DOI: 10.16538/j.cnki.jsufe.2025.05.010

具身智能体数据隐私风险的合作治理

李 智，陈盈盈
（上海大学 法学院，上海 200444）

摘 要：具身智能体具备感知交互、自主决策和信任匹配等核心特征，在展现强大能力的同时，也加剧了诸如数据泄露、物理侵犯和决策失控等风险。具身智能体将机器与数据结合的特性不仅会使个人信息控制机制难以发挥作用，还可能会让责任追究难以落实，无法有效保障相关主体权益。针对这些数据隐私风险和法律适用困境，需要从法律、技术等多层面协同发力：在权责体系层面，完善数据使用的监督体系和利益相关者的责任体系，形成与“权利束”相对应的“义务束”；在市场准入层面，健全符合具身伦理的算法影响评估机制和满足合规要素的信息保护认证机制，形成契合风险预防主义的准入制度；在技术嵌入层面，构建具有行业与制度约束力的匿名技术规则、数据分类分级规则及用户动态授权管理规则，促进技术与法律的融合。

关键词：具身智能体；数据隐私；法律治理；技术治理
中图分类号：D912.28 **文献标识码：**A **文章编号：**1009-0150(2025)05-0138-15

一、问题的提出

具身智能体（Embodied Intelligent Agent）[①]是指通过多模态传感器动态采集环境数据、基于算法自主完成行为决策的智能系统，其核心能力在于融合物理实体、感知交互与自主决策。根据应用场景和形态差异，具身智能体可分为工业机器人、服务机器人和特种机器人等类型。

技术迭代始终伴随着工具属性的拓展与治理逻辑的调适。[②]从工业自动化到具身智能体的演进，核心差异体现在“机械执行”“自动决策”与“自主决策”的分野中，三者既存在技术代际递进，又在风险形态与治理需求上呈现本质区别。“自动化”以预设程序实现机械动作标准化复现，其本质是工具的延伸，不具备自主判断能力。这一时期，工业机器人作为自动化技术的典型载体，聚焦机械结构优化（如关节灵活度、运动稳定性提升）。自动化通过替代人力实现生产效率跃升，但其风险仅限于物理安全范畴，此阶段的治理逻辑以技术安全为核心，依赖行业标准规范机械运动边界，尚未涉及数据与算法的深层规制。随着算法技术的发展，“自动决策”逐渐从自动化中分化，核心是通过数据分析实现非预设场景的适应性判断，但仍未突破“虚拟决

收稿日期：2025-06-12
基金项目：国家社会科学基金一般项目“企业数据确权和流转中的登记问题研究”（23BFX075）。
作者简介：李 智（1968— ），女，重庆涪陵人，上海大学法学院教授、博士生导师；
陈盈盈（2000— ），女，河南夏邑人，上海大学法学院硕士研究生。

①1950年，艾伦·图灵作为人工智能领域的奠基人，在其文章“Computing Machinery and Intelligence”中首次提出了具身智能（Embodied Intelligence）的概念，并提出了人工智能发展的前瞻性构想：一条路径聚焦符号逻辑运算（如棋局推演等抽象智能），另一条路径则主张为机器搭载先进感知设备，通过自然交互方式模拟人类婴幼儿认知发展模式。

②陈兵、董思琰：《人工智能通用大模型市场竞争治理架构探设》，《上海财经大学学报》2025年第3期。

策”范畴。其以“数据输入-模型运算-结果输出”为闭环，依赖大数据训练的算法模型优化决策，价值在于提升效率、拓展服务范围，风险集中于数据层面，治理焦点开始转向数据合规，但因无实体交互，尚未涉及物理伤害风险。具身智能时代的突破性在于将“算法自动决策”与“物理实体行动”深度绑定，形成“环境感知-算法决策-实体执行”的全链路闭环，其风险与价值均源于“虚实结合”的特有属性。相较于自动化与纯算法决策，其风险呈现叠加特征：一方面，其通过实体多模态交互隐蔽采集数据并分布式存储；另一方面，算法决策失误可能直接操纵智能体转化为实体伤害。因此，具身智能体的治理需超越单一维度，既要延续对物理安全的规制，又要强化数据全生命周期管控，更要针对“算法决策-实体行动”的耦合风险建立新型规则。

从自动化到具身智能的演进，本质是技术从“工具属性”向“类主体属性”的跨越，这种跨越在提升社会生产力与服务便利性的同时，也使风险从“单一领域”向“多维度耦合”升级，对法律规制与协同治理提出了挑战。从工业制造到家庭服务，[①]从应急救援到医疗辅助，具身智能体的应用场景不断拓展，其实体嵌入属性、自主决策能力和多元化的应用生态，给我国相关产业的监管与治理带来了独特的风险与挑战。[②]技术迭代倒逼规制逻辑从“事后救济”向“风险预控”转变。随着具身智能体从工具属性向准主体属性延伸的讨论，传统基于“人机分离”假设的治理框架面临挑战：一方面，多模态数据的实时采集与边缘存储打破了数据控制的时空边界；另一方面，自主决策的黑箱特性使得责任追溯难度陡增。这种转型在全球范围内呈现差异化路径：欧盟通过《人工智能法案》（Artificial Intelligence Act）构建“风险分级-义务递进”的刚性框架，按应用场景设定合规义务；我国则相继发布《具身智能发展报告》《人形机器人分类分级应用指南》等文件和研究报告，采用“场景类型-风险等级-规制强度”的动态适配模式，在工业、服务等场景中嵌入差异化合规要求。

在实践中，具身智能体的技术特性与法律规范之间仍存在显著张力，以我国《个人信息保护法》第47条为例，从法教义学视角观察，数据删除权的法律适用在具身智能场景中面临规范漏洞：该条款预设数据存储具有集中化、静态化特征，要求“个人要求删除的，数据处理者应当及时删除”，但具身智能体通过边缘计算实现的分布式存储形成数据碎片化留存，导致删除指令难以覆盖全链路。当前，我国对具身智能体的法律治理研究主要存在两种路径：一是对多模态数据采集合规、[③]算法安全风险控制、物理安全防护等具身智能聚合的技术与法律风险，选取其一展开讨论；[④]二是关注具身智能体应用落地中可能引发的新型数据泄露、物理侵犯隐私风险、人身安全责任认定等问题。[⑤]

从综合治理的角度看，具身智能体是集感知、决策、执行于一体的复杂系统，其运作涉及庞杂的数据流与物理交互，当前研究尚需对具身智能体引发的系统性风险（如多智能体协同数据流转、人机混合责任边界）进行整体梳理与体系化分析。本文试图结合我国具身智能产业发展及治理的实际状况，从具身智能体的核心特征与风险生成机理出发，系统解构其衍生的隐私侵

①叶劲松、方嘉彬、黄远浙、朱艺蕾:《工业机器人应用与家庭消费结构升级——基于创收水平与社会网络冲击视角的分析》,《上海财经大学学报》2024年第5期。

②陈兵、董思琰:《人工智能通用大模型市场竞争治理架构探设》,《上海财经大学学报》2025年第3期。

③陈柳钦:《 中国人形机器人产业的技术自主化、应用场景化与生态系统化》,《贵州师范大学学报（社会科学版）》2025年第4期; 赵精武:《论人工智能训练数据高质量供给的制度建构》,《中国法律评论》2025年第1期。

④曾聪:《生成式人工智能与“不可解释性”危害结果的客观归属》,《华中科技大学学报(社会科学版)》2024年第5期;余雅风、王朝夷:《由技术伦理向法律规范演进:国外人工智能应用规范研究综述》,《河北法学》2023年第2期。

⑤王雪琪:《数据要素赋能新质生产力的法律规制路径》,《广东财经大学学报》2024年第6期;祝高峰:《论人工智能领域个人信息安全法律保护》,《重庆大学学报(社会科学版)》2020年第4期。

权与安全挑战，检视现有法律规则在数据控制机制、责任分配规则、市场准入制度等方面的适配性困境，借鉴域外治理经验，以促进产业健康发展与保障公民权益为双重目标，提出融合“权责体系重塑-风险预防准入-技术规则嵌入”的三位一体协同治理框架，为平衡具身智能技术创新与隐私安全保护提供兼具前瞻性与可操作性的制度方案。

二、具身智能体利用数据隐私的主要风险维度

具身智能行业已从研发、适用逐步进入商用实践，相关数据隐私保护的潜在风险亟待厘清。探究具身智能体的数据隐私风险可以从两个维度入手：一是技术维度，基于具身智能体技术特征衍生的多元隐私风险；二是制度维度，现有数据与隐私保护规则因适配性不足导致的规制失灵风险。

（一）具身智能商用的主要风险维度

具身智能体作为具身智能技术的实体呈现，以人形机器人为代表，其特征体现在三方面：借助传感器直接感知环境与对象的感知交互能力，[①]基于具身认知分析环境信息并依此行动的自主决策能力，更便于融入人类社会环境的拟人化外观。这些特征使其区别于纯软件AI系统或非具身机器，衍生出特有隐私风险，其核心差异在于：具身智能体通过“物理实体-数据处理-社会交互”的三重耦合，将数据风险从虚拟空间延伸至物理空间，从被动收集升级为主动侵入，从技术层面渗透至伦理层面。

1.感知交互维度：实体移动性与多模态采集的叠加风险。具身智能体的动态感知特性，意味着它在持续运动和交互中不断积累数据，[②]数据的规模和敏感度都在不断增加，这无疑加大了数据泄露的风险。

（1）多传感器信息采集加剧隐私担忧。为实现与人类的高效交互，具身智能体集成了多种前沿传感器与高性能处理器。这些配置极大地增强了其对环境信息及个人信息的采集和存储能力。通过先进的传感器，具身智能体能够精准捕捉周围环境的细微变化，包括光线、声音、温度等信息；在与人类互动时，还可收集诸如语音内容、面部表情、肢体动作等个人信息，并利用处理器强大的运算能力进行记录和分析。这种软硬件协同配置在提升人机交互体验的同时，亦在法律层面引发数据保护与隐私保护的双重挑战。

（2）自主移动特性催生物理侵犯危机。具身智能体的自主移动特性使其能够突破传统设备的被动交互模式，实现对物理环境的主动适应与动态响应。然而，自主性在提升交互效率的同时，也在物理空间维度引发了新型法律风险。在隐私保护方面，具身智能体的自主移动能力使其得以突破传统物理空间的隐私保护边界。例如，家庭服务机器人在搭载摄像头或激光雷达系统时，其在自主清洁过程中可能未经授权进入卧室等私密空间，持续采集空间信息及自然人私密活动，实质上构成对《民法典》第1032条规定的“私密空间安宁权”的侵犯，打破了传统物理空间的隐私隔离状态。在安全威胁层面，具身智能体的自主移动与物理干预能力叠加，形成对人身与财产安全的潜在风险。由于路径规划算法的局限性或传感器感知误差，机器人在自主移动过程中可能发生碰撞、剐蹭等意外事件，或受恶意指令操控故意破坏设施。

2.自主决策维度：物理行动绑定与分布处理的失控风险。具身智能体在进行决策时，需要调用和分析大量的数据。自主决策能力虽然让具身智能体能够灵活应对各种情况，但也为数据

①张昌盛：《从具身智能到具身智能体》，《北京工业大学学报(社会科学版)》2024年第6期。

②陈兵、林思宇：《数字经济领域数据要素优化配置的法治进路——以推进平台互联互通为抓手》，《上海财经大学学报》2022年第3期。

隐私安全带来了不确定性。

（1）不可预测的决策行为产生侵权风险。关于机器学习算法的特征，瑞恩·卡洛（Ryan Calo）使用了“涌现（Emergence）”这一概念，即其往往会给出一些具有创新性或者出人预料的输出[①]。相较于传统工业品，以大模型为技术基底的人工智能在内容生产机制上缺乏明晰的内容生成规程与标准化内容产出，指令输入到结果生成的中间处理流程处于不透明状态。具体而言，每次内容生成均存在差异性，虽因算法逻辑呈现总体规律性与稳定性，但生成过程的具体路径仍具有不可知性，用户难以精准控制内容，“涌现性”输出无法通过预设程序实现精确预测和控制。这种涌现性赋予具身智能体自主能力，数据收集阶段可能违背用户预期收集隐私数据，数据使用环节若决策失灵，可能将收集的数据用于未经授权的目的，如将用户的生活习惯数据出售给第三方用于精准广告投放，侵犯用户的隐私权。当具身智能体的决策不可预测且失灵时，可能对人身和财产安全造成威胁。

（2）非线性处理模式加剧数据管控难度。具身人工智能的数据处理已突破传统“收集-分析-决策-使用”的线性模式，[②]呈现出动态耦合与非线性交互的特征。在数据生成阶段，其模糊了数据收集和处理的界限。例如，通过传感器收集环境数据时同步进行数据处理和特征提取等预处理操作，这种“采集即处理”的机制使得基于线性流程的传统数据管控方式难以介入。由于数据处理链条呈现高度交织的网状结构，各环节间的因果关系缺乏单向传导性，数据处理者难以追踪数据在不同阶段的完整流向与形态变化，[③]导致传统的数据溯源与过程监管机制失效，削弱了现有隐私保护法律的规制效能。同时，分布式架构下各智能体间的局部信息交互与自主决策，可能触发涌现效应——即使单个智能体的算法看似安全合理，但组合交互后可能产生超出预期的结果。在具身智能体执行任务时，各智能体间的数据交互和决策联动的复杂性以及外部环境的动态干扰，使数据被误处理、泄露或滥用风险指数级增长，而现有以单一算法、模型和主体为基础的治理框架难以适应多智能体系统中数据交互的复杂性与动态性，使得数据管控缺乏有效的应对策略。

3.信任匹配维度：信任诱导与情感交互的隐私泄露风险。相较于无形的系统，人们往往更愿意信赖并青睐以实体形式（多为人形机器人形态）呈现的人工智能。[④]拟人化外观使得具身智能体更容易获得人们的信任，人们在与它们交互时可能会不自觉地透露更多隐私信息。

（1）拟人外观与行为交互诱导的信息泄露风险。人类对于人形物件具有自然的心理投射和共情倾向。当具身智能体拥有高度拟人外观时，更容易被人类接受并被视为同类。研究表明，当机器人外观与人类高度相似时，用户会不自觉地将其视为社会实体，甚至产生情感依赖，从而更愿意分享敏感信息。[⑤]例如，在与陪伴机器人或亲密机器人互动时，人们可能会有意无意地放下防备，向其倾诉日常生活中的各种信息，包括个人隐私、情感经历、家庭状况等，从而导致信息泄露。心理治疗机器人（如Mirokaï[⑥]）通过拟人化外观和共情式对话，使用户误认为其具

①Ryan Calo,Robotics and the Lessons of Cyberlaw. California Law Review, 2015, 103(3), pp.532-542.

②Solove Daniel J, Artificial Intelligence and Privacy. Florida Law Review, 2025,77(1),p.35.

③马更新:《数据交易中个人信息保护制度之完善——以“知情-同意”规则为核心》,《河北学刊》2024年第2期。

④Asada M , Hosoda K , Kuniyoshi Y ,et al., Cognitive Developmental Robotics: A Survey. IEEE Transactions on Autonomous Mental Development, 2009, 1(1), pp.12-34.

⑤［美］理查德·扬克:《机器情人：当情感被算法操控》,布晚译,文汇出版社2020年版,第5页。

⑥由法国机器人初创公司Enchanted Tools推出的人形机器人Mirokaï是个有背景故事的可爱机器人，其有着半孩童半动物的形象，且被赋予了强大的背景故事和令人信服的角色设计。

有同理心，进而主动透露心理健康状况、家庭矛盾等私密信息。此类机器人利用柔性电子技术模拟真实面部表情（如微笑、皱眉）[①]，进一步强化用户的情感信任，形成“心理投射-信息共享”的闭环。具身智能体还能通过模仿人类社交行为（如点头、眼神接触）营造亲密感。

（2）情感计算驱动的个性化分析深度挖掘隐私。具身智能体基于算法驱动的个性化信息处理机制，依托“信息筛选-定向推送-行为引导”的技术链条，构建起符合“轻推（Nudge）”[②]理论的决策影响范式。从技术实现层面看，具身智能体通过自然语言处理、情感计算等技术，在对话场景中实施动态的信息策略干预。其运用语义理解模型识别用户话语中的敏感关键词，结合预设的引导脚本，通过开放式提问、共情回应等交互策略，诱导用户进行深度自我表露。例如，在健康咨询场景中，机器人可能以“最近睡眠质量是否影响工作效率？”等递进式提问，引导用户主动透露失眠频率、用药情况等医疗隐私；在消费场景中，则通过“您上次购买户外运动装备后，是否有尝试新的运动方式？”的关联性询问，挖掘用户的消费偏好与社交活动信息。这种交互模式利用人类认知中的自我表露互惠效应，在看似平等的对话中完成隐私信息的隐秘采集。从法律视角看，上述行为已突破传统隐私侵权的认定边界。具身智能体通过诱导式交互获取隐私信息，实质上构成新型刺探行为。其利用用户对智能体的信任错觉，规避了《个人信息保护法》第13条关于“取得个人同意”的合法性要件。特别是当具身智能体将采集的碎片化信息进行跨维度分析时，可能触发再识别风险——通过关联用户的消费记录、地理位置等数据，反向推导出医疗健康、政治倾向等高度敏感隐私。

（二）规则弱适配性的主要风险维度

人工智能对数据的持续依赖，使隐私法在数据收集、处理、救济等环节面临规则适配性风险。[③]既有隐私与数据保护规则以静态数据控制为核心框架，[④]但具身智能体的动态实体交互特征使其难以覆盖。

1.泛知情同意模式下的数据管理机制失衡。具身智能体场景下，知情同意原则的失灵并非单纯技术问题，其深层根源在于责任体系的结构性缺陷——现有规则未能针对具身智能的动态交互特性构建精准的责任分配与约束机制，导致“同意”的法律意义被技术特性架空。一方面，拟人化交互与动态场景导致的“同意有效性”责任真空。具身智能体通过情感反馈、类人行为获取用户信任，可能诱导老人、儿童等群体作出非理性授权，而现行责任体系未明确开发者对交互透明度的保障义务。同时，动态环境下的数据生成与处理界限模糊（如SLAM技术[⑤]同步采集与建模），用户无法预判数据收集范围，传统一次性授权本质上暴露了生产者对场景化授权机制的设计责任缺位——既未预设动态授权触发条件，也未建立用户实时干预通道，最终使“同意”沦为形式化的法律要件，而非实质的权利保障工具。另一方面，目的限制与最小必要原则的突破，折射责任约束的乏力。具身智能体依赖多模态传感器持续收集环境与行为数据，其

①Dai N , Zhang K , Zhang F ,et al., AI-assisted flexible electronics in humanoid robot heads for natural and authentic facial expressions. The Innovation, 2025, 6(2), pp.1-3.

②该理论由经济学家理查德·塞勒（Richard Thaler）和法律学者卡斯·桑斯坦（Cass Sunstein）于2008年提出。其指在不妨碍人们自由选择的条件下，对人们的决策进行鼓励或诱导，以引导人们做出特定选择。

③Solove Daniel J, Artificial Intelligence and Privacy. Florida Law Review, 2025,77(1),pp.1-73.

④任愿达:《〈民法典〉个人信息保护规定与数据资产治理观念的协调路径》,《西南民族大学学报(人文社会科学版)》2022年第6期。

⑤SLAM（Simultaneous Localization and Mapping，同步定位与地图构建）技术是机器人和自主移动设备在未知环境中进行自我定位和环境地图构建的关键技术。这项技术通过使用各种传感器收集周围环境的信息，并利用算法将这些信息融合起来，以确定机器人的位置并构建环境地图。这项技术为人形机器人提供了强大的环境感知能力，使它们能够在复杂的环境中自主导航和操作，极大地扩展了机器人的应用范围和灵活性。

自主行动需实时调整数据处理目标，导致数据处理目的超出初始设定，但责任体系未针对这种“功能必要性”扩张设置对应的责任门槛：其一，开发者未被施加目的变更预警义务，用户无法及时知晓数据用途的实质性调整；其二，生产者对冗余数据的清理责任缺失，机器人为算法优化留存的非必要数据缺乏强制删除机制。这种责任分配的失衡，使得《个人信息保护法》第6条的原则性规定因缺乏具体责任载体而难以落地，最终形成技术必要性对法律原则的实质架空。

2.技术“溯及力”背后的数据删除机制失序。现行数据删除机制因技术架构与法律设计缺陷难以有效落实。具体而言，具身智能体数据生成与处理动态交织，数据可能存储在本地、云端或在多个系统间流转共享，分布式存储使得用户难以追踪数据存储位置与流转路径，行使删除权时面临数据不可及困境，《个人信息保护法》第47条关于数据删除的规定无法实现；多智能体协作场景下，数据跨系统共享与融合使得删除请求难以覆盖全部副本，存在碎片化残留风险。这种技术实现与法律条文的脱节可通过“算法阴影（Algorithmic Shadow）[①]”理论解释：具身智能体的自主决策依赖于离线强化学习形成的数据库，其数据处理过程具有“感知-学习-遗忘”的动态性——算法会自动保留高频交互数据以优化决策模型，同时删除低价值数据，但这种技术层面的选择性留存不受法律上删除权的直接规制，形成规制盲区。当用户要求删除特定时段的交互数据时，机器人可能因算法记忆机制仍留存该数据的特征提取结果，导致形式上的删除行为与实质的隐私保护效果背离。此外，具身智能体的数据处理涉及复杂的技术架构和算法：一方面，为了实现智能化功能，数据可能被整合、加密或转化为其他形式存储。例如，一些人形机器人利用机器学习算法对用户行为数据进行训练，这些经过处理的数据难以简单地从系统中剥离和删除。另一方面，部分机器人可能存在技术漏洞或设计缺陷，使得数据删除操作异常，导致数据丢失或损坏。数据加密存储场景中若密钥管理不善，即使执行删除操作，数据可能无法真正从存储介质中清除，仍然存在被恢复和滥用的风险。

3.各主体法律适用中的数据责任机制失调。具身智能体的法律人格争议与责任分配规则缺位，导致侵权责任归属陷入“主体-工具”二元困境。2017年，沙特阿拉伯授予机器人“索菲亚”公民资格事件，促使法学界对具身智能体能否拥有独立主体地位这一关键问题，展开了更为深入、全面的思考[②]：肯定观点主张赋予其类主体地位[③]，否定观点则坚持其工具属性[④]。在法律主体责任说之外，学界还衍生出产品责任说、高度危险说、用人者责任说等具有代表性的观点。[⑤]而现行法律未对此明确界定，导致责任认定缺乏清晰依据。这种争议在责任主体识别与责任分配两方面尤为突出。

具身智能体的侵权行为可能涉及开发者、制造商、运营商、用户等多方主体，黑客攻击损害发生后，责任分配需穿透“算法黑箱（Algorithmic Black Box）[⑥]”与供应链复杂性。从责任主体类

①从法学领域来看，算法阴影是指算法在自主运行中产生的、超出人类直接控制和法律明确规制范围的隐蔽数据处理或决策行为。

②郑志峰：《人工智能立法的一般范畴》，《数字法治》2023年第6期；周详：《智能机器人“权利主体论”之提倡》，《法学》2019年第10期。

③张玉洁：《论人工智能时代的机器人权利及其风险规制》，《东方法学》2017年第6期；刘宪权、胡荷佳：《论人工智能时代智能机器人的刑事责任能力》，《法学》2018年第1期。

④张力、陈鹏：《机器人“人格”理论批判与人工智能物的法律规制》，《学术界》2018年第12期；解正山：《对机器人“法律人格论”的质疑——兼论机器人致害民事责任》，《暨南学报(哲学社会科学版)》2020年第8期。

⑤彭中礼：《新兴技术推动法理论变革的因素考量——以人工智能产品侵权责任分配理论为例的反思》，《甘肃社会科学》2022年第4期。

⑥算法黑箱（Algorithmic Black Box）的概念由美国法律学者弗兰克·帕斯奎尔（Frank Pasquale）在其2015年出版的著作《黑箱社会：控制金钱与信息的隐秘算法》中首次系统提出。这一概念揭示了当代社会中算法决策的核心困境：算法的运作逻辑和决策过程对公众甚至开发者而言高度不透明，导致其结果难以被理解、解释和监督。

别化来看，具身智能体的责任链条涉及全产业链多环节主体，需结合技术特性与应用场景动态界定。例如，若攻击源于硬件设计缺陷，责任可能归于制造商；若因软件漏洞引发，则开发者或运营商需担责。具身智能体区别于传统机器，其行为不仅受使用者直接控制，还依赖于机器学习算法的自主决策能力。这种人机混合控制的技术特性，进一步增加了过错认定的复杂性。现行法律体系的适配性缺陷主要体现在两方面：一是未区分产品与服务的责任差异，对销售企业的告知义务、服务企业的动态责任缺乏明确规范；二是未衔接风险分级理论，对高风险场景（如医疗）的无过错责任与低风险场景（如家庭清洁）的过错责任未作区分。欧盟《人工智能法》虽构建“风险分级-义务递进”框架，但未明确责任比例。我国现有规则仍依赖过错责任与产品责任双轨制，难以应对算法黑箱导致的过错认定难题，亟须针对不同主体与场景细化责任规则。

三、具身智能体利用数据隐私的法律保护困境

当前，具身智能体数据隐私风险的提炼聚焦两方面：一是具身智能与一般大模型在数据泄露可控性及规制方式上的差异，背后是检讨不同人工智能场景在法律规制维度内是否具有内在一致性；二是法律与技术价值对齐能否解决数据隐私风险。然而，当涉及技术的法律事实难以与法律要件对应时，单纯认知协调无法规避风险。对具身智能体实施法律保护，既是人格权保障的内在要求，因生物识别等数据关联隐私权与个人信息权益；也是应对技术风险的现实需要，其“虚实结合”特性使风险更复杂；更是平衡技术创新与权益保护的制度必然，否则可能加剧“技术失控-权利受损-创新受阻”的恶性循环。因此，需在法律与技术的耦合中深究治理困境。

（一）“情境+风险”的法律适用困境

具身智能体数据隐私保护的核心矛盾，集中体现为权义责配置失衡与利益相关者行为失序。前者源于现有法律规则对动态数据处理场景的适配不足，导致权利救济乏力与责任认定模糊；后者则因市场主体缺乏协同治理动力，引发系统漏洞与网络攻击交织的风险传导，二者共同构成数据隐私保护的基础性障碍。

1.权义责配置失衡的规则适配缺陷。在现有法律制度框架下，具身智能体数据治理与责任认定体系存在缺位，这一问题在数据删除义务履行与法律主体界定两个维度尤为突出。

在数据删除义务层面，现有法律虽确立了个人信息主体的删除权，但针对具身智能体的数据处理缺乏可操作的技术规范。《个人信息保护法》第47条虽规定了删除权，但未明确规定具身智能体在数据删除过程中应达到的匿名化标准。欧盟《通用数据保护条例》（GDPR）第4条对匿名化作出定义，要求数据匿名化需达到“无法通过合理手段重新识别数据主体”的程度，但我国尚未建立类似的技术验证体系，导致实践企业可能以“系统架构复杂”“数据关联度高”等理由规避删除义务，加剧了权利救济的不确定性。

在法律主体与侵权责任层面，具身智能体的责任链条因涉及多环节主体且场景差异显著，现有法律框架难以适配。传统法律体系基于“主体-客体”二元结构构建，自然人与法人作为法律主体需具备权利能力、行为能力与责任能力。伴随科技迅猛发展，具身智能体逐步走入大众视野，其智能化程度持续攀升，功能亦日趋多元，在社会中所扮演角色的复杂程度与日俱增。这类机器人不仅能够模拟人类的外形、动作，还能在一定程度上进行自主决策，与周围环境展开互动。但遗憾的是，我国现有法律并未对其法律地位予以明确规定，具身智能体到底是一种特殊的物，还是具有某种特殊法律属性仍在讨论中。我们认为，具身智能体虽具备一定程度的

环境感知、自主决策与执行能力，但其既无自然人的生物学属性，也不具备法人的组织形态与责任财产，难以纳入现有主体框架。当侵权事件发生时，责任链条呈现显著的复合性特征：算法设计者可能因训练数据偏差导致智能体产生歧视性决策，制造商可能因传感器精度缺陷等产品质量问题引发感知错误，所有者可能因指令设置不当或维护缺失扩大风险，实际使用者则可能因违规操作触发侵权行为。然而，现行法律体系在创设之际，并未全面且深入地考虑到因人工智能技术发展而涌现的这些新型责任主体。此外，特殊场景下的责任认定更凸显现有法律的局限性。当具身智能体销售后完全断联，责任主体限于使用者与生产企业——使用者对擅自破解数据采集限制导致的泄露负责，生产企业仅对出厂硬件缺陷担责，此时因无数据流转，通常不产生侵犯数据隐私的结果。而在多主体协同侵权场景中，算法设计者的逻辑缺陷、制造商的硬件瑕疵、销售商的告知不足可能共同导致损害，现行《民法典》侵权责任编的过错责任与产品责任双轨制难以界定各主体的责任比例，更无法应对算法黑箱导致的主观过错认定难题。

2.利益相关者逆向选择的风险传导。在现有市场环境下，具身智能体数据控制端面临系统固有漏洞与网络攻击双重隐患，二者交织可能引发多元风险，直接冲击数据隐私保护的完整性与安全性。

一是隐私数据泄露风险。不法分子可利用技术漏洞侵入系统，非法获取具身智能体存储或传输过程中的各类隐私数据（如用户基本信息、行为习惯、生物识别信息等）。由于大数据时代数据的强关联性，即便单一非敏感信息也可以通过交叉分析还原完整隐私画像。从法律角度看，此类行为不仅侵犯公民隐私权和个人信息权益，更违反了《民法典》《个人信息保护法》等规定，构成对个人信息处理规则的挑战，导致用户隐私在未授权情况下被暴露或滥用。二是隐私窥视与衍生侵权风险。具身智能体的视听功能在为用户提供服务的同时，也可能成为隐私泄露的窗口。黑客通过远程操控具身智能体的视听传感器，可非法采集私密空间影像，形成“非授权监控”，传感器被劫持后，私人生活场景暴露可能引发敲诈勒索等行为，既直接突破隐私保护的物理与数据边界，又对受害者的精神安宁与财产安全构成威胁，导致隐私权益在技术滥用中遭受多重损害。三是隐私关联的安全威胁风险。恶意代码植入可使具身智能体的行动模块（如机械臂、移动底盘）被操控，间接加剧隐私保护的脆弱性。2024年宾夕法尼亚大学的研究人员发现，一系列AI增强的机器人系统很容易受到越狱和黑客攻击，极端案例中甚至出现设备被用于威胁人身安全的情况。人形服务型机器人广泛应用于工业、家庭、医疗等领域，一旦被黑客操控，可能对用户的人身安全和财产造成直接威胁。在工业场景中，设备操控权被夺可能引发生产事故与巨额财产损失；在医疗场景中，若操控医疗辅助具身智能体的程序被黑客入侵，可能会影响其对患者的护理操作，危及患者生命健康。这不仅侵犯了公民的生命健康权和财产权，更通过破坏数据隐私的信任基础，扰乱数字社会的正常秩序与治理生态。

（二）大模型架构的技术治理难题

具身智能体对大型模型的依赖，使其从预训练阶段到模型运算，再到最终输出均存在数据泄露风险。

1.现有监管应对大模型机器幻觉乏力。大模型幻觉是指大型语言模型（LLMs）或生成式人工智能在输出内容时，生成与事实不符、逻辑混乱或虚构的虚假信息。从现行法律监管来看，以欧盟《人工智能法案》（Artificial Intelligence Act）为代表的“预设风险清单”监管模式，难以覆盖其高度的不可预测性与场景依赖性，传统分类监管模式难以覆盖。此外，我国《个人信息保护法》第24条要求的“算法解释权”在大模型场景中近乎虚置。大模型涉及的参数规模庞大（如

GPT-4的模型参数在1.8万亿左右）使得解释成本过高，用户与监管者均无法有效验证大模型的决策逻辑。人工智能产品在运行过程中对其运行轨迹的自主决策是基于概率得出的，由于这种决策模式的复杂性，其逻辑难以被解释，进而产生算法黑箱。依据传统过错责任框架，例如我国《民法典》第1165条规定，一般是基于行为人的主观过错来判定责任。但机器幻觉的成因难以追溯至开发者的主观过错，导致责任认定与追究陷入困境。由于难以确定开发者的主观过错，使得传统的过错责任框架难以适用。这就造成在出现因大模型幻觉导致的不良后果时，责任认定与追究面临困境，监管缺乏有效的法律责任判定依据。在算力与数据管控方面，大模型训练依赖海量数据与超算资源。一方面，数据来源广泛且复杂，包含大量用户个人信息、各类公开与非公开数据等。监管机构难以对如此庞大的数据在收集、存储、传输与使用的全流程进行实时监控，以确保数据的合法合规使用，防止因数据问题引发大模型幻觉。另一方面，大模型的迭代过程快速且复杂，模型不断优化升级，监管机构难以实时跟踪模型迭代过程中的变化。虽然在模型发布环节会实施参数优化与信息脱敏等操作，但仍无法确保所有敏感数据被完全清除。残留的数据信息会在模型执行推理任务生成文本内容时显现，进而形成数据泄露隐患。[①]《数据安全法》第21条对“重要数据”的定义尚未涵盖生成式数据，这使得在对大模型训练数据的监管上存在法律空白，进一步加大了监管机构管控数据流向与模型迭代过程的难度，无法有效从源头上遏制大模型幻觉的产生。

2.算法阴影加剧隐私保障与治理难度

数据具有独特的传播与增值属性，其低成本、零损耗的复制传播特性，打破了传统资源的使用限制；在流通演算中，数据不仅不会贬值，反而能通过加工分析衍生新价值。[②]正因数据价值需在流转交换中释放，一旦安全防线失守，数据泄露带来的损失将随其快速扩散与持续增值而被放大，严重威胁具身智能体技术生态的健康发展。在机器学习过程中，当特定数据被输入模型用于训练时，即便后续从初始训练数据集中删除该数据，模型依然会保留这些数据带来的影响，即产生难以消除的“算法阴影”。[③]这一特性使得传统的数据删除隐私救济手段失效。[④]例如，智能体可能超越用户授权范围收集数据，或者用户在与其互动时意外暴露隐私。在这种情况下，按照传统法律思维，删除相关数据应是保护隐私的重要方式。然而，由于算法阴影的存在，即使从训练数据集中删除了这些不当收集或意外暴露的数据，模型已经形成的持久印记仍会保留相关隐私数据。这意味着数据主体无法通过常规的数据删除操作来有效保护自己的隐私，隐私利益的损害难以得到弥补。并且，算法阴影的持久性在当前围绕机器学习和隐私的法律论述中尚未得到充分重视和规范。这使得在实际情况中，一旦出现因算法阴影导致的隐私问题，现有的法律框架难以提供有效的救济和应对措施。

（三）法律保护困境破解的方法论

新型技术治理需突破“政府主导”或“市场自治”的二元对立，构建“利益相关者参与-多元工具支撑-理论创新指引”的系统性方法，通过治理主体、方式与理论的多维融合，实现技术风险与法律规制的动态适配。

多元主体协同是解决复杂技术治理问题的根基所在。在具身智能体数据隐私保护场景下，

①赵梓羽:《生成式人工智能数据安全风险及其应对》,《情报资料工作》2024年第2期。

②李智、姚甜甜:《数据信托模式下受托人信义义务之规范》,《学术交流》2022年第2期。

③Li T C ,Algorithmic Destruction. SMU Law Review, 2022, 75(3), pp.479-506.

④李智、张津瑶:《数据信托本土化的现实困境与路径构建》,《学术交流》2023年第7期。

政府、企业、科研机构、社会组织及用户等理应形成紧密协作的网络体系。[①]政府作为规则制定者与监管者，需针对具身智能体数据收集、存储、传输和使用等全流程环节，制定详尽且具备前瞻性的法律规范。企业身为涵盖具身智能体的研发、生产、销售及服务等各个环节的主体，需要承担相应的保障责任。科研机构凭借其专业知识与技术研发能力，能够为隐私保护技术的创新提供支撑，如开发更先进的数据加密算法、隐私计算技术等。社会组织可以发挥监督与协调作用，推动行业自律规范的制定与实施，营造良好的行业隐私保护氛围。用户作为数据隐私的直接关联者，应积极参与到治理过程中，通过合理行使数据权利，如数据访问权、更正权、删除权等，对企业的数据处理行为进行监督与制衡。各方主体基于合作治理理念，围绕数据隐私保护这一共同目标，充分发挥各自优势，形成环环相扣、紧密协同的治理网络，从而有效提升具身智能体数据隐私保护的整体效能。

具身智能体数据隐私保护同样需要突破单一法律规制的局限，构建“法律规制-技术治理-行业自律-标准认证”的多元治理工具矩阵。法律规制为数据隐私保护提供坚实的底线保障，通过完善相关法律法规，明确数据处理各环节的权利义务关系，对侵权行为制定严格的惩处措施，形成强大的法律威慑力。技术治理则是利用先进的技术手段，将隐私保护要求融入具身智能体的技术架构中。例如，采用区块链技术对数据流转进行全程存证，确保数据的完整性与可追溯性。行业自律强调行业内部的自我约束与规范，行业协会可制定具身智能体数据隐私保护的自律公约，引导企业自觉遵守，对于违规企业实施行业内的惩戒措施，如公开谴责、限制市场准入等，以此提升行业的数据隐私保护水平。标准认证机制通过建立统一、科学的数据隐私保护标准，对符合标准的具身智能体产品或服务授予认证标识，为消费者提供明确的选择指引，同时也促使企业为获得认证而主动提升自身的数据隐私保护能力。

通过多元主体协同、多维度治理工具运用以及动态适应性治理，构建起全方位、多层次、动态灵活的具身智能体数据隐私保护体系，有效应对具身智能体发展过程中带来的复杂数据隐私风险挑战，切实保障用户的数据隐私权益，促进具身智能体技术的健康、可持续发展。

四、具身智能体利用数据隐私的合作治理路径

我国学界目前对具身智能体法律风险及其规制的研究思路虽已考虑立法论与解释论的双向维度，可仍未有效关注法律作为文本对技术规范的局限性，真正意义上的法律规范与技术规范相协调需要在权责体系架构的同时，健全与之相适应的“动静结合”的技术规范体系。合作治理是一种强调多元主体协同、多维度工具整合的治理范式，其核心在于打破单一治理主体或单一规制手段的局限，通过系统性整合政府、市场、社会等多方力量，融合法律、技术、行业自律等多种治理工具，形成应对复杂问题的合力。[②]在具身智能体等新兴技术领域，合作治理的必要性尤为突出——这类技术既涉及数据隐私、算法安全等法律问题，又关联技术标准、行业伦理等实践议题，单一的政府监管或企业自律难以覆盖其全链条风险。

（一）塑形适应具身场景的全周期权责体系

基于动态义务束理论，具身智能体的数据处理义务随场景、风险等级动态调整，需结合过程性规制理论，构建“实时审计-风险预警-合规追溯”的技术合规框架。

1.数据使用监督体系的过程性规制设计。过程性规制理论是应对复杂技术风险的重要治理工具，其核心在于突破传统“事前审批-事后处罚”的终端式规制局限，将监管嵌入被规制对象

①程雪军:《生成式AI下超大金融服务平台滥用算法权力的风险规制》,《上海财经大学学报》2024年第6期。

②许可:《论新兴科技法律治理的范式迭代——以人脸识别技术为例》,《社会科学辑刊》2023年第6期。

的全流程活动中，通过持续监测、动态反馈与实时调整实现风险前置防控。在具身智能产业迅猛发展的当下，数据使用安全问题愈发突出，构建适配的全生命周期数据治理体系已成为维护数字安全的重要议题。监管部门有必要对监管思维予以创新，借助合比例性分析[①]、成本收益分析等手段，拟定并推行审慎的监管举措。结合我国产业发展实际，鉴于人形机器人这类人工智能体风险的突发性与动态性，当务之急是汇聚各方力量，加快开发一套行之有效的监管反馈框架。[②]

具体展开，优化具身智能体的数据使用监督体系，需结合数据全生命周期的“收集-存储-处理-共享”各环节特性构建精细化框架。[③]在数据收集环节，针对多模态传感器实时采集特性，通过采集行为日志模块实时记录采集信息并上链存证，监管平台核查是否存在未经授权行为；同时预设场景化授权触发机制，进入私密区域时自动暂停采集并请求授权，未获同意则禁用采集功能，监管部门定期抽查匹配度。在存储环节聚焦分布式管理难题，要求采用合规加密算法保护本地数据并提交检测报告，委托第三方进行渗透测试；设定存储期限动态清单，敏感信息超期自动清理，监管部门监测留存时长并对超期行为预警整改。在处理环节确保决策合规，开发者需备案算法逻辑并提交评估报告，监管部门抽样回溯验证一致性；植入目的偏离监测模块，偏离授权范围时向监管平台与用户预警，按等级启动响应机制。在共享环节明确边界，实施共享白名单制度，审核通过后方可共享；利用数据血缘技术标记流转路径，监管部门追踪去向并追究违规流转责任。通过这种过程性设计，监督体系实现从结果合规到过程合规的转型，既将法律要求嵌入各环节，又通过互动式监管平衡创新与风险防控。

2.基于动态义务束的多元主体责任配置。具身智能体虽然在形态与交互上呈现类人特征，但其本质仍是缺乏自由意志的技术工具，因此不宜赋予独立法律主体地位。将其过度拟人化可能误导责任认定，弱化研发者、生产者等主体的义务，诚如相关警示所言：“我们务必竭尽全力规避人形机器人陷阱。”[④]因此，责任配置的核心应聚焦人类主体，结合技术链条与应用场景动态界定各环节的责任边界。

在具体责任划分上，需覆盖全产业链主体并适配场景差异。开发企业对核心算法的安全性与合规性负首要责任：初始阶段需履行算法透明度义务，公开数据处理逻辑；进入规模化应用后，需承担“算法可解释性维护义务”，对自主决策行为提供符合规范的说明。对于高风险应用（如医疗、军事），需承担无过错责任，即使已履行注意义务，仍需对算法涌现性缺陷导致的损害负责。[⑤]生产企业的责任与硬件及出厂设置绑定，需确保传感器加密、数据安全芯片预装等硬件合规，对因硬件缺陷导致的本地数据泄露承担产品责任；售后阶段需持续推送安全补丁，防范硬件漏洞引发的隐私风险。同时，生产企业在销售环节需提供完整合规手册，否则对销售端的误导性告知承担连带责任。销售企业负有场景化告知义务：面向个人用户时，需以可视化方式说明数据传输范围、第三方服务接入情况。若隐瞒“开机即启动麦克风采集”等默认设置，导致用户非自愿提供隐私，需承担信息不充分的过错责任。特殊场景下，若机器人销售后完全断联（无数据传输、共享或远程控制），责任主体限于使用者与生产企业：使用者对擅自破解采集限制导致的泄露负责，生产企业仅对出厂硬件缺陷担责。多主体共同侵权时，需按原因力大小划分责任，例如因“算法漏洞+服务端未拦截异常上传+销售端未提示风险”导致的隐私泄露，

①合比例性分析是指评估一个行为的措施与所要达到的目的之间是否相称、适当。

②张欣:《生成式人工智能的算法治理挑战与治理型监管》,《现代法学》2023年第3期。

③刘颖、刘佳璇:《数字经济中黑暗模式的法律规制：基本原理、域外方案与本土路径》,《上海财经大学学报》2024年第5期。

④［美］尼尔·M.理查兹、威廉·D.斯马特:《法律如何看待机器人》,陈吉栋、向梦涵译,《法治社会》2019年第1期。

⑤胡巧莉:《人工智能服务提供者侵权责任要件的类型构造——以风险区分为视角》,《比较法研究》2024年第6期。

可按一定的比例对不同主体进行追责。使用者则须依场景履行注意义务，家庭场景中不得擅自破解数据采集限制，工业场景中需对机器人采集的员工生物数据进行去标识化处理，因不当操作引发的隐私问题由使用者自行承担。通过这种全链条、场景化的责任配置，构建涵盖研发、生产、使用全链条的责任体系，既能明确各主体义务，又能形成多层级隐私保护屏障，避免责任真空或过度归责。

（二）健全风险预防主义的市场化准入机制

为从源头遏制具身智能体的技术与制度风险，须在合作治理框架下强化市场准入环节的风险筛查与合规约束，筑牢风险预防的第一道防线。

1.建立智能设备算法影响评估机制。为应对具身智能体数据隐私风险，建立智能设备算法影响评估机制可从以下几方面着手。首先，明确评估主体与流程。由专业的第三方评估机构、监管部门以及相关领域专家共同组成评估团队。在机器人投入市场前，开发者需提交算法设计文档、数据处理流程说明等资料。评估团队依据既定标准，对算法从数据收集源头开始，历经存储、传输、使用全流程进行审查。例如，审查数据收集环节是否遵循最小必要原则，避免过度采集用户信息。其次，设定全面的评估指标。在数据安全性方面，评估加密算法强度，确保数据在存储与传输中不被窃取或篡改。如采用高强度加密算法，定期检测加密有效性。在算法透明度上，要求算法具有可解释性，能向用户清晰说明决策依据。在合规性层面，对照《数据安全法》《个人信息保护法》等法规检查算法是否依法获取用户授权，是否保障用户的知情权与选择权。最后，建立动态评估与反馈机制。具身智能体使用过程中，持续监测算法运行，收集用户反馈与异常数据。若发现数据泄露风险或算法异常决策，则及时启动重新评估。同时，根据技术发展与新的安全问题，定期更新评估标准与流程。

2.构建合规性及信息保护认证制度。人类智能的运作核心在于思维，而人工智能则是依靠算法驱动，而算法构建的根基是数据。具身智能体若欲依托“大数据技术集成+高算力处理能力+强算法模型支撑”实现安全标准的提升，则构建完善的认证体系是市场主体合规运营及法律规制框架完善的重要环节。正如《人形机器人创新发展指导意见》指出的：“打造权威检验检测机构，完善评测配套工具”。我国现阶段对机器人的认证涵盖CR认证、CE认证、UL认证、FCC认证、CQC认证等多种类型，[①]目前认证市场正处于发展与培育阶段，存在众多有待持续改进和优化的地方。

制定严格的具身智能体的合规性及信息保护认证标准，需从多个维度发力，涵盖数据保护、隐私保护、算法透明度等多方面。与此同时，设立专业的认证机构或明确认证资质要求，确保认证的公正性和专业性。在认证内容方面，应细致审查数据处理流程、安全防护措施、用户隐私政策等关键模块。对于安全防护措施的检查，要求企业具备完善的数据备份、恢复和应急处理机制。在认证流程上，企业需按照标准对产品进行自我评估和改进，之后由权威认证机构进行严格审核。只有通过认证的具身智能体产品和服务才能进入市场，并向消费者明示认证标识，方便消费者识别和选择，促使企业重视数据隐私保护以获取市场竞争优势。对于未通过认证的产品，限制其生产和销售，以此激励企业主动提升产品的数据隐私保护水平，推动整个行业的规范化发展。

（三）构建技术与法律规范融合的规则体系

合作治理侧重多元利益相关者通过技术标准[②]、行业自律等柔性机制补位法律规制，核心

①王海霞、侯立本、张秋怡、郑柏恒、胡伟健:《工业机器人市场准入要求及认证标准浅析》,《中国标准化》2023年第3期。

②任愿达:《元宇宙供应商治理：标准与法律融合论的本土化进路》,《东方法学》2023年第3期。

在于以“代码即法律”理论为指导，将隐私保护的法律要求嵌入技术架构，使具身智能体的算法逻辑与数据处理行为符合合规性要求。这里的利益相关者既包括开发企业、销售企业等产业链主体，也涵盖监管机构、用户等关联方，各方通过参与技术规则制定、合规标准细化及风险动态监测，形成“技术嵌入合规要求、利益相关者共担治理责任”的闭环。

1.强化匿名技术防范数据滥用。在具身智能体设计阶段，需将隐私增强技术（PETs）作为核心安全模块，对其技术选型与应用效果需以《信息安全技术——个人信息安全规范》（以下简称《规范》）为基准。首先，在数据收集阶段，应采用满足《规范》中“匿名化处理后无法识别特定个人且不能复原”要求的匿名化技术，可运用差分隐私等技术，在数据中添加一定的噪声，[①]使收集到的数据在保持可用性的同时，难以直接关联到具体用户身份。数据收集需严格遵循关于信息保护的相关法律法规，以知情同意规则为准则，获取用户充分授权，确保数据采集的合法性，从源头守护数据来源者权益。[②]其次，在数据存储环节，需结合《规范》对敏感信息加密的要求，采用最新技术对用户的标识信息进行处理，将敏感数据转化为无意义的字符序列，同时结合高强度的加密算法对存储的数据进行加密。这样一来，即使存储系统被攻破，攻击者也难以解密和识别数据所对应的用户身份，有效防止数据被滥用。再者，在数据使用过程中，建立严格的匿名化数据访问和使用规则，通过技术手段固化数据用途，例如医疗辅助机器人的算法需预设“仅用于病情监测”的硬编码限制，若试图将数据用于广告推送等其他目的，系统自动触发权限锁定，呼应《规范》“目的限制原则”的技术实现。最后，持续监测和评估匿名技术的有效性，定期委托第三方机构依据《规范》对差分隐私、联邦学习等技术的实际效果进行测评，重点验证匿名化数据的不可识别性与不可复原性，若发现技术漏洞需立即启动升级，确保始终满足合规标准。随着技术的发展和攻击手段的更新，定期对所采用的匿名技术进行安全性评估和测试，及时发现潜在的漏洞和风险，并对技术进行升级和改进，确保其始终能够有效地防范数据滥用，保障具身智能体数据的安全性和隐私性。

2.健全数据分类分级保护体系。数据泄露事件频发，在一定程度上源于数据管理的过度集中化。[③]分类分级监管作为一种与传统粗放式监管截然不同的精细化监管手段，属于风险导向型的资源优化配置机制，核心理念在于以风险评估为基准，通过数据赋能实现监管资源的精准施策与动态调适，构建差异化的监管资源调配体系。[④]这种监管模式致力于依据风险状况，实现监管资源的高效配置。根据《数据安全法》第21条之规定，国家建立数据分类分级保护制度，针对不同数据实施差异化保护策略。数据分类是指以数据的内容、来源、特征及功能等属性为基准，把具备相同属性的数据归为同一类别，其核心价值在于为数据治理与开发利用提供规范化基础。从识别性维度界定，数据可分为个人信息与非个人信息两类：前者指能够单独或结合其他信息识别特定自然人身份、反映其活动轨迹的信息集合，作为人格权的法定客体受特别保护；后者则应不具备自然人身份识别属性，与人格权保护范畴无直接关联。这种法律层面的分类具有重要规范意义——二者在法律保护模式、义务配置及风险防控机制上存在显著差异。而数据分级制度则聚焦法益保护强度，通过评估数据所承载法益的重要程度，以及非法处理可能导致的权益损害后果严重程度，对数据进行重要性层级划分。这一机制旨在建立与数据风险等

①刘泽刚:《人工智能时代联邦学习隐私保护的局限及克服》,《中外法学》2025年第1期。

②李智、万欣怡:《数据信托的运行逻辑与路径选择》,《中南民族大学学报(人文社会科学版)》2025年第4期。

③李智、周智皓:《个人数据信托的发展困境与制度设计》,《学术交流》2024年第8期。

④孙志建:《怎样合理配置有限的政府监管资源——基于风险的监管模式的兴起及其潜在运行风险》,《上海行政学院学报》2022年第2期。

级相匹配的保护措施，实现监管资源的精准配置。为有效应对具身智能体在数据收集过程中的安全挑战，需依据数据的敏感程度、重要性和潜在风险，对其收集的数据展开科学合理的分类分级工作。[①]针对不同级别的数据，制定差异化的保护策略。对于敏感程度高的数据，如生物识别信息、金融数据等，采取更严格的访问控制、加密存储和传输措施；[②]对于一般数据，实施相对宽松但仍符合安全标准的保护方式。通过这种分类分级保护体系，合理分配安全资源，提高数据保护的针对性和有效性。

3.搭建用户动态授权管理模式。为强化用户数据自主权，应构建动态化、场景化的授权管理体系，实现用户对具身智能体数据访问与处理的全流程掌控。该体系依托技术架构与法律规范双重保障，通过权限动态配置、可视化管理及撤回机制，构建数据防护闭环。

在权限动态管理层面，系统支持用户基于场景需求进行精细化授权配置。例如，家庭场景下用户可设定具身智能体仅能访问与基础服务相关的环境数据（如温湿度、照明信息）；当有访客进入时，用户可通过移动端应用或语音指令，一键启用临时权限限制模式，暂停具身智能体的摄像头、麦克风等采集设备运行，确保隐私空间的物理与数据双重隔离。这种场景化授权机制通过预设规则与实时干预相结合的方式，既保障服务连续性，又有效防范隐私风险。系统设计需遵循透明化与可操作性原则，通过可视化界面清晰展示数据授权状态。用户可实时查看具身智能体当前访问的数据类型、使用目的及权限有效期，并通过图形化交互界面（如权限树状图、时间轴视图）进行动态调整。

在用户权益保障方面，采用“即时撤回+数据清除”的双轨制设计。依据《个人信息保护法》第15条规定的撤回权，系统需提供便捷的撤回操作入口，支持用户通过应用程序、语音指令等多种方式即时终止数据处理。当用户执行撤回操作后，机器人须在技术可行的最短时间内停止相关处理行为，并启动数据删除程序。对于无法立即删除的数据（如存储在分布式系统中的备份数据），应通过加密、去标识化等技术手段确保数据不可识别，并在规定期限内完成彻底清除。同时，系统须建立撤回操作日志，完整记录撤回时间、影响范围及数据处理后续状态，为监管审计与用户维权提供依据。这种全链路响应机制，将法律赋予的撤回权转化为可落地的技术实现，切实保障用户对个人数据的终极控制权。

五、结　语

具身智能体具有感知交互、自主决策和信任匹配三大特性，其在数据全生命周期中潜藏着独特的数据隐私侵权风险，而大模型的内生性技术瑕疵和法律规则的外源性适用困境相互交织，导致数据隐私保护陷入技术失控与制度失灵。对此，具身智能体的风险治理需突破传统法律范式，通过具身智能场景中多元主体的合作治理，构建覆盖数据全生命周期的保护体系，尤其需结合风险级别、原因力大小等要素，在断联、协同等场景中动态分配责任，明确开发企业、生产企业、销售企业、服务企业及使用者等多元主体的责任边界，形成权责清晰的治理网络。

在技术加速迭代的背景下，具身智能体的应用已然超越单一技术领域，成为关涉个人权益、产业发展与社会治理的复合命题。其中，强化数据隐私保护与责任主体的精准界定，是坚守“技术以人为本”价值取向的必然要求。未来研究可以进一步探索具身智能体在医疗、家庭、工业等特定场景中的隐私保护机制，以应对技术融合带来的新型法律与伦理挑战。唯有通过跨

①陈兵、顾丹丹:《数字经济下数据共享理路的反思与再造——以数据类型化考察为视角》,《上海财经大学学报》2020年第2期。

②赵海乐:《比较法视角下的我国“车联网”数据治理路径选择》,《上海财经大学学报》2021年第5期。

学科、跨领域的研究合力以及有效的制度设计与合作治理，明确各方权责，才能为技术创新划定制度边界，为数字社会构建安全可信的发展图景。

Cooperative Governance of Data Privacy Risks of Embodied Intelligent Agents

Li Zhi, Chen Yingying

(*Law School, Shanghai University, Shanghai* 200444, *China*)

Summary: Embodied intelligent agents, as the physical form of artificial intelligence, possess core features such as perception interaction, autonomous decision-making, and trust matching. While driving a leap in productivity, they also bring about complex data privacy risks. These risks present a three-dimensional superimposition: In the technical dimension, the physical mobility of entities breaks through the boundaries of physical space, the integration of multiple sensors intensifies unauthorized data acquisition, and algorithm black boxes and distributed processing lead to data control failure; in the legal dimension, the traditional "informed consent" framework fails due to the absence of dynamic scene authorization, the right to data deletion is difficult to achieve due to fragmented edge storage, and the multi-party responsibility is trapped in the attribution dilemma due to the autonomy of algorithmic decision-making; in the ethical dimension, anthropomorphic appearance induces emotional dependence, leading users to disclose sensitive information irrationally, and emotional computing technology further deepens the mining of personality profiles.

The triple coupling characteristics of embodied intelligent agents—physical entity, data processing, and social interaction—make data risks evolve in a trend of penetrating from virtual to physical space, upgrading from passive collection to active intrusion, and spreading from the technical layer to psychological cognition. To address these data privacy risks and legal application dilemmas, it is necessary to make concerted efforts from multiple levels such as law and technology: At the level of rights and responsibilities, improve the supervision system for data use and the responsibility system for stakeholders, and form a "bundle of obligations" corresponding to the "bundle of rights"; at the market access level, establish an algorithm impact assessment mechanism in line with embodied ethics and an information protection certification mechanism meeting compliance elements, and form an access system in line with the risk prevention principle; at the technology embedding level, construct anonymous technology rules, data classification and grading rules, and dynamic user authorization management rules with industrial and institutional binding force, promoting the integration of technology and law.

Key words: embodied intelligent agents; data privacy; legal governance; technological governance

（责任编辑：倪建文）