AI造假无孔不入 消费骗局盯上你的钱袋子

■中国城市报记者 朱俐娜

冒用医学专家名义宣传特效药、冒充知名演员推荐减肥药、克隆奥运冠军声音为产品站台——这些看似明星代言的带货场景，背后可能是AI换脸技术打造的虚假人设。

随着生成式 AI 技术的爆发式发展，造假已从传统的图文篡改升级为多感官欺骗，将消费骗局包装成“高科技陷阱”，不仅让消费者防不胜防，更严重扰乱了数字经济的信任基石。

破坏力远超以往

“他跟电视画面里的形象一模一样，耐心讲解着产品功效，怎么可能是假的呢？”72岁的王奶奶在社交平台刷到某位医学专家的视频，出于对专业人士的信任当即下单。直到子女把专家发布的打假声明给王奶奶看后，她才得知自己被AI换脸的虚假视频欺骗了。

类似的案例正在全国各地上演，且不再局限于老年群体。30岁的李女士因产后体重超标，一直寻找安全的减肥方法，偶然在短视频平台刷到一段知名演员推荐减肥药的视频。李女士平时喜欢该演员的影视作品，对她的“亲身推荐”深信不疑。后来，李女士刷到演员本人在社交平台发文称被AI复制其声音带货减肥药，她才知道自己被骗，原视频是商家利用AI技术，将该演员的形象与虚假宣传话术合成。

与传统消费骗局相比，AI造假行为呈现的破坏力远超以往。北京市社会科学院副研究员王鹏在接受中国城市报记者采访时指出，AI造假基于用户数据定制“人设”，利用信任感实施定向欺诈，具有精准心理操控的特点。

除了在消费欺诈中精准拿捏消费者心理，AI造假的滥用还延伸到了电商交易的纠纷场景中。技术的平民化让造假行为愈演愈烈，但识别能力却严重滞后。电商张先生之前碰到过通过AI修图申请“仅退款”的情况。“每件货物都是我亲自检查并打包的，完好无损。但是买家发来的图片显示榨菜包装袋严重破损，榨菜的汤汁洒在箱子底部，看着特别真实。”张先生说，买家利用这些图片申请“仅退款”，即便他提供了完整的封箱视频和物流记录，平台仍因无法判定图片真伪，最终支持买家仅退款。

天使投资人、资深人工智能专家郭涛在接受中国城市报记者采访时表示，对于普通人来说，要识别这类伪造内容极具挑战性。不过，一些专业机构和技术公司正在积极研发检测工具，试图利用AI自身的分析能力来识别异常模式。

AI造假泛滥的背后是成熟且低门槛的技术支撑。当被问及常见的AI造假技术及成熟度时，郭涛说，深度伪造（Deepfake）、语音合成、图像篡改等是当前最为常见的AI造假技术。这些技术已经相当成熟，尤其是深度伪造，通过深度学习算法可以生成极为逼真的人脸替换视频，让人难辨真伪，极大地增强了欺骗性。

郭涛进一步表示，借助深度学习等先进技术，AI能够迅速生成大量定制化内容，提高了制作效率。同时，利用互联网平台的广泛传播力，AI造假能迅速扩散影响范围。此外，技术门槛的降低使得更多人能够接触并使用这些工具，导致参与群体日益多元化，治理难度加大。而且，AI造假往往游走于法律边缘，给现行法律法规带来了前所未有的挑战。

治理面临多重难题

随着越来越多企业入局，虚拟数字人24小时带货早已不是什么新鲜事。AI技术带来的降本增效让企业难以拒绝，但风险也如影随形。

当谈到如何平衡“AI技术提升消费体验”与“防范AI造假风险”时，中国城市发展研究院投资部副主任袁帅向中国城市报记者分析称，核心矛盾在于技术工具的中立性与商业逐利本性之间的冲突。一方面，AI能为企业带来降本增效的革命性体验，放弃AI意味着在激烈的市场竞争中掉队；但另一方面，AI强大的生成能力极易越界，企业可能为了短期GMV增长而默许甚至主动利用AI制造虚假繁荣，如刷单、虚构用户评价、夸大产品功效，或者企业自身成为被攻击目标，品牌形象被AI换脸视频瞬间摧毁。

这种矛盾让企业在“用”与“管”之间左右为难。在袁帅看来，若严格管控，需要投入高昂的人力物力建立审核机制，甚至牺牲部分业务灵活性；若放任自流，则面临监管重罚和品牌崩塌的风险。

郭涛认为，解决这一矛盾的关键在于建立健全的数据保护机制，确保用户信息安全；同时加强内部监管，防止技术滥用。此外，公开透明地展示AI决策过程，增强用户对技术的理解和信任，也是缓解矛盾的有效途径。

此外，防范AI造假还需直面一个更核心的挑战：其治理难度，远超传统消费骗局，治理环节的复杂性也让问题解决难上加难。

当前AI造假治理面临多重突出难题。王鹏称，平台、技术服务商、商家之间的责任边界不清，出现问题时易出现推诿扯皮；黑产借助境外服务器规避监管，给跨境执法带来挑战；而个人维权又面临取证难度大、诉讼周期长的困境，增加了维权成本，这进一步加剧了治理的复杂性。

全联并购公会信用管理专委会专家安光勇向记者表达了相似观点，一方面是技术攻防失衡，AI生成技术快速迭代，防御手段滞后，尤其在深度伪造和数据篡改领域，现有的检测工具难以跟上技术发展的步伐；另一方面是责任链条碎片化，AI造假涉及从技术提供方到内容发布方的多个环节，导致责任难以明确追究；同时还存在跨平台监管难题，不同平台的审核标准和技术手段不统一，造成监管真空，黑灰产可以利用不同平台之间的监管漏洞进行跨平台传播。最终使得治理工作无法形成有效的统一监管和跨平台协作，整体仍处于分散应对的状态。

更好地服务消费市场

遏制AI造假，绝非单一力量可为。业内人士普遍认为，这需要消费者、企业、监管部门、技术领域形成多方联动的治理体系。

就在近期，国家市场监督管理总局、国家网信办联合印发的《直播电商监督管理办法》提到，直播间运营者使用人工智能等技术生成的人物图像、视频从事直播电商活动的，应当符合有关法律法规、规章和强制性国家标准的要求，并依照国家有关规定进行标识，持续向消费者提示该人物图像、视频由人工智能等技术生成。

这意味着，人工智能生成的数字人主播被纳入监管。

企业和技术领域层面，袁帅表示，企业作为技术应用的主体，必须承担“守门人”责任，不仅要自律不使用AI造假，还需投入资源研发反造假技术，建立内容溯源机制，对平台上的AI生成内容进行显著标注和严格审核，从源头切断AI造假的传播链。技术领域特别是AI研发机构和安全公司，应肩负起“技术赋能”责任，致力于开发更精准的AI生成内容检测算法，推广隐私计算和区块链存证技术确保内容可追溯，并在模型训练阶段就植入安全约束。

消费者层面，袁帅认为，消费者作为市场的最终裁判，需承担起“第一道防线”的责任，提升数字素养和风险意识，不轻信“完美”的宣传，遇到可疑内容主动举报，并在权益受损时积极利用法律武器维权，倒逼市场净化。

针对消费者识别造假内容的需求，多地陆续推出实际的技术手段，比如厦门、杭州、西宁等地已接入美亚鉴真平台，用户可以通过政务公众号跳转，上传疑似伪造内容进行快速核验，操作便捷且免费使用。

从长远来看，AI技术如何更好地服务消费市场，同时避免其成为“造假工具”？郭涛称，要让AI技术正面服务于消费市场，首先必须确立伦理指导原则，引导技术开发朝着有益于社会的方向发展。其次是完善相关法律法规，设立清晰的红线，严厉打击任何企图利用AI从事非法活动的行为。再者，鼓励行业内开展自律行动，比如成立联盟组织，共享最佳实践经验。最后，加大对公众科普力度，普及关于正确使用AI的知识点，培养大家理性看待这项新技术的态度。这样，我们就能最大限度地发挥出AI的优势，同时又能有效规避它所带来的负面影响。

袁帅强调，在技术层面，应大力发展“可解释性AI”和“可控生成技术”，研发具有自毁机制或溯源标记的生成模型，确保每一段AI生成内容都带有无法消除的“数字链痕迹”，实现从生成端到传播端的全生命周期可追溯。在市场生态层面，应培育“真实优先”的消费文化和行业标准，鼓励第三方认证机构对“无AI掺杂”的真实内容进行背书，利用区块链技术构建去中心化的信任体系，让真实数据产生溢价。

贵州黔西：普惠托育惠民生

近日，在贵州省黔西市莲城街道托育中心，老师带领托育班的孩子游戏。近年来，黔西市莲城街道聚焦“3岁以下婴幼儿照护”民生实事，积极推动社会力量共同兴办发展托育教育服务，通过落实落细健康保健、资金补贴、志愿服务等多层面扶持政策，加快建设普惠托育服务体系，全力打造“15分钟托育圈”，让群众就近享受多元、可及、优质的普惠托育服务。

人民图片