HANKUN

汉坤法律评述

2026 年 7 月 17 日

北京 | 上海 | 深圳 | 杭州 | 武汉 | 海口 | 香港 | 新加坡 | 纽约 | 硅谷 | 伦敦

汉坤具身智能数据合规与治理系列（一）：治理框架新范式

作者：陈文昊 | 原舒仪 | 李诗

摘要

当人工智能进入物理世界，大模型被嵌入机器人、智能汽车、工业设备和公共服务设施，合规审查的对象与边界也随之改变。传统互联网时代的数据合规主要回答“信息如何被收集和使用”；具身智能则必须进一步追问：信息如何进入模型，模型如何形成判断，判断如何转化为物理行动，以及行动结果如何反馈并再次用于训练。数据、模型、智能体、设备与现实环境由此构成一个持续运转的闭环。

具身智能的治理对象不再是静态数据或单一模型，而是“数据—模型—智能体—物理行动—生态反馈”构成的连续系统。企业由此需要从“管理数据”走向“治理智能”。针对具身智能企业面临的合规挑战，**本文提出四层数据合规治理模型：隐私保护（Privacy Protection）、重要数据治理（Important Data Governance）、AI 治理（AI Governance: Model & Agent）和数据生态治理（Data Ecosystem Governance）**。四层并非由低到高的风险等级，而是分别从个人权益、国家安全与公共利益、模型与自主行为以及产业链责任四个维度，对同一系统进行交叉审查。其价值在于，将分散在个人信息保护、数据安全、网络安全、算法与人工智能监管中的要求，转化为可以嵌入产品设计、研发评审、上线决策与商业合作的治理路径。

本系列将围绕“具身智能的数据合规与治理”展开，供具身智能企业、产业链合作方及相关投资人参考。作为系列首篇，本文主要阐释四层模型的理论基础，并说明传统互联网时代的数据合规框架为何需要在具身智能场景下重构。**之后，我们将通过后续系列文章依次从上述模型的四个方面来讨论具身智能的数据合规治理问题。**

一、大模型进入物理世界后，合规问题为何发生变化

（一）具身智能不只是“人形机器人”

具身智能通常是指智能系统借助物理载体，与真实或仿真环境持续交互，并在感知、理解、决策、执行与反馈的循环中学习和完成任务。其载体既可以是人形机器人，也可以是工业机械臂、仓储物流机器人、手术机器人、无人配送设备、智能车辆、巡检设备，以及其他具备自主感知和行动能力的设备。人形外观只是具身智能的一种产品形态，并不决定其技术属性或法律属性。

从软硬件系统结构看，具身智能通常由机器人本体（Robot）、智能体（Agent）、基础模型（Foundation Model）、世界模型（World Model）、动作模型或策略模型（Action/Policy Model）、传感器、云平台及外部工具共同构成。基础模型提供语言、视觉或多模态理解能力；世界模型用于表征环境状态并预测行动

后果；动作或策略模型将目标转化为可执行序列；智能体负责任务分解、记忆、规划与工具调用；机器人本体则将数字决策转化为移动、抓取、交互、调度或控制等现实行为。

因此，具身智能并非简单的“模型加外壳”。真正决定其风险性质和合规边界的，是模型、智能体、硬件控制器与部署环境之间的协同关系。同一基础模型被嵌入客服机器人、家庭陪伴机器人或工业机械臂时，其涉及的个人信息、行业数据、动作权限及安全后果可能完全不同。合规分析不能停留在模型名称或技术标签，而必须回到具体功能、应用场景与控制关系。

（二）数据从资源变为系统行为的组成部分

传统互联网产品的数据流通常被概括为“收集—存储—使用—共享—删除”；具身智能的数据流则更接近以下闭环：

环境（Environment）→传感器（Sensor）→机器人本体（Robot）→云端（Cloud）→基础模型（Foundation Model）→智能体（Agent）→动作执行（Execution）→环境（Environment）

在这一闭环中，摄像头、麦克风、激光雷达、力觉传感器和位置传感器持续获取环境信息；本地计算单元通常与云端模型共同完成识别、推理和规划；智能体调用地图、控制系统或其他第三方插件；执行机构据此实施移动、抓取、调度或设备控制；动作结果随后转化为新的环境信息和反馈数据，并可能进一步进入评测、微调、强化学习或持续学习流程。

这意味着，数据不仅是训练模型的原料，也是触发实时行为的条件。同一段图像既可能包含自然人的个人信息，也可能用于空间定位；同一组设备运行数据既可用于故障诊断，也可能揭示关键产线能力；同一条用户指令既可能被保存为交互记录，也可能直接触发外部工具调用或物理动作。由此，数据的合法性、质量、完整性与时效性都会直接影响系统行为。

更重要的是，具身智能将信息风险与物理风险直接连接。内容模型的错误输出可能造成误导，而具身智能系统的错误判断还可能转化为碰撞、误抓取、越权操作、错误调度或设备失控。一次过度采集可能同时引发隐私侵害与模型偏差；一次访问权限配置错误，也可能同时造成数据泄露与越权动作。合规、安全和产品可靠性因此不再是彼此割裂的三条工作线，而必须在同一治理框架下统筹处理。

二、具身智能数据的六项特征及其合规挑战

（一）多模态：单一数据合规不等于组合处理合规

具身智能往往同时处理图像、声音、空间、姿态、位置、设备状态和交互记录等多类数据。某些信息单独看并不敏感，但在时间序列或多模态关联后，可能足以识别个人身份、生活习惯、工作规律，甚至关键设施的运行状态。因此，企业不应仅按字段逐项判断数据属性，还应评估组合数据、推断信息及模型特征所产生的新增风险。

（二）连续感知：用户难以通过一次点击控制采集

传统互联网应用通常在用户主动打开页面或点击功能时启动数据处理；具身智能设备则可能长期待命、持续巡航，或在多人环境中持续运行。现场人员未必是设备的直接用户，也未必阅读过隐私政策，却可能在不经意间成为被采集数据的对象。企业因此需要综合运用机身标识、状态灯、语音提示、场所告示和移动端控制等方式履行透明度义务，并通过触发式采集、边缘计算和短周期缓存，减少默认的持续、全量采集。

（三）实时反馈：合规要求必须结合技术参数

具身智能的感知、推理与执行往往在毫秒或秒级完成，事后人工审查难以覆盖每一次动作。最小必要、权限控制和风险拦截等要求，必须进一步转化为可执行的设备规则与技术参数，例如摄像头采集区域、数据上传条件、动作白名单、速度和力度上限、人工确认阈值、地理围栏及紧急停止机制，而不能仅停留在政策文本和流程文件层面。

（四）长周期积累：记忆能力会改变数据性质

为提升个性化能力和任务连续性，具身智能系统可能长期保存用户偏好、家庭布局、工作流程、任务历史及异常样本。随着数据持续积累，原本零散的信息可能形成对个人、场所或业务的深度画像，并改变数据的敏感程度及合规属性：一般个人信息在聚合、关联后可能呈现敏感特征，原本分散的非重要数据在特定场景和规模下也可能产生重要数据风险。企业需要区分任务缓存、运行日志、由用户管理的长期记忆和模型训练数据，并分别设定保存期限、访问权限与删除机制。

（五）环境依赖：同一产品进入不同场所会触发不同规则

同一具身智能产品进入家庭、商场、医院、港口、工厂或公共道路等不同环境后，其处理的数据类型、相关主体的合理预期以及适用的行业监管要求都可能发生变化。产品获得通用认证或完成一次影响评估，并不意味着该结论可以自动覆盖所有部署场景。企业应将“进入新场景”设置为重新评审的触发条件，重新审查传感器配置、数据回传机制、重要数据可能性、物理安全措施及人员告知方式。

（六）人机交互：身份与信任本身成为治理对象

具身智能产品可能通过声音、名称、外观和记忆与用户建立持续关系，使用户产生该产品具备真实人格、专业资质或人类情感的“错觉”。拟人化设计在提升交互体验的同时，也可能放大信任、情感依赖和诱导风险，对未成年人、老年人及心理脆弱群体的影响尤为突出。因此，透明度义务不能止于说明“使用了人工智能技术”，还应清晰揭示系统身份、能力边界、关键限制，以及何种情况下需要人工介入或接管。

三、传统互联网数据合规框架的三项结构性不足

（一）只审查数据处理，无法覆盖模型和动作

传统隐私合规主要围绕个人信息及其他数据的收集、使用、存储、共享和删除等处理活动展开。企业通常通过数据处理清单、隐私政策、告知同意机制、数据处理协议以及删除和权利响应机制，识别并证明相关处理活动是否具备合法性、正当性与透明度。

然而，在具身智能系统中，数据处理合法并不当然意味着系统整体安全，也不当然意味着其行为结果合法、可控。即使训练数据和运行数据来源合法，系统仍可能因数据代表性不足或存在系统性偏差，形成错误甚至歧视性的模型判断；世界模型也可能因对物理环境、人物意图或因果关系理解不准确，作出偏离真实场景的决策；具备较高自主性的智能体还可能因权限范围过宽、目标设定不当或安全约束不足，执行超出预期的操作，进而影响人身、财产、生产经营秩序乃至公共安全。

因此，具身智能企业不能只证明“相关数据可以被合法地处理”，还需要进一步说明：数据是否适合用于特定模型的训练、微调、测试或推理；模型可以被部署于哪些业务场景、支持哪些功能，其能力边界和适用条件为何；智能体可以访问哪些系统、设备与数据资源，可以作出哪些决策并执行哪些物理

或数字动作；对于高风险动作，是否设置了权限分级、人工确认、实时监测、紧急停止和事后追溯机制。具身智能合规由此不再只是对数据处理活动的审查，而是对“数据输入—模型判断—智能体决策—动作执行”全过程的治理。

（二）只在上线前审查，无法覆盖持续学习和能力变化

具身智能的能力会随着模型升级、插件接入、长期记忆、在线反馈和部署场景变化而持续演进，一次性的上线前审查难以覆盖整个生命周期。新增传感器、切换基础模型、开放境外访问、提高动作权限，或者允许反馈数据进入训练集，均可能改变原有合规结论。企业需要将法律与合规评审嵌入产品变更管理，使其成为持续运行的控制机制，而不是上线后被归档的静态文件。

（三）只管理直接供应商，无法覆盖模型链和指令链

具身智能产品通常同时集成基础模型、云平台、传感器、地图、语音 SDK、插件和客户系统。传统供应商管理主要关注“谁能够访问数据”；具身智能还必须追问“谁能够改变模型能力、调用外部工具、下发控制指令或获取动作结果”。在此场景下，供应链同时也是数据链、模型链和指令链，任何一环的接口控制或合同安排存在缺口，都可能沿系统链路传导并放大。

四、四层数据治理模型：四个维度治理同一系统

基于上述特征，具身智能企业可以从四个相互交叉的维度，建立面向同一系统的治理模型。

治理层	主要保护对象	核心问题	典型风险	主要控制
隐私保护	自然人个人权益	个人信息处理是否具有合法性、正当性与必要性	持续监控、敏感信息滥用、权利难以实现、跨境访问失控	来源核验、隐私工程、影响评估、权利响应、跨境管理
重要数据治理	国家安全与公共利益	场景数据是否可能构成重要数据或核心数据	关键设施与运行数据汇聚、违规共享或出境	分类分级、场景识别、目录核验、审批、风险评估与审计
AI 治理	模型、智能体及受影响主体	模型与智能体是否安全、可信、可控	偏差、幻觉、越权调用、危险动作、身份误认	数据与模型溯源、评测、权限矩阵、人工确认、内容标识
数据生态治理	产业链及合作主体	数据、模型与动作责任如何分配	授权链断裂、第三方泄露、插件越权、跨境失控	角色定性、合同安排、接口控制、持续监测与退出机制

（一）第一层：隐私保护

隐私保护层解决的是企业能否合法、正当且必要地收集、使用、共享、保存和删除个人信息。审查对象既包括机器人现场采集、公开来源、第三方采购、网络抓取、开源数据集、合成数据和仿真数据，也覆盖训练、微调、检索增强生成（RAG）、远程运维与反馈学习等处理环节。其关键不在于取得一份概括性的同意，而在于将目的限定和最小必要等要求落实到传感器启用、数据存储、上传路径与删除机制之中。

（二）第二层：重要数据治理

重要数据治理层解决的是，企业如何识别并保护可能影响国家安全、经济运行、社会稳定、公共健康及重大公共利益的数据。工业、能源、港口、医疗、交通和城市运行等场景，尤其容易汇聚高密度的设施与运行信息。企业需要依据行业规则、重要数据目录、主管部门通知及国家标准，建立具有一致性和可解释性的识别流程。

（三）第三层：AI 治理

AI 治理层关注基础模型、世界模型、动作模型与智能体的全生命周期。治理重点包括训练数据的合法性与质量、模型和数据版本溯源、模型评测、智能体工具调用、长期记忆、动作权限、人工确认、身份标识及拟人化边界。对具身智能而言，AI 治理的重心应当从“输出内容是否安全”，进一步延伸至“系统行为是否安全”。

（四）第四层：数据生态治理

数据生态治理层解决的是机器人厂商、模型公司、云平台、供应商、开发者、OEM、ODM、集成商与运营商等主机厂与供应链主体之间的责任和权限配置。数据共享已不再只是单纯的文件传输，还包括 API、SDK、机器人云（Robot Cloud）、插件、联邦学习及外部工具调用。企业应通过角色定性、接口控制、合同安排、持续审计与退出机制，确保数据、模型和指令在生态链条中可识别、可追踪、可问责。

五、四层模型如何进入企业研发与商业化流程

四层模型不应被拆分为四套彼此独立的制度。实践中更有效的做法，是以具体产品用例为中心，将四个治理维度整合为一条贯穿研发、上线与商业化的统一路径。我们建议企业采取如下合规治理动作：

（一）第一步：建立合规治理清单

企业应以具体用例为基本单元，建立场景清单、数据清单、模型清单、动作清单和主体清单。五张清单共同回答五个基础问题：系统在何处运行，数据从何处取得并流向何处，数据进入哪些模型，模型与智能体能够完成哪些功能和动作，以及哪些主体可以接触数据或改变系统能力。

（二）第二步：绘制数据与指令流图

传统数据流图通常止于数据从设备流向云端；具身智能还需要标注模型推理、工具调用、权限校验、动作执行和反馈回流。只有同时呈现数据流与指令流，企业才能识别个人信息、重要数据、模型风险与第三方责任在何处交汇，并据此配置控制措施。

（三）第三步：设置风险门槛与变更触发器

企业可以将敏感个人信息、重要场所、关键动作、境外访问和高权限工具设置为上线门槛。新增传感器、启用长期记忆、接入新的模型或插件、开放新接口、提高动作权限，或者进入新的行业场景，均应自动触发法律、安全与技术复核。

（四）第四步：形成可验证证据链

企业应当保存并关联数据来源证明、告知与处理依据、个人信息保护影响评估、分类分级和重要数据识别记录、数据集及模型版本、评测报告、智能体权限矩阵、人工确认与紧急停止机制、备案及标识

判断、供应商尽调与合同、跨境合规材料、访问日志和事件处置记录。合规自证的关键不在于文件数量，而在于证据之间能够相互印证，并完整还原系统从设计到运行的决策过程。

（五）第五步：建立持续监测闭环

企业可通过日志审计、异常告警、用户投诉、模型评测、红队测试、漏洞管理、供应商审计及事件复盘，持续识别系统能力、部署场景和监管要求的变化。监测结果应及时反馈至模型版本管理、智能体权限、传感器设置和制度更新，形成真正的治理闭环，而不应止步于年度报告。

结语：从管理数据走向治理智能

具身智能并未使现有个人信息保护、数据安全与网络安全规则失去效力，而是显著强化了这些规则之间的关联。隐私保护决定数据能否进入系统；重要数据治理决定特定场景中的数据能否汇聚、共享与出境；AI 治理决定模型和智能体能否在既定边界内形成判断并实施动作；数据生态治理则决定多主体协作中的责任、权限与证据能否形成闭环。

“隐私保护—重要数据治理—AI 治理—数据生态治理”的四层模型，并非抽象的合规分类，而是一条从研发到商业化的治理路径。**后续系列文章将分别讨论：机器人持续感知环境时如何实现隐私保护；企业如何识别具身智能场景中的重要数据；模型与智能体如何从内容安全走向行为安全；以及产业链中数据、模型与指令责任应如何分配。**

特别声明

汉坤律师事务所编写《汉坤法律评述》的目的仅为帮助客户及时了解中国或其他相关司法管辖区法律及实务的最新动态和发展，仅供参考，不应被视为任何意义上的法律意见或法律依据。

如您对本期《汉坤法律评述》内容有任何问题或建议，请与汉坤律师事务所以下人员联系：

陈文昊

电话： +86 21 6080 0359
Email： wenhao.chen@hankunlaw.com