca3KkHVbOcUsA?scene=25#wechat_redirect

2. 最高人民法院、最高人民检察院、公安部发布《关于办理帮助信息网络犯罪活动等刑事案件有关问题的意见》

内容摘要：2025 年 7 月 28 日，最高人民法院、最高人民检察院、公安部

发布《关于办理帮助信息网络犯罪活动等刑事案件有关问题的意见》。《意见》明确指出，利用“深度合成”等人工智能技术实施帮助信息网络犯罪活动的，将被依法从严惩处。这显示了对利用新型技术手段犯罪进行重点打击的立场。

来源： https://www.court.gov.cn/zixun/xiangqing/472121.html

域外规范

1. 金砖国家签署人工智能全球治理声明

内容摘要： 2025 年 7 月 6 日，金砖国家签署《金砖国家领导人关于人工智能全球治理的声明》。该声明是在金砖国家领导人第十六次会晤期间达成的一项重要成果，标志着这个代表“全球南方”的重要国家集团在人工智能这一关键未来领域形成了统一的治理立场。

声明包含五部分内容:多边主义、合法性与数字主权:市场规范、数据治理与技术可及性;公平与可持续发展;符合伦理的可信的、负责的人工智能，造福所有人;以及未来之路。声明强调治理须在《联合国宪章》和各国监管框架下开展，尊重国家主权，循代表性、发展导向、可及性、包容性、动态更新和敏捷性原则。治理应切实注重个人数据保护，保障人类权益，确保安全、透明、可持续，且有利于弥合本国及国家间日益扩大的数字与数据鸿沟。声明的主要内容核心是倡导“以人为本”的人工智能发展理念。它强调人工智能技术的开发与应用必须服务于全人类的共同利益，并促进可持续发展。同时，声明特别强调了公平与包容性原则，指出人工智能的全球治理必须关注并吸纳发展中国家的需求和参与，确保它们不在新一轮技术革命中被边缘化，从而避免全球技术鸿沟进一步扩大。

此外，声明还强调了人工智能的安全与可控性，支持在联合国的主导下建立一个全面而公正的全球治理框架。这旨在确保人工智能技术能够在风险可控的前提下健康发展，防止其被滥用。

此次声明的签署不仅是金砖国家内部合作的深化，更是一次影响全球人工

智能治理格局的行动。它为未来在联合国等国际场合的相关讨论设定了议程，并彰显了发展中国家在塑造全球数字未来中不可或缺的角色。

来源：https://www.mfa.gov.cn/ziliao_674904/1179_674909/202507/t20250709_11668022.shtml

2. 欧洲议会发布《生成式人工智能与版权——训练、创造、监管》报告

内容摘要：7 月 9 日，欧洲议会法律事务委员会(JURI)发布《生成式人工智能与版权——训练、创造、监管》(Generative AI and Copyright - Training, Creation, Regulation)报告。报告的核心在于重新审视生成式 AI 模型训练阶段的合法性。它深刻指出，当前欧盟《数字单一市场版权指令》中的“文本与数据挖掘例外条款”存在被滥用的风险。AI 训练并非简单的“信息提取”，而是对受版权保护作品表达元素的全面复制与利用，这已超出了该例外条款的立法原意。因此，报告强烈建议澄清并修订现行法律，探索从当前的“选择退出”模式转变为“选择加入” 模式，即只有在获得权利人明确许可的前提下，才能将其作品用于 AI 训练，从而从根本上重构训练数据的合法性基础。

报告深入探讨了生成式 AI 带来的两大核心版权问题。首先是版权归属，报告明确了“完全由 AI 生成的内容不享有版权保护”的原则，同时主张法律应保护那些体现了充分人类智力贡献的“AI 辅助创作”。其次是公平的价值分配，报告揭示了一个关键矛盾：AI 开发者使用了高达 40%-70%的版权内容作为训练数据并从中获利，但原创作者和权利人却未获得相应的报酬。为此，报告提出了一系列机制，包括建立法定的公平补偿权、推广通过集体管理组织进行集体许可，以确保创作成果的价值能够回归创作者群体。

为解决 AI 训练数据的“黑箱”问题，报告将透明度义务置于关键位置，主张强制要求 AI 开发者公开其训练数据的详细摘要，并确保技术系统尊重权利人的权利保留声明。在监管架构上，报告建议在欧盟层面的 AI 办公室下设立一个专门的 AI 与版权监管单位，以系统性地负责监督合规、解决争议，并推动建立可持续的治理生态。

这份报告虽无法律约束力，但为欧盟下一步修订《数字单一市场版权指令》或制定专门法律提供了关键的政策蓝图和理论依据。它标志着欧盟正致力于将其强大的版权保护传统延伸至人工智能时代，旨在构建一个“责、权、利”分明的监管框架。其核心精神——“尊重版权是 AI 创新的前提”——预计将对全球的 AI 治理与版权立法产生深远影响，引导行业走向更加规范化和可持续的发展道路。

来源：https://www.europarl.europa.eu/thinktank/en/document/IUST_STU(2025)774095

3. 欧盟发布《通用人工智能实践准则》

内容概要：2025 年 7 月 10 日，欧盟发布了《通用人工智能实践准则》（General-Purpose AI Code of Practice）。该准则是一项自愿性工具，由独立专家在多方利益相关者参与下制定，旨在帮助行业遵守《人工智能法案》对通用 AI 模型提供商的义务。该实践准则有助于行业遵守《人工智能法案》关于通用 AI 模型的安全性、透明度和版权的法律义务。

通用人工智能（GPAI）模型能够执行广泛的任务，正逐渐成为欧盟众多人工智能系统的基础。其中部分模型若能力极强或应用广泛，可能会带来系统性风险。为确保人工智能的安全性和透明度，《人工智能法案》为此类模型的提供商制定了相关规则，包括透明度和版权方面的规定。对于可能存在系统性风险的模型，提供商应当对这些风险进行评估并采取缓解措施。该准则的核心内容分为三个独立章节，形成一个分层治理框架针对性应对不同风险层级的 AI 模型：第一章聚焦透明度要求，适用于所有 GPAI 模型提供者，强制其公开训练数据来源、技术局限性及版权合规措施，以履行 AI 法案第 53 条的义务；第二章关注版权问题，同样适用于所有模型提供者，强调数据使用中的知识产权保护；第三章则专为具有系统性风险的先进模型设计，仅适用于少数超大规模 AI 系统提供者，要求实施严格的风险评估、对抗测试和网络安全防护措施，以符合 AI 法案第 55 条的高标准要求。

在准则发布后，全球科技巨头积极响应签署，截至 2025 年 8 月 4 日，已有

包括 Google、Microsoft、Amazon、OpenAI、Anthropic、Mistral AI 等 32 家机构成为正式签署方，这些企业覆盖 AI 研发、云计算和内容生成等多个领域，彰显行业对欧盟监管框架的广泛认可；值得注意的是，xAI（由埃隆·马斯克创立）选择单独签署第三章的安全与安全章节，这突显其对高风险模型的特殊合规承诺，同时 xAI 需通过其他替代方式证明其在透明度和版权方面的合规性。

总体而言，《通用人工智能行为准则》不仅为全球 AI 治理树立了新标杆，还通过灵活的自愿框架平衡了创新激励与监管需求，其成功实施将推动 AI 产业向更负责任、透明和安全的方向发展，同时为其他地区提供可借鉴的监管模式。

来源： https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai

4. 荷兰数据保护局发布《人工智能和算法报告》

内容摘要： 2025 年 7 月 15 日，荷兰数据保护局发布《荷兰人工智能与算法报告》（第五版）（De Rapportage AI&Algoritmes Nederland）聚焦荷兰人工智能与算法的发展态势。报告指出，情感识别 AI 系统的发展正在增长，但其工作原理受到广泛质疑。这些系统声称能通过生物特征识别情感状态，例如分析面部表情来推断消费者情绪或辅助医疗服务，然而其背后的科学假设存在争议，实际测量的有效性并不明确。使用这类系统可能带来侵犯基本权利和公共价值观的风险，包括限制个人自由、导致歧视以及侵犯隐私。

在具体应用中，情感识别 AI 被用于客户支持、可穿戴设备和语言模型等多个领域，但其识别情感的方式往往不透明，有效性也难以验证。尽管应用增长迅速，人们常常不清楚系统是否在使用情感识别技术以及基于哪些数据运行。开发者和用户必须认识到不同应用场景下的风险，并考虑是否真正有必要使用这类系统。组织在使用情感识别应用时应保持批判态度，充分认识其技术局限性，谨慎评估使用的适当性和比例性。如果决定使用，组织必须保持透明度，告知被分析者相关情况，并获得其同意，这是负责任使用的基本要求。

在荷兰，社会各界正致力于发展可靠的 AI 系统，强调保护、安全和透明度的重要性。然而，AI 的快速应用可能导致许多问题未被及时发现或报告，因此从事件中学习并建立知识共享机制至关重要。监管机构需要与立法者合作，创造有利于负责任创新的环境。对于各类组织而言，现在可以利用多种工具来提升 AI 成熟度，包括建立算法注册、进行偏差测试、构建质量管理体系、执行基本权利评估和公开使用情况等。设定明确的组织目标，如对关键算法和 AI 系统进行定期审计，重点关注控制机制、公平性、网络安全等方面，是推动组织成熟度的重要方式。防止歧视仍然是重要挑战，无论是人类还是算法都可能存在偏见。减少歧视需要同时关注人类和算法在流程中的相互作用，并进行良好设计。荷兰议会近期通过的“盲评”动议体现了这一方向，监管机构强调需要审视整个决策链，确保风险选择的可解释性和可争议性。

AI 在地缘政治中扮演着日益重要的角色，被视为经济发展的关键。各国在创新与保护之间寻求平衡，AI 已成为从医疗到国防等多个领域的国家战略组成部分。欧洲也通过 AI 战略和相应资金支持推动发展。为了把握 AI 机遇并控制风险，荷兰需要制定全面的国家 AI 战略，强制要求（半）公共组织对具有影响力的算法进行注册和定期审计。监管机构需要获得足够资源，投资合作网络以跟上技术发展步伐。建议在公共和私人投资中预留专门资金用于风险管理和控制措施。AI 法规正在不断具体化，标准成为落实法规要求的重要工具。欧盟委员会已发布相关指南，但随着实践发展，仍需进一步细化和行业适配。荷兰政府正在制定相关法律框架，明确注册要求。即使在没有强制要求的情况下，算法注册表仍是提升透明度和控制力的实用工具。监管机构鼓励政府组织积极使用，并提供了具体实施建议。

来源：https://www.autoriteitpersoonsgegevens.nl/documenten/rapportage-ai-algoritmes-nederland-ran-juli-2025#:~:text=De%20Rapportage%20AI%20%26%20Algoritmes%20Nederland%20%28RAN%29%20gaat,de%20inzet%20van%20artifici%C3%ABle%20intelligentie%20%28AI%29%20en%20algoritmes.

5. 英国通过《计算路线图》

内容概要：2025 年 7 月 17 日，英国科学、创新与技术部通过《英国计算路线图》(UKCompute Roadmap)，核心目标是构建一个世界一流的计算生态系统，以确保英国在人工智能时代处于全球科技前沿。该路线图的总体愿景是将计算资源集中用于英国最高优先级的国家事项和最具变革性的机遇，以此推动全国的创新和经济增长。为了实现这一愿景，英国政府承诺进行大规模投资。自当前至 2030 年，将累计投入 20 亿英镑，用于部署新型超级计算机和构建现代化的计算生态系统。这笔投资的核心部分包括部署两台新型人工智能超级计算机，这是英国“人工智能研究资源”建设的第一阶段，目标是在 2030 年将人工智能研究资源扩大 20 倍。其中，将拨款 7.5 亿英镑在爱丁堡建设新的国家超级计算中心。

在具体行动计划上，路线图提出了多项关键措施。首先是构建一个多元化、互联互通且以用户为中心的现代公共计算生态系统。其次，将通过革新的算力分配模式，将资源优先用于医疗、国防和气候科学等最高影响力的国家优先事项，并保障政府人工智能部门和人工智能安全研究所的专用算力。此外，路线图还计划通过在全国设立“人工智能增长区”来支持大规模人工智能训练和推理，并探索可持续的能源解决方案以应对未来的能源需求。最后，为了创建自主和安全的能力，英国将支持本土企业开发安全、可持续的计算技术，并推动新型计算范式的研究创新，助力本土企业成长为全球领军者。

来源：https://www.gov.uk/government/publications/uk-compute-roadmap

6. 欧盟发布《通用人工智能模型提供者义务范围指南》

内容概要：2025 年 7 月 18 日，欧盟发布《关于通用人工智能模型提供商义务范围的指南》(Guidelines on the scope of obligations for providers of general-purpose AI models under the AI Act)，旨在明确在《人工智能法案》框架下，通用 AI 模型提供者必须遵守的核心合规义务。该指南的核心目的在于清晰界定何为“通用 AI 模型”，并为这些模型的提供者设定明确的合

规路线图。它详细规定了提供者必须履行的四大核心义务：第一是全面的透明度，包括撰写详细技术文档和发布受版权保护训练数据的摘要；第二是遵守欧盟版权法，要求建立机制允许权利所有者拒绝其数据被用于训练；第三是实施强大的网络安全措施；第四是识别、评估并缓解模型可能引发的系统性风险，例如对劳动力市场、公共舆论的潜在负面影响。

这份指南的出台，标志着全球首个综合性 AI 监管框架已从立法阶段正式转入实质性的执法准备阶段。它旨在为高速发展的 AI 技术划定“安全区”，在激励技术创新的同时，严格管控其可能带来的社会性风险。通过强调透明度、权责明确和风险管理，欧盟不仅期望保护公民的基本权利和民主价值，也意在为全球 AI 治理树立标杆。任何希望进入或维持在欧盟市场运营的通用 AI 模型提供者，都必须立即开始依据该指南调整其开发和运营实践，以适应即将到来的严格监管环境。

来源： https://digital-strategy.ec.europa.eu/en/library/guidelines-scope-obligations-providers-general-purpose-ai-models-under-ai-act

7. 美国白宫发布《美国 AI 行动计划》

内容概要： 2025 年 7 月 23 日，美白宫发布了名为《赢得 AI 竞赛：美国 AI 行动计划》（Winning the AI Race: America’ s AI Action Plan）的政策文件，旨在通过全面战略部署，确保美国在全球 AI 领域的绝对主导地位。该文件依据特朗普总统 1 月 23 日签署的《消除美国 AI 领导地位障碍》（Executive Order 14179）行政命令制定，围绕“加速 AI 创新、建设美国 AI 基础设施、领导国际 AI 外交与安全”三大支柱提出超过 90 项联邦政策行动。该计划特别关注与中国在 AI 领域的技术竞争，力求通过技术出口、出口管制和国际合作，确立美国在全球 AI 生态的核心地位。

根据计划，美国政府将在技术创新、基础设施建设和国际合作三大领域采取 90 余项具体措施。在技术创新方面，计划提出将简化人工智能研发的监管框架，清除过时法规，并建立企业意见反馈机制；基础设施建设方面将重点优化数据中心和半导体工厂的审批流程，同时启动电工、暖通技术等紧缺工种的专业人才培养项目；在国际合作领域，美国商务部和国务院将

联合产业界向国际伙伴输出包括硬件、算法、软件及应用在内的全套人工智能解决方案。

值得注意的是，计划特别强调要确保人工智能系统的客观公正性。为此，联邦政府将修订采购标准，要求合作企业必须保证其人工智能系统不受特定意识形态影响。白宫科技政策办公室主任表示，这一系列措施体现了美国政府巩固人工智能领域全球领导地位的决心，将通过促进技术创新、完善基础设施和深化国际合作等方式，确保美国在新一轮科技竞争中保持优势。

来源：https://www.whitehouse.gov/articles/2025/07/white-house-unveils-americas-ai-action-plan/

8. 美国总统特朗普签署系列人工智能领域行政命令

内容概要：7 月 23 日，美国总统特朗普签署《加快联邦对数据中心基础设施的许可》(Accelerating Federal Permitting ofData Center Infrastructure ）、《防止联邦政府中的“觉醒人工智能”》(Preventing Woke Al in the Federal Government)、《促进美国人工智能技术栈出口》(Promoting The Export of the American Al Technology Stack)系列行政命令。这一系列行政命令的核心目标是全方位强化美国在人工智能领域的领导与主导权。首先，通过《加快联邦对数据中心基础设施的许可》，意图在国内大幅简化数据中心建设的审批流程，将其视为关键基础设施，以快速提升美国的算力产能，为 AI 发展夯实本土基础。与此同时，《防止联邦政府中的“觉醒人工智能”》则着眼于国内意识形态领域，旨在禁止联邦机构使用被认为带有所谓“觉醒”偏见的 AI 系统，确保政府使用的 AI 技术符合其定义的“中立”和“美国传统价值观”。在对外层面，《促进美国人工智能技术栈出口》旨在将美国的 AI 技术、软件和服务作为重要的经济和地缘战略工具推向全球。计划通过放宽出口管制、简化审批流程并积极利用经济外交，来推动美国 AI 技术栈的出口，目的是在全球市场，扩大其技术影响力和地缘政治优势。

来源：https://www.whitehouse.gov/presidential-actions/2025/07/accelerating-federal-permitting-of-data-center-infrastructure/ // https://www.whitehouse.gov/p

residential-actions/2025/07/preventing-woke-ai-in-the-federal-government/ // https://www.whitehouse.gov/presidential-actions/2025/07/promoting-the-export-of-the-american-ai-technology-stack/

案例研讨

1. 广东阳江王某某利用 AI 生成虚假新闻扰乱公共秩序案

内容概要：2025 年 7 月 10 日，广东省阳江市江城公安机关在网络日常巡查工作中，发现一则题为“阳江一女子被拖入巷中暴打”的信息在多处互联网平台传播，引发部分网民关注。经公安机关深入核查，该信息内容纯属捏造，系网民王某某为博取关注、谋取流量利益，利用人工智能软件“无中生有”生成的虚假新闻，并伪造公安机关警情通报图片，在多家知名网络平台故意散布，严重扰乱社会公共秩序，造成不良社会影响。根据《中华人民共和国治安管理处罚法》相关规定，江城公安分局依法对王某某虚构事实扰乱公共秩序的违法行为作出行政拘留五日的处罚决定。该案是阳江警方在“净网 2025”专项行动中查处的又一典型案件，反映出利用 AI 技术编造、传播网络谣言的新型违法动向。公安机关重申，网络空间不是法外之地，任何利用新技术手段编造不实信息、扰乱社会秩序的行为，都将受到法律严惩。公众应自觉遵守法律法规，共同维护清朗网络环境。

来源：https://mp.weixin.qq.com/s/EG0xT505da-BrZvH37NjrA

2. 江西宜春网站运营者不履行网络信息安全管理义务案

内容概要：2025 年 7 月 19 日，江西省公安机关网络安全保卫部门在推进“护网-2025”专项整治工作过程中，依法查处一起未履行网络信息安全保护义务的案件。宜春市奉新县网安民警在日常巡查中发现，辖区内某科技公司运营的“AI 智能聊天机器人”服务平台，未按照国家相关规定建立信息内容审核机制，对人工智能自动生成的内容未进行有效过滤与监管，致

使部分违法信息通过该平台传播扩散，不仅严重扰乱互联网空间管理秩序，也存在较大社会风险隐患，极易造成不良社会影响。针对上述违法行为，公安机关立即启动执法程序，依法对涉事企业网站负责人进行警示约谈，明确指出其未落实网络安全主体责任的问题，并依据《中华人民共和国网络安全法》等相关法律法规，对该企业不履行网络信息安全管理义务的行为予以行政处罚。同时，公安机关责令涉事网站在规定期限内完成全面整改，切实加强 AI 生成内容的审核管理，严防违法信息传播，维护清朗网络环境。该案的查处，不仅体现了网安部门对新型网络安全隐患的高度警惕与依法监管，也为广大互联网企业特别是 AI 研发运营主体敲响了警钟——技术发展必须与安全责任同步推进，任何忽视内容安全、规避管理义务的行为，都将依法受到严肃追究。江西网警表示，将持续深化“护网-2025”专项行动，进一步压实企业主体责任，筑牢网络安全防线，为数字社会的健康发展提供坚实保障。

来源：https://mp.weixin.qq.com/s/Tqy4NfFsSrpS8aU3D5yMNQ

3. 王某彪被行拘，利用 AI 伪造“女儿走失”虚假警情

内容概要：7 月 23 日，浙江嘉兴男子王某彪为博取流量、快速增粉，使用 AI 大模型生成虚构的“寻人启事”，谎称女儿“王喵喵”走失。该虚假信息在短视频平台发布后迅速发酵，12 小时内播放量超 160 万次，点赞 2-3 万，转发 8000-9000 次，王某彪粉丝量从个位数暴涨至 600+。大量网民自发转发、组织“寻人”，多地警方接到群众报警，公共资源被严重挤占。嘉兴南湖警方调查发现信息异常：预留电话无法接通，且无相关报案记录。溯源锁定发布者王某彪，确认其未婚未育，不具备“父亲”身份。7 月 25 日，警方在嘉兴新丰镇将其抓获，王某彪对编造事实供认不讳。7 月 26 日，南湖公安分局依据《治安管理处罚法》第二十五条第一款，认定王某彪的行为构成虚构事实扰乱公共秩序，对王某彪依法处以行政拘留。

来源：https://mp.weixin.qq.com/s/Btux4TbI5i4DgDTmCrj1XQ

4. 国内首案：内容被判定为 AI 生成遭屏蔽，用户起诉平台获胜诉

内容概要：某科技公司运营的网络平台用户李某，发布了一段生活建议类即时性创作文本，平台通过自研算法将该内容标记为“包含 AI 生成内容但未标识”，随即采取隐藏内容、账号禁言 1 天的处理措施。李某主张内容为自身原创、未使用 AI 创作，申诉无果后诉至北京互联网法院，要求平台撤销处理措施并删除违规记录，认为平台行为构成服务合同违约。平台辩称，依据双方服务协议及相关规定，有权对未主动标识的 AI 生成内容进行治理，涉案内容经机器识别和人工复核（以“是否具备明显人类情感特征”为标准）判定为 AI 生成，且无需向用户披露算法逻辑。法院审理焦点集中在平台审查是否有合理依据、举证责任分配及算法解释义务上，认定李某因内容系即时创作，客观上无法提供创作底稿等佐证，而平台提交的算法备案信息与 AI 识别无关联性，人工复核标准主观缺乏科学依据，未对判定结果作出合理解释。北京互联网法院一审判决某科技公司构成服务合同违约，判令其撤销对涉案内容的隐藏措施，删除后台系统中的违规处理记录。该案判决已生效，成为国内首起因平台判定用户内容为 AI 生成引发的标志性判例。

来源：https://xinwen.bjd.com.cn/content/s687782b2e4b0aabe0a03b5b4.html

5. 最高人民法院发布依法惩治帮助信息网络犯罪活动及相关犯罪典型案例

内容概要：2025 年 7 月 28 日，最高人民法院、最高人民检察院、公安部联合召开新闻发布会，发布《关于办理帮助信息网络犯罪活动等刑事案件有关问题的意见》、依法惩治帮助信息网络犯罪活动及相关犯罪典型案例。最高人民法院刑三庭副庭长王鲁发布了 7 起依法惩治帮助信息网络犯罪活动及相关犯罪典型案例，其中包括：依法严惩有组织提供“账号解封”等技术支持行为，斩断“输血供能”犯罪链条——被告人张某某帮助信息网络犯罪活动案；依法严惩利用 GOIP 设备提供通讯传输支持行为，依法保

护公民个人信息安全——被告人邓某某、王某某帮助信息网络犯罪活动、被告人黄某帮助信息网络犯罪活动、侵犯公民个人信息案；对行业“内鬼”依法宣告职业禁止，制发司法建议推进综合治理——被告人薛某帮助信息网络犯罪活动案；依法严惩通过虚拟币交易转移赃款犯罪——被告人王某等人掩饰、隐瞒犯罪所得案；与电信诈骗团伙事先通谋或者形成较为稳定配合关系，以诈骗罪共犯论处——被告人付某诈骗案；综合认定涉“两卡”犯罪的主客观情节，做好行刑衔接——朱某某掩饰、隐瞒犯罪所得案；对未成年人、在校学生全面准确贯彻宽严相济刑事政策——被告人高某等人帮助信息网络犯罪活动案。

来源：https://www.court.gov.cn/zixun/xiangqing/472111.html

6. Clorox 诉 Cognizant 网络攻击案

内容概要：7 月 23 日，Clorox 以违约、严重疏忽等为由向美国加州高等法院起诉 Cognizant，要求赔偿 3.8 亿美元及惩罚性赔偿。全球清洁用品巨头 Clorox 与合作超 10 年的 IT 服务台提供商 Cognizant，因 2023 年 8 月的“灾难性”网络攻击引发纠纷。当时，疑为 Scattered Spider 的黑客组织通过社会工程学手段，冒充 Clorox 员工致电 Cognizant 服务台，以各类理由索要网络访问权限。Cognizant 客服违反 Clorox 明确的安全流程，未做任何身份核验，便直接提供核心认证平台账户密码、重置多因素认证凭据、更改绑定手机号，且未发送任何变更通知。最终导致 Clorox 生产暂停、供应链效率骤降，产生 3.8 亿美元总损失（含 4900 万美元直接修复成本），销售额也受影响，可持续发展目标需重新评估。目前，该案仍处于诉讼阶段，尚未有法院最终判决结果。现阶段案件进展：Clorox 已向法院提交完整证据链（含通话录音、服务协议、损失核算报告等），明确寻求 3.8 亿美元直接赔偿及额外惩罚性赔偿。

来源：https://therecord.media/clorox-cyberattack-lawsuit-cognizant-it-contractor

行业动态

1. 2025 世界人工智能大会发布人工智能全球治理行动计划

内容摘要：2025 年 7 月 26 日，2025 世界人工智能大会暨人工智能全球治理高级别会议发表《人工智能全球治理行动计划》，提出“六大原则”与“十三项具体行动”，标志着全球人工智能治理从理念共识迈向实际行动。该计划凝聚国际共识，强调“智能时代 同球共济”理念，以“向善为民”为首要原则，将发展作为治理核心目标，反对将技术治理工具化。其创新性地提出“可持续人工智能”理念，倡导制定能效水效标准、推广绿色计算技术；在开源创新领域推动合规体系建设与开发资源共享。务实举措方面，计划聚焦人工智能在工业制造、医疗教育、农业减贫等领域的赋能应用，强调数字基础设施建设与高质量数据集合作，为构建普惠包容、安全可控的人工智能全球治理体系提供了中国智慧与方案。

来源：https://world.people.com.cn/n1/2025/0801/c1002-40534788.html

2. 上海 AI lab 发布 DeepLink 超大规模跨域混训技术方案

内容摘要：上海人工智能实验室于 7 月 19 日正式发布了其创新的“DeepLink 超大规模跨域混训技术方案”，该方案已成功完成多个落地项目，实现了支持千公里距离下、跨越多个智算中心稳定联合训练千亿参数大模型的突破。这一成就标志着我国在超大规模智算跨省互联领域取得了重要进展。具体而言，实验室联合中国联通，跨越 1500 公里连接了上海和济南的智算中心，高效完成了千亿参数大模型的混合训练，算力利用率高达单集群的 95%以上；同时与中国电科合作，实现了北京、上海与贵州三地智算中心的互联与混训。这些成功实践为解决当前国内智算中心建设分散、算力资源碎片化导致难以灵活低成本获取大算力的行业难题探索出了一条新路径。DeepLink 方案的核心在于攻克了大规模跨域异构集群调度、高性能通信整合和高可靠容错等一系列技术难题：它创新性地采用“3D 并行+PS”架构，通过算法优化减少通信开销，使得普通专线即可满足远程训练需求，并确保了即使单个节点故障也不影响整体进程；同时，通过改进的异构流水线

并行策略和自研框架，动态调节不同硬件任务量，有效解决了因芯片种类繁多、性能不一导致的异构混训效率低下问题。该方案的实用价值在于，它验证了无需依赖高算力芯片集中部署，即可通过灵活组合不同性能的芯片来获取大算力，从而降低了对外部特定硬件的路径依赖，使得各地分散的算力集群能够互联形成“合力”，突破单集群性能上限。目前，DeepLink 已深度集成到多家合作伙伴的平台中，为实现全国算力资源的共建、共营、共享，优化全国算力一体化布局提供了关键的技术支撑和核心动能。

来源：https://mp.weixin.qq.com/s/U2hMkdDsbXuKybR8gvLS4Q

3. 阿里开源泛音频生成和视频生成模型

内容摘要：2025 年 7 月，阿里巴巴通义实验室正式开源了其首款音频生成模型 ThinkSound，这是一款能够基于视频、文本或音频输入来生成高保真音效与音景的多模态 AI 模型，为视频内容创作带来了革命性突破。ThinkSound 堪称一位“专业 AI 音效师”，它采用先进的链式推理技术，能够深入分析视频画面的场景、动作与情感，从而生成与之高度匹配且同步的音效，无论是自然界的风声水声，还是都市中的车辆鸣笛，都能逼真还原。其核心优势在于多模态融合与高精度同步，通过先进的计算机视觉算法逐帧解析视频内容，确保生成的音频与视频帧精准对齐，并兼容从标清到 4K 的多种主流视频格式，在行业基准测试中表现优异。

作为阿里巴巴开源战略的重要一环，ThinkSound 的模型权重与推理脚本已全面开放，开发者可通过 Hugging Face、ModelScope 及 GitHub 等平台免费获取，这一举措极大地降低了专业音效生成的技术门槛，使中小型创作者和独立开发者也能轻松使用，并通过其提供的交互式编辑功能对音效进行精细调整。该模型的应用场景极为广泛，不仅能显著提升影视后期制作的效率，快速为无声视频配乐，还能为游戏开发生成动态音效以增强沉浸感，其语音合成技术甚至能为虚拟角色生成带情感和唇部同步的多语言对话。ThinkSound 的开源不仅为内容创作领域注入了全新活力，其多模态融合与链式推理的应用也为 AI 音效生成行业树立了新标杆，结合阿里巴巴在视频与语音生成领域的持续创新，预示着多模态 AI 技术未来在真实感与交互性上拥有无限的潜力。

来源：https://baijiahao.baidu.com/s?id=1837245756091549756&wfr=spider&for=pc、https://hub.baai.ac.cn/view/47698

4. 月之暗面发布并开源 Kimi K2 模型

内容摘要：7 月 11 日，月之暗面正式发布并开源了其新一代基础模型 Kimi K2，该模型采用混合专家架构，拥有高达 1 万亿的总参数，每次推理激活 320 亿参数，在代码能力与通用智能体任务处理方面表现尤为突出。在技术性能上，Kimi K2 展现出强劲实力，在 SWE Bench Verified、Tau2、AceBench 等多项权威基准测试中均取得了开源模型中的最优成绩，充分证明了其在自主编程、工具调用和数学推理三大核心领域的领先水平。更值得关注的是，在预训练阶段，Kimi K2 采用了创新的 MuonClip 优化器，有效解决了当前业界面临的人类高质量数据瓶颈问题，实现了万亿参数模型的稳定高效训练，并显著提升了数据利用效率，为模型性能的持续提升找到了新的增长空间。除了基准测试成绩亮眼，Kimi K2 在多个实际应用场景中也体现出强大的能力泛化性和实用价值。为方便开发者与用户体验，月之暗面同步开放了多种使用渠道：用户即可通过访问官网 kimi.com 或下载官方 App 直接体验全新模型，也可通过其新上线的 API 服务进行集成，该 API 兼容 OpenAI 和 Anthropic 的接口标准，支持长达 128K 的上下文，并具备更强的通用性与工具调用能力，使得用户能够便捷地将现有的大模型工具链切换至 Kimi K2，以满足多样化的开发与应用需求。

来源：https://baijiahao.baidu.com/s?id=1837405587776178197&wfr=spider&for=pc

5. 阶跃星辰上线新一代基础大模型 Step 3

内容摘要：7 月 25 日，2025 世界人工智能大会开幕前一天，上海 AI 独角兽阶跃星辰正式发布了新一代基础大模型 Step 3。这是阶跃星辰首个全尺寸、原生多模态推理模型，解码效率为同类顶尖产品的三倍。对于大模型技术的发展趋势，阶跃星辰创始人兼 CEO 姜大昕提出了他的见解。他认为，随着技术迈向“推理时代”，优秀的大模型必须同时具备四个特征：强智能、低成本、可开源和多模态。单一的能力亮点已无法满足用户的综合需

求。他用“多、开、好、省”四个字精炼地概括了 Step 3 的核心优势。

具体来说，“多”指的是多模态能力。Step 3 继承了公司“多模态卷王”的基因，能够深入理解图片和视频内容，而不仅仅局限于文本对话和解题。“开”代表可开源。模型计划于 7 月 31 日向全球开源，这将极大方便用户进行私有化部署，并在此基础上进行后续训练和微调，以提升模型在特定垂直领域的性能。“好”是指强智能。Step 3 支持多模态推理，结合了多模态和复杂推理的双重优势，使其不仅能识别图像内容，更能理解图像背后蕴含的逻辑关系。姜大昕用了一个生动的例子来说明这一点：一位正在减肥的用户，只需将一桌饭菜的照片发送给 Step 3，模型不仅能识别出所有菜品，还能估算每道菜的热量，并最终为用户推荐一套总热量在 300 大卡以内的饮食搭配方案。

来源：https://export.shobserver.com/baijiahao/html/952723.html

6. 上海 AI lab 开源「书生」科学多模态大模型 Intern-S1

内容摘要：7 月 28 日，2025 世界人工智能大会科学前沿全体会议现场，上海 AI 实验室发布并开源“书生”科学多模态大模型 Intern-S1。Intern-S1 在书生大模型家族基础上，重点强化了科学能力，为首个融合专业科学能力的开源通用模型。Intern-S1 首创“跨模态科学解析引擎”，在化学、材料、地球等多学科专业任务基准上超越了顶尖闭源模型 Grok-4。介绍称，在多模态综合能力方面，Intern-S1 全面领先 InternVL3、Qwen2.5-VL 等主流开源模型。此外，上海 AI 实验室一并推出“书生”科学发现平台 Intern-Discovery，为全球研究者提供从假设到验证的一站式科研支撑，目前平台已开放全球试用申请。

来源：https://www.tmtpost.com/nictation/7640179.html

7. 字节发布端到端同声传译模型 Seed LiveInterpret 2.0

内容摘要：7 月 24 日，字节跳动宣布正式发布端到端同声传译模型 Seed LiveInterpret 2.0。这是首个延迟&准确率接近人类水平的产品级中英语音同传系统。在多人会议等复杂场景中英双向翻译准确率超 70%，单人演讲翻译准确率超 80%，接近真人专业同传水平。翻译延迟可低至 2-3 秒，较

传统机器同传系统降低超 60%。

来源：https://baijiahao.baidu.com/s?id=1838511588531751615&wfr=spider&for=pc

8. 百度多模态大模型 MuseSteamer 携「绘想」平台上线

内容摘要：7 月 2 日，百度商业研发团队正式发布了其自主研发的视频生成模型 MuseSteamer 及其配套的视频产品平台“绘想”。这一创新旨在通过“生成式 AI+多模态技术”打造全面的视频生成解决方案，以满足搜索、广告、推荐等场景对原生化内容生产的强劲需求。MuseSteamer 视频生成模型系列丰富，目前包含 Turbo、Lite、Pro 以及全系列有声版本。其中，Turbo 版已率先上线“绘想”平台并开启限时免费公测，其余版本预计将于今年 8 月陆续面向用户开放。据悉，MuseSteamer 的核心亮点在于其强大的功能特性，包括支持音视频一体化生成，具备极强的可控性，能够实现电影级别的制作效果，支持生成连续 10 秒的动态视频，同时兼具极致的性价比和极速的生成速度。MuseSteamer 的发布将显著帮助客户和内容创作者突破传统视频创作的瓶颈，极大地激发内容的多样性与创意空间。用户现在仅需上传一张图片，便能生成专业级的视频内容，极大地简化了视频制作流程，降低了创作门槛。

来源：https://baijiahao.baidu.com/s?id=1836520155173592286&wfr=spider&for=pc

9. 马斯克旗下 xAI 发布 Grok 4

内容摘要：7 月 10 日，马斯克旗下的人工智能公司 xAI 正式发布了 Grok 4，这是该公司自 2023 年推出首代大模型以来的第四次重要迭代，号称是“世界上最强 AI 模型”。Grok 4 系列包含 Grok 4 和 Grok 4 Heavy 两个版本，均为纯推理模型。Grok 4 是单代理版本，Grok 4 Heavy 是多代理版本，支持四个代理同时工作，上下文窗口最高支持 256k tokens。根据 xAI 的内部基准测试结果，Grok 4 在 Humanity’s Last Exam（一个涵盖数学、科学与语言等多学科、约 2500 道题目的权威考试）中，通过文本输入达到了约 25%的正确率。这一数字与 OpenAI 在今年 2 月公布的 Deep Research

工具约 26%的正确率不相上下。但 xAI 方面强调，两者测试方式并非完全相同。此次 Grok 4 的一个看点是新增五种声音模式，以及反应速度较早期几乎缩短了一半。为方便开发者使用，Grok 4 加入了专门的编程模型 Grok 4 Code，能在代码生成、调试等方面提供更专业的支持。xAI 表示，这一变化将明显提升开发者效率，以应对 GPT- 4 系列和 Meta Gemini 在编码领域的竞争。在 xAI 的描述中，Grok 4 不仅具备“逻辑推理更强、语言理解更准确”的优势，还首次尝试加入多模态信息处理，如图像、视频的潜力。这使其在社交媒体内容上的应用更具时效性和文化敏感性。目前 xAI 累计融资额已超过 200 亿美元。4 月底有报道称，xAI 控股公司正在与投资者洽谈，计划为其整合后的 xAI 及社交媒体业务 X 平台筹集约 200 亿美元资金。若交易达成，该公司估值将突破 1200 亿美元。

来源：https://baijiahao.baidu.com/s?id=1837242355054604515&wfr=spider&for=pc

10. OpenAI 发布 ChatGPT Agent

内容摘要：7 月 18 日，OpenAI 如约发布了其最新力作——ChatGPT Agent。根据 CEO Sam Altman 和四位 OpenAI 研究员介绍，ChatGPT Agent 是一个具备自主执行复杂任务能力的 AI Agent，它不再仅仅“对话”，而是可以打开虚拟机，完成搜索、筛选、判断、执行等一整套流程，最终输出可交付的结果。ChatGPT Agent 的定位非常“简单直接”：一个拥有终端、图形浏览器、文本浏览器的多工具整合智能体系统。功能上，几乎等于一个受控的远程虚拟操作系统。值得注意的是，ChatGPT Agent 可以说是 OpenAI 自今年以来推出产品的一次阶段性整合与释放：Operator 和 Deep Research，一个偏执行，一个偏思考，如今彻底融合。

来源：https://baijiahao.baidu.com/s?id=1837958448060628052&wfr=spider&for=pc

11. 亚马逊推出 Agent 全家桶

内容摘要：7 月 16 日，在亚马逊云科技纽约峰会上，亚马逊云科技发布 Amazon Bedrock AgentCore，一次性推出七大功能模块，直指 Agents 落地

的痛点——安全、可观测、身份识别、长期记忆、工具连接等等，几乎覆盖了工程化 Agent 的全链路能力。亚马逊云科技的愿景是成为“构建全球最实用 AI agents 的最佳平台”，赋能组织大规模部署可靠且安全的 agent。

来源：https://baijiahao.baidu.com/s?id=1837861387483846456&wfr=spider&for=pc