

室内空间智能

从数字感知到物理行动

—从感知闭环到行动闭环的技术变革与产业实践

联合发布

亚信科技 × 联通在线 × 清华 AIR × VERYK GROUP

前言

当生成式AI重构数字世界，Physical AI 正推动智能系统走出屏幕、进入空间并直接参与现实生产与服务。面向室内这一高频、复杂且强约束的物理环境，机器人不再只需要“看见人、货、设备与通道”，更需要理解人的自然语言意图、识别空间语义和业务规则、在动态环境中自主规划与安全行动。以 VLN（Vision-Language Navigation，视觉—语言导航）为代表的空间智能，正成为连接数字感知与物理行动的关键能力：它使机器能够将“把客厅茶几上的水送到书房”、“去儿童房确认窗户是否关好”、“避开正在玩耍的孩子，把扫地机器人带回充电座”等人类任务，转化为可执行、可验证、可恢复的真实移动行为。

这意味着，室内智能化正在经历一次根本性的技术变革：从以摄像头、传感器和平台为核心的“感知闭环”，迈向由意图理解、空间认知、路径规划、具身执行、结果验证和持续学习共同构成的“行动闭环”。在前一阶段，系统能够识别异常、定位目标、生成告警；在新的阶段，系统还必须判断任务目标、理解环境变化、选择可行路线、完成物理操作，并对行动结果负责。对于智能家居、物流仓储、工业制造与电力巡检等典型场景而言，这一变化将推动自动化能力从依赖固定路线、预设规则和单点设备控制的“任务执行”，升级为能够理解自然语言指令、感知动态环境、统筹多设备协同，并在复杂空间中持续完成任务的“自主空间服务”。

白皮书使命

本白皮书面向产业决策者、技术架构师与具身智能行业从业者，聚焦一个核心命题：**如何将企业的空间数据资产、智能算法、自动化装备与业务流程，转化为可验证、可复制、可持续运营的物理行动能力。**

本白皮书由**亚信科技（中国）有限公司（以下简称“亚信科技”）、联通在线信息科技有限公司（以下简称“联通在线”）、清华大学智能产业研究院（AIR，以下简称“清华AIR”）、VERYK GROUP（综合性物流电商服务企业）**联合编制。四方在产业分工中各具禀赋：**联通在线为场景运营方**，依托中国联通 5G-A 专网、家庭 FTTR 全光网络、覆盖全国的云资源池体系与“智家通通”家庭算力中枢，提供场景入口、网络底座与线下服务网络；**亚信科技为技术底座方**，提供 Physical AI 工程化技术底座，覆盖仿真数据工厂、模型后训练、云端协同架构与 Agent 产品化交付；**清华 AIR 为模型理论方**，攻坚世界模型基础理论与可信推理方法；**VERYK GROUP 为仓储物流业务落地方**，依托自身仓储物流产业积累，负责把 Physical AI 技术落地到真实仓储场景，输出物流仓储业务需求，完成场景验证与业务价值闭环。四方构成“场景运营 + 工程落地 + 底层研究 + 行业场景落地”的协同体系。

本白皮书的核心判断是：室内空间智能正在从以感知和告警为主，向具备任务理解、空间认知、自主导航、受约束执行和结果验证能力的系统演进。全书围绕这一命题展开：第 1 章分析产业窗口与典型需求，第 2 章梳理数据、认知、安全、成本四类规模化约束；第 3 章提出三大核心主张、两个横切机制与两个闭环；第 4 章映射为工程架构；第 5 章讨论家庭、机房、仓储与楼宇的验证路径，第 6 章展望长期演进方向。

需要向读者说明的是本白皮书的定位：这是一份**技术主张与产业路径型**白皮书。其核心价值不在于罗列既有商用案例，而在于基于四方各自已完成的技术验证、理论与场景储备，给出一条可工程化、可商业化、可复制的完整路径，并面向产业界发出联合验证与试点共建的开放倡议。文中涉及的企业实测数据均为限定条件下的技术验证结果，其测试对象、数据集、Baseline、指标定义与实验环境遵循统一披露规范，不构成对任意场景下性能表现的普遍承诺。

联合发布方



亚信科技：Physical AI 工程化底座，打通从算法验证到商业运营的全链路

构建覆盖仿真数据工厂、模型后训练与云端协同架构的 Physical AI 工程化底座，打通从实验室算法到真实场景规模化交付的完整链路，让空间智能从“演示可行”走向“商业可运营”。



联通在线：家庭入口与网络底座，Physical AI 规模化普惠的战略支点

中国联通正聚焦“连接、算力、服务、安全”四大核心赛道，加快从基础管道提供商向数字化综合服务商转型。对于室内空间智能而言，“连接”构成机器人、摄像头、IoT 与云边端系统协同的网络底座，“算力”支撑多模态感知、空间理解、世界模型推演与实时决策，“服务”推动能力从单点设备交付走向可运营的场景化解决方案，“安全”则为数据合规、权限控制、行动约束与可信运营划定边界。四大赛道共同构成 Physical AI 从技术验证走向规模化应用的基础能力体系。



清华AIR：世界模型的保真表征与可信推理——Physical AI 认知底座的理论根基

世界模型的保真表征与可信推理是 Physical AI 的认知理论根基，必须保障机器对物理世界的推演符合物理规则、结果可解释可验证，为“理解世界”提供可审计的底层认知底座，降低室内低容错或安全敏感场景场景下的推演风险。



VERYK GROUP: 面向仓储物流场景落地，VLN实现空间真实理解导航

仓储库房动态多变、人机混场，传统机器人依赖预建地图与预设路线，布局变动便失效；VLN 可以结合视觉观测、语言指令以及地图或空间记忆完成目标驱动导航。相比仅依赖固定坐标与预设路线的传统方案，其优势在于能够利用语义信息理解目标和环境约束，并在一定范围内适应场景变化。

目录

前言..... i

白皮书使命..... ii

联合发布方..... iii

第 1 章 产业发展现状：室内空间智能的规模化窗口..... 1

第 2 章 关键挑战：室内空间智能规模化的四道关..... 5

第 3 章 核心观点：三大核心能力 + 两个横切机制 + 两个闭环..... 8

第 4 章 整体解决方案：从单点能力转向系统闭环..... 11

第 5 章 典型应用场景与验证路径..... 18

第 6 章 未来展望..... 23

致谢..... 25

参考文献..... 26

附录 A 28

附录 B 29

第1章

产业发展现状

室内空间智能的规模化窗口

室内空间智能的规模化元年已至



政策、市场、技术多重因素正在交汇，为室内 Physical AI 打开产业化窗口期。本章从政策环境、市场格局、技术演进与真实需求出发，剖析室内空间智能所处的产业阶段，明确家庭作为首发商业场景的价值，同时为后续分析落地挑战做好铺垫。

1.1 政策与产业环境：从顶层设计到常态部署

顶层设计持续加码。2025 年，具身智能首次写入政府工作报告；2026 年政府工作报告继续将具身智能列为未来产业重点方向^{[8][9]}。工业和信息化部已建立人形机器人与具身智能标准化技术委员会，标准化范围覆盖应用与安全等领域；2026 年实景实训专项行动进一步将仓储物流、维修维护、医疗康养、安全生产等列为重点场景，为室内移动服务与巡检类应用提供了政策和标准化环境。

消费侧存量升级被激活。消费品以旧换新政策带动智能门锁、智能照明、智能中枢等存量改造产品显著增长，家庭智能设备的保有量和在线率持续提升，为空间智能能力"寄生"于既有设备生态、低成本进入家庭创造了条件。

标准从制定走向前置性规范。智能家用电器相关国家标准 GB/T 28219.1—2025《智能家用电器的智能化技术要求和评价 第 1 部分：通用要求》、GB/T 46505.1—2025《智能家用电器应用场景 第 1 部分：通用要求》均于 2025 年发布、2026 年 5 月 1 日实施^{[11][12]}，从标准层面推动跨品牌、跨品类的互联互通，从根本上打破生态壁垒；国际上 Matter 等互联互通标准持续演进，为跨品牌设备接入提供生态基础。智能家居的政策红利与互联互通基础，正在为室内空间智能的规模化部署铺路。

1.2 市场格局：室内空间智能的市场空间

室内空间智能覆盖四大核心场景，呈现出"行业侧放量、消费侧蓝海"的鲜明结构：

- **智慧家庭：**安防看护、适老监测。这是全球规模最大的室内空间。截至 2025 年末，我国 60 岁及以上人口为 3.2338 亿，占总人口 23.0%；其中 65 岁及以上人口为 2.2365 亿，占 15.9%^[13]。随着老龄人口规模扩大，跌倒、异常状态发现、紧急联络等居家安全需求受到持续关注，为家庭侧主动看护类服务提供了潜在应用空间。
- **机房/数据中心：**设备状态与环境巡检。算力基础设施建设高速增长，带动机房巡检需求同步放量；机房属于非结构化程度较低但 7×24 值守要求极高的室内场景，增量明确。
- **仓储物流：**库存盘点、消防隐患监测。盘点是高频刚需任务，与空间自主移动能力天然匹配。
- **商业楼宇：**安防巡视、设备设施巡检。存量楼宇体量庞大，物业人力成本持续上升，替代需求明确。

与此同时，中国智能家居设备市场已形成较大规模的设备与渠道基础。IDC 预计 2026 年中国智能家居设备市场出货量约为 3 亿台，同比增长 8.8%^[14]。这为室内空间智能提供了设备接入、用户教育与渠道基础，但移动巡检与主动看护仍处于早期探索阶段，产品形态、用户付费意愿和规模化商业模式尚待进一步验证。

1.3 技术发展：世界模型认知底座与工程架构逐渐成熟

三条技术曲线的历史性交汇，为大模型向室内空间智能场景的后训练迁移提供了技术与数据基础：

曲线一：世界模型从潜空间动态预测，正在扩展到生成式环境建模和 Physical AI 仿真/预测平台。从早期基于潜空间动力学的世界模型^[4]，到可交互 3D 环境生成^[5]、面向 Physical AI 的世界基础模型平台^[6]，世界模型正在成为 Physical AI 中用于环境表征、状态预测、仿真和规划的重要技术方向。现阶段，其在真实机器人闭环系统中的作用和成熟度仍取决于任务类型、训练数据、预测时域、计算时延以及安全约束。

曲线二：高保真仿真使合成数据规模化生产成为可能。物理仿真引擎、神经渲染与域随机化技术的成熟，使室内场景的高拟真合成数据可以工业化批量生产，为受隐私红线约束的室内场景提供了合规的数据供给路径。

曲线三：视觉-语言导航验证“预测—行动”范式。视觉-语言导航（VLN）与视觉-语言-动作模型（VLA）研究在近年快速发展^{[7][16][17]}，验证了“语言指令—视觉感知—行动执行”端到端对齐的可行性，为空间智能任务的自然语言交互与自主执行提供了方法论基础。

数据、模型、架构三条曲线的交汇，意味着室内空间智能所需的三条技术曲线首次进入可工程化调用的交集区间。

1.4 需求缺口：从“看得见”到“看得懂、能行动、可信赖”

需求侧的升级与供给侧的滞后之间，存在清晰的结构性缺口。

行业侧：高频空间服务与人力约束的硬矛盾。机房、仓储、楼宇的巡检频次要求与人力成本、值守质量之间的矛盾日益尖锐；大量场景处于“有监控、无服务”状态——摄像头只负责记录，发现问题仍依赖人盯屏幕，事后追溯多于事前预警。

消费侧：看护刚需与供给错配。家庭用户不再满足于被动查看摄像头画面，而是需要系统主动发现问题、理解状况、采取行动并形成处置闭环——老人跌倒要被“发现”更要被“响应”，燃气泄漏要被“预警”更要被“处置”。

由此可以凝练本白皮书的核心矛盾判断：**空间智能需求正从“看得见”升级为“看得懂、能行动、可信赖”，但现有方案仍停留在“有监控、无服务”阶段。**“看得懂”要求空间语义理解而非图像匹配，“能行动”要求自主导航与任务执行而非固定轨道，“可信赖”要求隐私安全、低误判与处置闭环而非单纯告警。这三重升级共同定义了新一代室内空间智能的能力标准，而“巡检”正是验证这三重能力的最小完备商业单元。

1.5 场景选择：为什么从智能家居开始

四大场景之中，家庭是室内空间智能的重要消费侧验证场景；机房和仓储则更适合作为工程可靠性与商业价值的先行验证场景。两条路径分别验证用户侧复杂性与行业侧可运营性，并逐步沉淀可迁移的通用能力。

- **规模逻辑：** 家庭是全球最大的室内空间，家庭场景中验证的室内理解、导航与判断能力，可迁移至几乎一切室内场景；
- **刚需逻辑：** 3.2 亿老龄人口^[13]的居家看护是确定性刚需，安全监测与预警位居养老需求首位；
- **禀赋逻辑：** 运营商掌握家庭宽带入口、智家渠道与装维网络，具备触达、部署与运维家庭的现成能力；
- **标准逻辑：** 空间自主服务列入具身智能标准体系重点方向，互联互通国标与家庭服务机器人安全标准为能力迁移提供合规保障，Matter 等国际标准为跨品牌设备接入提供生态基础。

由此，本白皮书确立“家庭技术首发验证与机房商业造血并行→仓储放量→楼宇普惠→千行百业复制”的双轨演进路径：将家庭场景作为算法与交互模型首发验证的试验田，借助智能家居产业生态沉淀核心空间智能能力；同时以机房与仓储等高价值、强结构化场景作为商业变现与工程落地的造血基座，随后沿场景复杂度逐级迁移，最终走向千行百业的室内空间智能。这是“从智能家居到千行百业”兼顾技术前瞻与商业自洽的闭环路径。

第2章

关键挑战

室内空间智能规模化的四道关

商用没有及格线，只有生死线



尽管当前行业核心需求已全面转向具备看得懂、能行动、可信赖综合能力的一体化室内空间智能解决方案，但从技术验证走向大规模商用仍存在多重刚性约束。真实部署更关心以下问题：

- 是否有足够且可持续获得的数据；
- 是否能够理解任务和空间状态；
- 是否能够形成可靠规划并完成行动；
- 是否可以约束错误或越权动作；
- 异常后能否恢复或由人工接管；
- 是否能够在可接受的算力、硬件和运营成本下持续运行；

本章聚焦室内空间智能规模化落地必须跨越的数据、认知、安全、成本四道核心关卡，解析各瓶颈的内在矛盾与相互耦合关系。这些现实约束，也是后续提出整套解决方案的出发点。需要强调的是，四道关不是加法关系，而是**乘法制约关系**：规模化 = 造得出 × 巡得准 × 守得住 × 算得来。任一维度趋近于零，商业化即不成立。

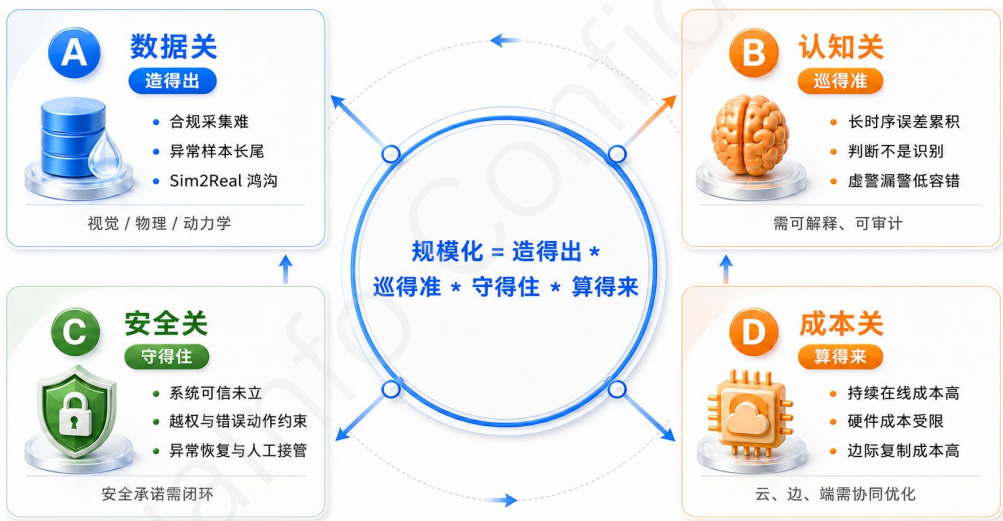


图 1：室内空间智能规模化的四大关键制约

2.1 数据关：训练数据匮乏，尤其是家庭场景

数据是第一道关，且难点具有场景特殊性：

真实数据合规和治理成本高。家庭空间通常涉及人脸、声纹、行为轨迹、未成年人信息等个人或敏感个人信息。大规模真实数据采集受到目的限定、最小必要、授权同意、存储安全和数据使用范围等约束，合规和治理成本较高^{[21][22]}。机房和楼宇数据通常受到商业秘密、数据权属、租户隔离及安全域管理要求约束，跨主体集中汇聚存在较高治理与授权成本。

异常样本天然长尾。跌倒、热失控、设备异响皆为低频事件，真机采集“等不到也不够”；叠加室内非结构化特性——光照多变（日间自然光/夜间弱光/遮挡阴影）、物体材

质多样（反光/透明/柔性）、空间动态变化（家具移动/人员活动）——场景组合数呈指数级爆炸。

仿真合成存在室内特有的三层 Sim2Real 鸿沟：视觉鸿沟（弱光/反光/遮挡的渲染偏差）、物理鸿沟（家具碰撞/柔性形变的仿真失真）、动力学鸿沟（窄空间运动约束的近似偏差）。相关研究表明，域随机化、动力学随机化与域适配等方法可用于缩小部分现实差距，但“仿真生成即可用”并不成立，仍需真实环境校准与系统性质量审计^{[23][24][25]}。

2.2 认知关：空间服务是“判断”不是“识别”，且低容错

长时序误差累积。一次完整空间任务包含几十上百个连续决策步骤，单步准确率的微小瑕疵在长序列中累积放大，系统会“忘记巡过哪”、重复巡查或遗漏点位。

“状态异常”是判断题而非识别题。判断“灶上锅烧干了”“设备指示灯状态异常”“地面有水渍且临近插座”，需要空间语义与常识推理，而非图像模板匹配；通用视觉模型在此类开放判断上误判率不可接受。

虚警漏警双零容忍。家庭场景误报一次，用户即弃用；机房场景误报引发工单风暴、漏报酿成事故。误判还带来责任界定问题——这要求推理过程可解释、可审计，而非黑箱输出。

2.3 安全关：系统可信未立与安全承诺无闭环

空间智能的核心价值是安全保障，但现有方案存在双重短板：

其一，系统自身可信度不足——隐私保护机制、推理追责机制缺失，用户不敢把家门钥匙交给一个不透明的系统；

其二，安全承诺无法兑现——异常预警仅停留在消息通知层面，缺乏线下处置链路，“告警了但没人管”使安全承诺落空。一个以安全为价值的系统，自身必须先值得信任，且承诺必须闭环。

2.4 成本关：价值密度低 vs 持续在线

算力账单矛盾。云侧全程推理的持续算力成本，与家庭百元级月费、行业项目制订阅价之间存在量级矛盾。

硬件成本约束。家庭本体须控制在百元级，行业侧须满足项目制 TCO 要求，高成本本体路线在两端都不成立。

边际复制成本高企。家庭与行业点位高度分散，若每个点位都依赖定制调试与人工部署，规模化即被边际成本锁死。

第3章

核心观点

三大核心能力 + 两个横切机制 + 两个闭环

室内空间智能，不在单点在体系



针对前文提出的数据、认知、安全、成本四道规模化关卡，需要一套体系化思路，而非依赖单点模型或单一硬件能力。本章提出**三项核心能力主张与两个横切机制**：以数据能力解决持续数据供给问题，以空间认知与决策解决任务理解与规划问题，以安全与治理建立物理行动边界；同时，通过云—边—端运行架构平衡性能、实时性与成本，通过 Agent 实现跨模型、跨设备与跨业务流程的任务编排。

两类运行闭环贯穿其中：**在线任务闭环**保障一次任务从感知、规划到执行、验证和恢复；**离线数据闭环**通过真实数据、仿真、训练、评测、部署与 Badcase 回流推动系统持续迭代。产业竞争的重点正在从单模型能力，转向数据、认知、安全和运行体系能否形成稳定、可验证、可持续优化的系统闭环。

3.1 观点一：数据先行，建立可持续的数据供给与迭代体系

针对室内真实数据合规采集受限、异常样本长尾稀缺、仿真存在三层 Sim2Real 鸿沟的三重约束，本白皮书主张搭建**分层数据供给体系**：以高拟真仿真批量生成基础样本，解决"量"的问题；以少量真机数据完成域迁移校准，解决"真"的问题；以无本体采集补充人类巡检示范，解决"教"的问题。三者之上配套全流程质检流水线，并依托 Real2Sim2Real 闭环持续沉淀可复用的室内数据资产——数据不是一次性投入，而是在闭环中增值的资产。

3.2 观点二：分层认知与决策，实现高层理解与实时控制解耦

空间认知与决策采用分层分频控制架构：VLN / Task Planner 负责语言任务理解、空间语义判断与长时程规划，并将任务分解为局部目标；Local Planner / VLA Policy^{[16][17]}根据实时环境完成局部路径规划、避障或动作策略生成；Safety Controller 与 Motor Controller 分别承担安全校验和底层执行。世界模型可作为较低频的环境表征、状态预测和风险推演模块，为规划与异常重规划提供辅助信息。不确定性度量用于识别需要更高推理深度或人工介入的情况，相关阈值与效果应通过目标场景评测确定。

3.3 观点三：安全内生，建立从语义约束到物理控制的纵深防护

安全必须成体系而非单点加固。机器人平台设置不可绕过的物理安全边界；语义本体（Ontology）通过显式描述实体、属性、关系、状态、约束和业务规则，为 Agent 的推理与行动提供结构化知识边界，并可结合规则检查、状态机和 Action Masking 对候选动作进行运行时约束^[15]。相关规则检查的时延、覆盖率和误拦截率需在目标硬件与任务环境中实测。隐私数据本地优先处理，并通过审计与人工接管机制形成可追溯的安全治理链路。

3.4 横切机制一：云—边—端协同，平衡性能、实时性与运行成本

云、边、端根据任务复杂度、实时性、安全要求和成本预算进行分层部署：云端承担复杂推理与集中服务，边缘侧承担本地感知和任务级实时处理，终端负责传感器采集、局部规划、安全控制和运动控制。在当前试点中，常规巡检主要由边端闭环完成，复杂场景按需调用云端能力；具体比例受场景、模型、硬件与网络条件影响，不作为固定部署假设。FTTR、5G-A 与边缘云为云边协同提供连接基础，实际端到端响应时间仍取决于接入、传输、算力调度、模型推理及控制链路，应以 P95/P99 E2E 时延、可用性和丢包率等指标验证。

3.5 横切机制二：Agent 负责任务编排，实现能力统一交付

Agent 不是四道关的解决方案本身，而是将数据、认知、算力、安全能力编排成业务闭环的软件智能体。依托 Agent 架构打通感知、导航、研判、处置全流程，统一多品牌硬件接入标准；采用**脑体解耦**设计，模型迭代无需更换巡检硬件，消除多设备碎片化割裂问题。作为横切主张，Agent 承载数据工厂成果、调度世界模型推理、编排云端算力、发起安全闭环处置，是四条逐关主张的统一交付形态——用户面对的不再是“一堆设备加一堆 App”，而是“一个入口、一个承诺、一个闭环”。

第4章

整体解决方案

从单点能力转向系统闭环

感知只是起点，行动才是交付



针对数据、认知、安全、成本四道规模化关卡，提出了"三大核心能力 + 两个横切机制 + 两个闭环"的体系化观点；本章将这一观点落实为一套可工程化实施的整体解决方案。方案的核心不是为机器人叠加一个大模型，而是**建立数据形成能力、认知产生决策、决策形成受约束行动、执行结果反向改进系统**的完整链路——这正是从"单点能力"走向"系统闭环"的实质所在。

需要强调的是，本章给出的仍是架构与机制层面的方案，其有效性最终要接受真实业务场景的检验。第 5 章将把这套架构带入家庭、机房、仓储、楼宇四大室内场景，说明各场景的任务、部署模式与已完成的工程验证结果，并给出"验证-造血-规模化"的阶梯式落地路径。

4.1 总体架构

室内 Physical AI 的核心并不是让机器人增加一个大模型，而是建立从**数据形成能力、从认知产生决策、从决策形成受约束行动**，并由**真实执行结果反向改进系统**的完整链路。

整体关系可以概括为：**数据能力 → 空间认知与决策 → 安全与治理** 三类能力构成系统主体。在其上，**Agent / 任务编排** 承担任务级状态管理、Skill 调用和异常处理；在其下，**云—边—端运行架构** 根据任务复杂度、实时性和安全要求配置计算与执行位置。

系统同时运行两个不同时间尺度的闭环：

- **在线任务闭环：**感知 → 理解 → 规划 → 控制 → 验证 → 恢复
- **离线数据闭环：**真实 → 仿真 → 训练 → 评测 → 部署 → 回流

前者解决“当前任务如何可靠完成”，后者解决“系统如何通过真实运行持续改进”。



图 2：三大能力 × 两个横切机制 × 两个闭环

4.2 数据能力：物理仿真数据生成与规模化数据闭环

数据能力解决的是 Physical AI 的训练、评测和持续迭代基础问题。室内环境具有光照变化、遮挡、反光、人员活动、空间布局变化以及异常事件长尾等特点，仅依赖真实数据采集往往面临采集成本、隐私约束和样本覆盖不足。因此，需要构建由**真实数据**、**仿真数据**与**运行反馈**共同组成的数据体系。

数据能力主要包括三个环节：

- **真实与仿真数据供给。**基于真实场景采集、数字场景资产和物理仿真生成不同空间、光照、材质、遮挡及异常状态的数据。仿真数据主要用于扩大场景覆盖范围和补充低频、长尾样本；真实数据用于校准仿真分布和验证模型在目标环境中的实际性能。
- **数据治理与训练支撑。**建立数据版本、标注规范、质量检查和训练集/验证集管理机制，对物理合理性、标注一致性、场景覆盖度和数据分布进行检查。神经渲染、域随机化和风格适配可以用于降低部分视觉域差异，但最终效果仍需要通过目标场景中的真实数据进行验证。
- **Badcase 回流与持续优化。**真实部署过程中产生的失败案例、人工接管记录和异常事件进入数据闭环，用于补充训练集、调整仿真参数并开展回归评测。

因此，数据体系不是一次性交付物，而是伴随系统运行持续更新的工程资产。

4.3 认知能力：室内空间智能专属世界模型与可信推理

空间认知与决策负责把人的任务意图和环境状态转化为可执行的局部目标与行动策略。这一能力不应由单一世界模型、VLN 或 VLA 独立承担，而应采用**分层、分频的规划与控制体系**。

运行时的基本控制链路为：**语言任务 → VLN / Task Planner → 局部目标 → Local Planner / VLA Policy → Safety Controller → Motor Controller。**

- **感知、定位与空间状态维护。**系统首先融合视觉、定位、地图或空间记忆等信息，维护机器人自身位置、目标对象、区域状态和任务上下文。在复杂环境中，空间认知不仅需要回答“机器人在哪里”，还需要描述“目标在哪里、哪些区域可以进入、环境发生了什么变化”。
- **VLN / Task Planner：任务理解与全局规划。**VLN 或任务规划器根据语言任务、空间状态和业务约束理解任务目标，并将长时程任务分解为可执行的局部目标。VLN 是连接**语言目标**、**视觉理解**与**导航决策**的重要能力，但并不替代定位、局部规划和底层控制。
- **Local Planner / VLA Policy：局部决策与动作生成。**局部规划器或 VLA Policy 依据实时环境观测和局部目标，以较高频率完成路径调整、避障和动作策略生成。不同任务可采用不同实现方式：导航类任务可以主要依赖局部规划与运动控制，复杂交互或操作任务则可引入 VLA Policy。

- **世界模型：辅助预测与风险推演。**世界模型可用于环境状态表征、未来状态预测以及潜在行动结果推演，为长时程规划、风险评估和异常重规划提供辅助信息。原方案已将世界模型用于状态、动态及行动后果预测。世界模型原则上不直接承担安全敏感的高频底层控制，而是作为较低频的认知和预测模块，与实时控制链路解耦。

这种分层设计的目的，是将**任务理解、全局规划、局部策略、安全控制和底层执行**分别放在适合的时间尺度和计算位置运行。

4.4 安全能力：与数智本体 Ontology 结合的多层安全防护

当系统从“提供判断”进入“驱动物理设备”，安全机制必须从模型输出后的附加检查，转变为贯穿认知、规划和执行过程的系统能力。安全与治理主要包括三个方面：

(1) 权限与动作约束

通过语义本体（Ontology）、业务规则、区域权限和动作白名单等机制，明确：

- 什么对象可以被操作；
- 什么区域可以进入；
- 什么状态允许执行；
- 哪些动作需要人工确认。

数智本体（Ontology）通过显式的实体、属性、关系、状态、约束和业务规则，为 Agent 的推理与行动提供结构化知识边界，在端侧运行期构建神经符号（Neuro-Symbolic）约束护栏，毫秒级过滤越权指令并抑制超范围判断，同时完整支持决策链追溯^[15]；但语义约束不能替代底层功能安全。

(2) Safety Controller 独立实时控制

在 Local Planner 或 VLA Policy 与 Motor Controller 之间设置独立 **Safety Controller**。其根据碰撞距离、速度限制、禁行区域、机器人状态以及硬件安全规则，对候选动作进行校验、限制或拦截。对于急停、碰撞保护等安全敏感能力，应尽可能由本地控制器独立完成，避免依赖云端模型或网络连接。

相关安全设计与验证可参考服务机器人 ISO 13482、ISO/TR 23482-1，以及仓储 AMR 等无人驾驶工业车辆适用的 ISO 3691-4:2023^{[18][19][20]}。

(3) 隐私、审计与人工接管

针对家庭、楼宇和机房等场景，对数据实施分级授权、本地优先处理和最小必要原则；记录关键输入、状态、规划结果、安全拦截和最终动作，为异常分析、责任追踪和模型回归提供依据。隐私数据本地优先存储，从源头降低泄露面；对于断电、关阀、门锁控制以及紧急事件处置等高风险动作，应依据设备接口、用户授权、风险等级和属

地运营流程明确执行边界，而不将所有异常默认设计为自动物理处置。

4.5 横切机制一：云—边—端运行架构

Physical AI 同时存在复杂推理和实时物理控制两类计算需求，两者在时延、算力和可靠性要求上存在明显差异，因此需要采用云—边—端分层运行方式。

(1) 云：复杂认知与集中服务

云端适合承担：

- 复杂多模态推理；
- 长时程任务规划；
- 世界模型预测；
- 跨设备状态管理；
- 模型训练与版本管理。

(2) 边：实时认知与任务自闭环

边缘节点承担对时延较敏感、同时需要一定计算资源的任务，例如：

- 本地感知与识别；
- 局部任务推理；
- 地图与状态维护；
- 部分导航策略；
- 本地任务闭环。

(3) 端：确定性执行与安全控制

机器人或智能终端负责：

- 传感器采集；
- Local Planner；
- Safety Controller；
- Motor Controller；
- 急停及基础安全机制。

计算位置不是固定的，应根据任务复杂度、网络状态、硬件算力和时延预算动态配置。网络异常时，安全相关能力应继续在本地运行。

4.6 横切机制二：Agent / 任务编排

Agent 不是数据、认知、安全之外的第四种基础能力，也不直接替代机器人控制系统。它的主要作用是作为**任务级软件编排层**，把多个模型、服务、设备和业务流程组织成一项完整任务。

Agent 主要负责：

- **状态管理**。维护任务目标、当前进度、上下文、历史动作以及执行状态。
- **Skill 调用**。根据任务调用感知、VLN、地图、规划、设备控制、工单或通知等标准化能力。
- **异常处理**。根据任务结果触发重试、重规划、降级、人工接管或业务处置流程。

因此，Agent 解决的不是“机器人如何实时转弯”，而是：**任务目标是什么、下一步调用什么能力、任务是否已经完成，以及失败以后如何处理**。通过任务 API、Skill API 和机器人能力描述接口，可以降低业务层与具体硬件之间的耦合，并为多种机器人或智能终端接入提供统一的任务编排方式。

4.7 闭环一：在线任务闭环

在线任务闭环描述的是一次**真实任务**如何从意图进入物理世界并得到结果确认。

其基本链路为：**感知 → 理解 → 规划 → 控制 → 验证 → 恢复**

- **感知**，获取视觉、位置、设备状态及环境信息。
- **理解**，结合空间语义、任务上下文和业务规则形成当前状态判断。
- **规划**，由 Task Planner / VLN 生成任务路径和局部目标。
- **控制**，Local Planner / VLA Policy 生成候选动作，经 Safety Controller 校验后，由 Motor Controller 执行。
- **验证**，通过传感器和业务状态检查任务是否真正完成，而不是仅确认指令是否发送。
- **恢复**，发生路径阻塞、目标丢失、执行失败或环境变化时，进入重新感知、重新规划、降级执行或人工接管。

因此，Physical AI 的任务完成定义应从：“**机器人执行了一条命令**”转变为：“**系统完成目标，并验证实际结果**”。这是从“感知闭环”进入“行动闭环”的关键区别。

4.8 闭环二：离线数据闭环

在线任务解决当前问题，离线闭环负责让系统从真实运行经验中持续改进。

其基本路径为：**真实 → 仿真 → 训练 → 评测 → 部署 → 回流**

- **Real**：真实数据，收集经过治理的真实环境数据、失败案例和人工接管记录。

- **Sim:** 仿真生成，根据真实数据校准场景参数，并针对难例、长尾和风险场景补充仿真数据。
- **Train:** 训练与后训练，结合真实与仿真数据完成任务模型、策略模型或相关认知模型的训练与适配。
- **Eval:** 系统评测，不仅评价离线模型指标，还应评价导航成功率、任务完成率、安全违规率、人工介入率、恢复成功率以及端到端时延。
- **Deploy:** 受控部署，通过模型、地图、规则和安全策略版本管理，将验证通过的能力部署至目标场景。
- **Feedback:** 运行回流，将新的 Badcase、异常行为和任务失败重新进入数据体系，形成下一轮优化输入。

这一机制不再只是“数据能力”中的一个环节，而成为贯穿整个 Physical AI 生命周期的离线持续改进闭环。

4.9 两个闭环如何协同

两个闭环分别运行在不同时间尺度。

在线任务闭环关注“秒到分钟”：当前任务是否安全、可靠地完成。

离线数据闭环关注“天到版本周期”：系统是否因为真实运行而持续提高能力。

在线任务产生真实执行数据和 Badcase，进入离线闭环；离线训练与评测形成新的模型、地图、规则和策略版本，再重新进入在线系统。由此形成：**执行产生数据 → 数据推动改进 → 改进重新进入执行** 的持续演进机制。

这也是 Real2Sim2Real 与 Agent 任务闭环需要明确区分的原因：前者属于**模型与系统生命周期闭环**，后者属于**实时业务执行闭环**，二者通过真实运行数据相互连接，但并不是同一个闭环。

第5章

典型应用场景与验证路径

在场景中验证，在验证中造血



架构方案需要在真实业务场景中完成检验迭代。本章围绕家庭、机房、仓储、楼宇四大室内场景，阐述各场景任务、部署模式与已完成的工程验证结果。同时说明试点共建机制，确立“验证-造血-规模化”的阶梯式落地路径。

5.1 智慧家庭安防巡检（首发验证）

任务内容：三类任务形态——主动巡检（定时/事件触发的全屋例行巡查）、指令巡检（“去看看厨房的火关了没有”等自然语言指令）、风险预警（跌倒、烟雾、燃气、异响、陌生人等异常主动上报）。



图 3：智慧家庭安防巡检场景

方案形态与部署模式：百元级家用移动相机本体 + 边端自闭环 + 云端按需复核，单目 RGB 输入；依托运营商智家渠道与装维网络完成入户部署与运维，能力边界显性声明——不承诺超出传感器与算力配置的能力。

环境验证：本方案已在客户指定的真实家庭环境中完成限定条件下的工程验证，已验证基本导航与端侧推理链路。

与智能家居产业基础的关联：全屋智能已积累数千万级用户的教育基础；互联互通标准（国标 + Matter）降低设备接入门槛；运营商智家渠道与装维网络提供部署与运维触点。家庭巡检是智能家居从“静态智能”走向“主动巡检”的自然延伸。

5.2 数据中心与通信机房巡检（行业级桥梁）

任务内容： 设备状态检测（指示灯、开关、屏显）、温湿度与热异常监测、仪表读数识别。



图 4：数据中心与通信机房巡检场景

方案形态与部署模式： 轨道/轮式本体 + 边端实时巡检 + 云端集中研判，与运维系统打通形成"发现—工单—销缺"完整闭环；采用"联通自用在先、对外输出在后"策略，依托中国联通自有机房与数据中心体系先行部署，打磨成熟后作为标准化能力对外输出。机房场景结构化程度高于家庭、容错要求高于家庭，恰是消费侧能力向行业侧迁移的中间台阶。

环境验证： 本方案已在客户指定机房环境中完成工程化部署验证，已验证仪表识别与自动巡检流程。

与智能家居产业基础的关联： 机房巡检复用家庭场景中锤炼成熟的室内导航、状态判读与三级响应闭环能力，并共享运营商在智家业务中建立的装维与工单体系——家庭首发的认知能力由此迈出向行业侧迁移的第一步。

5.3 仓储巡检（规模化延伸）

任务内容： 库存盘点、货物状态监测、消防隐患预警、环境温湿度监测。

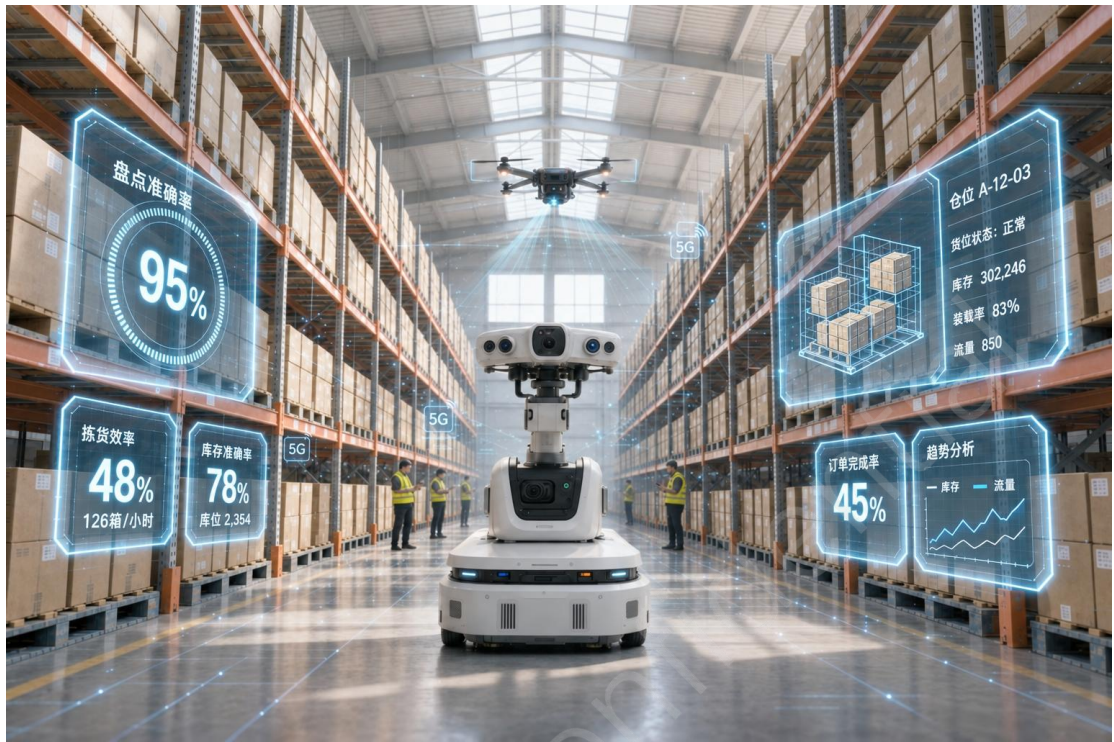


图 5： 仓储巡检场景

方案形态与部署模式： 同一套系统白天盘点、夜间安防，盘点作为高频任务让巡检体"天天有活干"，直接对冲"低频价值 vs 持续在线"的成本矛盾，设备利用率最大化使 TCO 模型成立；依托运营商 5G 专网（在仓储物流行业已规模落地）部署，盘点数据与 WMS 系统对接，巡检从"安全成本项"转化为"运营增效项"。

环境验证： 本方案已在客户指定仓储环境中完成工程化部署验证，已验证库存盘点与 WMS 对接流程。

与智能家居产业基础的关联： 仓储盘点所需的空

5.4 商业楼宇巡检（普惠延伸）

任务内容： 公共区域安防巡视、消防设施状态检查、设备机房定期巡检、环境异常（漏水/异响/异味）预警。



图 6：商业楼宇巡检场景

方案形态与部署模式： 复用已有物业巡更路线与点位体系，与楼控及物业管理系统对接，部署门槛低、改造成本小；按"楼层分区 + 定时路线 + 事件触发"组织巡检任务，适配楼宇存量运营的实际情况。

环境验证： 本方案已在客户指定楼宇环境中完成工程化部署验证。单套系统稳定覆盖 200 个以上巡更点位、替代约三成例行人工巡更工作量，现场部署周期由行业惯常的周级缩短至天级。

与智能家居产业基础的关联： 楼宇场景与家庭共享"非结构化室内空间"的深层共性——光照多变、材质多样、人员动态、空间语义复杂、异常长尾。家庭首发验证的认知能力可低成本迁移，使楼宇成为家庭能力向泛商业空间普惠延伸的最典型场景。

第6章

未来展望

从单体巡检，走向全域智能



当前室内空间智能仍处在发展早期，现有巡检能力只是 Physical AI 发展的阶段性起点。本章从技术演进、产业商业模式两个维度研判长期发展方向，梳理多智能体、服务化交付等重要趋势。最后发出产业倡议，呼吁行业共建标准、评测与普惠服务生态。

6.1 技术趋势：从单体巡检智能体走向全域空间智能

近中期，多智能体协同将使空间服务从“一台设备巡一个空间”走向“多本体协同覆盖整栋建筑”；系统对空间的理解将从点位级走向建筑级——“理解整栋建筑”的空间状态与动态。远期，任务形态将探索从观测型向介入型的渐进演化（从“发现燃气泄漏并关阀”到更复杂的物理干预），届时需同步重新评估安全边界与责任框架，技术进步与治理框架必须同频演进。

VLN 是室内空间智能中负责“语言目标—视觉理解—导航决策”衔接的一类关键能力，将从“单房间导航”演进为“跨楼层、跨建筑的长程任务规划”，并与世界模型深度融合，实现“先推演、再行动”的预测性空间服务。

6.2 产业趋势：硬件标准化、能力服务化、数据资产化

本体接口标准化将使硬件成为可替换的执行载体；空间智能能力将以机器人即服务（Robot as a Service, RaaS）模式订阅化交付；在室内数据闭环中沉淀的场景数据资产将形成数据即服务（Data as a Service, DaaS）价值层。能力外溢路径清晰：从家庭向工业园区、商业楼宇持续延伸，远期进一步外溢至工业园区室外能源电力设施等增量场景。

6.3 产业倡议

1. 开放室内场景数据，共建评测基准。呼吁产业界在合规前提下开放室内场景数据与评测任务，建立行业公认的室内空间智能能力评测基准，让技术进步可测量、可比较；
2. 共研标准与评估体系。依托具身智能标准化技术委员会框架，共同推进室内空间智能的接口标准、安全标准与能力评估体系，以标准先行护航产业规模化；
3. 依托运营商云网，让空间智能像通信服务一样普惠可及。以运营商网络与云资源池为底座，把室内空间智能能力做成像宽带、像话费套餐一样可订阅、可依赖的基础服务，让空间智能进入千家万户、千行百业。

致谢

本白皮书的研究与编撰，得到各方机构与专家的大力支持。感谢联合编制单位亚信科技、联通在线、清华 AIR、VERYK GROUP 的研究与技术团队，为报告输出贡献理论研究、工程实践与行业场景经验。

感谢工业界、学术界各位专家给予的宝贵意见，同时感谢参与技术验证、试点测试的相关合作伙伴。欢迎产业同仁开展交流共建，共同推进室内空间智能产业发展。

AsiaInfo Confidential

参考文献

- [1] Miriyev A, Kovač M. Skills for physical artificial intelligence[J]. Nature Machine Intelligence, 2020, 2: 658–660. DOI: 10.1038/s42256-020-00258-y.
<https://doi.org/10.1038/s42256-020-00258-y>
- [2] Miriyev A, Kovač M. Materializing artificial intelligence[J]. Nature Machine Intelligence, 2020, 2: 653 (评论/Comment). DOI: 10.1038/s42256-020-00262-2.
<https://doi.org/10.1038/s42256-020-00262-2>
- [3] 工业和信息化部. 三问具身智能[EB/OL]. (2025-03-26).
https://www.miit.gov.cn/xwfb/mtbd/twbd/art/2025/art_e183ac0452d448b2ba554ca1e706b790.html
- [4] Ha D, Schmidhuber J. World models[EB/OL]. arXiv:1803.10122, 2018.
<https://arxiv.org/abs/1803.10122>
- [5] Parker-Holder J, et al. Genie 2: A large-scale foundation world model[R/OL]. Google DeepMind (官方博客/技术报告), 2024-12-04. <https://deepmind.google/blog/genie-2-a-large-scale-foundation-world-model/>
- [6] NVIDIA. Cosmos world foundation model platform for physical AI[R/OL]. NVIDIA Research, 2025-01-06. arXiv:2501.03575. <https://arxiv.org/abs/2501.03575>
- [7] Zhang Y, Ma Z, Li J, et al. Vision-and-language navigation today and tomorrow: A survey in the era of foundation models[EB/OL]. arXiv:2407.07035, 2024.
<https://arxiv.org/abs/2407.07035>
- [8] 国务院. 2025 年政府工作报告[R]. 北京, 2025-03-05.
https://www.gov.cn/yaowen/liebiao/202503/content_7013163.htm
- [9] 国务院. 2026 年政府工作报告[R]. 北京, 2026-03-05; 工业和信息化部. 中华人民共和国工业和信息化部公告 (2025 年第 44 号) [EB/OL]. 2025-12-31.
https://www.miit.gov.cn/zwgk/zcwj/wjfb/gg/art/2025/art_c157671d4b3a4f5bb6a1bf4335adb2c4.html
- [10] 工业和信息化部办公厅, 国务院国有资产监督管理委员会办公厅. 关于联合开展 2026 年度人形机器人与具身智能实景实训专项行动的通知 (工信厅联科函 (2026) 256 号) [EB/OL]. 2026-06-08.
https://www.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2026/art_f291ccd3da4c47ce95741de63cc088e6.html
- [11] 国家市场监督管理总局, 国家标准化管理委员会. GB/T 28219.1—2025 智能家居电器的智能化技术要求和评价 第 1 部分: 通用要求[S]. 2025-10-31 发布, 2026-05-01 实施.
<https://openstd.samr.gov.cn/bzgk/std/newGbInfo?hcno=DC9445D7B15E16AC6C967CA34FDA9CE8>
- [12] 国家市场监督管理总局, 国家标准化管理委员会. GB/T 46505.1—2025 家用电器应用场景 第 1 部分: 通用要求[S]. 2025-10-31 发布, 2026-05-01 实施.
<http://openstd.samr.gov.cn/bzgk/std/newGbInfo?hcno=D1F24EC14107C7F8962A0BF24474A5A3>

- [13] 国家统计局. 中华人民共和国 2025 年国民经济和社会发展统计公报[R/OL]. 2026-02-28. https://www.stats.gov.cn/sj/zxfbhjd/202602/t20260228_1962662.html
- [14] IDC. 中国智能家居设备市场季度跟踪报告[R/OL]. IDC 中国, 2025–2026. 注：出货量口径以 IDC 正式发布为准。 https://my.idc.com/getdoc.jsp?containerId=IDC_P37480
- [15] W3C. OWL 2 Web Ontology Language document overview (Second Edition)[EB/OL]. W3C Recommendation, 2012-12-11. <https://www.w3.org/TR/owl2-overview/>.
- [16] Brohan A, Brown N, Carbajal J, et al. RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control[EB/OL]. arXiv:2307.15818, 2023. <https://arxiv.org/abs/2307.15818>
- [17] Kim M J, Pertsch K, Karamcheti S, et al. OpenVLA: An Open-Source Vision-Language-Action Model[EB/OL]. arXiv:2406.09246, 2024. <https://arxiv.org/abs/2406.09246>
- [18] ISO. ISO 13482:2014 Robots and robotic devices — Safety requirements for personal care robots[S]. 2014. <https://www.iso.org/standard/53820.html>
- [19] ISO. ISO/TR 23482-1:2020 Robotics — Application of ISO 13482 — Part 1: Safety-related test methods[TR]. 2020. <https://www.iso.org/standard/71564.html>
- [20] ISO. ISO 3691-4:2023 Industrial trucks — Safety requirements and verification — Part 4: Driverless industrial trucks and their systems[S]. 2023. <https://www.iso.org/standard/83545.html>
- [21] 全国人民代表大会常务委员会. 中华人民共和国个人信息保护法[Z]. 2021-08-20. https://www.miit.gov.cn/zwgk/zcwj/flfg/art/2022/art_04a0f1fb5df244e39688fd5372623a8d.html
- [22] 国家市场监督管理总局, 国家标准化管理委员会. GB/T 45574-2025 数据安全敏感个人信息处理安全要求[S]. 2025. <https://openstd.samr.gov.cn/bzgk/std/newGbInfo?hcn=F9F3A2EBF49E9B4D73AD8C8912986D5A>
- [23] Tobin J, Fong R, Ray A, et al. Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World[C/OL]. IROS 2017; arXiv:1703.06907. <https://arxiv.org/abs/1703.06907>
- [24] Peng X B, Andrychowicz M, Zaremba W, Abbeel P. Sim-to-Real Transfer of Robotic Control with Dynamics Randomization[C/OL]. ICRA 2018; arXiv:1710.06537. <https://arxiv.org/abs/1710.06537>
- [25] Zhao W, Peña Queralta J, Westerlund T. Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: a Survey[EB/OL]. arXiv:2009.13303, 2020. <https://arxiv.org/abs/2009.13303>

附录 A 数据与实测指标披露规范

为保证白皮书的严谨性与可核验性，全书数据引用遵循以下规则：

1. 政策、标准、人口等外部事实，优先引用政府部门、国家标准平台或正式学术论文；
2. 市场规模、设备出货等商业数据，注明机构、报告名称、统计范围与预测口径；
3. 中国联通网络与云资源相关表述以中国联通官方口径为准；
4. 四方联合验证数据按统一实验环境、统一 Baseline、统一指标定义披露，限定条件在正文中明确标注。

附录 B 参编单位

亚信科技	联通在线	清华 AIR	VERYK GROUP
欧阳晔 亚信科技首席执行官、 IEEE Fellow	刘立国 云智产品支撑中心	张亚勤 院士 中国工程院院士、清华 大学智能产业研究院 (AIR) 创始院长、 IEEE Fellow	Yun Li
陈胜 亚信科技首席技术官、研 创中心(2030 实验室) 总 经理	郑朴汉 云智产品支撑中心	刘云新 教授 清华大学智能产业研究 院(AIR) 副院长、教 授、IEEE Fellow	
孟祥德 亚信科技副总裁、研创中 心(2030 实验室) 执行总 经理	张晓东 云智产品支撑中心	曹婷 教授 清华大学智能产业研究 院(AIR) 教授	
陈果 亚信科技研创中心(2030 实验室) 副总经理	许凯 云智产品支撑中心	詹仙园 教授 清华大学智能产业研究 院(AIR) 教授	
曹阳 亚信科技研创中心(2030 实验室) Physical AI 创新 中心总经理	王凯迪 云智产品支撑中心		
张唯 亚信科技研创中心(2030 实验室) 标准与创新中心 总经理			
王首峰 亚信科技研创中心(2030 实验室) 标准与创新中心 标准化与科研项目负责人			
刘子光 亚信科技研创中心(2030 实验室) Physical AI 创新 中心规划专家			
王雨澄 亚信科技研创中心(2030 实验室) Physical AI 创新 中心产品总监			
张泽生 亚信科技研创中心(2030 实验室) Physical AI 创新 中心规划专家			



版权声明

本白皮书由亚信科技、联通在线、清华 AIR、VERYK GROUP 联合编制并联合发布。
文中技术观点与方案设计自四方共有，转载请注明出处。

免责声明

本白皮书所载内容基于编制时点的公开信息与四方技术验证结果，不构成任何商业承诺或投资建议。文中实测数据均为限定条件下的技术验证结果，实际部署效果因场景条件而异。市场预测数据来源于第三方机构。四方不对其准确性承担担保责任。



亚信科技