从“可测”到“可保”：专属保险何以兜住AI风险

■中国城市报记者 邢 灿

人工智能带来前所未有的发展机遇，也带来前所未遇的风险挑战，如算法偏见、内容侵权、数据投毒等。如何补全人工智能安全治理的闭环？人工智能“成长的烦恼”如何破解？

在近期结束的2026金融街论坛年会上，全国首个人工智能应用风险综合保障方案发布。通过保险产品分别对人工智能服务给第三方造成的损失和人工智能供应方自身费用损失进行保障，有效弥补国内AI保险市场在产品供给上的不足，为化解AI风险提供实打实的兜底支撑。

AI风险正从理论走向现实

当前，各类智能产品和服务正加快普及应用，AI手机进入“口袋中”，智能网联汽车跑在“公路上”，人形机器人加快走进“工厂里”。工业和信息化部数据显示，2025年我国人工智能核心产业规模超过1.2万亿元，企业数量超过6200家。

人工智能带来前所未有的发展机遇，也带来不容忽视的风险。屿策智库秘书长侯文杰向记者介绍，这些风险包括：模型输出错误导致业务决策失误，如金融风控模型误判；生成内容引发侵权或虚假信息传播；算法歧视带来合规和声誉损失；系统遭攻击或数据投毒导致服务中断；自动化决策造成第三方人身或财产损害，如自动驾驶事故。

事实上，随着AI技术加速落地应用，AI风险导致经济损失的案例已经出现。中国信息通信研究院人工智能研究所所长魏凯在接受记者采访时举例介绍，国际上曾发生智能体误删生产数据库的事故，也有过自动驾驶事故赔付2.43亿美元的案例。这表明相关风险已经实质化，亟须把“看得见的风险”变成“保得了的产品”。

在此背景下，保险成为化解AI风险的重要途径。中国企业资本联盟副理事长柏文喜告诉记者，AI保险的价值不在“全赔”，而在于通过信用增级、风险定价和保前减损，让模型方敢商用、采购方敢集成。短期看，AI保险是刚需；长期看，它是AI商业化的基础设施。

侯文杰对此持相似观点。在他看来，AI保险价值体现在三个方面：一是填补保障空白，给AI企业提供风险分担渠道；二是通过承保风控，倒逼企业完善AI安全治理；三是增强社会对AI应用的信任，为产业规模化落地提供托底。

“从趋势看，随着AI应用深入和监管要求趋严，AI保险正在从可选配置向准刚需演变，尤其在自动驾驶、医疗AI、金融AI等高风险场景，需求更为迫切。”侯文杰说。

AI保险落地面临三大难题

今年3月，科技部等4部门联合发布《关于加快推动科技保险高质量发展有力支撑高水平科技自立自强的若干意见》，提出在人工智能、集成电路、量子科技、生物制造、氢能和核聚变能、脑机接口、具身智能等重点科技和产业发展领域，鼓励开发科技保险专属产品，探索建立专项风险准备金制度。

政策方向已经明确，但AI保险要真正实现落地开花，仍面临三大难题。

风险评估难。魏凯介绍，AI模型行为随版本快速漂移，导致应用程序可能产生不一致或不可预测的结果。研究显示，仅间隔三个月的两个模型版本，在同一数学任务的准确率从84%降到51%。承保时点的静态评估结论本质上是一张“快照”，难以充分识别AI动态风险。

精准定价难。AI保险绕不开一个核心难题：如何定价？受访专家提到，尽管AI安全评测已有产业基础，但距离保险精算的输入要求仍有差距：一方面通用基准评测结果难以反映模型在真实场景下的能力；另一方面，AI属于新兴风险，存在历史损失样本匮乏、数据碎片化的问题。

定责定损难。一旦发生AI事故，责任归属和损失界定都极其复杂。“当前AI应用层产品百花齐放，模型产品却在快速收敛，单个基座模型可以支撑越来越多的应用。”魏凯提醒，基座模型一旦存在缺陷，可能跨行业、跨地域同时致害，呈现类似巨灾的“聚合风险”特征。

在侯文杰看来，传统责任险以明确过错和可界定损失为基础，网络安全险聚焦数据泄露和网络攻击，都无法覆盖AI特有的模型缺陷、算法歧视、生成内容侵权、自动化决策失误等新型风险。这些风险既不属于典型的网络安全事件，也难以套用产品责任或职业责任的框架，保障范围和定价基础都不匹配。

北京市社会科学院副研究员王鹏告诉记者，AI风险本质上是技术内生的概率性风险，不同于传统保险聚焦的外生意外、硬件故障或人为过失，它自带持续演化、跨主体传导、黑箱不可解释的属性，打破了传统保险依赖的稳定风险分布假设。

王鹏认为，AI风险特点直接改变了承保逻辑。传统静态核保、标准化定价模式不再完全适用，保险人无法再沿用“追溯事故原因”的思路，必须转向以“输出结果是否合规”为核心的判定框架，同时要建立全流程动态风控机制，才能适配AI风险持续迭代的特征。

多方协同筑牢AI“安全网”

如何破解AI保险落地堵点？相关机构已经开展积极探索。

自2025年起，中国信息通信研究院人工智能所联合中国人民财产保险股份有限公司（以下简称“人保财险”）、中国财产再保险有限责任公司探索合作机制，并逐步建立了“技术评估—保险承保—再保分散”的三方协同框架：直保公司据此判断可保性、做差异化定价，再保公司探索统筹分散跨行业聚合敞口的路径，评估则以技术视角把住承保的关口。2026金融街论坛年会上发布的全国首个人工智能应用风险综合保障方案便是三方共同合作的成果。

魏凯介绍，中国信息通信研究院人工智能所围绕“从可测到可保”做了四方面工作。一是建评估底座。自主研发大模型、智能体自动化安全测试平台等技术工具，把评估体系、自动化测试、持续监测和数据沉淀的能力建立在自主底座上。二是做语言转换。把每项技术风险转化为触发条件、发生频率、危害程度等保险关注的要素，让技术语言和保险语言在同一份评估结论里完成转换，通过这种方式直接回应风险看不清、算不出价的问题。三是搭建合作机制。与其他两家单位建立“评估、承保、分散”各司其职的三方合作机制，形成贯穿保前核保、保中监测、出险理赔的服务能力。四是打通全流程能力。保前开展风险评估与分级，支撑核保准入和差异化定价；保中跟踪模型版本变更与行为基线漂移，动态预警；保后复现事故场景，区分模型缺陷、部署不当与使用错误，辅助责任判定与损失核定。

“今年9月以来，我们联合人保财险，为一家企业智能体数字员工产品，率先跑通人工智能应用责任险从风险评估到服务投保的保前全流程，目前该笔业务已完成核保出单。”魏凯说。

魏凯介绍，具体成效体现在三个层面：对投保企业，把难以预测的技术风险转化为可量化、可管理的经营成本，评估报告已反馈企业用于安全整改，客户逐条采纳、起保前已完成整改，这表明评估不再是产品上市的一道手续，而是伴随保险全周期的基础设施；对保险行业，每一次承保合作中评估结论和风险画像都在持续入库，行业最缺的精算数据正在业务过程中逐步沉淀；对保险模式，验证了先评估、后承保在国内可行，保中监测、保后归因机制已建立，将随保单存续持续运转。（刘振峰、侯竞瑜对本文亦有贡献）

第30届北京国际钱币博览会举办

10月30日至11月1日，第30届北京国际钱币博览会在京举办。本届展会吸引332家境内外展商参展，其中境外展商58家，参展规模再创新高；展会期间举办62场活动，为中外钱币机构、行业从业者和收藏爱好者搭建专业展示、交流合作与文化共享平台。

中国城市报记者 全亚军摄



《国家信息化“十五五”规划》发布

■中国城市报记者 邢 灿

近日，中央网络安全和信息化委员会印发《国家信息化“十五五”规划》（以下简称《规划》），对体系化深层次推进“十五五”时期国家信息化发展作出部署安排。

《规划》提出，到2030年，我国信息化取得重大进展，在推进中国式现代化中的作用更加凸显。信息基础设施巩固拓展领先优势，陆海空天立体化网络体系基本形成，新一代通信网全面建成，适时启动第六代移动通信（6G）商用，国家区块链网络基本建成，智能互联网加快布局发展。网信科技创新能力显著增强，科技创新体系更加健全，重点领域关键核心技术取得突破。网信产业生态繁荣发展，数字技术和实体经济深度融合，网信企业综合实力显著增强。经济社会数字化网络化智能化应用水平迈上新台阶，赋能经济社会发展和治理能力明显提升。网络安全保障坚实有力，法规标准体系更加完善，网络空间命运共同体理念深入人心。

《规划》围绕“十五五”发展目标，提出“建设三个体系、提升五个水平、夯实两个支撑”的总体布局，部署了10个方面具体任务，包括建设新型信息基础设施体系、建设网信科技创新体系、建设网信产业生态体系、提升现代治理水平、提升信息惠民水平、提升数字化绿色化协同水平、提升区域协调水平、提升开放发展水平、夯实安全保障支撑、夯实法规标准支撑。