HANKUN

汉坤法律评述

2026 年 8 月 14 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（三）：重要数据治理

作者：陈文昊 | 原舒仪 | 李诗

摘要

当具身智能设备进入工厂、医院、港口、能源设施、公共空间和道路后，便会持续记录设施布局、生产流程、物流调度、交通运行和城市状态。这些数据可能因覆盖范围扩大、精度提高或者相互关联，逐渐具有反映关键设施、行业运行乃至区域态势的能力。企业由此需要回答一个更具系统性的问题：如何识别和管理可能直接影响国家安全、经济运行、社会稳定、公共健康和安全的重要数据。

长期以来，企业往往将重要数据识别理解为一项“等待主管部门目录”的工作。随着《网络数据安全管理条例》《促进和规范数据跨境流动规定》、《数据安全技术 数据分类分级规则》（GB/T 43697—2024）以及行业、地方规则陆续出台，这种做法已经难以满足合规要求。企业仍应以重要数据目录和主管部门认定为核心依据，但在目录尚不明确时，也需要结合部署场景、数据属性和潜在危害开展内部分析，并对检索范围、判断依据和复核过程留痕。2026 年 6 月公布、将于同年 8 月 20 日起施行的《网络数据安全风险评估办法》，进一步把重要数据识别、风险评估、报告报送和整改闭环连接起来。对于具身智能企业而言，真正需要避免的不是某一次判断中存在合理不确定性，而是既未检索适用规则，也无法解释其判断过程。

一、为什么具身智能企业必须前置识别重要数据

（一）具身智能设备将数据采集延伸至关键物理空间

传统互联网平台所处理的数据，主要来自用户的线上行为。具身智能设备则把持续感知带入以往难以被完整数字化的物理空间。工业机器人能够记录产线节拍、设备状态和故障模式，巡检机器人会接触能源设施的位置、运行参数和安全缺陷，物流机器人则不断积累仓储布局、货物流向和调度信息。随着设备进一步进入医疗和公共空间，诊疗活动、道路状况、建筑结构及城市设施状态也可能进入同一数据链路。数据来源由线上行为扩展至现实环境，是具身智能领域需要前置考虑重要数据问题的第一个原因。

单台设备在局部场所、短时间内采集的数据，未必足以产生重大影响。然而，企业通过机器人云平台集中管理大量设备后，数据会沿时间、地域、客户和行业等维度持续汇聚。原本分散的局部信息经过关联分析，可能逐渐呈现关键设施布局、生产能力、供应链薄弱环节或者城市运行规律。因此，判断重点不能停留在某个字段是否“敏感”，而应进一步考察数据汇聚后的规模效应和组合效应。

（二）模型训练放大数据汇聚与复用效应

这种汇聚效应在模型训练阶段会进一步放大。为了训练世界模型、动作模型和策略模型，企业通常需要将不同客户、不同地区和不同时段的运行数据接入统一的数据管线。数据经过清洗、标注和特征提取后，可能脱离最初的部署场景，被重复用于通用能力训练。如果企业只在设备交付时判断数据性质，却不审查训练集的来源、汇聚范围和后续用途，就可能遗漏数据在二次利用过程中发生的风险变化。

与此同时，治理对象也不限于原始文件。模型更新、特征库、向量表示和分析结果虽然经过技术处理，但仍可能保留或者反映特定场景的信息。企业不能仅凭数据已经“模型化”就排除其重要数据属性，而应结合可还原程度、表达精度、覆盖范围和实际用途作进一步判断。当然，模型参数是否构成重要数据仍需结合具体技术路径和危害后果审慎分析，不能作一概而论。

（三）全球化研发使识别结论直接影响系统架构

数据性质一旦发生变化，影响会直接传导至全球研发和系统架构。具身智能企业常采用跨境协同研发、境外云服务和远程运维模式，而境内运营中收集和产生的重要数据确需向境外提供的，依法应当通过数据出境安全评估。识别结论因此会影响数据存储地点、研发人员权限、训练任务分工和远程访问方式。重要数据识别并非出境申报前才启动的文书工作，而应成为本地化处理、境内训练、分域存储和跨境隔离等架构决策的前提。

上述特点共同说明，具身智能企业不宜对某类数据作脱离场景的抽象判断。更可行的做法，是沿着设备进入的物理空间、数据如何汇聚以及数据用于何种处理活动，识别更可能触及重要数据的业务场景。以下七类场景并非封闭目录，而是企业开展初步筛查时需要优先核查的场景。

二、具身智能领域重要数据识别的重点场景示例

场景分类的意义，不在于为每类数据预设结论，而在于帮助企业找到影响判断的事实变量。同一类数据在不同精度、覆盖范围和汇聚方式下，可能得出不同结论；同一套机器人系统也可能同时跨越多个场景。需要注意的是，我们例举的以下分类场景，应当与主管部门公布的重要数据目录和危害分析结合使用。

（一）地图与空间数据

机器人导航可能使用高精度地图、点云、视觉定位、建筑平面图和道路信息。判断时不能只看数据名称，还要分析空间精度、覆盖区域、更新频率、是否涉及敏感地点、能否还原关键设施，以及是否适用测绘和地理信息领域的专项规则。需要注意的是，普通室内避障地图与大范围、高精度、持续更新的空间数据，在法律风险上可能存在明显差异。

（二）工业数据

当机器人从导航进入生产环节，判断重点随之转向工业数据。数字化工厂会持续形成产线布局、设备参数、产能、良率、故障情况、工艺流程和控制指令等数据。这些数据是否构成重要数据，取决于其能否反映重点行业整体运行、关键产品供给能力、重要工业控制系统或者产业链薄弱环节。单一客户的普通设备日志可能不会仅因属于“工业数据”就当然构成重要数据，但跨企业、跨区域汇聚后，结论可能发生变化。

（三）能源数据

能源场景与一般工业场景相比，往往更直接关联关键基础设施和公共供给。巡检、维护和应急机器

人可能接触电力、石油、天然气等设施的位置、运行状态、负荷、故障和安全缺陷。企业除应检索能源行业规则外，还需结合设施等级、数据覆盖范围和实时性进行判断，并对远程访问、云端汇聚和境外运维采取更严格的管理措施。

（四）医疗与公共卫生数据

在医疗和公共卫生场景中，机器人可能形成反映医疗资源配置、诊疗能力、设备运行、疾病趋势和公共卫生响应能力的数据。单一机构、局部时段的数据与覆盖较大区域、能够呈现关键医疗服务能力的数据，影响范围并不相同。企业应当把关注点从单条诊疗记录扩展到汇聚后能否反映公共卫生状况和医疗保障能力。

（五）港口、物流与供应链数据

类似的汇聚效应也会出现在港口、物流和供应链场景。港口机器人、无人装卸设备和仓储系统可能记录货物流向、吞吐能力、关键物资、调度规则和供应链状态。判断时应当结合覆盖范围和实时性，分析数据是否涉及战略物资、关键节点或者重要运输通道，以及一旦泄露或被篡改，是否可能影响经济运行和供应链安全。

（六）交通与自动驾驶数据

设备从相对封闭的场所进入开放道路后，数据的区域覆盖和实时关联能力会进一步增强。自动驾驶系统和道路作业机器人会采集道路、交通流、车辆轨迹、地理空间和基础设施信息，企业需要重点关注大规模实时交通运行数据、重点区域道路信息和智能驾驶训练数据的安全。企业当前可以参照《汽车数据安全管理若干规定（试行）》《汽车数据出境安全指引（2026 版）》以及相关自贸区负面清单，对该领域的重要数据进行识别。

（七）公共空间与城市运行数据

承担安防、清洁、巡检和城市服务任务的机器人可能横跨多个公共空间，并持续汇聚公共设施、人员流动、应急资源和城市运行状态等数据。与局部、低精度、短周期数据相比，跨区域、实时且可关联的数据更可能形成城市级运行态势，或者暴露关键公共系统的运行规律。因此，公共空间数据的判断尤其需要关注覆盖范围扩大和多源关联所带来的性质变化。

三、识别依据：目录认定与企业内部分析如何衔接

（一）主管部门和地区认定仍是核心依据

在识别依据上，重要数据目录和主管部门认定仍然具有优先地位。企业应当从行业规则、公开目录、地方要求和主管部门通知入手，确认相关数据是否已经被明确纳入重要数据范围。具身智能业务往往同时涉及制造、汽车、医疗、能源、测绘、通信和公共服务等领域，因此不能只按照工商登记的行业类别查找目录，而应根据具体部署场景、数据来源和处理目的识别主管领域。

尽管《促进和规范数据跨境流动规定》规定，数据处理者未被相关部门、地区告知，且相关数据未被公开发布为重要数据的，不需要将其作为重要数据申报数据出境安全评估，但数据处理者不应忽视《网络数据安全管理条例》的明确规定 — 网络数据处理者应当按照国家有关规定识别、申报重要数据。企业若想开展重要数据识别工作，数据分类分级、建立企业数据处理地图是第一步，只有这样才能对企业所处理的数据有全面的了解，并对不同类型、级别的数据施加不同的管理与技术保护措施。

（二）建立重要数据内部分析框架

在目录尚不明确的领域，《数据安全技术 数据分类分级规则》（GB/T 43697—2024）可以提供内部分析工具。企业可围绕数据覆盖对象、规模、精度、实时性、可替代性、关联组合能力以及潜在危害开展判断，并记录事实依据和分析过程。该标准属于推荐性国家标准，不能取代有关地区、部门和行业领域的最终认定，但可以作为内部识别、风险评估和合规自证的方法依据。

四、从场景梳理到动态复核：六步重要数据识别流程

以下六个步骤并非彼此孤立的检查项目，而是一条可以复核的证据链。企业需要先还原业务场景和数据流，再查明适用的外部规则；在此基础上分析数据属性与潜在危害，形成内部结论，并在业务发生变化时重新判断。只有前后步骤能够相互印证，内部识别结果才具有稳定性和可解释性。

第一步：围绕部署场景绘制数据底图

识别工作应从具身智能设备部署场景开始，而不是从数据库字段开始。企业可按家庭、商业、工业、医疗、能源、港口、交通和公共空间等场景绘制数据底图，列明设备类型、客户、地理范围、传感器、处理位置、联网方式和远程访问安排。只有先还原数据在真实业务中的来源和流向，后续的目录检索与危害分析才有事实基础。

第二步：检索行业与地区规则

形成场景底图后，企业才能有针对性地检索行业和地区规则。检索范围应覆盖可能涉及的主管部门、重要数据目录、地方要求和公开通知，并记录检索日期、规则名称及必要的咨询结果。即使尚未发现适用目录，也应保留检索和分析记录，以证明企业已经履行合理识别义务，而不是简单写下“无重要数据”。

第三步：评估数据属性和组合效应

目录检索解决的是外部规则依据，不能代替对数据本身的分析。企业还应考察数据内容、精度、规模、覆盖范围、时间跨度、实时性、可替代性和关联能力，尤其要判断数据跨客户、跨区域、跨时间汇聚后，是否可能还原例如关键设施、产能、调度能力或者城市运行规律等敏感信息。

第四步：分析不同风险事件的潜在危害

在事实属性基本明确后，判断才进入潜在危害层面。企业需要分别分析数据遭到篡改、破坏、泄露、非法获取或者非法利用后可能产生的后果，因为不同事件路径对应的风险并不相同。例如，控制参数被篡改可能影响生产安全，设施位置泄露可能增加遭受攻击的风险，大规模运行数据外泄则可能暴露经济运行和公共服务能力。

第五步：形成初步结论并开展分级复核

完成事实梳理和危害分析后，业务团队与数据团队可以先形成初步结论，再由法务、数据安全团队和行业专家对高风险场景进行复核，必要时与主管部门沟通。在内部管理上，结论可区分为“已经目录或者主管部门明确”“具有较高可能性，需进一步确认”“暂未识别为重要数据，但需持续监测”和“一般数据”。这种分层并不改变最终法律认定，而是为了在目录和业务持续变化的情况下配置相应的内部控制。

第六步：将识别机制纳入变更管理

识别流程的最后一步，是将结论纳入变更管理。设备进入新地区、接入新客户、提高采集精度、扩大数

据规模、启用云端汇聚、开放境外访问，或者将运行数据用于通用模型训练时，都可能改变原有判断。企业应设置重新识别触发条件，使重要数据管理随业务变化持续更新，而不是停留在项目启动时填写的一张表格。

五、识别之后：分类分级结论如何落到系统控制

识别重要数据之后，企业需要解决的不是再给整个数据库贴上一个统一标签，而是把法律结论落实到具体数据集、处理活动和信息系统。《数据安全法》要求建立数据分类分级保护制度，并对重要数据实行重点保护；关系国家安全、国民经济命脉、重要民生和重大公共利益的数据属于国家核心数据，适用更加严格的管理制度。一般数据也并非无需保护，而应根据其业务价值、可能造成的损害以及合同约定配置相应措施。

对于具身智能企业，更可行的做法是按照数据集、数据项和处理场景分别判断。例如，单台工厂机器人为完成即时避障而临时处理的数据，通常影响有限，但如果企业长期汇聚多台机器人的数据，并由此还原整座厂区的布局、产能和设备运行情况，数据的影响范围就可能发生变化，需要采取更严格的安全措施。分级结论最终应映射到存储区域、访问角色、共享审批、接口调用频率、加密或者脱敏措施、日志留存和删除机制，只有完成这一步，分类分级才真正转化为产品和系统能力。

六、从数据台账到风险评估：重要数据处理者的合规闭环

一旦相关数据被确认或者高度可能被认定为重要数据，治理重点就会从“如何判断”转向“如何持续履责”。数据目录、访问控制、风险评估、出境管理和事件响应并不是五项相互分离的工作。企业应当先通过目录界定管理对象，再以技术和管理措施控制日常风险，并通过风险评估检验措施是否有效；涉及出境或者安全事件时，还需要启动相应的专项程序。上述环节应当围绕同一数据范围和处理活动相互衔接。

（一）建立数据目录，明确安全负责人和管理机构

合规闭环的起点，是建立能够与实际系统相互印证的数据目录。企业应明确重要数据的名称、内容、规模、来源、处理目的、存储位置以及共享和出境情况，并依法设置网络数据安全负责人和网络数据安全管理机构。目录不能只停留在抽象的业务名称层面，而应当能够对应具体系统、数据表、接口、访问账户和责任团队，为后续权限控制和风险评估提供基础。

（二）加强访问、共享与传输控制

在数据范围和责任主体明确后，企业需要把保护要求落实到日常处理活动。对于重要数据，应根据风险采取最小权限、强化身份认证、加密、备份、分域存储和完整日志等措施。向客户、供应商或者合作方提供、委托处理或者共同处理重要数据前，还应依法开展风险评估，并通过合同约定处理目的、方式、范围和安全保护义务。通过 API 接口或者远程运维方式访问数据，同样可能构成数据提供，不能只管理文件下载。

（三）按照《网络数据安全风险评估办法》建立持续评估机制

《网络数据安全管理条例》第三十三条已经要求重要数据处理者每年度开展风险评估并报送报告。2026 年 6 月 18 日，国家网信办、工业和信息化部、公安部联合公布《网络数据安全风险评估办法》，该办法将于 2026 年 8 月 20 日起施行，并进一步明确风险评估的周期、方式、报告和监督要求。企业因此不能再把年度评估理解为一次漏洞扫描或者年末材料汇总，而应将其建设为覆盖网络数据和数据处理活动的持续治理机制。

按照《办法》，重要数据处理者应当每年度开展风险评估；重要数据安全状态发生重大变化，并可能对数据安全造成不利影响的，还应及时评估发生变化及其影响的部分。评估可以由企业自行开展，也可以委托第三方机构。自行评估时应指定专人负责；委托评估时则应通过合同或者其他具有法律效力的文件明确双方权利义务。同一评估机构及其关联机构不得连续三次以上为同一网络数据处理者开展年度评估。

评估结束并不意味着程序完成。重要数据处理者应按照主管部门规定编制报告；主管部门没有具体规定的，可以参照《数据安全技术 数据安全风险评估方法》（GB/T 45577 — 2025）等国家标准。报告至少保存三年，并应在年度评估完成后 20 个工作日内向有关主管部门报送；主管部门不明确的，则向省级网信部门或者国家网信部门报送。在数据处理活动存在较大安全风险并可能危害国家安全、公共利益，或者发生导致重要数据泄露、被窃取的安全事件等情形下，有关部门还可以要求企业委托通过认证的评估机构开展评估。

对具身智能企业而言，年度周期只是最低频率，不能替代对重大变化的即时判断。模型训练范围扩大、数据跨客户汇聚、感知精度提升、境外远程访问开放以及云端架构调整，都可能改变重要数据的安全状态。企业应当把这些事件纳入变更管理，并确保评估结论能够转化为访问权限、数据分域、训练任务、接口调用和日志策略的实际调整。

（四）依法管理重要数据出境

风险评估的结果还需要与数据出境安排衔接。境内运营中收集和产生的重要数据确需向境外提供的，应当通过国家网信部门组织的数据出境安全评估。企业宜在系统设计阶段优先减少原始数据出境，并评估境内训练、联邦学习、仅传输必要结果、访问隔离和本地运维等替代方案。《促进和规范数据跨境流动规定》关于“未被告知或者公开发布为重要数据”的安排，是判断是否需要将相关数据作为重要数据申报安全评估的重要依据，但不能被理解为对已经明确认定的重要数据提供豁免。

（五）建立事件响应机制并保留审计证据

即使企业已经建立预防性控制，也仍需为泄露、篡改、破坏和相关服务中断预留应急处置机制。此类事件可能同时影响客户运营和公共利益，企业应当预先明确事件分级、内部报告、客户协作、监管沟通、证据保全和业务恢复安排，并定期核查数据目录、访问权限、接口和日志是否与制度要求一致。事件处置结束后，相关原因和整改结果还应反馈至风险评估和控制措施，形成持续改进闭环。

七、企业自查：重要数据治理的十个问题

下面十个问题适合作为管理层、业务团队、法务和数据安全团队的共同核查框架。它们并非可以分别打勾的形式化清单，而是用于检验企业能否从业务事实出发，形成可追溯的识别依据，并把结论落实到系统控制和风险评估。

- 机器人将进入哪些行业和物理场所，其中是否涉及关键设施或者公共服务？
- 数据是否会跨设备、跨客户、跨区域或者长期集中汇聚，汇聚后能够呈现哪些新的信息？
- 企业是否已经检索并记录行业、地方和主管部门发布的重要数据目录、规则或者通知？
- 从数据精度、规模、覆盖范围、实时性和关联组合能力看，相关数据可能造成何种影响？
- 识别结论是否记录了事实依据、适用规则、复核人员和持续监测安排？

- 数据进入训练集、特征库、向量表示、模型更新或者分析结果后，是否仍可能反映关键场景？
- 重要数据目录是否与实际系统、接口、数据表、访问账户和责任团队逐一对应？
- 境外总部、境外云服务商或者境外远程运维人员能否访问相关数据，是否已经评估替代方案？
- 模型升级、部署范围扩大、数据精度提高或者汇聚方式改变时，是否会触发重新识别和专项评估？
- 企业是否已经建立年度风险评估、重大变化触发评估、报告保存和按期报送机制？

重要数据识别并不要求企业在缺乏目录和认定依据时一律采取最保守结论，但企业必须能够说明自己查阅了哪些规则、掌握了哪些事实、采用了何种判断方法，并在业务发生变化时重新评估。监管风险往往不只来自某一次结论是否准确，更来自企业是否建立了可以复核、可以更新并能够落实到系统控制的治理过程。

结语：让重要数据识别进入业务决策和系统架构

具身智能领域的重要数据风险并非由数据名称单独决定，而是在具体物理场景、持续汇聚和关联使用中逐步形成。企业既要以地区、部门和行业领域的重要数据目录及主管部门认定为核心依据，也应借助《数据安全技术 数据分类分级规则》（GB/T 43697—2024）等标准建立内部识别能力。识别完成后，还需要把结论传导至权限配置、训练数据管理、第三方提供、跨境访问和事件响应，避免重要数据治理停留在数据目录或者合规报告中。

《网络数据安全风险评估办法》的公布进一步强化了这一治理逻辑：重要数据管理不再是一次性的识别和申报，而是由年度评估、重大变化触发评估、报告报送、风险整改和持续复核构成的闭环。对于具身智能企业，越早把这种闭环嵌入设备部署、云端架构和模型训练流程，越能降低在跨境研发、客户交付或者监管检查时被动重构系统的成本。

在具身智能治理的四层框架中，本篇主要讨论关系国家安全和公共利益的重要数据治理问题；下一篇我们将进一步进入模型与智能体（Agent）层面，分析训练、评测、工具调用、长期记忆、身份标识和多主体责任。

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