

具身智能训练场研究报告

(2026 年)

中国信息通信研究院人工智能研究所
人形机器人（上海）有限公司
具身智能测试实验室

2026年8月

版权声明

本报告版权属于中国信息通信研究院、人形机器人(上海)有限公司和具身智能测试实验室，并受法律保护。转载、摘编或利用其他方式使用本报告文字或者观点的，应注明“来源：中国信息通信研究院、人形机器人(上海)有限公司和具身智能测试实验室”。违反上述声明者，编者将追究其相关法律责任。

前 言

具身智能作为与物理实体融合的人工智能，正加速从实验室走向产业落地，成为改变人类生产生活方式、推动社会智能跃升的重要未来产业方向。其发展的底层逻辑在于，智能的进化必须实现“知行合一”：既要“读万卷书”，通过海量知识学习理解世界；更要“行万里路”，通过真实物理交互习得行动经验。然而知易行难。大语言模型、多模态大模型依托上千年文字积累和海量互联网数据，实现了“知”的跨越式发展；而具身智能面向物理世界的“行”，则高度依赖真实环境中的实际作业数据。这类数据此前从未被系统地采集、处理和存储，导致具身智能面临巨大的数据缺口。高质量真实数据的缺失，已经成为制约具身智能发展的首要因素。

在此背景下，具身智能训练场作为关键基础设施应运而生，在全国范围内掀起一轮具身智能训练场建设热潮。具身智能训练场是一类面向具身智能研发和应用搭建的关键基础设施，在真实物理环境中支撑具身智能数据生产、模型训练、系统应用能力验证和迭代优化。然而，在快速推进过程中，训练场场景任务建设同质化、数据实际产能有限、数据多样性和流通性不足以及可持续商业模式尚未形成等问题逐渐凸显。训练场建起来仅是第一步，更重要的是用起来、转起来，真正成为数据资源基地和实景实训平台，支撑具身智能产业规模化发展。

本报告全面梳理我国具身智能数据发展背景，以及训练场建设现状，并结合行业实践，分析训练场共性挑战、发展阶段、实施落

地路径和标准化需求。本报告可为政策制定方、具身智能技术提供方、数据需求方、应用需求方、训练场建设方和运营方等提供全面的行业洞察、路径参考与决策依据，助力相关方准确把握具身智能训练场的发展方向，协同推动我国具身智能基础设施体系完善，加速具身智能技术成果转化与产业落地。因技术与产业发展迅速，内容难免存在疏漏，恳请各位专家和读者不吝指正。

目 录

一、 具身智能数据需求催生训练场模式	1
（一）规模定律驱动，数据是具身智能发展重要支撑	1
（二）数据瓶颈凸显，探索真实数据供给和利用能力	3
（三）场景应用牵引，训练场成为当前闭环关键路径	12
二、 我国具身智能训练场建设情况	15
（一）政策牵引地方支持，各地训练场加速建设进程	15
（二）政企建设多方发力，综合行业双模式协同发展	19
（三）应用需求建设导向，训练场行业场景呈梯次分布	23
（四）数据方案多元应用，工程化支撑训练场产能升级	25
（五）产业生态初步构建，训练场探索可持续商业模式	29
三、 我国具身智能训练场实践分析	33
（一）训练场核心建设要素与挑战	33
（二）训练场发展阶段及典型路径	36
（三）训练场标准化实践	42
四、 具身智能训练场未来展望	47

图 目 录

图 1 具身智能训练场相关概念梳理13

图 2 各省（市）已建成训练场的数量分布及统计17

图 3 已披露训练场场景分布23

图 4 训练场产业链上下游分布31

图 5 具身智能训练场能力发展阶段图38

图 6 具身智能训练场标准化需求43

图 7 2026-0074T-YD《人工智能 关键基础技术 具身智能训练场综合评估技术规范》框架46

表 目 录

表 1 不同真实数据采集技术对比4

一、具身智能数据需求催生训练场模式

（一）规模定律驱动，数据是具身智能发展重要支撑

图灵奖得主理查德·萨顿在《惨痛的教训》一文中指出，回顾人工智能七十余年的发展，随着规模扩展而持续受益的通用学习方法，最终总是更为有效，而且优势巨大。谷歌团队于 2009 年发表的论文《The Unreasonable Effectiveness of Data》¹中指出，在机器学习中，数据的规模往往比复杂的模型设计更重要。大语言模型的发展进一步证明了：数据世界的人工智能存在“规模定律（Scaling Law）”，即随着数据规模、模型规模和计算资源的持续扩展，模型能力会不断提升。规模定律已经成为推动人工智能能力突破的重要底层范式。

具身智能作为人工智能与多学科融合的前沿技术，其演进路径的探索主要建立在规模定律成立的假设之上。业界认为，数据驱动的通用学习范式在物理世界依然有效。一方面，视觉语言动作模型、世界模型等通用具身基础模型快速发展，试图将视觉、语言、动作、触觉、力觉等多模态信息统一到同一学习框架中，使模型能够从大规模数据中学习任务语义、物理规律和行动策略。另一方面，产业界和科研界开始遵循“大力出奇迹”的简约扩展策略，用数据驱动的方式，持续整合不同场景、不同任务、不同本体的数据，并依托大规模数据，训练更强大的具身基础模型。

从已有实践看，具身智能模型性能的提升持续受益于数据规模扩展。预训练阶段数据规模的扩大，能够支撑具身智能认知和行动能力

¹ <https://ieeexplore.ieee.org/document/4804817>

的显著提升。2025 年 11 月，美国 Generalist AI 公司在 GEN-0²的技术报告中，围绕具身智能领域的规模定律进行了详细讨论，发现预训练数据越多，模型在下游后训练任务上的平均成功率越高，两者之间存在幂律关系。2026 年 4 月，Generalist AI 推出了新一代模型 GEN-1³，使用超过 50 万小时的真实数据作为预训练数据，在折叠 T 恤、折叠纸盒、打包手机等任务上的平均成功率达到 99%。对比来看，GEN-1 的预训练数据较前代 GEN-0 模型多近 1 倍，平均任务成功率也提升了约 1.5 倍。英伟达高级研究科学家 Jim Fan 认为，具身智能领域正在复刻大语言模型的发展路径。其团队发布的 EgoScale⁴论文研究表明，大规模数据能帮助模型学习操作技能，且数据规模与模型验证损失之间存在清晰的对数线性关系。行业越来越倾向于认为，足够多的数据，是具身智能享受规模定律红利的前提。

然而，在数据从何而来这一问题上，具身智能与大语言模型存在天然差异，前者无法复制后者的数据积累路径。对大语言模型而言，有上千年的文字沉淀和全球超 60 亿网民持续贡献数据，相关机构预测互联网上的文本资料约有 300 万亿词元（token）⁵。对具身智能来说，所需的视觉、动作、关节扭矩、力触觉等多模态交互数据，此前从未被系统性记录，也无法从互联网直接获取，只能从零逐条生产。行业正沿着仿真和真实两条路径缩小数据差距。一方面，依托物理仿

² <https://generalistai.com/blog/gen-0>

³ <https://generalistai.com/blog/gen-1>

⁴ <https://arxiv.org/abs/2602.16710>

⁵

<https://www.bvp.com/atlas/bessemer-predicts-robotics-and-physical-ai#Prediction-1-The-ChatGPT-moment-for-robotics-is-coming-but-we-re-not-there-yet>

真、数字孪生、世界模型和其他生成式方法，能够快速、低成本地获取仿真合成数据；另一方面，不断完善数据采集系统和本体硬件，持续采集真实数据。人工智能科学家、World Lab 公司创始人兼 CEO 李飞飞近期的访谈⁶中提到，仿真数据与真实数据是互补关系，仿真能够提供可扩展、可控制的大规模训练数据，但无法替代真实数据。真实数据能够提供真实世界的物理约束和分布特性，是模型最终落地不可缺少的部分。然而当前真实数据面临巨大的资源缺口。据相关预测，假设实现机器人“ChatGPT”时刻的模型参数量级为 55B，按大语言模型的缩放定律，训练所需数据量约 1.1 万亿 token⁷。若全部由真机回流提供，按控制频率 8Hz、每步约 3 个预测 token、每天有效交互 2 小时、每年工作 300 天测算，约需 21200 台机器人连续工作一年，产出千万小时级有效数据。而当前全球可用的高质量真实数据仅数十万至百万小时级，远未达到支撑具身基础模型智能涌现的门槛。

（二）数据瓶颈凸显，探索真实数据供给和利用能力

1. 多路线探索真实数据采集技术

为解决真实数据短缺瓶颈，行业多路线探索真实数据采集技术。

一是无本体采集，凭借人类身体的高灵活性与环境适应性，通过人类操作员手持末端执行器或穿戴采集设备，在真实场景或生产工作中执行任务，以采集人类运动视频、轨迹、手部动作，以及人类通过自身判断对末端执行器的操作信息等。这类数据从操作员及其周围环境中得到，经常需要多种设备组合及识别算法处理，并从人体向具身智能

⁶ https://a16z.com/podcast/the-frontier-of-spatial-intelligence-with-fei-fei-li/?utm_source=chatgpt.com

⁷ <https://rayansmalik.substack.com/p/scaling-to-a-trillion-tokens-in-robotics>

本体映射，以覆盖多模态动作信息和训练具身智能模型；**二是真机采集**，依托具身智能本体获得数据，例如遥操作采集，主要包含本体视觉、触觉、运动轨迹、关节角度、末端状态等多模态信息。这类数据从本体及其周围环境中得到，因此与本体结构一致性高。此外，真机部署在真实场景后回流数据，也成为真机数据新的来源。不同真实数据采集技术在数据获取成本、规模化程度和数据价值上各不相同，形成多元协同互补的数据供给格局。如表 1 所示。

表 1 不同真实数据采集技术对比

路线	数据采集方法	成本	规模化程度	真机映射难度	数据用途
无本体采集	人类第一视角视频数据采集	低	高	高	支持场景理解、人类行为理解、空间预测
	末端操作数据采集	低	高	高	支持多场景下末端精细交互操作任务
	动作捕捉数据采集	中等	中等	中	支持全身或手部复杂运动交互
真机采集	遥操作数据采集	高	低	低	无域偏移问题，与本体运动、力控、结构一致
	自主作业回流数据	最高	低	低	具备真实环境信息、真实扰动、失败案例等高价值数据

来源：中国信息通信研究院

一是人类第一视角视频数据采集，利用头戴式相机采集人类第一视角的行为数据，用于提升模型的场景理解、人类行为理解以及空间预测等能力。这类数据规模最大，采集成本低，单台套售价在数万元，但通常缺少机器人本体状态和可执行动作标签，无法直接使用，需要进行动作识别、轨迹估计和本体映射等后处理。例如英伟达 EgoScale

大规模人类第一视角数据，需结合同步定位与地图构建（Simultaneous Localization and Mapping, SLAM）轨迹恢复、三维手部姿态估计、逆运动学与动作重定向等进行数据后处理，以构建统一的末端执行器动作空间。京东开源的 EgoLive 大规模人类第一视角数据集，原始数据包含 RGB 视频、相机信息以及惯性测量单元（Inertial Measurement Unit, IMU）数据，但需要配合后续数据标注流程中的手部重建、手势识别、手-物分割、密集深度场景重建等后处理流程，并估计腕部与手部关节的 6D 轨迹，与相机同步，建立行为参考系。超维动力 KAI Halo 通过超广角相机和 IMU 采集数据，重建全身关节点运动轨迹与双手位姿，以满足具身智能数据需求。

二是末端操作数据采集，通常由操作员手持或穿戴通用操作接口（Universal Manipulation Interface, UMI）夹爪、数据采集手套或腕部相机作为采集工具，采集末端操作信息，多用于模型学习真实场景下的末端精细操作技能。UMI 夹爪利用与机器人末端执行器形态相近的夹爪作为操作接口，通过传感器实时采集操作员的夹持开合、位姿变化及运动轨迹数据；数据采集手套由操作员穿戴于手部，实时采集手部动作视频、运动轨迹和力触觉信息等多模态数据；腕部相机通常与人类第一视角视频数据采集方式相配合，针对操作员的手部操作动作形成多视角视频。末端操作数据采集能够获取操作员手部运动轨迹、操作细节以及力触觉等多模态真实物理交互数据，且结构轻量、部署灵活、单套成本在数万元，在数据质量、采集效率和成本方面取得较好平衡，已成为行业目前最受关注的数据采集方案之一。例如觅蜂科

技 MEgoGripper⁸，可采集 RGB 图像、深度信息、IMU 数据、音频，以及夹爪开合与触觉信息。鹿明机器人 FastUMI Pro⁹，计划 2026 年与训练场合作实现 100 万小时数据产能。帕西尼自研 P MEC 数据手套，可同步采集手部姿态、触觉信息，部署在天津帕西尼超级数据采集工厂中。

三是动作捕捉数据采集，主要包括光学捕捉和惯性捕捉两种方式，用于采集人体行为数据，多用于全身复杂运动或手部灵巧操作技能学习。这类数据需要经骨骼解算和动作重定向后，才能映射到机器人本体，映射过程中可能出现动作失真、约束不匹配和信息损失等问题。同时，光学动捕方案对场地要求严格，需布置多视角高精度相机，成本在数十万到百万不等，具体取决于相机数量、动作捕捉场地空间和精度要求等。惯性动捕方案环境适应性强、便携性强，但惯性传感器本身不能进行空间绝对定位，存在漂移问题，成本在数万到数十万元不等。例如虚拟动点 OptiTrack 光学运动捕捉系统可实现高精度三维重建与全身骨骼动作数据，已在北京石景山人形机器人训练中心二期部署。Xsens Link 惯性动捕系统可实时捕捉人体运动姿态数据，覆盖关节角度、角速度、加速度等关键参数，可与 Manus 数据手套协同实现全身动作捕捉。

四是遥操作数据采集，由数采人员通过虚拟现实（Virtual Reality, VR）设备、主机械臂、外骨骼、动捕设备等主端控制设备，控制从端具身智能本体执行真实任务，同步记录本体产生的视觉、力觉、触

⁸ <https://www.maniformer.ai/en/mego>

⁹ <https://lumosumi.lumosbot.tech/pro/>

觉等多模态数据的采集方式，与机器人本体结构一致性最强，用于提升真机物理交互能力。设备方面，VR 遥操作套装例如 Meta Quest 3、Pico 4 Ultra 等，通过头戴式显示设备和遥控操作设备，能够遥操作多品牌、多型号本体。方舟无限 AC One 可移动主从机械臂平台，搭载 2 台 ARX X5 机械臂，采集物体位置、姿态及操作过程数据。然而遥操作数据采集投入成本较高，需要配置本体、操作人员、采集场地和任务物料等。据 IT 桔子披露数据¹⁰，真机遥操作单小时有效数据成本在规模化运营后约 275 元（含设备折旧成本、人工成本、场景成本等），小规模采集时单小时成本可达数千元。

此外，真实场景部署后的自主作业回流数据，是对模型持续迭代优化有益的高价值真机数据。这类数据自动采集本体在真实任务执行过程中的成功样本、失败案例、人工接管记录和环境长尾问题等信息。但受限于现阶段的部署规模，短期内难以快速形成大规模高质量供给，更适合作为小规模高价值场景数据，用于对模型开展评估、验证和持续优化。

2.探索挖掘真实数据训练应用价值

为充分高效组织多源采集的真实物理交互数据，释放数据价值，提升模型性能，业界始终在探索真实数据在模型训练中的应用方式。一是无本体采集真实物理交互数据，具备任务逻辑与本体结构解耦特性，不依赖特定机器人本体，为模型提供跨硬件的通用学习样本，根据“环境如何随行动变化”相关信息提升空间智能；二是真机数据能

¹⁰ https://mp.weixin.qq.com/s/depWwWo2TcnTKWzkAEjs_Q

够在预训练中为模型提供真实执行信息，同时自主作业回流数据进一步为模型补充各类长尾情况与失败纠错经验。

无本体数据主要用于模型预训练阶段学习大规模行为先验，其中数据对齐是弥合跨本体行为差异的关键。一是物理动作空间显式对齐，多用于具备包括但不限于末端位姿、六自由度轨迹、夹爪开合等通用表达的 UMI 数据、动作捕捉数据或人类第一视角数据，再迁移到不同本体。例如英伟达 GR00T N1.7¹¹，在预训练阶段使用 2 万小时的第一视角人类视频数据集 EgoScale，凭借该数据集处理后人类动作空间与本体对齐的特性，将从人类视频中学习到的操作先验直接转移到机器人控制中。阿里巴巴通义千问联合中国人民大学、上海科技大学、北京通用人工智能研究院和北京航空航天大学发布的 Ego2Robot¹²，从人类第一视角操作视频的手部动作中恢复轨迹，映射到机械臂，并在画面中合成具身智能本体外观，进一步处理视觉差异。二是硬件层面物理对齐，多用于 UMI 采集数据，从末端相机、末端执行器和数据采集动作空间直接对齐，减少数据转换带来的误差。清华大学 RDT2¹³采用超 1 万小时家庭场景 UMI 数据集进行预训练，并通过采用与机器人末端相机同型号的 UMI 相机、执行器夹爪和 6-DoF 位姿，实现免后训练的跨本体部署，在叠衣服、整理餐桌、拉拉链等复杂任务上表现优于 π 0.5。

真机数据与实际执行任务的本体动作空间一致，支持模型快速实现任务对齐、本体适配和迭代优化。一是在后训练阶段使用真机数据

¹¹ <https://github.com/NVIDIA/Isaac-GR00T#whats-new-in-gr00t-n17>

¹² <https://arxiv.org/pdf/2608.02580>

¹³ <https://arxiv.org/abs/2602.03310>

学会具体任务。例如 NVIDIA 的 GR00T N1 系列模型、Physical Intelligence 的 π 系列模型等，均采用真机数据进行后训练。自变量机器人 XRZero-G0¹⁴ 实验证明在后训练阶段可借助少量真机数据让模型适应底层控制策略。高德 ABot-M0.5¹⁵ 世界动作模型，在基础监督微调阶段引入真机数据，使模型学会将真实场景的未来状态映射到正确的动作。**二是**用真机自主执行任务的回流数据不断优化面向具体任务的作业能力。真机自主执行任务回流的执行成功、执行失败或任务恢复等多种情况的数据，相比人工采集数据更具备长尾效应，因此对于模型来说具有高学习价值。Physical Intelligence 2026 年 4 月推出的 π 0.7¹⁶，除人类视频、开源机器人数据、遥操作真机数据外，引入 π *0.6¹⁷ 强化学习训练后自主执行任务的高价值回流数据，为模型制造出从高质量动作到低质量动作的数据分布，提高推理策略的执行质量。原力灵机 DW0.5¹⁸ 世界模型通过真机数据持续校准，为模型提供接近真实部署分布的动作预测。上海创智学院联合智元发布 LWD¹⁹，实时回流机器人在真实环境中的任务执行数据更新模型，在长程任务上的平均子步骤得分相比监督微调方法提升近 50%。

3.探索真实数据持续供给可行路线

随着具身智能加速从技术验证走向产业化，全球具身智能真实数据供给逐步从科研驱动转向产业驱动。早期真实数据采集由高校和科研机构主导，通过多方参与的科研协作路径，汇聚多种数据采集形式，

¹⁴ <https://arxiv.org/abs/2604.13001>

¹⁵ <https://arxiv.org/abs/2607.00678>

¹⁶ <https://arxiv.org/abs/2604.15483>

¹⁷ <https://arxiv.org/abs/2511.14759>

¹⁸ <https://github.com/dexmal/opendm>

¹⁹ <https://arxiv.org/abs/2605.00416>

并形成开源数据集，主要服务于模型技术迭代和前沿探索，广泛支撑具身智能领域技术研究。例如斯坦福大学联合加州大学伯克利分校、丰田研究院等 13 家机构共同推出的 DROID 数据集²⁰，在北美、亚洲、欧洲等全球多地组织 50 名采集者，使用统一的 Franka Panda 机械臂平台，历时 12 个月累计收集了 7.6 万条、约 350 小时真机操作数据，截至 2026 年 6 月，该数据集技术报告在谷歌学术已被引用近 900 次。2022 年 2 月，CMU、Meta、Google 等全球 13 家顶尖机构联合发布的 Ego4D 数据集，由 900 余名参与者在 9 个国家采集形成 3670 小时第一视角视频，截至 2026 年 6 月，该数据集技术报告在谷歌学术被引量超 2200 次。

现阶段，具身智能数据供给已从理论研究导向加速转向产业落地导向。产业侧对真实数据的规模体量、采集效率和场景适配性提出了更高标准。当前，行业重点关注的数据供给方式主要有以下两种：**一是分布式真实数据采集**，组织采集人员在不同真实场景采集真实物理交互任务数据。**国外方面**，特斯拉在其加州工厂雇佣数据采集工人，穿戴视频采集头盔，采集生产操作数据。**Sunday Robotics** 自研技能捕捉手套，记录用户在真实家庭环境中做家务的数据，目前已向超过 2000 名分布式数据采集者发放手套，在超 500 个真实家庭中采集约 1000 万个家务操作片段。美国 **Human Archive** 公司已与印度的工厂、酒店、餐饮、家政服务、建筑等行业的 120 多个公司签约，给工人配备头戴式设备以及触觉手套、腕部摄像头等，记录工人在工作中的所

²⁰ <https://arxiv.org/abs/2403.12945>

有动作。**国内方面**，京东与江苏宿迁湖滨新区联合建设具身智能数据采集社区，用头戴的 JoyEgoCam 数据采集终端收集居民做家务数据。戴盟机器人与中国移动正式达成合作，依托中国移动线下营业网点，共建覆盖外发式数据采集网络。**二是真实场景数据回流**，随着机器人进入工厂、仓储、商业服务、家庭等真实应用环境，数据获取方式正在从一次性采集转向持续回流。**国外方面**，特斯拉依赖自有工厂开展数据飞轮闭环，Optimus Gen3 机器人在德州 Giga 工厂车间运送 Model Y 的电池托盘、零件分类、质检，回流任务数据。Figure AI 与宝马集团合作，将 Figure 02 部署在斯帕坦堡汽车工厂执行 BMW X3 车身钣金零件焊接任务，并进一步在该工厂部署全球首个 Figure 03 人形机器人试点，用于汽车生产物流中的组件排序环节。**国内方面**，具身智能企业开始逐步落地工业、服务等场景，如银河通用进入宁德时代、博世集团、丰田汽车等国内外制造企业生产线，承担自动质检、搬运与分拣等任务，通过真机作业回流数据形成自增强。智元在江西南昌龙旗科技园完成了流水线全闭环操作。

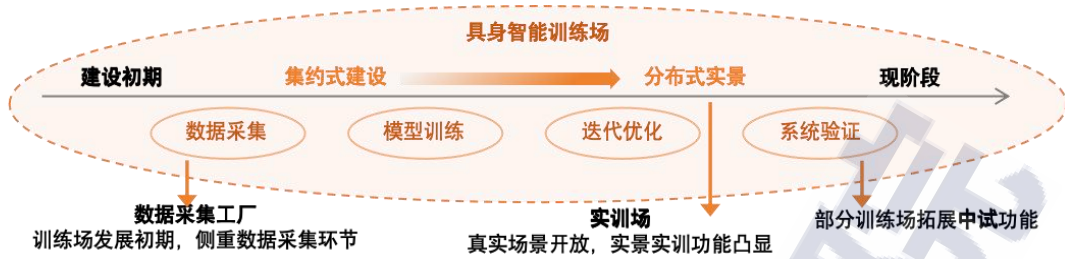
然而，分布式数据采集与真机回流仅能在数据规模上累积，仍面临数据质量和规模问题。**一是**分布式数据采集多以众包方式，雇佣非专业数据采集人员以摊薄成本，在人口规模大、物价工资水平低的地区招工，例如印度泰米尔纳德邦、古吉拉特邦多家纺织代工厂的工人兼职使用头戴式相机采集数据。这类方式关键问题在于采集地点分散、采集人员众多，短期的上岗培训难以保障数据采集员专业性，导致产出数据存在如动作不流畅、任务理解错误、场景杂乱、重复采集同一

场景等问题²¹，数据有效率低。**二是**当前阶段具身智能产品可靠性、商业部署规模、客户现场安全约束和数据回流机制仍不成熟，仅在小规模试点中产生回流，远不足以支撑行业规模落地。以工业和家庭场景为例，工厂流水线对节拍和可靠性要求极高，当前具身智能容错率远远不够；家庭场景则面临隐私、安全、非结构化环境与家务规则数字化等刚性难题。

（三）场景应用牵引，训练场成为当前闭环关键路径

真实物理交互数据积累仅是起点，数据要服务于模型研发和场景落地。具身智能训练场成为现阶段推动实景实训的关键路径。它是一类面向具身智能研发和应用搭建的关键基础设施，在真实物理环境中支撑具身智能数据生产、模型训练、系统应用能力验证和迭代优化。行业针对具身智能训练场有多种不同叫法，常见包括具身智能数据采集工厂（采集场、采集中心）、具身智能实训场等，本质上反映了训练场在不同发展阶段的功能定位和演进路径。具身智能训练场发展初期侧重数据采集，随着真实场景逐步开放、下游需求进一步释放，驱动训练场向集数据生产、模型训练、系统验证、迭代优化于一体的长链路综合性基础设施演进，形态上也从集约式建设向分布式真实场景拓展，实景实训功能凸显。此外，部分具身智能训练场能够拓展中试功能，验证具身智能在真实场景下的实际工作能力。

²¹ <https://mp.weixin.qq.com/s/uqImvthA6w4Y8mNBR6JVUg>



来源：中国信息通信研究院

图 1 具身智能训练场相关概念梳理

具体来看，具身智能训练场主要承载数据供给、模型研发和真实场景下的系统验证等共性技术服务，并通过场景开放与资源协同带动产业生态集聚。一是**高质量真实数据规模化供给**。训练场通过统一的人员培训、场景布置、任务流程、采集方式和质检规则，形成完整的数据生产流水线。训练场产生的数据能完整反映真实的物理接触、真实的不确定性、真实的失败和恢复。二是**高频次地开展模型验证、数据补采和模型迭代**。具身智能模型的能力提升不仅依赖离线训练，还需要在真实物理环境中反复执行任务、暴露失败、分析原因并定向补采数据。训练场可以围绕模型短板和任务难点，系统设置物体变化、位置变化、光照变化、人员干扰、指令变化、任务难度和异常情况等，形成“模型评测—失败分析—定向补采—再训练—再验证”的闭环。相较于真实客户现场，训练场允许更高频率、更低风险地复现失败场景和边界条件，从而加快模型迭代和能力提升。三是在**部署到真实场景前，开展能力验证**。具身智能产品最终要进入工厂、仓储、商超、家庭、康养等真实应用场景，但客户现场通常受到安全、隐私、生产节拍、运维责任和试错成本等约束，难以在早期阶段直接承载大规模训练和高频试错。训练场能够在进入客户现场之前，提前复刻典型行

业场景，对机器人开展任务成功率、执行效率、稳定性、鲁棒性、安全性和场景适配性验证，识别系统在真实应用中的失败模式和部署风险。**四是促进产业资源协同和生态集聚。**各类具身智能研发主体、数据采集服务商、模型开发团队、整机制造企业和场景需求方都可以依托训练场开展合作，打通从场景需求到技术落地的对接渠道，共享基础设施降低中小团队的研发门槛，推动产学研用各方资源协同，加速具身智能技术迭代和场景落地的整体进程。例如上海人形机器人麒麟训练场带动浦东张江模力社区汇聚超过 200 家具身智能、人工智能领域上下游企业，逐步形成“上下楼就是上下游”的产业生态。

国内外训练场模式在数据需求与场景落地驱动下快速发展，并在中国率先形成规模化建设。**国外方面**，训练场主要分布于美国、韩国、欧洲等国家和地区。例如韩国 LG 电子将其位于首尔江南区良才洞的研发园区，改造为家庭机器人真实环境训练中心²²，复刻真实家庭与工厂环境，部署 LG CLOiD 家用机器人采集物理世界操作数据。美国 Figure AI 公司依托自建 BotQ 工厂开展研发、数据采集、测试、数据回流。Apptronik 公司在得克萨斯州奥斯汀启动机器人园区²³，建设物流、制造、零售服务和家庭场景，用于人形机器人数据采集、训练和迭代，以及真实场景部署。Tutor Intelligence 在马萨诸塞州建设 Data Factory²⁴，设置 100 台 Sonny 机器人集中在一个场地里，用于真实世界数据采集、遥操作、纠错、评价和模型训练。德国慕尼黑工业大学

²²

<https://www.lg.com/global/newsroom/news/corporate/lg-electronics-establishes-robotics-business-center-to-accelerate-future-growth/>

²³ <http://apptronik.com/news-collection/welcome-to-robot-park-where-apptroniks-apollo-goes-to-work>

²⁴ <https://tutorintelligence.com/blog/building-a-100-robot-data-factory-toward-factory-ready-ai>

（TUM）²⁵与 NEURA Robotics 共建 RoboGym，采集工业、家庭场景人形机器人数据，训练人形机器人。

国内方面，我国在数据需求、政策牵引、场景开放和产业集聚的共同推动下，率先将训练场作为面向产业生态的基础设施开展规模化建设。早在 2024 年，具身智能企业如智元等已经开始自发主导建设大规模真实场景下的数据采集场地，聚焦构建面向产品迭代和商业落地的核心数据资产。2025 年 1 月，上海人形机器人麒麟训练场启用，成为全国首个异构人形机器人训练场，是集数据采集、训练、验证与流通于一体的具身智能基础设施。随后，在政策牵引、地方场景开放和产业集聚推动下，全国加快建设具身智能训练场，通过复刻典型应用场景、集中部署多类型机器人本体、统一采集流程和质检规则，推动真实数据供给从零散采集走向标准化、工程化、流水线式生产，目前已建、在建和规划中的训练场已超百家，初步形成以标准化数据生产为基础、以模型训练验证和场景适配为延伸的建设路径。

二、我国具身智能训练场建设情况

（一）政策牵引地方支持，各地训练场加速建设进程

具身智能训练场通过提供面向行业需求的一系列真实物理环境，为具身智能的真实数据采集、模型训练、系统应用能力验证和迭代优化提供有力支撑。近两年，在政策支持和产业需求的牵引下，我国训练场建设加速推进。

²⁵

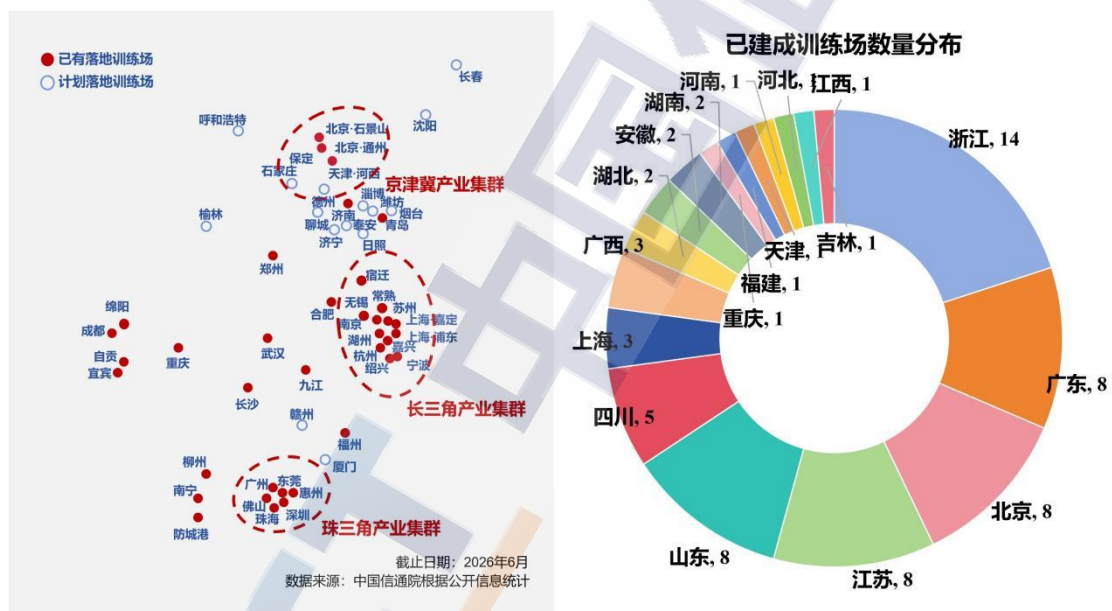
<https://www.tum.de/en/news-and-events/all-news/press-releases/details/partners-plan-worlds-largest-training-center-for-ai-powered-robotics>

国家层面，高度重视训练场建设。工业和信息化部等八部门在《“人工智能+制造”专项行动实施意见》²⁶中提到，“推动具身智能产品创新，**建设人形机器人中试基地和训练场**，打造人形机器人标杆产线，在典型制造场景率先应用”。工业和信息化部、国务院国资委联合印发的《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》²⁷中提出，“打造实景实训空间。围绕工业、服务、特种领域，聚焦生产制造、检测分析、维修维护、仓储物流、餐饮零售、医疗康养、安全生产、应急救援、防灾减灾等重点场景的人形机器人、四足机器人应用需求……**打造可训练、可测试、可验证的实景实训空间。**”国家发展改革委指出**加快具身智能方向应用中试基地建设**，健全完善具身智能软硬件生态，加强与训练基础设施的协同联动，加速面向应用落地的技术创新。国家数据局《关于推进行业高质量数据集建设行动的实施方案》指出，“顺应人工智能加速向行业渗透，从对话向多模态生成、决策执行、**具身智能**、物理交互等范式跃迁的趋势，拓宽数据供给渠道，丰富数据供给类型，**加快建设行业高质量数据集**”。**地方层面**，各地政府也积极出台相关政策，2024 年至 2026 年 6 月期间已发布超 40 份政策文件，覆盖北京、天津、河北、上海、江苏、浙江、广东、广西、福建、安徽、河南、山东、湖南、湖北、重庆、四川、吉林、黑龙江等 18 个省（自治区、直辖市），加大资金支持与要素保障力度，多措并举推进具身智能训练场建设，构建协同联动的训练基础设施体系。

²⁶ https://www.nda.gov.cn/sjj/zwgk/zcfb/0112/20260107214358696030895_pc.html

²⁷ https://www.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2026/art_f291ccd3da4c47ce95741de63cc088e6.html

据不完全统计，截至 2026 年 6 月底，全国建成启用超 70 家训练场，在建或规划中的训练场 46 家。从区域分布看，当前训练场建设已覆盖全国过半的省级行政区，形成以长三角、京津冀和珠三角为核心的三大集群。截至 2026 年 6 月，国内已启用和在建的训练场分布在浙江、广东、江苏、四川、北京、上海、广西、山东、湖北、天津、重庆、福建、安徽、湖南、河南、江西、河北、吉林等 18 个省（自治区、直辖市），如图 2 所示。训练场主要集中在长三角、京津冀和珠三角地区，和国内具身智能产业发展集群高度吻合。



来源：中国信息通信研究院

图 2 各省（市）已建成训练场的数量分布及统计

长三角地区训练场建设最密集，依托上海的龙头引领作用，带动江苏、浙江、安徽等地同步建设，产业配套完善、龙头企业集聚，如上海人形机器人麒麟训练场、智元机器人数据采集工厂、银河通用苏州具身智能数据采集中心、杭州市人形机器人中试基地、合肥市具身智能机器人数据采集预训练场等，总占地面积超过 7 万平方米。京津

冀地区依托北京的科研资源优势，形成了从基础研发到场景落地的完整推进链条，如北京石景山区梯度建设三期人形机器人训练场，总面积超 1 万平方米，数据年产量近千万级；天津帕西尼具身智能数据采集工厂面积达 1.2 万平方米，数据年产量预计数亿条。北京通州、北京朝阳、石家庄、保定等地也有训练场建设。**珠三角地区**依托庞大的制造业基础和应用场景需求在广州、佛山、深圳、珠海、东莞、惠州等地建设多家训练场，面向通用场景和垂域场景各自布局。如广州白云建科分训练场依托广州大学城与白云区产业优势，聚焦人形机器人与工业巡检场景开展实训示范；深圳玉树具身机器人智创中心面向清洁类场景垂域数据，广东佛山多模态大模型全品类机器人训练场，训练场总面积超 1 万平方米。**此外**，我国东北地区如吉林、辽宁，中西部如河南、山西，西南地区如四川、广西、重庆，以及东南沿海地区如福建，也陆续建设训练场，面向本地工业、餐饮、零售等产业需求，发挥场景优势。

在具体建设城市上，当前具身智能训练场**突破传统人工智能产业高度集聚于省会城市、直辖市和一线城市的发展模式**，不仅分布于北京、上海、深圳、成都等创新资源密集地区，也**逐步向三线、四线乃至五线城市延伸**，呈现出多层次城市协同布局的发展态势。这一布局特征反映出，训练场的选址逻辑正由传统的人才、资本积聚驱动，逐步转向以真实场景、数据供给和运营效率为核心的资源配置模式。**一方面**，三、四、五线城市拥有丰富的工业制造、物流仓储、商业服务、家庭生活等真实场景资源，场地租赁、场景改造及数据采集等运营成

本相对较低，能够为长期、规模化开展真实数据采集提供有利条件。例如广东东莞，作为全国制造业重地，已形成电子制造、模具注塑、家具制造、现代物流等特色产业集群，为具身智能提供了丰富场景与落地需求，已联合百度智能云推进具身智能数据采集场。另一方面，随着各地将具身智能纳入未来产业重点布局方向，地方政府积极通过产业基金、专项政策、场景开放、园区载体等方式招引训练场项目落地，希望以训练场为切入点，带动机器人研发测试、数据服务、装备制造等产业链上下游企业集聚，培育新的产业增长点。在产业需求与政策引导的共同作用下，训练场正逐步形成由核心城市向多层级城市扩散的发展格局，构建多区域、多场景真实具身智能数据基础设施。

（二）政企建设多方发力，综合行业双模式协同发展

当前具身智能训练场在建设主体、建设内容和组织模式上呈现政府、企业双线推进的鲜明特征，逐步形成“政府引导+企业主导”“综合平台+行业垂直”双轮驱动的发展格局。

从建设主体看，在产业需求与政策牵引驱动下，当前具身智能训练场主要有**政府引导、企业自建两类模式**。一是政府引导，提供场地、资金或政策支持，由具身智能企业、创新中心、科研院所或具身智能产业链上下游企业等联合建设运营。如上海人形机器人麒麟训练场由政府指导，国家地方共建人形机器人创新中心牵头建设。北京人形具身智能机器人数据采集与训练基地由国家与地方共建具身智能机器人创新中心建设运营，该机构由政府指导、多家企事业单位联合组建。二是由企业基于自身研发需求，或场景数字化转型需求，自主投资建

设。**一方面**，由具身智能模型、本体企业自建，这类训练场主要服务于企业自身的模型迭代和产品研发，部分企业也会逐步对外开放能力，向行业供给数据资源。如上海浦东智元数据采集工厂由智元机器人建设，构建家居、餐饮、工业、商超和办公等真实场景，除供给自身研发数据外，也对外开源 Agibot World²⁸系列开放场景真实数据集。京东机器人数据采集中心则由京东集团建设，覆盖物流仓储、工业制造、健康医疗、家庭服务、城市运维等多个核心场景，并于 2026 年 4 月开源时长 1680 小时的第一人称视角数据 EgoLive。**另一方面**，由场景数字化转型需求方与技术提供方联合建设，在自有场地、业务方向上探索面向场景的具身智能数据采集与能力适配。例如深圳玉树具身智能机器人智创中心，由玉禾田集团子公司深圳玉树智能机器人产业集团与智元机器人联合建设，围绕母公司长期深耕的环卫与物业保洁领域，设置捡拾垃圾、卫生保洁、家政服务，模拟家庭服务、城市公共设施维护等数十种场景。

从建设内容看，训练场呈现两大建设方向。一是**综合类训练场**，面向多行业场景应用需求搭建的具身智能训练场，支持跨场景泛化能力验证。二是**行业垂类训练场**，面向某一垂直行业场景应用需求的具身智能训练场，支持关键作业环节的能力训练和验证。据不完全统计，当前训练场场景建设在 3 类以上行业、多细分场景的综合类训练场占 60%以上，面向单行业，深入多场景的行业垂类训练场不足 20%。

综合类训练场建设主要受多样化数据需求和地方产业转型需求

²⁸ <https://agibot-world.cn/>

共同驱动。一方面，面向通用具身智能泛化能力需求，构建覆盖多场景、多任务、多品类的多样化数据产品。数据多样性越高、数据量越大，性能越好，而数据多样性比数据的绝对数量对提升模型泛化性更为重要²⁹。据不完全统计，同时建设包含“工业场景、家庭场景、商业零售”场景组合的训练场比例占全部综合类训练场的 44%，体现出训练场建设通用性数据平台的布局思路。另一方面，围绕本地支柱产业，开展训练场场景建设。多数大型综合类训练场由地方政策、产业集群和公共平台共同驱动，具有一定普惠属性和产业服务属性。地方政府建设训练场，不仅是为了支撑机器人企业采集数据和训练模型，也希望依托训练场服务本地制造、能源、港口、农业、医疗、文旅等多个支柱产业，推动具身智能技术与地方特色场景融合。例如广西防城港市人形机器人数据采集与测试中心，由防城港市投资集团联合优必选建设运营，部署了钢铜铝制造、电力配电、边境口岸、医疗医药等多个特色行业场景。广西东盟具身智能中试训练场，在广西壮族自治区科技厅指导下，由广西产业技术研究院联合银河通用等企业共建，围绕广西特色优势产业设置矿山矿洞巡检排险、农业作业、柔性制造等行业场景。

行业垂类训练场的建设主要受应用需求方牵引，是具身智能向具体产业深入渗透的重要载体。从当前建设情况看，行业垂类训练场的占比约在 18%，主要聚焦某一行业的典型任务、关键流程和高价值作业环节，建设模式包括：**一是**应用需求方作为场景合作方，开放行业

²⁹ <https://arxiv.org/abs/2410.18647>

场景，牵引训练场扎根行业关键环节。例如北京医康养机器人场景验证中心，由北京奥程运营管理有限公司与数泽动力科技（北京）有限公司合作共建，解放军总医院第一医学中心率先开放首批应用场景，覆盖术前宣教、术后康复等多个关键环节。**二是**应用需求方作为建设运营方，依托自身业务场景开展建设，将长期积累的行业场景、业务流程经验导入到训练场的建设中。例如青岛市康养领域具身机器人暨智能康复类产品训练验证中心，由青岛市民政局牵头，由民政部门提供真实康养场景资源，构建康养机构、医疗机构、社区养老、居家养老、园区服务五大类 11 个细分场景，进行失能失智预防筛查、残疾人和孤独症儿童康复训练、生活照料、精神慰藉、心理干预、情感陪护、预警响应、安全巡检、无人驾驶车辆协同等模块化科目训练³⁰。

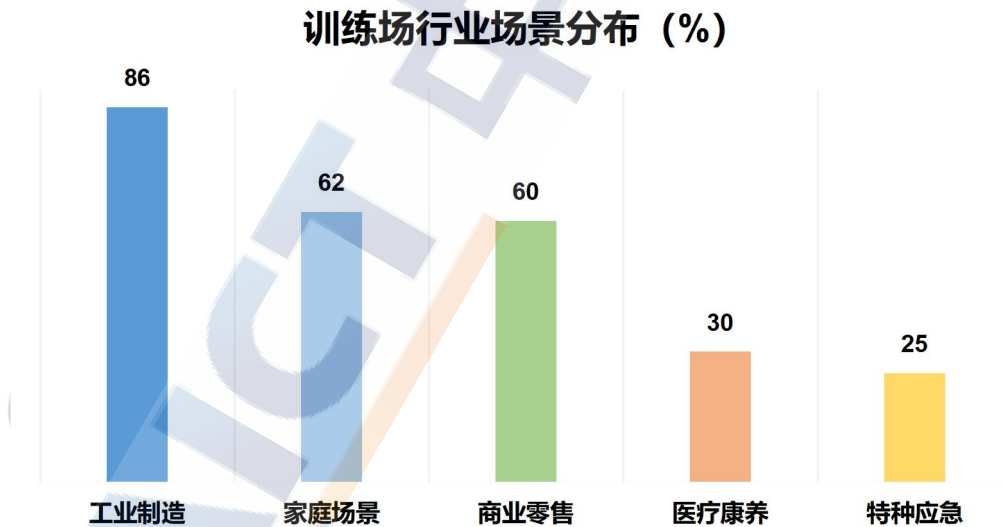
从组织模式上看，具身智能训练场向**区域协同、行业合作方协同**演进。**一是区域协同**，在全国各地协同建设训练场，汇聚数据实现统一能力与标准。例如上海人形机器人麒麟训练场与河南、江苏、浙江、北京、山东、陕西等地的分训练场，形成“1+N”区域协同模式，实现全国范围内的数据归集、标准统一和能力共享。广东提出建设“1+1+N”具身智能训练场体系，以省级综合训练场作为全省能力中枢，布局多个地市和行业分训练场，推动全省训练资源共享、数据互通和标准统一，形成覆盖全省的具身智能训练网络。**二是行业合作方协同**，以智元为例，以其建设在上海张江的数据采集工厂为中心，向全国多家合作训练场方输送训练场数据采集能力，统一回流采集数据，

³⁰ <https://mp.weixin.qq.com/s/ZLwCG78eZN6pmvqm4z1a1A>

实现行业合作方数据标准统一互通。

（三）应用需求建设导向，训练场行业场景呈梯次分布

场景建设将直接决定数据生产方向、能力训练重点和运营服务边界。当前，我国具身智能训练场建设总体呈现“以应用为牵引、以综合能力提升为目标”的特征，围绕工业制造、家庭服务、商业零售、医疗康养、特种应急等行业场景加快布局，形成由可控场景向开放场景、由高频任务向复杂任务、由通用操作向行业作业逐步拓展的梯次分布格局，如图 3 所示，主要建设工业制造、家庭场景、商业零售、医疗康养、特种应急五大类行业场景。其中含工业制造场景的训练场最多，占 86%，家庭服务和商业零售次之，建设最少的是特种应急行业场景，占比约 25%。此外，在农业、文旅等行业也有零星场景分布。



来源：中国信息通信研究院

图 3 已披露训练场场景分布

工业制造是当前训练场建设最集中的方向，主要涵盖汽车制造、家电制造、电池制造、3C 电子制造、半导体制造等细分场景，重点

训练物料分拣、搬运、装配、质检等作业能力。如北京石景山人形机器人训练场还原中兴的智能生产线、一汽集团的汽车车间等实际作业环境。江苏省具身智能机器人工业数据采集与实训中心，模拟工业工厂物料分拣、电动汽车插电充电功能检测及电池电芯装包等场景。其中，**物流仓储、物料搬运等任务布局最为集中，占已披露工业制造场景的近 40%**，已成为训练场面向工业制造场景开展数据采集和能力训练的重要切入点。京东具身智能数据采集中心依托自有仓库场景开展物流仓储数据采集。无锡市具身智能机器人工业数据采集与实训中心围绕场景合作方长城汽车的物料搬运任务建设数据采集能力。工业制造场景任务边界相对清晰、环境可控度较高、商业价值明确，既符合具身智能当前的技术成熟度，也具备率先规模化落地的应用基础，适合作为具身智能训练场建设的优先方向。

家庭服务和商业零售是训练场建设的重点拓展方向。家庭服务涉及**家庭住宅**场景，侧重卫生保洁、衣服整理、房间整理等作业任务。如深圳玉树具身机器人智创中心数十台机器人采集家庭场景中的倒水、擦桌子、捡垃圾、叠衣服等任务。这类场景物体种类多、空间布局变化大、任务链条较长，能够服务于提升机器人在非结构化环境中的感知理解、物体操作、任务规划和长程执行能力。**商业零售**涵盖**商超、餐厅、办公**等细分场景，聚焦导览、导购、制作咖啡、打包、送餐、拣货等作业环节。例如西南具身智能训练中心，采集商超场景的理货、补货数据，餐饮厂家的传菜、收餐数据，以及办公场景的访客接待、文件派送、桌面整理数据。相比工业制造场景，商业服务场景

的任务流程相对明确、用户交互频次更高、容错空间相对更大，适合作为具身智能产品从受控作业走向开放服务的重要过渡场景。随着多任务执行稳定性、环境适应性和通用操作能力持续提升，具身智能产品在商业服务场景下的市场价值有望进一步释放。

医疗康养和特种应急则代表更高难度、更高安全要求的前沿探索。

医疗康养涵盖医院、养老、康复等细分场景，聚焦健康监测、按摩、辅助诊断等任务。例如福州创新中心的医疗场景人形机器人训练场构建了高还原、多任务的训练环境，通过“大模型+机器人”协同，开展试剂分装、医疗垃圾分拣等训练。医疗康养类场景对安全性、可靠性和人机协作能力要求更高，同时商业闭环推进难度大，目前多面向**辅助、陪伴类任务建设场景**。**特种应急**主要涵盖野外、仓库、矿山、电厂等细分场景，聚焦运维、巡检、救援、监测等任务。如杭州市人形机器人中试基地的双浦机器人测试训练场，采用沙地、石子地、崎岖石块区等模拟野外，训练机器人复杂地形行走能力。广西东盟具身智能中试训练场（一期）建设矿山矿洞场景，采集巡检排险任务数据。**特种应急场景建设与政府需求推动高度相关**。高危险、人员替代价值高、人工进入风险大的特种应急场景，国家对代替人工的具身智能产品落地需求迫切。但该类场景环境还原难度大，场景建设成本高，数据采集危险性较高，因此整体布局数量最少，目前仅围绕巡检、安防等需求开展小规模场景测试建设。

（四）数据方案多元应用，工程化支撑训练场产能升级

训练场要产出大规模高质量数据产品，提升用户数据交易意愿，

从而实现商业落地，需要形成一套完善的数据工程化“流水线”闭环，从生产任务确定、数据采集、清洗标注到合规质检、存储管理，最后开展基于采集数据的训练与场景验证，并对数据生成任务进行反馈迭代。

一是任务下达，确定数据生产目标。一方面，根据数据使用目的，确定数据采集场景、任务及规模。科研用途通常对数据场景多样性、时长具有要求；真实场景落地用途通常不仅以数据时长为衡量目标，而是在具体场景中确定任务指标，以达成指标为数据最终目的。另一方面，根据具体任务确定数据采集方式、模态以及本体型号。高精度、长程复杂任务，适合通过遥操作方式采集，数据物理一致性强；精度要求不高的抓取、放置等任务，适合通过 UMI 或配套头戴式相机采集，满足数据快速积累、跨机型训练需求。对于长尾场景、危险工况或极端场景，可采用仿真合成方式生成。

二是真实数据采集，形成原始数据。采集员根据定义好的任务操作流程，重复执行多次采集任务。其难度不仅在于重复，而且需要操作标准有效。当前具身智能所需数据，在动作路径和执行速度等方面与人类日常操作仍有较大差异。例如，使用 UMI 采集叠衣服数据，必须有技巧地操作衣物，不能完全照搬人的动作，否则难以用于训练模型。《Business Insider》对特斯拉数采人员的采访报道中披露，员工采用遥操作、触觉手套或者头戴式传感设备执行汽车零部件整理、擦桌子等任务，有时一个简单的动作需要连续做几周，所有任务执行均有详细的操作手册指导。

三是数据清洗标注，形成有效数据。与传统语言模型不同，具身智能数据清洗标注要处理空间坐标、时间序列、动作轨迹、视觉数据、力触觉数据等多模态信息，不仅要进行时间戳同步，还要对动作进行标注。阿里巴巴千问团队 Qwen-RobotManip³¹指出，开源数据集的动作指令与状态趋势对齐不足，例如 RoboMIND UR 机械臂数据有 81% 没有通过对齐检验。**标注时**，根据训练场内制定的规则标注，如按照原子技能拆分、按照时间戳与语言描述拆分、按照谓词逻辑拆分等。Meta 主导的第一人称多模态数据集 Ego4D，标注员不使用固定词汇，而是按时间顺序进行事件描述，如“人类右手拿起菜刀”“人类左手稳住土豆，右手开始切片”。此外，训练场每日产出数据量在数百小时到数千小时不等，依靠小规模人力标注难以支撑，因此部分训练场寻求**众包标注**的解决方案。例如京东机器人数据采集中心可容纳近万人同步依托京东众智平台进行多模态数据标注。

四是数据合规质检，形成合格数据。有效数据并非合格数据，需要经过合规处理和质检等多重流程。**合规方面**，具身智能通过机器人等物理载体直接作用于现实物理世界，涉及的安全合规问题远多于传统人工智能。数据作为具身智能系统输入，涉及的人、事、物、环直接与具身智能合规相关。具身智能数据采集内容包含视频、声音及环境信息，可能涉及人员或场所隐私信息，同时部分场景的数据涉及企业核心生产流程，因此需要对关键内容进行脱敏处理。**质检方面**，训练场根据各自的数据质量评价规则进行数据筛选。例如筛除完全采集

³¹ <https://arxiv.org/abs/2606.17846>

失败的数据，但对动作不标准、采集操作存在偏差的数据，可直接保留用于训练，也可要求所有不符合操作标准的数据全部剔除。当前，业界逐步认识到“失败数据”对模型学习真实复杂物理世界规律的重要性，在质检环节回收失败数据以及失败前后的状态、人工接管及恢复记录等。

五是对数据开展存储管理，做到“存得下、查得快”。训练场进行真机数据采集，从机器人搭载的多目摄像头、激光雷达、IMU、力触觉传感器等设备记录多模态数据，单台机器人单次运行数据采集量可达到 GB 级，当训练场数十台机器人同时采集数据时，日均数据增量可突破 TB 级。因此**存储架构方面**，对于大规模、多模态具身智能数据，可通过多架构，构建“存储底座+高性能加速”分层结构，例如阿里云对象存储 OSS 与文件存储 CPFS 等多种存储组合，腾讯云以对象存储作为数据底座，在计算和业务之间构建高性能存储缓存层或热数据层。**数据组织方面**，当前多种方式并存，有以项目、数据类型、生命周期涉及目录层级统一管理的方式，有进行向量化与原始数据统一管理的方式，也有通过时序流、图结构或多模态特征统一表示的数据组织方式。

六是训练场内数据训练与场景验证。具身智能训练场的数据产品是具身智能落地闭环的中间环节，其最终价值仍然体现在真实场景应用侧。因此，需要在训练场内体系化建立数据应用闭环模拟，检验、迭代优化数据产品应用能力，才有可能在真实场景应用侧具备使用价值。利用生产数据及开源数据，根据预训练、后训练配方完成初步训

练，测试数据对模型的性能提升效果，再将模型部署到本体，进入训练场搭建的真实场景，在真机环境下验证具体任务指标，排查数据分布偏差、场景覆盖缺失等问题。将验证结果反馈到数据生产前端，帮助调整下一阶段数据生产的任务目标、采集标准和标注规则，形成“生产-验证-迭代”的完整闭环，从而匹配下游具身智能模型研发和产业落地需求。

（五）产业生态初步构建，训练场探索可持续商业模式

1. 产业链初步成型，带动上下游供需

训练场以数据-模型-本体-场景闭环为核心，正在成为连接具身智能技术研发与场景落地的重要产业枢纽。作为场景数据生产方和训练验证平台，训练场一方面将真实行业场景转化为可采集、可训练、可验证的数据资源，向模型企业、本体企业和科研机构提供数据产品、训练环境和测试验证服务；另一方面，训练场也将模型和本体能力反馈到具体行业任务中，推动具身智能系统围绕真实应用需求持续迭代优化。现阶段，训练场已初步形成对技术研发供给侧与场景应用需求侧的连接，为具身智能技术创新、产品中试和产业化落地提供关键基础支撑。

当前训练场产业链结构如图 4 所示。训练场上游主要由软件技术提供方、硬件技术提供方和基础支撑方构成。其中，硬件技术提供方，涵盖数据采集设备企业、传感器企业以及各类形态的具身智能本体企业等，为训练场提供数据采集工具以及训练、验证载体；软件技术提供方，涵盖具身智能模型算法企业、仿真合成平台企业，为训练场提

供数据仿真合成、训练与验证工具；基础支撑方，包括网络、通信、算力和存储等企业，以及训练场数据平台、运营平台等管理工具开发方，为训练场场地建设与运行提供所需的技术保障服务。

中游主要由训练场的建设运营方和行业场景合作方构成。行业场景合作方覆盖工业、家庭、商业、应急、医疗等各大行业，为训练场场景设计提供行业建议及需求，助力训练场建设。训练场建设运营方则承担平台建设及运营职责，一方面负责场地、软硬件设备以及基础设施建设，另一方面负责训练场的生产数据、模型训练、系统验证以及商业运营。

下游主要是具身智能领域的各类能力需求方，既包括模型企业、本体企业、高校、科研机构等技术研发主体，也包括工业、物流、商业、医疗、应急等行业应用主体。**对于技术研发主体**，训练场可提供真实物理交互数据、仿真合成数据、模型或本体代训平台，以及测试评估和验证服务；**对于行业应用主体**，训练场可依托已构建的真实场景环境，开展产品部署前的场景适配、任务验证、效果评估和风险筛查，降低具身智能产品进入真实场景的试错成本。

目前产业链各环节分工已初步清晰，整体呈现上游供给活跃，下游应用待拓展的特征。上游随着具身智能模型、本体、传感器、遥操作、仿真平台和算力基础设施快速发展，供给能力持续增强；**中游**正在围绕高价值场景合作、数据产品设计、训练验证服务和商业模式进行探索；**下游**应用需求虽在加快释放，但受限于高价值场景尚待挖掘、数据产品标准化和流通机制尚不成熟、训练验证结果与真实部署转化

链条尚未完全打通，整体仍相对薄弱。



来源：中国信息通信研究院

图 4 训练场产业链上下游分布

2.重资产特征明显，聚焦探索盈利模式

具身智能综合类训练场属于重资产投入，根据已披露的信息，数千平方米以上的训练场，其建设成本在数千万至上亿元。具体来说，训练场建设成本涵盖场景建设、数采及本体设备购买、算力建设、网络建设、存储能力建设以及各类管理系统构建等多方面，各个要素环环相扣。训练场建设的场景越丰富，配套的数据采集设备、本体等数量越多，对算力、网络、存储性能要求越高，系统性推高建设成本。

从建设成本结构看，本体采购是占比最高的单项支出。当前已披露信息中，大多数训练场部署本体规模在百台套，以部署 100 台人形机器人计算，按照当前市场价格（单台数十万元），仅本体采购成本就可达数千万元。数采设备费用在百万元级别，例如无本体数采设备单价数万元，光学动捕系统价格在数十万元等。场景建设费用在百万

至千万元级别，包括场地改造、场景道具、环境模拟设备等，费用根据预设场景的丰富度和还原度要求浮动较大，如果要 1:1 还原工业生产线、医疗护理病房、野外复杂环境等真实场景，场景改造和道具购置成本会进一步提升。例如户外环境的高低坡道调节系统，单价在数十万元；温度-湿度-震动综合环境试验设备，单价在数百万元。此外，还需考虑算力、网络、存储和管理系统、场地运维等建设和运营成本。算力配套成本在百万元级，若选择自建算力中心，仅硬件采购和机房改造百万元起步；长期租赁公有云算力，每年运营成本也会达到百万元级别。网络配套成本也在百万元以上。网络带宽、覆盖率以及基站建设数量，都需根据场地面积选择性建设。除建设成本投入外，训练场日常运营涉及团队人力、场地运维、设备维保、办公、水电等，年运营费用可达数百万元。

当前国内大多数训练场建设时间不足一年，整体处于起步阶段，商业模式尚未成熟。从已披露情况看，训练场的商业服务有数据产品出售、模型及本体训练、联合研发、测试验证服务等。数据产品出售是训练场最直接、最快形成收入的业务类型，包括成品数据出售，以及定制化产品出售等。从数据定价看，训练场数据受采集和标注复杂度、定制化程度等因素影响，价格差异较大。简单非定制化无本体数据采集费用在每小时百元左右，真机数据因质量与成本原因在每小时数百元到上千元，复杂数据则根据需求定价更高。以年产量 10 万小时数据、建设投入 5000 万元估算，如果是年产数据全部售出的最理想情况下，仅依靠数据产品出售业务，回本周期在 3 年以上。从用户

群体看，目前训练场采集数据主要面向国内的头部具身智能模型和本体企业，其次是面向高校及科研机构。此外，也有训练场达成对外数据出口业务。2026 年 5 月，帕西尼感知科技（天津）有限公司东疆综合保税区分公司遵循《中国（天津）自由贸易试验区数据出境负面清单管理办法（试行）》管理机制要求及相关法律法规规定，在确保业务全链条合规、安全、可控的前提下，将不含个人信息和重要数据的高价值数据产品推向国际市场³²。

三、我国具身智能训练场实践分析

（一）训练场核心建设要素与挑战

训练场的建设目标，不仅是提供具身智能数据采集场景，更是建设集数据生产、模型训练、系统验证和商业运营服务等关键核心环节于一体的综合性产业基础设施。训练场一方面通过“数据采集—模型训练—系统验证”回路提升技术能力，另一方面通过商业运营深入产业需求，对接下游真实痛点，将行业任务目标转化为可量化、可测试的指标，持续反馈推动产业落地。**场景**是训练场的物理载体和能力验证基础。高质量的训练场系统性复刻真实世界中的典型任务环境，并具备可配置、可扩展、可迭代的场景重构能力，为数据生产、系统验证提供多样化土壤。**数据**是训练场的核心产出，其生产流程贯穿采集、标注、合规、质检、存储、训练与验证全环节，形成标准化、工程化、可复用的数据资产体系。**训练验证**是训练场的核心功能，依托采集的规模化真实数据开展模型训练、系统验证与诊断，打通从供给数据到

³² <https://mp.weixin.qq.com/s/jeWxsLDxBmTdyKLCzYO4NA>

输出能力的转化路径。**商业运营**决定了训练场的可持续性。面对重资产投入与长回报周期的现实约束，综合运营能够提供多元化收入来源，支撑训练场稳定运转。**安全合规**是训练场运行的底线保障，贯穿场景建设、数据生产、训练验证与商业运营全过程，建立完善安全规范与合规体系，能够为训练场服务筑牢可信基础。当前，大多数训练场仍处于基础设施建设和数据产能爬坡阶段，产业服务效用尚未释放，训练场建设仍面临共性问题。

一是场景任务同质化，数据泛化价值有待提升。当前综合类训练场普遍布局工业、家庭和商业等多个行业场景，但每个行业内部都有大量作业环节尚未被充分挖掘。任务设计多集中在少数易采集、易展示、易标准化的任务上，导致同一行业场景内任务趋同。如，家庭场景中可覆盖衣物清洗、收纳整理、厨房备餐、卫生清洁、照护协助等多个环节，但当前仍集中采集叠衣服、叠毛巾、桌面整理等任务，仍有更多真实复杂环节尚待挖掘，例如“从洗衣机中取出成团湿衣物并展开晾晒”等长程高质量真机数据。同时，同一任务内部的数据分布多样性不足，如桌面整理任务往往只覆盖少量桌面类型、物体组合和摆放方式，难以反映真实应用中的多样变化，导致数据对模型泛化能力的支撑有限。

二是高质量数据供给不足，数据工程问题凸显。据不完全统计，当前已启用的训练场 2026 年数据年产量将累计达到千万小时级。但从实际生产过程看，训练场“可采集数据规模”并不等同于“可训练数据规模”，大量原始数据距离真正进入模型训练仍需经过清洗、对

齐、标注、质检、格式转换和适配等多道工程环节。每个环节都涉及不同的数据处理工具链和质量控制要求，任何一个环节效率不足或标准不统一，都可能造成数据流转卡点，导致数据虽然已经采集完成，却难以及时转化为可用、可训、可复用的数据产品。在任务定义环节，重生产规模、轻应用指标，缺乏以应用指导数据采集的产品研发思路。在数据采集环节，不同数据采集工具的传感器类型、精度、采样率、本体结构和控制接口存在差异，采集员的动作节奏和操作习惯也不完全一致，这些差异可能被模型学习为额外分布噪声。在清洗标注环节，不同企业对任务拆解、动作阶段划分、成功失败判定和异常情况记录的标准不一，导致不同来源数据难以直接融合。在合规质检环节，行业尚未形成针对具身智能数据的合规标准，缺乏统一的数据质量评价体系，失败样本是否保留、异常数据如何标记、低质量轨迹如何剔除，仍主要依赖各训练场自身规则，影响数据质量的一致性和可比性。部分训练场仍缺少对生产数据充分的训练、验证环节，将数据的落地交给下游模型企业。此类短链条的数据开环，一方面难以保障数据真实场景应用质量，造成资源浪费，另一方面难以形成长期运营能力，仅依靠数据出售，在大规模真机落地产生回流数据后，将失去自身竞争力。

三是数据难整合、难互通、难共享，阻碍模型训练与系统验证。跨本体方面，现阶段模型和本体难解耦，真实数据传感器类型、精度、采样率上存在差异，在一个数据集上训练好的模型部署到不同的本体型号时，性能表现会大打折扣，难以跨本体迁移。跨场景方面，不同

训练场对同一任务的定义、执行流程和评价标准缺乏统一规范，导致数据语义不一致，即便物理环境相似，也难以直接拼接或联合训练。**跨机构方面**，几乎每个具身智能创新企业都有自建的数据采集工厂，试图在重复“造轮子”中形成自己的数据方法论和技术壁垒，导致数据难以跨机构共享、训练和验证。数据构建链路中的方法、流程、动作标准等是企业最重要的经验沉淀，共享数据可能会失去对数据的掌控力，减弱其技术竞争优势。对于数据需求方来说，缺少事前验证数据质量与价值的手段和能力，也增加了数据采购风险，需通过第三方机构或平台帮助供需双方建立双边信任。

四是数据交易机制尚不成熟，商业价值转化仍受制约。平台方面，国内专门面向具身智能数据交易的专业平台仍较少，现有数据交易平台大多服务于通用数据交易，难以匹配客户对具身智能数据提前看样、质量检测、权限管控等方面的需求，数据供需对接效率低。**价格方面**，数据定价不透明。行业尚未形成统一的定价标准，不同训练场因数据采集方式、场景覆盖、标注质量差异较大，导致定价维度和区间差异明显，需求方难以快速对比选型，也增加了供需双方的交易沟通成本。**确权方面**，数据确权不明晰。不同场景涉及场地、本体、参与采集人员等多方主体，数据采集生成过程中权利归属划分不清晰，若后续产生知识产权纠纷，训练场可能面临潜在法律风险，同时也制约了数据的大规模自由交易。

（二）训练场发展阶段及典型路径

1. 训练场多阶段波动发展

当前具身智能技术本身仍处在快速迭代中，技术路线尚未定型，应用场景仍在探索，作为关键基础设施的训练场也随之呈现多阶段波动发展的特征，如图 5 所示。**第一阶段为建设热潮期**，训练场大规模建设，入驻大批量本体，但存在建设场景宽泛不深入、硬件本体堆砌、数据作坊式生产、产品开环缺乏验证、交易途径及确权不明晰等问题，导致训练场输出数据产品质量参差不齐，难流通、难共享。这一阶段虽处在发展初期，但数据工程化相对成熟的训练场，仍通过大规模数据生产积累了一批数据，推动如视觉语言动作（Vision Language Action, VLA）模型、世界动作模型（World Action Model, WAM）、世界模型（World Model, WM）等具身智能基础模型性能提升，同时发现数据缺口，为能力沉淀打下基础。**当前国内训练场正进入第二阶段能力沉淀期**。这一阶段，场景建设向综合类、行业垂类分化，部分训练场形成“1+N”训练场体系统筹规划；多种数据采集方式并行，数据工程体系初步成形，场内数据闭环带动场外应用闭环，数据产能与质量逐步提升；同时，探索多形式合规交易路径，逐步形成可持续运营模式。未来训练场将发展到**第三阶段生态繁荣期**，届时训练场标准化体系基本完善，分级分类建设格局成熟，数据工程化能力成熟，具备数据合规交易路径，充分链接行业，向行业输送数据、技术与人才。

三个阶段没有明确界限，由于具身智能技术快速迭代发展，且推动场景建设、数据工程化、商业运营路径闭环的动能不同，因此**未来训练场或将在能力沉淀中长期波动演进**。此外，由于当前训练场处于

建设热潮期向能力沉淀期过渡阶段，标准规范、商业模式和运营机制仍在完善过程中，未来可能有部分训练场因建设运营粗放，场景定位不清、数据产品能力不足，未形成良好业态，导致成本覆盖难、持续运营难等问题。这类训练场或不再追求全链条综合服务，而是进一步围绕训练场能力栈分化到多环节开展服务，例如提供数据采集工具链、模型本体测试等，通过缩短业务链条，做深局部能力，在产业链细分环节形成可持续闭环。



来源：中国信息通信研究院

图 5 具身智能训练场能力发展阶段图

2.面向核心要素推动建设完善

为支撑训练场从建设热潮向能力沉淀平稳过渡，训练场沿“场景差异化、数据工程化、数-训-推闭环化、商业多元化”路径有序建设。

（1）锚定特色场景需求，开展差异化设计

为避免场景任务同质化建设带来的任务广而不深，需在训练场建

设初期进行差异化定位。**一是锚定特色优势场景。**立足地区产业优势、建设运营方业务基础或场景合作方需求开展场景建设，有助于资源快速对接。例如京东具身智能数据采集中心，依托京东自身供应链优势，及其零售、物流、健康、工业、外卖、家政等真实业务，规划训练场建设场景及任务。深圳玉树具身机器人智创中心由玉禾田子公司玉树智能与智元机器人联合建设，围绕玉禾田深耕多年的服务业，聚焦擦洗台面、马桶保洁、叠放衣物、收拾垃圾、清洁桌面等多个立体清洁类关键作业环节，打造垂域数采场景。**二是探索构建多复杂度场景任务库。**围绕场景构建从基础任务到长程感知、决策、行动任务的场景任务库，从单一物品、固定位置、无干扰，逐步过渡到多物品、随机位置、动态遮挡等复杂条件，提升数据多样性与泛化引导能力。例如智元数据采集工厂，采用场景多维变体以及三级难度覆盖的方法体系，设置超过 200 个细分场景任务，以及超过 3000 类操作物品，实现复杂操作流程的标准化建模与多场景适应性增强。

（2）打通工程化流水线，减少数据全链路环节卡点

为以可控成本获得大规模高质量数据，需通过建设工程化数据生产体系，保障训练场稳定产出。**一方面**，统一生产各环节的平台规范。**针对数据采集**，统一定义传感器选型参数、采集动作要求、场景环境变量等内容，减少非任务本质的分布差异；**针对数据标注**，建立统一的任务拆解逻辑和标注规范，明确不同类型数据的标注粒度、格式要求；**针对数据质检**，分层级判定质量，明确合格数据、不合格数据的划分规则，从全链路保障不同批次、不同来源产出数据的一致性，提

升数据兼容性与可复用性，降低下游企业的二次处理成本。北京人形具身智能机器人数据采集与训练基地建立工业化数据生产流程，设置标准化采集、多层复核标注，以及动作完整性、标注准确性和场景泛化性“三审”质检，并将质检问题反馈回前端进行优化。另一方面，搭建自动化生产管线，引入自动化标注、智能质检工具，替代人工完成基础重复性工作，在提升生产效率的同时，稳定数据产出质量，实现数据生产的规模化、标准化。上海人形机器人麒麟训练场打造涵盖任务管理、数据采集、数据处理、数据标注、数据集管理全链路的数据平台，其人形机器人具身操作数据集入选国家数据局《高质量数据集典型案例》。智元依托一站式开发平台 Genie Studio，提供覆盖数据全周期解决方案，能够回流训练场行业合作方数据，并提供模型训练微调、仿真评测和真机部署验证能力。百度智能云具身智能数据采集服务则通过真实场景数据定义任务价值，仿真与合成数据补充长尾样本、高风险场景和低频失败案例，在规模、质量和成本之间寻找更优平衡。

（3）延长能力链条，发展多环节能力支撑点

具身智能训练场能力链条正在从数据生产单环节向“数据生产—模型训练—应用验证—反馈优化”闭环演化。多数具身智能训练场已具备真实世界数据采集、清洗、标注等基础能力，但整体仍以数据生产和交付为主要定位，模型训练、性能验证、部署评测及数据反馈优化等能力尚未全面覆盖，完整的数据闭环仍处于探索阶段。随着具身智能模型迭代速度不断加快，训练场的功能定位已从“数据采集中心”

向“具身智能基础设施”升级，将训练、验证、中试、评测等环节纳入统一平台，通过模型训练发现能力短板，以应用验证识别失效场景，再将问题反馈至数据采集与数据配方优化环节，形成持续迭代机制，实现数据价值和落地能力的协同提升。

（4）多路径探索运营，多元化对外服务

运营设计需解决覆盖重资产投入、实现可持续收益的核心问题。

一是分阶段推进变现，例如**短期**以数据采集为主要建设能力，同步依托政府产业补贴、研学培训、活动表演等降低前期投入压力。**中期**依托积累的数据资产和场景能力，规模化对外出售标准化数据产品，推进解决方案定制与联合研发服务。例如四川省人形机器人训练场探索文旅场景解决方案，与成都自然博物馆、哈尔滨冰雪大世界等景点合作，提供机器人讲解服务。**长期**向“训练即服务”推进数算训一体化转型，并发展面向场景落地能力。前者是引导用户在训练场内闭环，面向不具备自建场景训练场能力的模型企业开放场地、数采设备、本体、算力等全套资源，提供数据采集、数据清洗标注、模型训练调优、场景验证的全流程代训服务，按训练时长、算力使用量、数据使用量等方式收取服务费。后者是引导具身智能技术在真实行业场景闭环，可依托多品类真实场景环境，承接不同具身智能产品的落地适配测试，帮助企业在产品推向市场前完成场景适配打磨，降低真实场景落地的试错成本，收取场景验证与适配服务费用。**二是**构建多元收入结构，拓展中试、测试认证、人才培养等增值服务，对接具身智能产品上市前的验证需求以及产业端的人才实训需求，进一步挖掘场地、设备与

技术资源的潜在收益。例如北京石景山人形机器人数据训练中心，一期集核心技术研发、场景化应用测试、操作技能培训及生态合作于一体；二期构建了“训练—应用—孵化—科普”四位一体的生态模式。

（5）探索平台化交易方案，应对数据交易潜在问题

一是上线数据交易平台，开展境内数据交易。2026 年年初，全国首单“具身智能数据集”在江苏省数据交易所上架并完成交易，采购方为上海头部大模型企业，数据提供方为箬境智能科技。北京、贵阳等地数据交易所也陆续推动具身智能数据上架。**二是**推动建设数据超市，汇聚多训练场、多数据厂商产品，便捷选购。2026 年 4 月，京东发布京东云具身智能数据交易平台，在平台上明确数据类型、标识码、采集方式、场景、任务分级以及购买方使用权限，进一步避免确权争议。鹿明机器人推出 FastUMI Pro 数据超市，将操作数据按场景和具体任务分类，并透明定价。百度智能云联合多家具身智能企业推出具身智能数据超市。**三是**建设可跨境流通的数据平台。针对不同合规要求的数据探索构建跨境交易路径，针对脱敏后的通用场景数据推动合规出境，满足不同地区研发主体的数据需求，扩大盈利市场空间。例如北京石景山科技创新集团有限公司联合中国电信北京公司，推进合规框架下的具身智能数据要素跨境流通技术服务平台建设。

（三）训练场标准化实践

具身智能训练场逐步进入能力沉淀期，亟需标准化建设引导。具身智能训练场标准化需求主要包括共性基础要求、核心要素要求以及安全合规要求三大方面，如图 6 所示。



来源：中国信息通信研究院

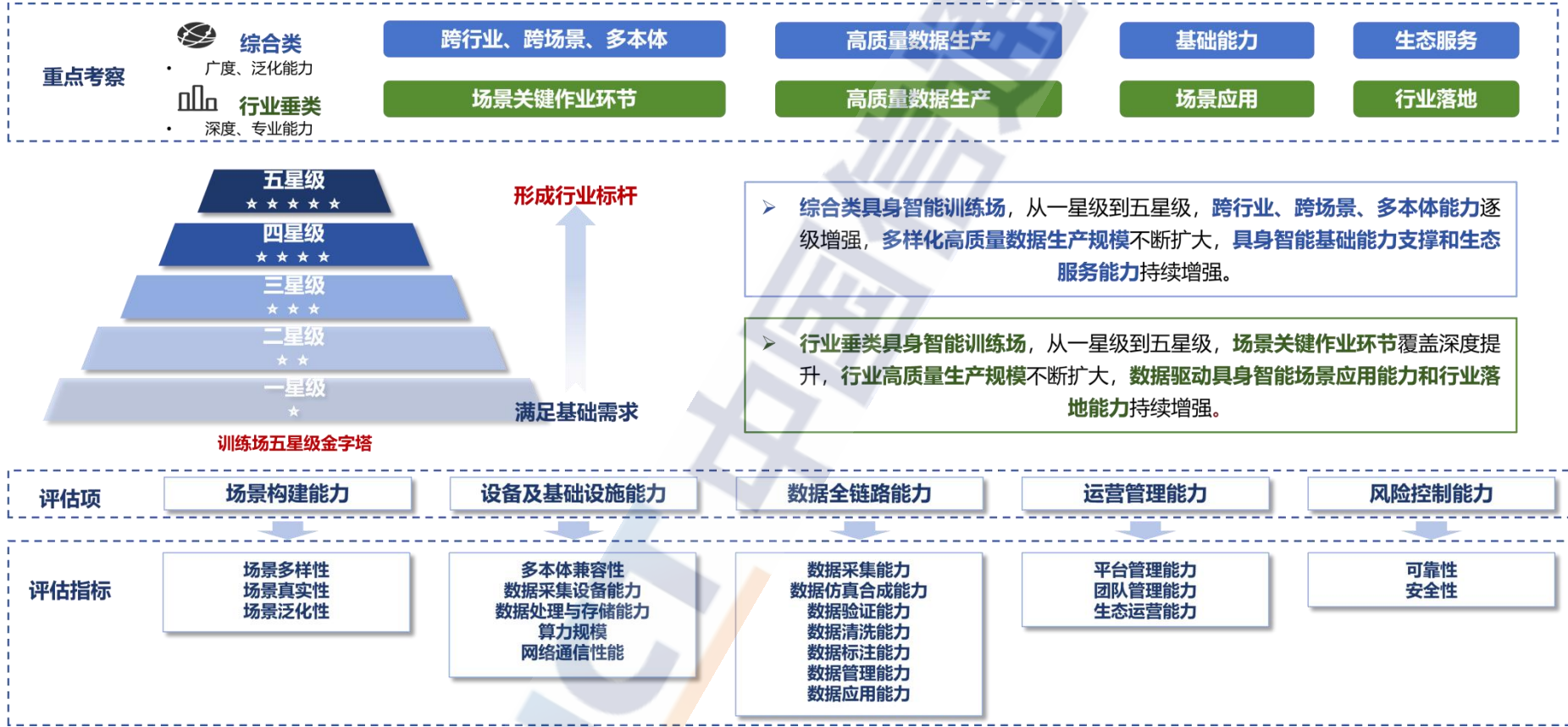
图 6 具身智能训练场标准化需求

一是场景建设标准化。当前各训练场对场景的类型、任务划分、环境要素、复杂度分级等内容标准不一，影响后续场景中数据采集的一致性，不利于数据流通共享。需面向不同应用场景，构建标准化场景建设规范、分类目录和任务定义方法，明确不同复杂度场景的环境要素划分规则，建立场景复杂度分级标准，引导数据生产有序开展，同时帮助数据需求方快速匹配适配自身训练需求的场景数据。二是数据生产标准化。当前产业在具身智能数据采集、数据质量评价、数据存储管理、数据确权和数据流通等方面缺乏统一规范，亟需推进标准建设，保障数据全链路一致性、兼容性和可复用性，降低跨主体数据融合和二次处理成本。三是模型训练服务标准化。亟需对场内训练环境、训练任务、模型以及硬件适配进行标准化，方便不同训练主体快速接入使用，减少训练环境适配的额外成本，提升训练资源利用效率。

四是中试服务标准化。延长训练场服务链条，拓展中试验证服务，亟需对系统集成、工程验证、小批量一致性验证等服务进行规范，形成可量化、可复现、可比对的中试服务标准，打通训练场“数-算-训-用”全流程。**五是训练场建设运营标准化。**亟需明确训练场综合建设、管理、运营规范，建立分类分级与评估规则，规范化统筹多类型、多层级训练场的多维度服务。**六是安全合规标准化。**物理安全、系统安全、数据安全、运营安全等方面需形成覆盖训练场全生命周期的安全管理体系，明确各环节责任边界与防护要求，制定可落地的安全技术标准与合规操作指引。

面向具身智能训练场标准化的多维度需求，国内外标准化组织均已在多维度部署推进。**国际方面**，2026 年 2 月 19 日，标准化组织 ITU-T SG21 通过具身智能焦点组（FG-EAI）WG2 数据组，推动统一数据格式规范与使用准则。SG21 同步推进数据采集标准化，包括面向真实场景的具身智能遥操作数据采集，以及具身智能系统多模态数据采集等。**行业方面**，依托工业和信息化部人工智能标准化委员会 MIIT/TC1 具身智能组，推动发布首个具身智能数据集质量行业标准 YD/T 6771-2026《人工智能 关键基础技术 具身智能数据集质量要求及评价方法》，并推动 2026-0074T-YD《人工智能 关键基础技术 具身智能训练场综合评估技术规范》标准研制。两份行业标准均同步在人工智能产业发展联盟（AIIA）发布并推进先行先试。此外，AIIA 也在联合产业力量，推进遥操作数据采集系统、无本体数据采集系统等具身智能数据采集设备规范化。

2026-0074T-YD《人工智能 关键基础技术 具身智能训练场综合评估技术规范》框架如图 7 所示，聚焦训练场在场景建设、设备与基础设施、数据全链路、运营管理和风险控制等核心要素，对现阶段形成的综合类训练场与行业垂类训练场开展针对性评估。对于综合类训练场，考察其跨行业、跨场景、多本体泛化能力，以及数据产出质量，同时对提升具身智能通用基础能力、服务行业生态的效果开展评估，以保障综合类训练场能够横向支撑具身智能行业；对于行业垂类训练场，考察其深入行业场景关键作业环节的数据采-训-用能力，以保障行业垂类训练场能够纵向支撑特定行业。同时，针对两类训练场各方面能力进行五星级评级，以帮助各训练场明确自身能力及定位，旨在以评促建，引导训练场规范化建设，为具身智能训练场发展提供参考依据，助力产业健康有序发展。



来源：中国信息通信研究院

图 7 2026-0074T-YD《人工智能 关键基础技术 具身智能训练场综合评估技术规范》框架

四、具身智能训练场未来展望

具身智能训练场是具身智能发展过程中涌现的新型基础设施形态。随着具身智能技术逐步成熟、产业应用不断深化，训练场作为支撑具身智能数据生产、模型训练与系统验证的关键基础设施，将在场景、数据、服务、运营、标准化等多维度持续演进，推动我国训练场从建设热潮期向能力沉淀期稳步过渡，呈现五大发展趋势。

一是场景方面，实景实训成为当前关键特征。训练场建设正从集约式建设向分布式实景发展，在真实场景中铺开训练场业务。集中式场地虽便于统一管理，但难以完整覆盖真实环境的复杂性与多样性，数据的真实性和任务的泛化价值受到制约。未来训练场将走出封闭场地，把数据采集、模型训练与系统验证能力下沉到工厂生产线、家庭住宅、商业门店、医疗康养等真实场景之中，形成“中心训练场+分布式实景点位”的协同布局，既保障数据与真实应用需求的一致性，也有助于摊薄重资产投入、贴近行业一线需求。

二是数据方面，数据工程化成为必然趋势。数据全链路向工具化、自动化方向发展，人工密集投入的数据生产方式难以为继，任务下达、数据采集、清洗标注、合规质检、存储管理、训练验证等全链路环节将加快平台化、工具化和自动化改造，以标准化流水线替代分散人工作业，在可控成本下稳定产出大规模、高质量数据产品，为训练场产能升级和商业落地提供工程化支撑。

三是交易流通方面，合规化是必经之路。数据出售业务作为当前训练场的快速变现业务之一，其可持续发展离不开合规化交易体系的

支撑。训练场需在数据确权、定价评估、隐私保护、安全管理、场内交易等环节建立健全合规机制，积极依托数据交易所等规范渠道开展数据流通，明确数据权责边界、防范法律与安全风险，合规化交易在一定程度上能够保障训练场稳定运营，推动数据要素价值有序释放。

四是服务方面，一体化是训练场发展的长期目标。训练场将探索延长服务链条，从单一的数据供给向“训练即服务”转型，对外提供覆盖数据生产、模型训练、应用验证一体化服务能力；同时依托既有场地与技术资源，拓展中试、测试认证、人才实训等增值服务，挖掘场地和技术资源潜在价值，构建多元收入结构，增强可持续运营能力。

五是标准方面，全周期标准化是训练场走向成熟、推广经验、增加影响力的关键。随着训练场从建设热潮步入能力沉淀期，将围绕场景建设、数据生产、模型与整机服务、中试、安全合规、运营与评估等环节，逐步建立覆盖训练场全生命周期的标准体系，统一技术要求与评估口径，沉淀可复制、可推广的建设运营经验，以标准化引领训练场规范高效发展，提升我国具身智能基础设施建设话语权与影响力。具身智能训练场的发展将始终与具身智能产业的落地进程同频共振，也将持续为技术突破和产业落地托底。短期来看，训练场建设可能会在路径探索中经历些许波折与调整；长期来看，这套依托物理场景沉淀高质量数据、支撑技术迭代验证的基础设施模式，将会逐步走向成熟，持续为具身智能技术迭代提供真实场景养分，加速推动具身智能从技术验证概念走向大规模产业应用，为人工智能与实体经济的深度融合注入强劲发展动力。

编制说明

本研究报告自 2026 年 4 月启动编制，分为前期研究、框架设计、文稿起草、征求意见和修改完善五个阶段，面向具身智能训练场软硬件技术提供方、基础支撑技术提供方、建设方、运营方、场景合作方、服务应用方开展了深度访谈和问卷调查等工作。本报告由中国信息通信研究院人工智能研究所联合人形机器人（上海）有限公司、具身智能测试实验室牵头撰写。

参编单位：具身智能测试实验室长三角中心、具身智能测试实验室（北京朝阳）、中国人工智能产业发展联盟、北京人形机器人创新中心有限公司、北京中关村石景山园发展有限公司、北京银河通用机器人股份有限公司、安徽聆动通用机器人科技有限公司、智元创新（上海）科技股份有限公司、上海觅蜂具身智能科技有限公司。

中国信息通信研究院 人工智能研究所

地址：北京市海淀区花园北路 52 号

邮编：100191

电话：010-62308338

传真：010-62308338

网址：www.caict.ac.cn

