DARKNAVY

CIIPA

数说安全
CYBERSECURITY REVIEWS

具身智能安全技术白皮书

——机器人篇

2026.04

目录

导语

一部最新的旗舰智能手机，专业网络安全研究团队实现远程完整攻破，通常至少需要数月；一辆成熟的智能汽车，攻击者实现多域系统的全面破解控制，周期甚至更长。而本团队对一台市面在售的知名品牌具身智能机器人实施渗透测试，在获取设备后，从漏洞识别到远程完整攻破，整个攻击周期不足8小时。

这一悬殊数据折射出了具身智能行业安全体系建设的严重滞后。更值得警惕的是，具身智能安全与传统网络安全存在本质差异：过去数十年的网络安全风险多局限于数字空间，直接后果往往是数据泄露或服务中断；而当智能体获得物理实体并具备自主作业能力后，数字世界的漏洞可直接转化为现实空间的物理伤害。一条精心构造的语音指令，或局域网内的一次无线信号注入，便足以劫持机器人的决策链路，引发失控乃至对周边人员构成直接威胁。

当前，具身智能产业正处于规模化部署前夕。2025年全球市场规模达44.4亿美元，人形机器人出货突破1.3万台，预计2035年部署量将超260万台。技术能力的高速演进与安全体系的缺失严重脱节，行业正面临一段高危的安全真空期。为此，本白皮书基于长期攻防实证研究系统梳理行业安全现状，旨在绘制切实可行的安全风险地图与防御指南，核心贡献包含三个方面：

1. 系统还原真实攻击路径。从攻击者视角解构系统架构，推演突破控制平面、篡改感知数据并最终劫持执行单元的完整渗透链路。

2. 发布RoboSec Top 10关键风险清单。归纳具身智能机器人当前最具代表性的十类高危问题，覆盖端侧权限、云端通信、控制逻辑、感知欺骗与AI资产完整性等核心环节。

3. 构建面向现实后果的风险评估框架。结合物理危害分级与防护优先级，为企业风险识别与安全基线建设提供分析依据。

本白皮书希望通过对具身智能机器人关键风险、攻击路径与防护重点的系统梳理，为行业提供有价值的分析参考，推动安全能力建设与技术发展同步推进。我们呼吁产业界、研究界与监管相关方共同关注具身智能在真实世界中的安全挑战，加快形成与之相适应的安全评估、治理与防护体系。鉴于相关技术与风险形态仍在快速演进，相关研究仍需不断完善，本白皮书中疏漏不足之处，恳请大家批评指正。

本白皮书由 DARKNAVY 主持撰写，**CIIPA中关村华安关键信息基础设施安全保护联盟、数说安全**协助编写并联合发布。同时，向在本白皮书研究、撰写与完善过程中提供指导、支持与帮助的各行业单位及专家，致以衷心感谢。

声明：本文所涉案例均基于公开研究成果，严格遵循行业负责任披露惯例，不包含任何可直接复现的攻击代码或未公开漏洞细节。

0. 具身智能进入现实世界，风险从数字走向物理

具身智能是指赋予智能体物理实体，使其能在真实环境中运行，形成感知、决策、执行的闭环，通过物理交互完成各类任务。其范畴涵盖具身智能机器人、智能无人车与无人机等系统；本篇作为具身智能系列安全白皮书的首篇内容，重点聚焦具身智能机器人。

过去数年，具身智能机器人正从实验室与展示性应用，快速走向真实部署场景。机器人数量持续增长，能力日益增强，从简单的路径规划到复杂的多模态感知和自主决策，并逐步呈现出规模化部署特征。根据行业数据，2025年全球人形机器人市场迈入快速增长阶段，全年总出货量预计达 1.3万台。全球具身机器人部署量预计到2035年将超过260万台。

与传统软件系统或单一自动化设备不同，具身智能机器人具备持续感知环境、自主决策并作用于现实的能力。一旦系统失控或遭受攻击，风险将从数字空间外溢，对现实世界造成直接影响或物理伤害。

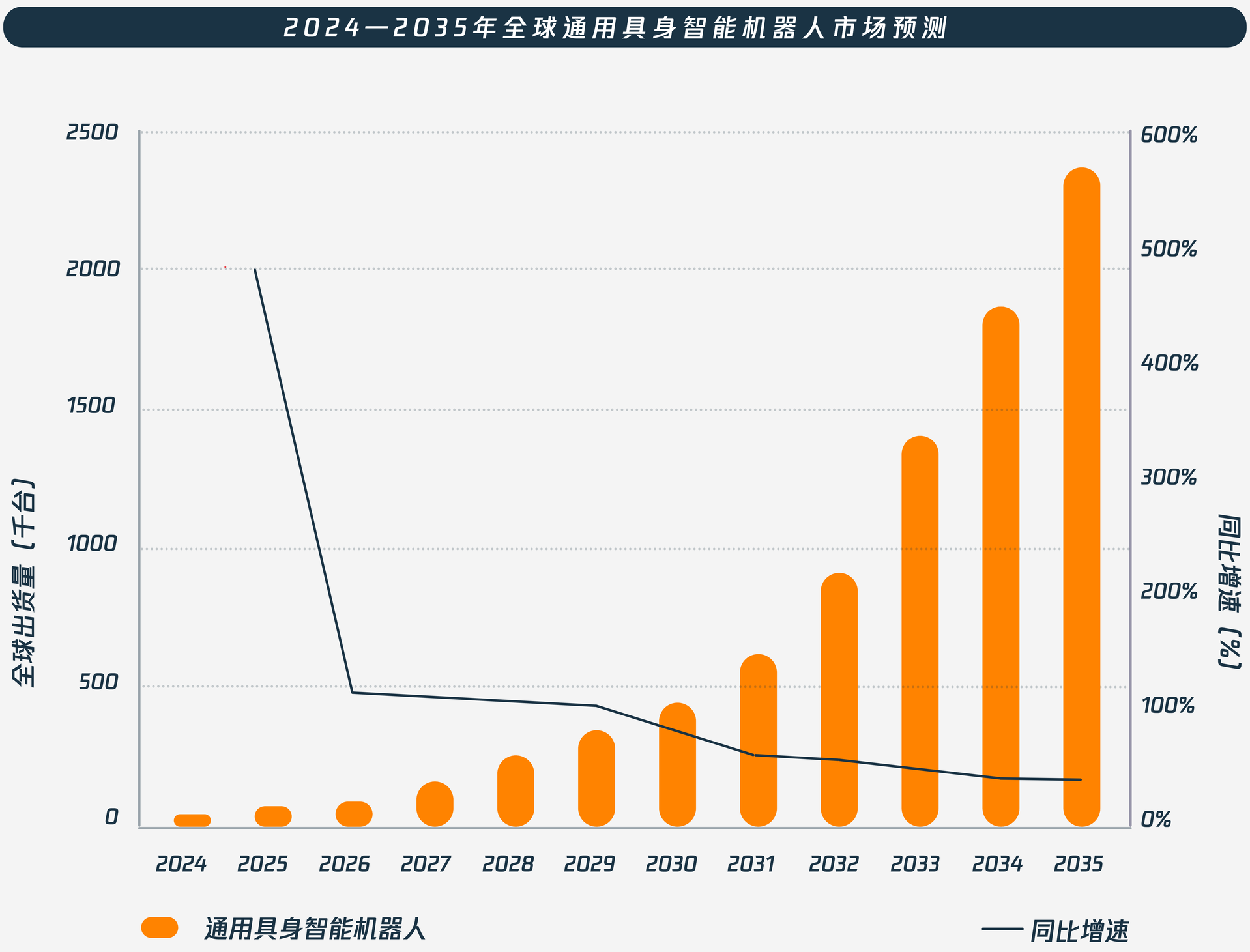


图1 具身智能机器人市场规模与部署趋势示意

资料来源：Omdia，2026

资料来源：Omdia《Omdia Market Radar: General-purpose Embodied Intelligent Robots, 2026》，见参考资料[1]。

1. 具身智能机器人安全现状

1.1 具身智能安全水位低于早期智能终端

当前，具身智能行业正处于类似于早期智能终端及物联网（IoT）行业的快速扩张期。厂商的核心竞争焦点普遍集中于算法精度、硬件性能、任务完成度及成本优化。安全防护往往被视为影响用户体验或增加成本的非核心需求，在设计与开发流程中存在严重的后置或缺失现象。

根据 DARKNAVY 对国内主流品牌具身智能机器人安全现状的深度调研与对比分析，研究显示：当前国内主流产品的整体安全水平尚未达到早期智能终端及物联网设备的安全基准。

实测验证表明，以某知名品牌具身智能机器人为例，从获取物理设备、识别系统漏洞到实现完全控制，整个攻击周期不足8小时。相比之下，传统智能终端的防护强度显著更高，例如市场主流智能手机和智能汽车的破解，往往需要数月甚至一年以上。这一对比反映出具身智能领域网络安全防护能力普遍处于较低水平，远逊于传统智能终端。

同时，具身智能系统融合了云端协同、模型能力、通信网络及运动控制等多重技术栈，其攻击面较传统智能终端、IoT系统及功能单一的工业机器人更为广泛，安全设计与实现的复杂程度也大幅提升。

此外，具身智能系统具备感知、决策、执行的闭环特征，面临更为多样的外部输入与安全挑战。例如，攻击者可通过劫持具身智能体获取对机器人的完全控制权。同时，部分具身智能机器人具备群体性管理与运营的特点，这也使蠕虫式攻击可能在设备间快速传播，演变为大规模系统性风险。

由于具身智能系统具备多维度的物理世界感知能力与深度的物理干预能力，一旦发生滥用或失控，其潜在后果将远超传统终端。这不仅会对数字世界的安全与隐私（Security & Privacy）构成重大威胁，更将直接危及现实世界的物理安全与人身安全（Safety），甚至可能引发公共安全事件，导致公众对人工智能系统的信任危机。

表1 不同智能设备安全属性对比

设备	功能复杂度	攻击场景复杂度	安全强度	安全机制成熟度	攻击后果严重程度	相关标准成熟度	责任复杂度
具身智能机器人	感知-决策-执行耦合；能力开放	应用场景多样化；新型AI攻击面	🔴	🔴	🔴	🔴	🔴
智能手机	平台系统；生态与安全域并存	攻击入口广；攻击链路长	🟢	🟢	🔴	🟢	🟠
智能汽车	多ECU/总线/车云互联	多域耦合；运维与供应链参与	🟠	🔴	🔴	🟠	🔴
智能门锁	单功能闭环；依赖链短	入口集中；场景相对固定	🟠	🟢	🟠	🟢	🟢

注：颜色表示风险或能力缺口程度：

🟢 表示成熟度较高或风险较低　🟠 表示中等风险　🔴 表示风险高或安全能力缺口明显

1.2 具身智能机器人安全风险

机器人相关的现实安全事故早已屡见不鲜。例如，2023年特斯拉工厂中，工业机器人因传感器故障导致安全边界内的操作员重伤。此类工业机器人在失控、误操作或防护不足情况下造成人员伤害的案例，反复证明了物理系统一旦失效，失控的机器人将对物理世界造成严重破坏。

而具身智能机器人一旦失控或被劫持，其强大的行动能力便会转化为破坏物理世界的武器。在2025年GEEKCON大赛中，研究人员演示了如何通过远程网络攻击和人传人式攻击影响离线机器人，实现劫持并使其做出危险动作。

此外，DARKNAVY团队对当前针对具身智能机器人的公开攻击案例进行了梳理，发现这些攻击案例存在以下特征：

针对性强且后果严重：大量攻击聚焦于应用较广的四足机器狗，并能实现完全控制。攻击者借此具备了对电力网络、巡线管道等基础设施进行物理破坏的能力。

攻击门槛低：绝大部分攻击可从远程发起，无需接触设备或用户配合即可夺取控制权。多数案例要求攻击者与目标处于邻接环境〔物理靠近或网络邻接〕，少数案例甚至能通过网络实现大规模、群体性控制。

漏洞类型多样：攻击手段以传统智能终端通信漏洞为主，同时也开始出现具身智能独有的安全问题。

表2 公开披露的具身智能机器人攻击案例（2022–2025）

时间	目标设备	团队	攻击入口	攻击条件	攻击后果
2022	某商用四足机器人	GeekPwn	UWB 模块	处于 UWB 覆盖范围内 无接触攻击	获取完整控制权 任意移动操控
2024	某主流商用四足机器人	Bin4ry 团队	云端	网络可达 设备连接云端	完全控制权 理论上可控制移动
2025	某主流商用四足机器人	TheRoboverse 社区	机器狗 Web 服务	同一局域网	完全控制权 理论上可控制移动
	某主流商用四足机器人 某主流商用人形机器人	DARKNAVY / Bin4ry 团队	蓝牙模块	蓝牙信号可达的物理范围内	完全控制权 可操控机器人 可蠕虫式感染
	某主流商用人形机器人	DARKNAVY	智能体模块	互联网可达	完全控制权 可操控行为逻辑
	某主流商用人形机器人	DARKNAVY	SDR 无线通讯模块	远距离无线电可达（约2-3km）	完全控制权 可操控行为逻辑
	某主流商用四足机器人	Pwno	蓝牙协议线	蓝牙可达	完全控制权 理论上可控制行动
	某主流商用四足机器人	百度	基带服务	同一局域网	完全控制权 控制移动

注：本表旨在反映具身智能机器人中普遍存在的安全问题，具体相关品牌已隐去

这些近年公开的安全研究进一步表明，具身智能系统不仅可能因故障造成伤害，更可能在恶意控制或安全缺陷被利用的情况下，执行非预期的危险物理行为。这标志着具身智能的安全风险已从理论推演阶段进入实证验证阶段，必须引起高度重视。

然而，目前部分具身智能企业仍热衷于通过机器人的预设攻击性表演来博取公众关注，其安全体系的建设进度普遍远滞后于机器人能力的提升。

1.3 国内外具身智能安全成熟度对比

通过对国内外三十家知名具身智能机器人研发与销售企业开展安全建设及投入调研，可以发现，国内外企业在具身智能安全上的投入重点存在客观差异。海外领先企业遵循安全设计优先原则，将安全考量置于快速迭代之前；而国内企业多以商业化落地为先，部分企业的安全体系建设相对滞后。

海外企业以波士顿动力（Boston Dynamics）为代表，其于2024年针对具身智能四足机器人发布了《Spot与Site Hub安全白皮书》，公开了其安全设计目标、架构及具体的防护措施（如加密通信、权限分离、安全状态监控等），传递出将安全视为核心产品属性并主动进行透明化沟通的负责任态度。此外，特斯拉（Tesla）多次强调安全对机器人的重要性，并在团队组建时设立了专门的安全岗位。这标志着海外顶尖厂商已开始系统性地构建覆盖具身智能机器人硬件、软件和运营的全生命周期安全体系。

相比之下，国内大多数机器人厂商在追求技术突破与商业化落地速度的同时，安全能力建设仍处于早期阶段。尽管部分厂商（如宇树、优必选等）已开始建立安全应急响应中心（SRC）并招募安全人才，但整体而言，大量企业尚未对安全予以足够重视，在产品设计阶段缺乏系统性的安全体系考量，安全团队及运营机制尚不完善。

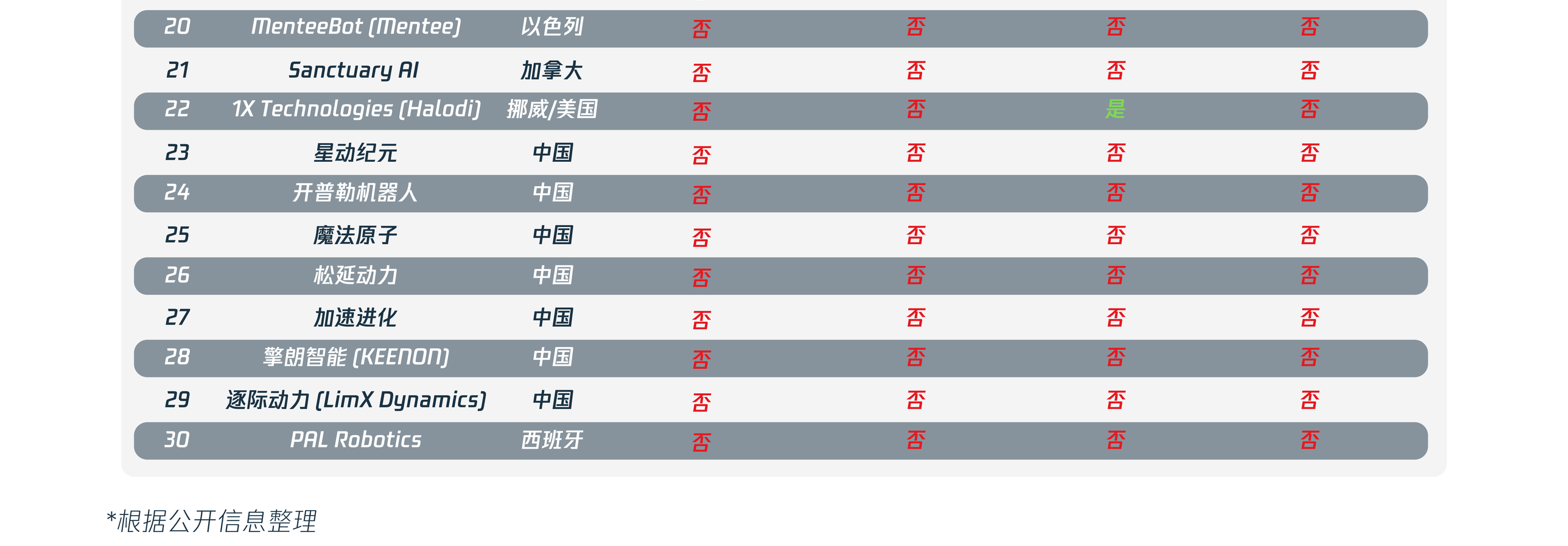
相较于功能的快速迭代，将安全深度融入产品设计、开发、测试、运营全生命周期的成体系安全能力在多数企业中仍然缺失。根据头部企业的安全建设经验，若缺乏完善的安全响应机制，在面对突发安全事件时将难以有效应对。此外，行业普遍缺乏针对具身智能特性的通用安全标准、测试基准和认证体系。

综合观察发现，当前具身智能行业整体正处于安全能力逐步觉醒但尚未成熟的阶段，安全正在从事后补救转向系统设计必备的核心能力。

表3 国内外具身智能机器人厂商安全能力成熟度对比（截止2026年2月）

序号	厂商	国家	是否公开具身智能产品安全白皮书/安全考量	是否拥有SRC/漏洞报告渠道	招聘范围是否包含Security 岗位	招聘范围是否包含Safety 岗位
1	Tesla (Optimus)	美国	否	是	是	是
2	小米机器人	中国	否	是	否	否
3	Figure AI	美国	否	否	是	是
4	追觅机器人 (Dreame)	中国	否	否	否	否
5	智元机器人	中国	否	否	是	是
6	小鹏机器人	中国	否	是	是	是
7	乐聚机器人	中国	否	否	否	否
8	NEURA Robotics	德国	否	否	是	是
9	优必选 (UBTECH)	中国	否	是	否	否
10	Apptronik	美国	否	否	否	是
11	银河通用机器人	中国	否	否	否	否
12	越疆机器人 (DOBOT)	中国	否	否	否	否
13	Agility Robotics	美国	否	否	否	是
14	宇树科技 (Unitree)	中国	否	是	是	否
15	傅利叶智能 (Fourier)	中国	否	否	否	否
16	Boston Dynamics	美国	是	是	是	是
17	众擎机器人	中国	否	否	是	否
18	非夕机器人 (Flexiv)	中国	否	否	否	否
19	云深处科技 (DeepRobotics)	中国	否	否	否	否
20	MenteeBot (Mentee)	以色列	否	否	否	否
21	Sanctuary AI	加拿大	否	否	否	否
22	1X Technologies (Halodi)	挪威/美国	否	否	是	否
23	星动纪元	中国	否	否	否	否
24	开普勒机器人	中国	否	否	否	否
25	魔法原子	中国	否	否	否	否
26	松延动力	中国	否	否	否	否
27	加速进化	中国	否	否	否	否
28	擎朗智能 (KEENON)	中国	否	否	否	否
29	逐际动力 (LimX Dynamics)	中国	否	否	否	否
30	PAL Robotics	西班牙	否	否	否	否

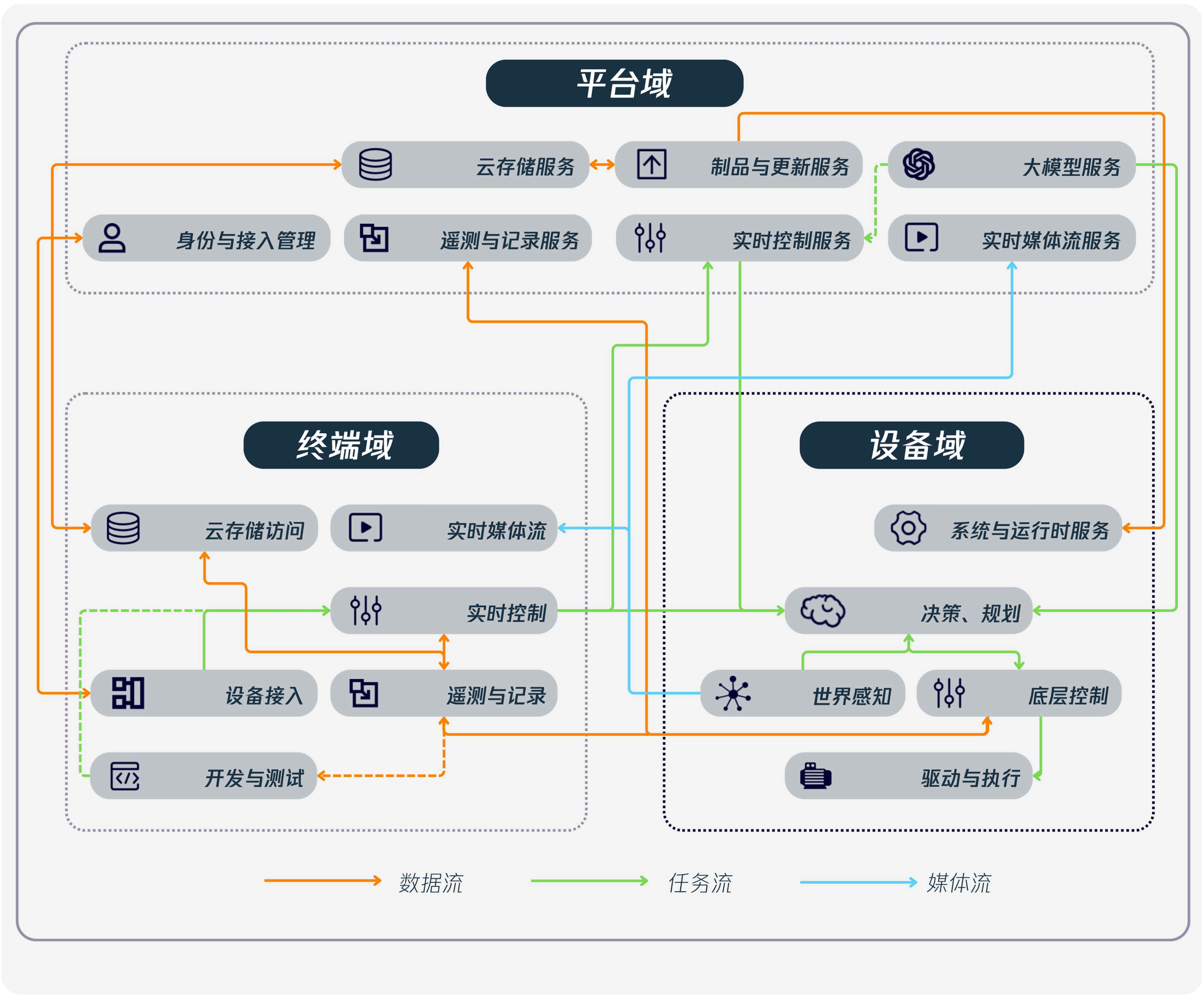
*根据公开信息整理



2. 具身智能机器人安全架构梳理

2.1 典型架构

具身智能系统的实现架构与职责分工可分为平台域、终端域、设备域三个部分：平台域需要向终端域与设备域提供基础服务与平台支撑，终端域主要承载用户交互以及提供必要开发测试的功能，设备域负责任务可信执行。三者通过数据面、任务面、媒体面协同形成可管、可控、可观测、可更新的整体系统能力，为后续安全边界划分与防护机制设计提供架构基础。



平台域

围绕账号体系，平台域需要向终端域与机器人域提供设备接入、设备绑定/解绑与共享关系管理，以及面向审计与合规的日志存储与风控基础能力。

在具身智能控制方面，平台域负责为终端以及设备建立远程网络环境下的可信连接。具体而言，平台域负责任务下发与任务状态机管理，提供高层任务创建/开始/暂停/取消、执行编排与结果回传的能力；另一方面，平台域将运控、音频、灯光等控制流进行协议适配与下行分发，确保指令在不同型号与版本间的一致语义与可追踪执行。此外，平台域还提供遥测与记录服务，并为实时媒体流建立信令与传输能力，支撑预览、录制、回放与跨端访问。平台域还可引入大模型能力，作为可选的推理与编排组件接入，提供对话式交互、任务生成规划等服务。

此外，平台域为设备提供固件更新服务与云存储服务：覆盖版本管理、更新分发与校验策略下发以及媒体、日志、地图/模型等对象的存取与生命周期管理。

终端域

终端域承担用户交互的主要职责，面向用户提供设备接入全流程能力：完成设备发现、配网激活与首次注册，引导绑定/解绑与共享授权，并对账号登录态与会话进行管理。

在控制与交互层面，用户终端承载实时控制任务，完成具身智能终端媒体流、对讲、截图/录制等输出以及将用户意图转化为可执行的控制与任务请求，实现对具身智能终端的控制。

此外，用户终端还提供状态与遥测可视化能力以及云存储访问能力，在研发或运维场景中，用户终端在严格的权限与审计约束下提供开发与调试能力，例如诊断视图、日志导出、配置开关与测试接口。

设备域

设备域是承担任务可信执行的核心角色，提供基础系统与运行时服务以保证设备可控、可更新、可追溯：

设备域首先需要为其他服务提供世界感知的底层能力，管理摄像头、激光雷达、麦克风、IMU 与各类传感器的数据采集、时间同步与预处理；并通过决策与规划模块实现从任务输入到控制指令输出的转化，包括任务接收与本地校验、路径/轨迹规划、约束条件与安全策略执行。

在具体执行方面，设备提供稳定的控制与驱动体系：将控制指令落实到运动控制、音频控制、灯光控制等底层控制接口，完成关节电机、执行器、扬声器等硬件的驱动与控制，并向终端域以及平台域上报运控状态与执行反馈。

2.2 安全目标

与传统信息系统不同，具身智能不仅需要在复杂环境中进行自主感知、推理与决策，还能够直接作用于物理世界，其行为结果具有现实可见、可触达、可放大的安全影响。因此，具身智能安全不应仅停留在数据与网络层面的保护，更应覆盖具身智能系统任务下发、决策、规划、执行全链路的可信与可控，确保系统在面对攻击、缺陷或异常状态时仍能保持行为可控、风险受限与责任可追溯。基于上述认识，本文从具身智能安全的视角提出以下安全目标：

具身智能行为可约束：具身智能系统应具备清晰的授权边界与职责范围，确保指令来源、决策链路与执行路径的完整性与抗干扰能力，具身智能系统在任何状态下均不得扩权、越界执行或绕过既定安全限制。

系统性风险不扩散：具身智能系统应具备面向系统性风险的隔离抑制能力，确保任意单点缺陷、单点攻击、单次错误更新或关键依赖失效，不会引发跨设备、跨区域、跨场景的批量异常或危险行为扩散。

用户控制权与隐私数据可控：具身智能系统应确保机器人控制、管理、所有权变更与数据访问严格归属于合法所有者，传感数据、运行数据与用户隐私信息应在采集、传输、存储、使用与共享各环节保持机密性与完整性。

事件可追溯：具身智能系统应满足重大风险事件可追溯、可取证、可复盘能力。事故或争议发生后，应能够以可信证据还原关键事实并可被验证，包括：发生了哪些关键行为、由谁触发、在何种系统状态与策略约束下发生、影响范围与传播路径为何，以及相关证据链的完整性与可信度。

在安全目标设定上，本文参考了机器人与功能安全相关的既有标准体系，包括工业机器人安全要求 ISO 10218 [3]、服务/个人护理机器人安全要求 ISO 13482 [4]、协作机器人安全技术规范 ISO/TS 15066 [5] 以及功能安全标准 IEC 61508 [6]。这些标准为机器人设计、系统集成、人机协作与安全功能实现提供了重要基础。与此同时，从适用范围与关注重点来看，现有标准主要覆盖机械安全、作业风险控制与功能安全等问题，对具身智能系统在云端控制、端侧权限、模型资产、世界感知与智能决策链路等方面的新型安全风险仍难形成完整覆盖。基于此，本文在借鉴既有标准思路的基础上，进一步从感知、决策、执行这一闭环出发，对具身智能系统的系统性安全问题进行补充分析。

2.3 典型攻击场景

具身智能具备自主决策并作用于物理世界的能力，不同能力的攻击者可在不同阶段、不同层级选择切入点并进行组合利用，从而形成目标各异、能力分层的多样化攻击场景。基于此，本文归纳并总结了7类典型攻击场景，用以刻画具身智能系统在真实部署环境下的主要威胁形态与风险表现。

场景 1：云端未授权访问（突破用户—设备绑定/授权边界）

当云端服务在权限隔离与功能访问存在缺陷时，即使攻击者不具备目标设备的合法绑定身份，也可能通过云端开放接口直接触达目标设备的控制、调度、媒体与遥测等关键能力。此类攻击常表现为对目标设备控制面与数据面的越权访问：越权下发任务、切换模式、建立遥操作通道，或读取媒体流、遥测数据、云存储对象与运行日志。

一旦云端服务存在此类风险，其影响通常具有放大性：通过枚举设备标识、滥用调度与连接资源、或诱发级联故障，攻击者可将单设备风险扩展为系统性风险；同时，越权访问获得的会话与令牌信息又可以为更深层的端侧突破奠定基础。

场景 2：公网合法身份接入并向端侧扩权（账号接管/令牌泄露/被授权滥用）

在攻击者已具备合法控制通道的前提下，风险重心转向具身智能终端侧执行权限的获取与提升。由于指令与数据流符合授权路径，攻击者能够绕过权限控制，沿任务编排、控制栈、通信中间件与端侧执行链路持续深入，最终触达驱动执行层、设备更新、设备策略文件等高风险目标。

该场景的直接后果是具身智能行为完整性与安全约束的系统性削弱：通过篡改轨迹与关键控制参数、旁路模式限制、污染配置与日志，可导致目标设备的执行异常；若进一步获得端侧更高权限，还可能实现持久化与横向扩散，使单一缺陷演化为跨设备的长期控制能力。

场景 3：邻接网络接入（同局域网）

当攻击者处于目标设备局域网环境内时，本地发现、首次配网/绑定、局域网控制与媒体/遥测直连链路将成为暴露的攻击面。通过对链路进行嗅探、重放、逆向与篡改，攻击者可以破坏会话机密性与指令完整性，进而实现会话接管、指令注入、设备与控制端伪装，或造成控制不可用与任务异常。

邻接网络攻击通常同时具有门槛低、联动强的特点：既可能直接窃取媒体与遥测数据，也可能在首次绑定窗口获取关键凭据或植入恶意配置，从而将邻接突破转化为远程网络层的持续控制；同时，若存在本地控制枢纽或边缘网关，单点被控还可能带来同域设备间的横向移动能力。

场景 4：近场无线接入（蓝牙/遥控器/其他近场链路）

近场无线链路承担配网、配对、遥控器控制、重连与应急控制等关键职能，其安全性高度依赖攻击者需身处现场这一前提。一旦配对协议、重绑定流程或控制报文在认证、执行、抗重放、鉴权与密钥管理上存在缺陷，攻击者即可在覆盖范围内实施伪装或重放，造成未授权控制、配对劫持、控制权争夺以及关键控制链路不可用。

近场突破的危险性在于，攻击路径往往可以进一步延伸：一旦通过配对劫持建立稳定连接，攻击路径便可能扩展至邻接网络与云端控制面，形成更长的攻击链；同时，近场链路还可能成为局部扩散的媒介，被攻击者利用对周边设备发起近场攻击。

场景 5：物理接入（接口级接触与硬件级占有）

当攻击者获得物理接触设备接口能力时，维护口、调试口、可插拔介质与启动链成为主要突破点。接口接触可导致配置与数据导出、指令注入与维护功能滥用；硬件级占有则可能进一步触及调试接口与存储介质，实现离线提取敏感信息、克隆镜像、替换固件、篡改启动参数或植入持久化后门。

物理攻陷对体系安全的破坏具有结构性：一方面可直接获取设备身份材料与密钥，反向削弱云端对设备的信任基础；另一方面可通过返修回流、二次流转形成扩散效应。

场景 6：物理环境影响（对多模态输入与定位的对抗）

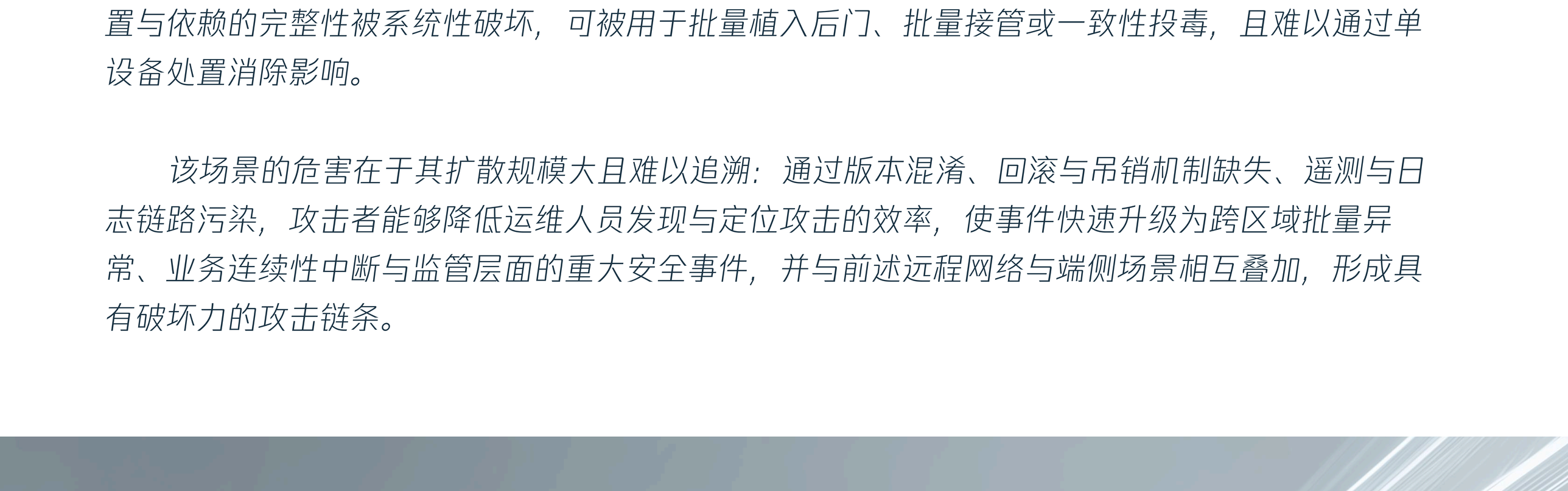
即便不接触网络与设备本体，攻击者仍可通过操控环境输入影响机器人对世界的感知与理解，从而在规划与控制闭环中引入系统性偏差。多模态感知、定位建图与场景理解一旦被干扰，错误将沿链路放大为错误决策与危险执行，表现为误识别、误定位、越界运动、错误抓取与安全距离失效等安全后果。

此类攻击同时具备隐蔽性与持续性：持续扰动可触发反复失败与重规划，造成可用性下降；诱导机器人接近敏感区域则可能扩大数据采集边界，带来隐私与合规风险，并在空间层面为后续的邻接接入与物理接触创造条件。

场景 7：运维与供应链链路影响（更新/模型/依赖被污染，触发跨设备扩散）

当攻击者能够影响发布与运维体系以及供应链服务，如云端大模型供应链、固件发布等服务，更新与下发机制会快速将此类攻击扩散到每一台目标设备上。在此条件下，固件、模型、策略、配置与依赖的完整性被系统性破坏，可被用于批量植入后门、批量接管或一致性投毒，且难以通过单设备处置消除影响。

该场景的危害在于其扩散规模大且难以追溯：通过版本混淆、回滚与吊销机制缺失、遥测与日志链路污染，攻击者能够降低运维人员发现与定位攻击的效率，使事件快速升级为跨区域批量异常、业务连续性中断与监管层面的重大安全事件，并与前述远程网络与端侧场景相互叠加，形成具有破坏力的攻击链条。



3. 具身智能机器人风险评估框架

3.1 框架概览

具身智能机器人具备持续感知、决策与物理执行能力，同类技术缺陷在具身系统中可能进一步影响控制链路与执行机构，使风险从数字域外溢为现实世界后果。为便于在不同部署场景下开展风险识别、分级与处置优先级讨论，本章给出一套风险评估参考框架，包含三部分：〔1〕以现实世界后果为锚点的五级危害分级〔L1–L5〕及放大/升级原则；〔2〕RoboSec Top 10 关键风险清单及其对应危害等级；〔3〕结合公开研究与演示信息，对典型攻击路径进行结构化拆解。

为描述缺陷、利用与后果之间的传导关系，本章采用三要素表述：

攻击途径：威胁行为体介入系统的入口与初始能力〔如公网远程、局域网邻接、近场通信、物理维护接口、传感器输入欺骗等〕。

系统薄弱点：在既定途径下可被利用的缺陷或不当设计/配置〔如认证与权限、通信可信机制、更新与供应链、模型与相关资产完整性治理、默认配置、OS/中间件/驱动与运控组件漏洞等〕。

潜在后果：薄弱点被利用后可能引发的现实世界影响〔从信息泄露与可用性下降，至财产损失、环境破坏与人身安全风险〕。

本章内容基于独立第三方视角与公开信息整理，用于提供可复用的分析参照，不构成强制性标准；具体结论需结合威胁模型、部署场景与安全边界进行评估。

3.2 安全影响分级

分级以**攻击可稳定复现的最严重现实后果**为准。

表4 具身智能机器人安全影响分级（L1–L5）定义与判定标准

等级	名称	判定标准	典型数据/场景	备注
L1	信息泄露与隐私风险	攻击者可读取、导出或推断敏感信息，但无法改变机器人行为	视频/音频流、SLAM 、位姿轨迹、家庭/工厂布局、账号密钥、场景语义信息等	以数据可见性为主，不涉及控制面影响
L2	功能滥用与非预期行为	攻击者可触发未授权功能、耗尽资源、造成服务中断或异常任务执行，但不产生危险物理动作	拒绝服务攻击、异常任务堆积、误导性提示、低风险动作被滥用等	关注可用性与“非危险行为”的偏离
L3	设备控制权获取	攻击者可稳定影响关键控制面，形成持续或可重复的越权能力（包括提权、持久化、策略/配置/任务篡改、遥操作接管等），为更高等级后果铺路	权限提升后篡改任务队列、接管遥控接口、持久化植入后门以反复执行未授权命令等	只要能操纵关键策略或控制入口即视为 L3
L4	物理环境实质性破坏	攻击导致明确的财产损失、环境破坏、越界进入受限区域或生产中断，但不以直接伤人为主要后果	撞毁物体、扰乱现场运行、误入危险区域、设备永久性损坏等	以可观测的物理/运营损失为主
L5	人身安全威胁	攻击可引发与人体接触相关的高风险行为，或显著提升伤害概率（尤其在有人协同或近人场景）	绕过安全限制执行高能量动作、近人高速移动、危险末端动作、失控冲撞等	直接涉及人身伤害风险

3.2.1 分级评估原则与风险放大因子

在应用上述分级时，需遵循以下原则，以适应具身智能风险的复杂性与潜在放大效应：

后果导向：等级判定应基于最终物理后果，而非攻击的技术复杂度。一个简单的配置错误若可直接引发L5后果，其风险等级即为L5。

安全兜底失效视为最高风险：任何能够导致或利用硬件或固件级最终安全兜底机制（如硬线急停、不可旁路的硬件限速）失效的攻击，无论其初始入口如何，均应被视为至少 L4级风险，在近人场景下即为 L5级。

规模放大因子：当单个风险点可通过云端控制平台、集群调度系统或同构软件供应链影响整个机器人群体时，其整体风险与处置优先级应显著提升。例如，单设备的L2级服务中断，在车队规模下可能意味着城市级服务的瘫痪。

场景敏感因子：同一技术风险在不同部署场景下的危害差异巨大。评估需依据明确的场景假设，例如无人仓库与开放式养老院，并据此调整风险等级。

3.3 RoboSec Top 10 关键风险清单

RoboSec Top 10 是面向具身智能安全的关键风险清单，旨在归纳当前阶段最具代表性的十类安全问题，为行业提供统一的风险认知框架与评估参照。该清单综合公开研究成果、案例披露信息与实际攻防观察形成，在方法上借鉴风险清单类工作的通行做法，并结合具身智能系统感知、决策、执行的闭环特征，重点纳入物理后果、系统扩散性与部署环境等关键因素。

表5 RoboSec Top 10 关键风险清单（2026）概览

编号	风险类别	核心描述（简洁版）
RS01:2026	端侧内部权限模型缺陷	端侧设备内部组件权限边界混乱，低权限进程可越权访问运动控制等高危功能
RS02:2026	端侧外部暴露面管理缺陷	设备存在调试接口开放、默认密码、非必要服务端口等开发阶段遗留问题
RS03:2026	内部通信安全机制缺失	端侧设备内部组件间通信缺乏认证与加密，内部网络被过度信任
RS04:2026	云端控制平面风险	云端群控管理/任务编排/远程协助等管理平台存在越权或隔离缺陷，可引发大规模设备群体性失控
RS05:2026	感知与决策层欺骗	Agent/VLA等决策链路被对抗样本或提示注入等攻击欺骗，输出错误或危险的行为指令
RS06:2026	控制栈软件漏洞	存在内存破坏等软件漏洞，导致控制链路被破坏
RS07:2026	软件供应链污染风险	第三方软件依赖缺乏来源验证与完整性保障，被植入恶意代码可导致大规模设备失控
RS08:2026	AI资产完整性失控	AI 模型、API服务、训练/评测数据及技能/动作库等核心资产缺乏防篡改校验，存在被投毒或恶意替换的风险
RS09:2026	更新与启动链完整性缺陷	固件更新与系统启动过程缺乏有效签名验证，易受恶意更新包攻击，可导致恶意更新包刷写篡改系统行为
RS10:2026	安全兜底机制失效	缺乏硬件级、不可旁路的最终安全措施，使数字攻击可直接造成物理伤害

*其作用主要在于支持风险识别、优先级研判与防护基线建设，而不将其作为严格意义上的严重性排序。

详细说明

以下为各风险类别的详细说明，用于深入理解其成因、影响与关联的危害等级范围。

1. RS01:2026 – 端侧内部权限模型缺陷

详细描述：指机器人设备端操作系统或运行时环境未能严格遵循最小权限安全原则。表现为不同进程、服务或用户账号之间的权限边界模糊或隔离失效，致使原本低权限的软件实体能够越级调用或访问关键高危功能模块，例如直接发送运动控制指令、篡改任务队列或修改限速限力、碰撞阈值等安全参数。

危害说明：此缺陷是攻击者在突破外部防御后，在设备内部进行**权限提升**和**横向移动**的关键跳板。一旦攻击者控制了一个普通进程，便可利用此缺陷直接夺取核心控制权，从而为实施L4〔环境破坏〕或L5〔人身伤害〕级别的攻击铺平道路。

2.RS02:2026 – 端侧外部暴露面管理缺陷

详细描述：指机器人产品在出厂或现场部署时，未能完成基本的安全加固，遗留了不必要的攻击入口。典型情况包括未更改默认用户名/密码、开启了非业务必需的远程访问服务〔如FTP、Telnet〕、以及将用于开发的调试接口〔如SSH、ADB、UART串口〕暴露在外且缺乏访问控制；在具身系统中还常见将**标定、运控调参、运动单元测试**等高风险维护能力与常规接口混用。

危害说明：这些疏漏**极大降低了攻击者的初始入侵难度**，使其无需利用复杂漏洞即可获得设备的一定访问权限。该风险常与RS01或RS06结合，使攻击链得以启动，因此其危害等级常从初始的L3〔获得入口〕向更高级别演变。

3.RS03:2026 – 内部通信安全机制缺失

详细描述：在采用ROS 2、DDS等框架的系统中，各功能模块〔节点〕间通过消息总线进行通信。本风险指此类内部通信**默认缺乏身份认证、消息加密、完整性校验和防重放攻击机制**，且网络域隔离不足。在具身系统中尤需关注**控制话题、状态话题、安全相关话题**在端内被默认可信处理的情况。

危害说明：这使得攻击者一旦通过某种方式接入机器人内部网络〔如利用ROS 2〕，就能**监听、伪造或篡改**任意模块间的指令与数据。危害范围取决于可操控的消息类型：窃取感知数据为L2，篡改导航目标为L3-L4，若直接注入运动控制指令则可能直达L5。

4. RS04:2026 – 云端控制平面风险

详细描述：指用于集中管理机器人集群的云端平台存在的安全漏洞，主要包括：用户越权访问他人设备、不同客户〔租户〕的数据与控制指令未有效隔离、以及提供可批量操作大量机器人的接口而未实施严格的风险管控与审批流程。在具身智能场景中，云端往往承担**群体管理、任务编排、远程操作与策略下发**等能力，其缺陷更易产生群体影响。

危害说明：此风险具有明显的放大效应。攻击者一旦攻破某个云端账户或发现平台漏洞，便可能对多台设备形成规模化影响，导致大面积设备异常〔L4〕；若批量指令涉及近人作业或高危动作，则可能上升至L5。

5. RS05:2026 – 感知与决策层欺骗

详细描述：针对机器人依赖的AI模型发起的非传统攻击。包括：在摄像头前放置特殊对抗图案误导视觉识别，或向语音/文本交互接口输入精心设计的指令，诱使大模型等决策系统突破预设的安全规则。在具身智能系统中，此类输入往往进一步进入**VLA/Agent 的任务分解、工具调用与动作生成链路**，从而间接影响控制输出。

危害说明：此类攻击未必需要借助传统软件漏洞，仅通过误导决策即可实现。其危害起初可能表现为功能异常〔L2〕，但若成功诱导系统规划出一条穿越安全边界的路径〔L4〕或在人附近执行高速动作〔L5〕，后果将十分严重。危害是否升级，关键取决于系统是否具备冗余的安全校验机制〔RS10〕。

6. RSO6:2026 – 控制栈软件漏洞

详细描述：指机器人软件栈底层，包括操作系统内核、硬件驱动程序、通信中间件〔如ROS 2底层〕、运动控制服务等关键组件中，存在的各类经典软件漏洞。例如内存缓冲区溢出、命令注入、反序列化漏洞等。

危害说明：这些漏洞是攻击者实现**代码执行、权限提升或拒绝服务**的直接技术手段。利用此类漏洞，攻击者可以从一个较低权限的访问点，逐步夺取系统的最高控制权〔L3〕。若漏洞直接影响实时控制环路或安全监控功能，则可能直接导致物理失控〔L4/L5〕。

7. RSO7:2026 – 软件供应链污染风险

详细描述：指机器人的软件开发高度依赖第三方开源库、商业SDK、容器基础镜像等，但缺乏对所有这些依赖项的**来源可信验证、完整性签名校验和全生命周期追溯**能力。攻击者可通过污染上游源码、劫持软件仓库或植入恶意构建工具来注入后门。

危害说明：这是一种具有**先天隐蔽性**和**广泛传播性**的高风险。恶意代码在开发阶段就已被植入，并随合法软件分发给所有用户。其初始危害等级为L3〔潜伏后门〕，一旦被远程激活，可同时控制海量设备，执行数据窃取、大规模破坏或持久化驻留等动作，危害可升至L4/L5。

8.RSO8:2026 – AI资产完整性失控

详细描述：机器人的核心智能资产〔模型权重、训练/评估数据、知识库、决策策略文件、 模型服务接口等〕在存储、传输、加载和更新过程中缺乏有效手段确保不被篡改、替换或发生非预期变更；在具身系统的运行过程中，还包括**策略/规则配置、动作库、技能插件**等对行为边界有直接影响的资产。

危害说明：模型或数据被投毒后，可能导致机器人行为出现难以察觉的偏差，或在特定条件下发生恶性失效。危害通常表现为决策错误〔L2-L4〕。最危险的情况是模型被植入后门，在特定场景下执行预定的危险动作，例如识别到特定标志后触发异常行为，从而构成L5级精准威胁。

9. RS09:2026 – 更新与启动链完整性缺陷

详细描述： 指机器人固件/软件的在线更新流程以及从硬件上电到操作系统加载的整个启动链，**缺乏端到端的完整性保护**。包括：更新包未经验证签名、更新服务器可能被篡改或冒充、系统允许回滚到有已知漏洞的旧版本、以及启动过程中每个环节的代码未被度量与验证。

危害说明： 这是攻击者实现**持久化控制**和**批量固件破坏**的主要途径。攻击者可分发恶意更新包，从而在设备上永久植入后门〔L3〕。如果恶意更新能覆盖安全模块或控制逻辑，则可直接让大量设备停摆〔L4〕或丧失安全功能〔L5〕。

10. RS10:2026 – 安全兜底机制失效

详细描述： 指机器人系统在设计上，**缺乏在核心软件系统完全被入侵或发生致命故障时，能够确保物理安全的最终硬件保障**。例如，没有独立于主控芯片的安全微处理器来监控行为并执行急停；被攻陷软件可以禁用限速、限力功能；紧急停止按钮的信号可以被软件层拦截或忽略。

危害说明： 此风险意味着任何软件层面的安全防线被突破后，将没有任何物理屏障来阻止伤害发生。因此，只要存在近人作业场景，该风险直接关联最高级别的L5人身安全威胁。即便在无人场景，往往也可以导致L4级的财产损失。

3.4 典型攻击路径与风险传导模式

基于近年公开的安全研究、竞赛挑战及行业披露的案例，具身智能系统的攻击路径已呈现出从传统IT入侵向智能决策欺骗与物理执行劫持深度融合的演进趋势。本节通过剖析三类已被完整验证的攻击场景，旨在具体阐明RoboSec Top 10中的系统性风险根因如何被串联利用，从而构成从数字渗透到物理危害的完整传导链。

路径一：AI决策链路的欺骗与漏洞利用链

关联风险根因： RS05〔感知与决策层欺骗〕、RS06〔控制栈软件漏洞〕、RS10〔安全兜底机制失效〕

攻击模式： 该路径攻击机器人依赖高级语义输入〔自然语言、任务描述〕进行决策的流程。攻击者通过构造对抗性输入，欺骗AI模型或决策系统生成非预期或越权的行动指令，进而穿透软件层的逻辑安全边界。

案例实证： DARKNAVY安全团队的公开研究演示了此类链式攻击。研究人员通过精心设计的提示词注入〔Prompt Injection〕，绕过服务机器人云端大模型的安全限制，诱导其输出包含恶意代码的任务指令。该指令在机器人端侧的任务执行代理〔Agent〕中被解析时，触发了已有的命令注入漏洞，最终实现远程代码执行，完全夺取设备控制权。此案例清晰呈现了从认知层误导到控制层漏洞利用的跨层攻击链。若系统缺乏硬件安全兜底〔RS10〕，此类攻击可直接导致物理失控。

路径二：近场通信缺陷触发的规模化劫持链

关联风险根因： RS02〔端侧外部暴露面管理缺陷〕、RS06〔控制栈软件漏洞〕、RS04〔云端控制平面风险〕

攻击模式： 此路径针对机器人的近场无线通信功能〔如蓝牙〕与内部网络。攻击者利用弱认证、固定密钥或缺乏隔离等缺陷，在物理邻近范围内实现非授权接入、会话劫持，并进一步利用内部通信的信任缺失进行横向移动，形成跨设备扩散。

案例实证： 在GeekCon 2025等安全赛事/会议的公开演示中，研究人员成功展示了针对机器人集群的人传人式攻击。攻击者首先利用某个漏洞获取单台机器人的控制权，随后将其脆弱的蓝牙服务作为中继，将恶意负载传播至邻近的其他机器人，在短时间内构建起受控网络。这表明，近场通信与内部网络的安全缺陷一旦叠加〔RS02、RS03〕，便可能在规模化部署场景下引发快速扩散的系统性风险，其影响还可通过云端管理平面〔RS04〕进一步放大。

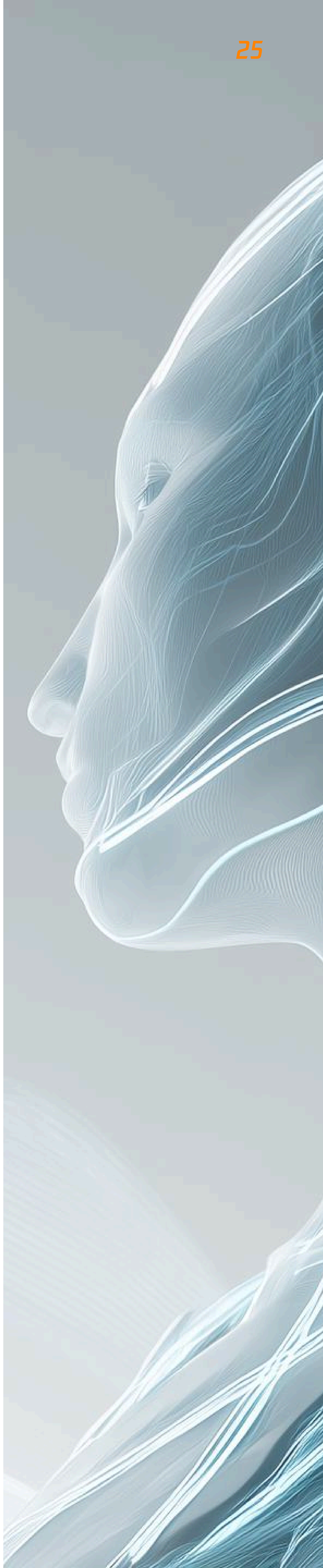
路径三：传统软件漏洞的利用与物理影响放大链

关联风险根因： RS01〔端侧内部权限模型缺陷〕、RS02、RS06〔控制栈软件漏洞〕、RS09〔更新与启动链完整性缺陷〕

攻击模式： 该路径表明，许多具身智能系统仍广泛存在经典的信息安全漏洞。攻击者利用未授权访问、文件上传、命令注入等传统手段获取初始立足点后，通过权限提升漏洞〔RS01〕深入系统，最终触及底层的运动控制模块，将单纯的代码执行漏洞转化为物理动作。

案例实证： Theroboverse安全社区的PawRoot研究提供了典型示例。攻击者在同一局域网下，访问了机器人未鉴权的Web服务接口，利用文件上传漏洞覆盖了关键系统脚本，从而获得远程代码执行能力。随后，通过进一步利用系统权限模型的缺陷，攻击者最终直接操控了运动控制系统。此案例证明，即便未涉及AI攻击面，传统软件漏洞〔RS06〕与不当的权限设计〔RS01〕结合，其潜在危害会因设备的物理执行能力而被急剧放大。

上述攻击路径的共同规律在于，**数字世界的单一薄弱点可作为跳板，触发跨架构层（感知、决策、控制、执行）的连锁反应，最终导致物理后果**。攻击链的构成往往是混合式的，可能始于AI欺骗，转而利用软件漏洞，并通过不安全的通信进行扩散。RoboSec Top 10清单与L1-L5分级框架的结合，为系统性地识别这些风险传播路径、评估其潜在的危害等级、并确定应优先加固以切断高风险攻击链的环节，提供了结构化的分析工具。



4. 未来展望

4.1 行业趋势展望

包括宇树科技、优必选、波士顿动力在内的国内外部分机器人厂商已开始建立安全应急响应中心并招募安全人才，这一动向表明，具身智能行业的安全意识正在由缺位走向初步觉醒。当前阶段，行业整体仍处于安全能力建设已经启动、但体系化建设尚未成形的过渡期，安全投入多以点状方式分布在头部企业的研发架构中，主要体现为漏洞接收渠道、应急响应团队或基础安全测试，尚未深度融入产品定义、架构设计、供应链管控与售后运维的全生命周期。绝大多数中小企业仍处于安全能力空白状态，行业层面也缺乏针对具身智能特点的攻击面定义、风险评估基准与检验认证体系。

这一转型期的核心特征在于，安全正从可选投入逐步转向必要投入，但尚未成为行业普遍认可的能力基线。随着具身智能机器人从封闭测试环境进入半开放乃至全开放的真实部署场景，安全事故正由偶发个案演变为可复现攻击，产业与公众对安全失能的容忍度也在快速下降。可以预期，未来3~5年将是具身智能安全能力从起步期迈向成长期的关键窗口。

本白皮书所提出的RoboSec分级框架与风险清单，旨在为这一转型进程提供可供行业对齐的风险认知坐标系与行动参照，帮助各方在共通的术语与评估尺度下，识别关键风险、明确改进方向、积累体系化能力。

4.2 什么样的具身智能机器人，才可称为安全？

从安全属性角度出发，具身智能机器人应当在以下三个相互关联的维度上具备可验证的能力：

Security（抗攻击能力）

机器人应能有效抵御来自网络、近场或物理接口的恶意攻击。这意味着系统在设计之初即遵循纵深防御与最小权限原则，关键控制链路具备身份认证与完整性保护，固件与软件更新具有端到端签名验证，且任何安全缺陷都能通过负责任的披露渠道得到及时修复。

Safety（物理安全能力）

机器人即使在软件故障、感知错误或被部分入侵的情况下，仍应具备避免对人身和环境造成伤害的最终保障。这体现为硬件级、不可旁路的急停与限速机制，独立于主控系统的安全监控模块，以及在异常状态下可自动降级至安全姿态的行为策略。

Privacy（隐私保护能力）

机器人采集的视音频、空间地图、行为轨迹及用户交互数据，应遵循数据最小化原则，具备传输加密、存储隔离与访问审计能力。用户应能清晰感知并控制自身数据的收集范围与留存周期，防止数据被滥用或未经授权披露。

4.3 未来方向

为实现上述多维愿景，产业必须协同行动，推动安全从可选投入转变为核心基础能力。结合过往产品安全的发展历程，以下为未来可能的关键演进方向：

安全设计前置

安全将从事后补救与被动加固，进一步前移至系统架构设计与能力规划阶段，成为与感知、决策、执行并列的核心设计维度，并在需求定义与方案选择阶段即形成约束与取舍依据。

智能决策安全与物理设备约束的协同

单纯依赖模型输出进行决策与执行的路径正暴露出边界与失控风险。未来系统需在 AI 决策层与物理安全约束层之间建立更紧密的协同与闭环关系，使高层智能在任何情况下都受到可验证的物理安全边界约束，从而避免危险决策被直接放大为现实世界后果。

行业标准与测试体系逐步成形

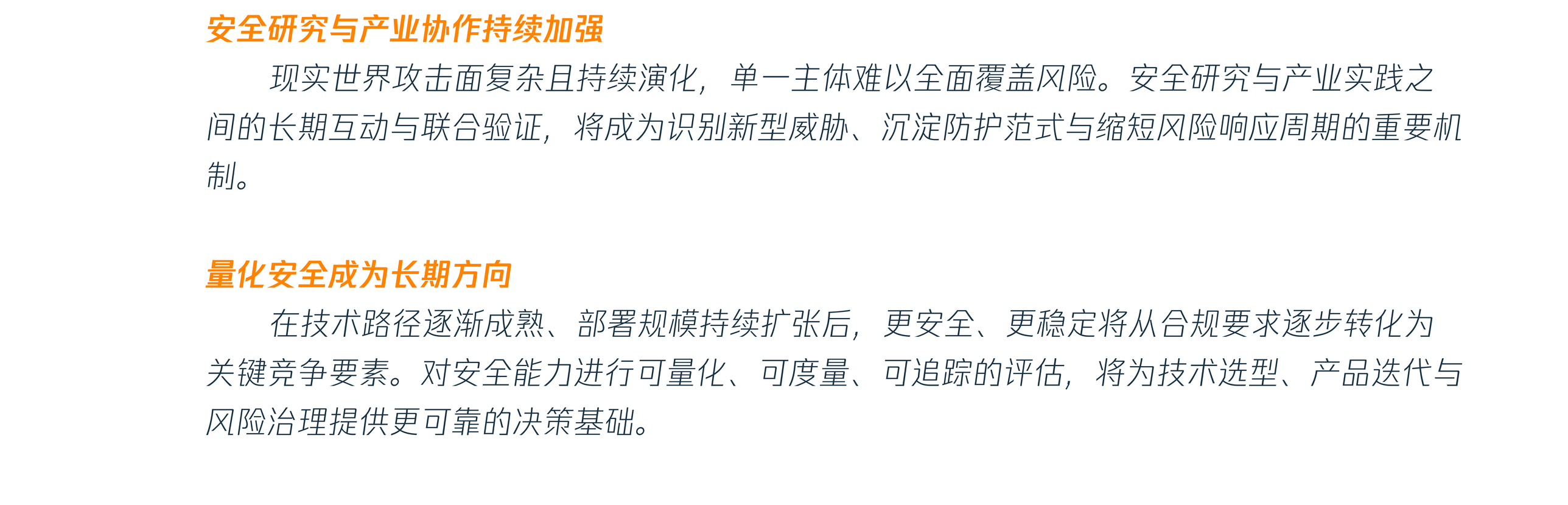
随着攻击路径与风险模式逐步清晰，具身智能安全的评估需求将从个案分析走向体系化测试。未来将更需要可比较、可复用、可迁移的测评框架与指标体系，以支撑安全能力的工程化建设、横向对标与持续改进。

安全研究与产业协作持续加强

现实世界攻击面复杂且持续演化，单一主体难以全面覆盖风险。安全研究与产业实践之间的长期互动与联合验证，将成为识别新型威胁、沉淀防护范式与缩短风险响应周期的重要机制。

量化安全成为长期方向

在技术路径逐渐成熟、部署规模持续扩张后，更安全、更稳定将从合规要求逐步转化为关键竞争要素。对安全能力进行可量化、可度量、可追踪的评估，将为技术选型、产品迭代与风险治理提供更可靠的决策基础。



参考标准与公开资料

[1] Omdia Market Radar: General-purpose Embodied Intelligent Robots, 2026. https://omdia.tech.informa.com/om143809/omdia-market-radar-generalpurpose-embodied-intelligent-robots-2026
[2] Boston Dynamics Spot and Site Hub Security. https://bostondynamics.com/wp-content/uploads/2024/03/spot-and-site-hub-security-white-paper.pdf
[3] ISO. ISO 10218-1:2025 Robotics — Safety requirements — Part 1: Industrial robots; and ISO 10218-2:2025 Robotics — Safety requirements — Part 2: Industrial robot applications and robot cells.
[4] ISO. ISO 13482:2014 Robots and robotic devices — Safety requirements for personal care robots.
[5] ISO. ISO/TS 15066:2016 Robots and robotic devices — Collaborative robots.
[6] IEC. IEC 61508-1:2010 Functional safety of electrical/electronic/programmable electronic safety-related systems — Part 1: General requirements.

附录A: 术语表

为便于不同背景读者理解本文所使用的机器人、安全与人工智能相关概念，现对文中高频术语作简要说明。以下释义仅服务于本文语境下的理解，不构成严格的标准定义。

术语	英文/缩写	释义
具身智能	Embodied AI	指具备物理实体、能够通过感知、决策、执行闭环与真实环境持续交互的智能系统。本文主要聚焦具身智能机器人。
控制面	Control Plane	指负责身份、授权、任务下发、调度、策略配置与远程管理的系统平面，决定由谁控制、如何控制以及控制什么。
数据面	Data Plane	指承载业务数据与运行数据传输的系统平面，如媒体流、遥测数据、传感数据与状态回传链路。
智能体	Agent	指能够基于输入目标、环境信息与工具调用能力，自主完成任务拆解、决策与执行编排的软件实体。
视觉—语言—动作模型	VLA（Vision-Language-Action）	指将视觉输入、语言理解与动作生成统一到同一模型框架中的具身智能模型，用于支持机器人从感知到动作的端到端决策。
提示注入	Prompt Injection	指攻击者通过构造恶意输入操纵大模型或智能体的行为，使其偏离原有约束、执行非预期指令或泄露敏感信息。
大模型越狱	LLM Jailbreak	指通过特定提示或上下文设计绕过模型既有安全限制，使模型输出本不应提供的内容或执行越权任务；通常可视为提示注入的一种表现形式。
机器人操作系统	ROS / ROS 2	ROS 是用于构建机器人应用的软件库与工具集合；ROS 2 在此基础上演进，更适应现代机器人系统的通信、中间件与工程化需求。
数据分发服务	DDS（Data Distribution Service）	一种面向实时和嵌入式分布式系统的开放中间件标准，直接支持发布订阅通信模式，常作为 ROS 2 底层通信机制的重要基础
节点	Node	机器人系统中的基本运行单元。一个机器人应用通常由多个节点构成，不同节点通过消息交换实现感知、规划、控制等功能协同。
同时定位与地图构建	SLAM（Simultaneous Localization and Mapping）	指机器人在未知或部分未知环境中，一边估计自身位置，一边构建环境地图的技术过程。
空中下载更新	OTA（Over-the-Air Update）	指通过网络对设备软件或固件进行远程更新的机制，通常涉及版本管理、分发、校验、安装与回滚控制。
遥测	Telemetry	指设备运行过程中持续上报的状态与监测数据，如位置、速度、电量、故障状态、传感器读数与系统健康信息。
远程代码执行	RCE（Remote Code Execution）	指攻击者能够通过网络或其他远程入口，在目标系统上执行其控制的代码，是高危安全后果之一。
超宽带	UWB（Ultra-Wideband）	一种短距离无线通信技术，可用于测距、定位和设备间通信；若相关协议或实现存在缺陷，可能成为近场攻击入口。
软件定义无线电	SDR（Software Defined Radio）	一种通过软件实现无线电信号处理与通信控制的技术路线，常用于无线协议分析、调试与安全研究。
惯性测量单元	IMU（Inertial Measurement Unit）	用于测量加速度、角速度等运动信息的传感器组件，是机器人姿态估计、运动控制与定位的重要输入来源。
安全兜底机制	–	指在主控制系统异常、失效或被攻击时，仍能限制危险动作、将系统带入安全状态的最终防护措施，如硬件急停、不可旁路限速限力、独立安全监控模块等。
个人护理机器人	Personal Care Robot	指面向个人服务场景、用于改善用户生活质量的机器人类型，典型包括移动服务机器人、物理辅助机器人和载人机器人。

DARKNAVY

新加坡·上海

独立自由的

安全研究服务机构