代表委员热议“AI向善”

一张贴在手机背后的卡片，能进行录音和多语种翻译；一只感音空气架子鼓，只需敲击空气就能发出美妙声音：一款“AI庄稼医院”，只需拍照就能识别庄稼病虫害，并精确开出处方……

人工智能（AI）不断与时代同频共振，在技术演进与商业化落地双轮驱动下，正展现出前所未有的发展潜力，同时也引发了人们对未来趋势的更多思考。工业和信息化部部长李乐成3月5日在十四届全国人大四次会议首场“部长通道”上表示，“要努力推动AI电脑、AI手机、智能家居更好满足人民群众对美好生活的需求。”

技术创新步履不停，智能如何向善而行？从今年政府工作报告中明确部署“完善人工智能治理”，到两会会场内外关于“AI+教育”“AI+安全”的热烈讨论，“AI向善”成为今年全国两会科技领域的高频词。

法治筑基 划定AI发展清晰边界

去年10月中旬，不少爱茶人士刷到一条视频，镜头里的老人自称“活到108岁，试茶80多年”，倾情推荐一款茶叶，吸引了众多茶友下单。殊不知，视频中的老人是毕生推广茶学的专家张天福，早已在2017年离世。企业这一不当行为引发广泛争议，逝者亲属也已通过法律途径介入。今年3月，某社交平台发布打击利用AI仿冒名人引流虚假宣传的公告，处置违规账号1200余个。

这些案例正是AI技术快速发展下，各类风险问题的缩影。全国人大代表、海尔集团董事局主席周云杰认为，AI技术爆发式增长伴生的安全、隐私与伦理挑战日益凸显，已成为制约人工智能健康发展的关键瓶颈。

当前，我国已出台网络安全法、数据安全法、个人信息保护法等法律，还颁布了《人工智能科技伦理管理服务办法（试行）》等针对性举措，但如何在进一步提升社会生产效率的同时降低风险，成为巩固AI产业先发优势、提升国际竞争力的必经之路。

“AI合成名人虚假音视频进行虚假宣传等现象，严重侵犯了公民的肖像权、名誉权。”全国政协委员、中国地震局原党组书记闵宜仁在小组讨论中直言，针对此类乱象，应当及时完善政策规范和监管体系。

周云杰代表认为，除了技术滥用，还有责任主体缺失的问题。比如在自动驾驶、医疗机器人等物理交互场景，算法的安全容错率极低，事故发生后的责任归属如何在开发者、部署者、使用者之间界定，目前还是法律与伦理的真空地带。

全国人大代表、中国工程院院士张伯礼对此深有同感，他表示：“一旦AI形成自主编程、修改程序的能力，更将带来不可预估的安全隐患。”张伯礼代表建议，要加快推动国家层面人工智能综合性立法，构建以相关法律为统领，部门规章、技术标准、伦理指南相互支撑的制度体系；针对使用人工智能的侵权行为，强化流程追溯，尽快出台认定标准和赔偿细则，凝聚社会共治合力。

技术赋能 筑牢AI安全防控防线

AI风险的技术属性，决定了技术创新是防控工作的核心支撑。

“没有安全的创新走不远、也走不稳。从这两年大模型、智能体、具身智能等各类AI形态遭受网络攻击的情况来看，为AI预设‘安全护栏’十分必要。”全国政协委员、奇安信科技集团董事长齐向东表示，其带领的团队围绕AI安全防范形成了大模型卫士、大模型安全护栏等创新成果，还探索构建了“端—网—云—机”一体化的防御体系。

齐向东委员建议，要将安全能力嵌入AI应用的全生命周期，实现纵深防御；明确合规红线，压实安全主体责任，强化权限与内容管控；坚持用AI对抗AI，让安全能力始终领先于安全风险。

如今，“用AI对抗AI”已在多地落地实践。上海外滩举办的Deepfake攻防挑战赛，聚焦AI换脸问题，汇聚了一群年轻的“AI打假师”；世界互联网大会上的AI大模型攻防赛，通过双向竞技的模式，让“漏洞穷尽式发现”成为可能。

关于数据安全与保护问题，全国政协委员、中国市场主体研究院院长屈庆超认为，在行业尚未成熟的早期探索阶段，不应过早锁定或限制特定技术路线，应鼓励企业开展差异化探索。但无论如何选择，隐私、合规和责任都是不可逾越的底线。

当前，国内已有多个“用AI对抗AI”的成熟案例落地，尤其在金融风控、身份核验、内容鉴伪三大领域形成了技术闭环，还持续推动着行业标准完善与监管协同发展。围绕这一思路，齐向东委员还建议大力发展安全大模型与安全智能体，将稀缺的专家经验转化为可复制的数字能力，实现7×24小时自动化防护，全方位筑牢AI安全防线。

素养培育 夯实AI向善人才根基

“技术都是双刃剑，重要的是引导好使用它们的人。”“我们科协界别可以和教育界别等联合开展调研。”在北京会议中心，政协委员们围绕AI的使用与引导展开热烈讨论。

全国政协委员、360集团创始人周鸿祎表示，人工智能本质是人类智能的延伸，是工具而非主体。“AI向善”，就是要在人的指导下，让人工智能的设计、开发与应用全过程，始终坚守服务社会公共利益的导向，确保技术发展方向与人类福祉保持一致。

周鸿祎委员以视频制作为例，生动阐释了人工智能与人的关系：AI只是创作者能力的“放大器”，创作者依然主导创意和艺术表达，AI则协助完成大量技术性工作，让创意更快转化为作品。当技术与人的智慧协同发力，人工智能不仅能提升生产效率，更在不断拓展人类创造力的边界。

而实现“AI向善”，全民AI安全素养的系统提升是重要基础。全国人大代表、天津职业技术师范大学副校长王劲松表示，提升全民AI安全素养，需要从三个层面系统推进。

面向公职人员，作为政策制定、实施与落地的核心主体，必须具备扎实的AI安全意识与素养，既要为技术创新留出包容的试错空间，又能敏锐识别、有效防范AI带来的各类风险；面向技术人员，要引导AI领域从业者正向运用技术，倡导“技术向上”的理念，确保技术研发与应用始终沿着服务经济社会发展的正向轨道前行，平衡好技术发展与行业规制的关系；面向广大公众，每一位AI使用者都要建立对应的安全素养，清醒认识AI个性化服务背后的数据隐私风险，增强个人信息保护意识与自我防护能力，守住AI使用的安全底线。

平衡发展 让AI在正轨释放核心价值

如今，脑机接口、具身智能、AI教育、AI手机……一批优秀的AI应用切实改变着人们的生活，带来可喜变化。AI的终极价值，在于服务实体经济、改善民生福祉、推动社会进步。为AI划底线、立规矩，引导技术在合规中落地、在规范中创新，才能让其真正成为“十五五”时期高质量发展的新动能。

如何平衡AI发展与安全的关系？全国人大代表、招商局资本投资有限责任公司董事李引泉说：“监管政策要平衡好规范安全和创新发展的关系，管得过严可能不利于新技术的发展，对于在人工智能技术方面具备赶超国外能力的企业，应予以保护和扶持。”

齐向东委员则认为：“安全和创新是一体两翼。安全从来不是约束，而是助力新技术走向成熟的重要标志。”他以智能网联车为例，作为具身智能的重要形态，为保障其安全发展，我国已形成涵盖法律法规、路政设施等方面的全面安全体系，如今智能网联汽车发展迅速、市场份额持续扩大。“AI时代，我们势必要建立同样完善的体系和机制。‘安全带’越牢靠，AI这辆‘疾驰的快车’跑得就越稳健。”

政府工作报告提出深化拓展“人工智能+”，促进新一代智能终端和智能体加快推广。对此，工信部部长李乐成在接受采访时表示，要努力推动AI电脑、AI手机、智能家居等产品发展，更好满足人民群众对美好生活的需求。

周鸿祎委员建议，要加快出台行业统一标准，明确数据采集、模型调用、内容生成的合规边界，强化算法透明度与可解释性；将安全智能体嵌入终端，主动防御AI化攻击、恶意利用等新型风险，让安全成为AI终端的标配，让AI手机真正服务于民，成为智能经济新形态下可靠的普惠终端。

作为终端智能体的代表性应用，AI手机助手的安全实践为行业提供了可参考的范本。据悉，AI手机助手严格遵循用户授权与合规原则，仅在用户明确授权下调用必要能力；云端处理屏幕内容严格执行“不存储、不训练”准则，数据全程加密传输；涉及支付、密码等高危操作，采用用户授权+手动确认的双重机制，从制度设计与技术实现上，牢牢守住用户权益与安全底线。

今年的政府工作报告，首次写入“完善适应人工智能技术发展促进就业创业的措施”。全国政协委员、中国民营经济研究会副会长李连祥说，目前AI手机等行业处于发展初期，对于新技术形态，应重视人才培养，秉持宽容的态度，但在监管和法律保障上，更要精准防控风险，为技术创新预留足够的空间。

AI越来越聪明的时代，更应是制度有温度、技术有边界、公众有判断的时代。“人工智能的未来，归根结底取决于人类如何使用它。”周鸿祎委员表示，只有坚持以人为本、智能向善，让技术创新与伦理治理同步推进，人工智能才能真正成为推动社会进步、增进人类福祉的重要力量。

据3月11日《新华每日电讯》

AI声音滥用 新华社发 徐骏 作