

TCAR

机器人检测认证联盟认证技术规范

CR-3-0603TS:2025

人形机器人数据和标注质量评估及互联互通技术规范

Evaluation of Humanoid Robot Data and Annotation Quality and Interconnection
Technology Specifications

2025-09-16 发布

2025-09-16 实施

机器人检测认证联盟发布

目 录

前 言	III
引 言	V
1 范围	6
2 规范性引用文件	6
3 术语、定义	6
4 应用场景	6
5 数据集质量要求	6
5.1 基础性能	6
5.2 数据来源要求	7
6 数据标注要求	7
6.1 数据标注任务及说明文档要求	7
6.2 数据标注工具要求	7
7 数据互联互通	7
7.1 统一数据模型	8
7.2 元数据规范	8
7.3 数据经纬架构要求	8
7.4 接口与协议	9
7.5 数据集全生命周期管理与角色权限	9
8 数据质量评价方法	9
8.1 基础性能评价	9
8.2 多样性评价	9
8.3 数据来源要求评价	10
9 数据标注评测方法	10
9.1 评测流程概述	10
9.2 评测方法	11
9.3 评测报告与质量控制	11
10 数据互联互通评价方法	11
10.1 评价流程	11
10.2 评价测试	12
参 考 文 献	13

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件起草单位：上海机器人产业技术研究院、睿尔曼智能科技（北京）有限公司、江淮前沿技术协同创新中心、中国科学院上海微系统与信息技术研究所、上海电器科学研究所（集团）有限公司、上海电器设备检测所有限公司、上海添唯认证有限公司、中国信息通信研究院、北京人形机器人创新中心有限公司、人形机器人(上海)有限公司、中汽检测技术有限公司、芜湖赛宝机器人产业技术研究院有限公司、中国软件评测中心、中国科学院沈阳自动化研究所、杭州宇树科技有限公司、北京小米机器人公司、深圳市优必选科技股份有限公司、五八智能科技（杭州）有限公司、上海智元创新技术有限公司、上海库帕思科技有限公司、上汽通用动力科技(上海)有限公司、节卡机器人股份有限公司、上海ABB工程有限公司、中晶新源（上海）半导体有限公司、中国电子科技集团公司第二十一研究所、中移（杭州）信息技术有限公司、深圳市人工智能与机器人研究院、上海电气集团股份有限公司中央研究院、上汽集团北京有限公司、北京工业大学、上海工业控制安全创新科技有限公司、具身智人（北京）科技有限公司、四川具身人形机器人科技有限公司、合肥科大智能机器人技术有限公司、智昌科技集团股份有限公司、科沃斯机器人股份有限公司、苏州泰斯特智能科技有限公司、上海思岚科技有限公司、苏州大学、中国计量大学、工业和信息化部电子第五研究所、江苏省产品质量监督检验研究院、福建省特种设备检验研究院、上海仪器仪表自控系统检验测试所有限公司、北京市产品质量监督检验研究院、威凯检测技术有限公司、常州检验检测标准认证研究院、浙江华电器材检测研究院有限公司、中数（深圳）时代科技有限公司、UL美华认证有限公司、上海识度数字科技有限公司、杭州汉德质量认证服务有限公司、深圳市北测检测技术有限公司。

本文件主要起草人：郑军奇、朱晓鹏、郑随兵、友清、李强、曹天翔、朱滨，李一朵，刘亚楠、朱晓俊、熊友军、刘宇飞、辛强、陈仟、卫能、李峰、梁学修、李志海、任容玮、全琪、王启舟、王俊芳、郭文平、陈利钦、梁乔玲、张国亮、闫维新、曹宇、徐啸顺、许雄、李顺冲、张文敏、张楠、王洪武、蒲琪然、刘佳璐、陈忠、李博程、石照耀、李嘉茂、张璇、刘宇龙、冯振宇、汪中原、甘中学、严海、宋伟、安伯廷、张丽萍、陈良、张斌、董成举、陈胜利、陈照春、江旭强、李海涛、张传甲、王直荣、张浩、王光、章韧、张弋、郑威、宛江坤、陆迪康。

本文件为首次发布。

引 言

人形机器人作为人工智能与机器人技术深度融合的典范，正逐步从实验室走向产业化应用，其发展高度依赖于高质量、大规模、多模态的训练数据。数据的质量直接决定了机器人感知、认知、决策与执行能力的上限，是实现其智能化、通用化的关键基石。然而，当前人形机器人领域在数据的采集、标注、评估、管理及交换等方面仍缺乏统一的标准规范，导致数据质量参差不齐、格式各异、互操作性差，形成了严重的数据孤岛，制约了技术的快速迭代与产业协同发展。

为破解这一瓶颈，引导产业健康有序发展，特制定《人形机器人数据质量标注评估及互联互通规范》。本文件旨在建立一套科学、系统、可操作的标准体系，规定人形机器人数据集在质量维度、标注要求、评估方法、互联互通的数据模型与接口协议以及全生命周期管理等方面的统一要求。本规范聚焦于从原始数据到训练数据的转化过程，涵盖语音、视觉、触觉、力觉、本体感知等多模态数据，致力于推动数据资源的高效生产、质量评估、可信流通与协同利用，为人形机器人技术的创新突破和规模化应用提供坚实的数据基础支撑。

TCAR

人形机器人数据和标注质量评估及互联互通技术规范

1 范围

本文件规定人形机器人数据集的质量维度、评估方法、互联互通的数据模型、接口与协议，以及全生命周期管理要求。

本文件规定人形机器人原始数据到训练数据的标注任务分类、标注内容、质量评估方法、工具平台要求及自动化标注实践，适用于语音、视觉、触觉等多模态数据。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 36344-2018 信息技术 数据质量评价指标
T/SAIAS 025-2025 人形机器人数据集数据标注规范
T/SAIAS 026-2025 人形机器人数据集格式要求
T/SAIAS 027-2025 人形机器人数据集质量评价
T/CESA 1040-2019 信息技术 人工智能 面向机器学习的数据标注规程

3 术语、定义

下列术语和定义适用于本文件。

3.1

数据质量 data quality

在指定条件下使用时，数据的特性满足明确的和隐含的要求的程度。

[来源：GB/T 36344-2018 信息技术 数据质量评价指标 2.3]

3.2

标注任务 annotation task

按照数据标注规范对指定数据集进行标注的过程。

[来源：T/CESA 1040-2019 信息技术 人工智能 面向机器学习的数据标注规程 3.3]

4 应用场景

主要应用场景包括但不限于：

- a) 具身智能大模型训练与微调：为训练参数量 $\geq 10B$ （100亿）级别的具身智能大模型提供数据支撑。单个模型训练周期需处理 $\geq 10000h$ 的高质量、时序对齐的多模态标注数据（包括但不限于RGB图像、深度信息、点云、力觉、音频、本体感知数据等）。这些数据用于训练模型理解物理规则、生成复杂技能链、实现与环境的安全与柔性交互。
- b) 仿真环境构建与验证：为高拟真度的物理仿真环境（如Isaac Sim, Mujoco等）提供真实世界的行为数据与物理参数，用于进行大规模的安全测试、算法验证和“模拟到现实”的迁移学习。

5 数据集质量要求

5.1 基础性能

基础性能包括完整性、一致性、可访问性、安全性、应符合T/SAIAS 027—2025的规定。

5.1.1 多样性

多样性指的是数据集在多个维度上的丰富性和覆盖广度，是模型获得良好泛化能力的关键。

5.1.2 环境多样性

数据集所具备的环境多样性应满足以下要求：

- a) 光照条件：应覆盖 ≥ 5 类，包括但不限于：晴天自然光、阴天漫射光、夜间室内光、强光直射（如逆光）、弱光环境。
- b) 背景场景：应覆盖 ≥ 5 类，包括但不限于：家庭、办公室、商场、户外、仓库等多种结构化与非结构化场景。

5.1.3 对象多样性

物体类别应覆盖 ≥ 200 类日常物品。物体材质必须包含刚性（金属、塑料）、柔性（织物、玩偶）、透明（玻璃杯）、反光（镜面、不锈钢）等具有挑战性的材质，且每类材质至少有数十个实例。应覆盖静态操作、短时序动作和长时序复杂任务链。

5.2 数据来源要求

5.2.1 数据集数据应来源于数据集采集系统及数据采集训练场。

5.2.2 数据集采集系统宜符合《人形机器人数据集采集系统通用要求》。

5.2.3 数据采集训练场宜符合《人形机器人数据采集训练场技术规范》。

6 数据标注要求

6.1 数据标注任务及说明文档要求

数据标注任务分类及标注任务描述、图像数据标注、视频数据标注、听觉数据标注、运动学数据标注、动力学数据标注、数据标注的准确性、一致性、时效性、可访问性、可移植性、保密性、可追溯性均应符合 T/SAIAS 025—2025 的规定。

6.2 数据标注工具要求

6.2.1 一般要求

标注工具一般要求应符合T/SAIAS 025—2025的规定。

6.2.2 选取原则

标注工具选取原则应考虑以下内容：

- 1) 数据集具体应用场景。
- 2) 标注任务类型。
- 3) 标注数据格式。
- 4) 满足功能目的要求，经使用证明、验证或确认的工具，常见工具见T/SAIAS 025—2025规定中的附录B。

6.2.3 安全要求

应确保标注工具和平台的安全，要求应符合T/SAIAS 025—2025的规定。

6.2.4 标注环境

标注实施方给出标注所需要的软件环境信息，如网络带宽、数据存储、数据计算等。

7 数据互联互通

7.1 统一数据模型

为消除数据孤岛，所有数据应遵循一个统一的核心数据模型。该模型定义五大核心实体及其关系。

7.1.1 机器人状态 (RobotState)

机器人状态是描述机器人本体实时状态的数据实体，其技术规格包含关节角度、关节扭矩、末端执行器位姿、电池电量、系统状态码等。所有位姿数据必须基于统一的坐标系框架。

7.1.2 传感器数据 (SensorData)

传感器数据是描述各类原始传感器数据的采集信息，其技术规格包含数据本身（如图像帧、点云、音频波形、力觉读数）、传感器内在参数（如相机内参、畸变系数、麦克风阵列几何）、时间戳（同步到高精度时钟）、坐标系信息。

7.1.3 任务上下文 (TaskContext)

任务上下文用于描述机器人正在执行或即将执行的任务的抽象信息，其技术规格包含任务ID、高级语言指令（如“泡一杯茶”）、当前任务步骤、成功/失败状态、任务参数等。

7.1.4 环境信息 (Environment)

环境信息用于描述机器人所处的静态和动态环境信息，其技术规格包含预先构建的静态地图（如OctoMap）、检测到的动态物体列表及其属性、环境温度、光照强度等。

7.1.5 交互对象 (InteractionObject)

交互对象用于描述与机器人发生交互的物体信息，其技术规格包含物体类别ID、名称、物理属性（质量、摩擦系数）、当前位姿、抓取点集合、3D模型URI等。

7.2 元数据规范

元数据规范格式应符合T/SAIAS 026—2025的规定。

7.3 数据经纬架构要求

为应对大规模数据挑战，建议采用数据网格去中心化架构，将数据作为产品进行管理。

7.3.1 数据资产目录

必须建立全局统一的数据资产目录，对所有数据集进行编目。每个数据产品应被打上丰富的标签，包括：

- a) 时间标签：数据采集时间戳、更新频率、有效期。
- b) 空间标签：采集地点、场景类型、机器人位置、环境栅格地图编码。
- c) 业务标签：关联的任务等级、涉及的技能（如抓取、导航）、标注状态、质量等级。

7.3.2 权限中心

7.3.2.1 差分隐私

通过数学框架保护数据隐私的技术，其核心目标是确保在数据分析或统计结果中，无法通过任何背景知识推断出个体的具体信息。通过添加受控噪声实现隐私保护。

7.3.3 数据流通市场

为促进数据要素流通，应建立内部或公开的数据市场。

7.3.3.1 API 接口定义：

- a) 数据查询API：遵循RESTful架构，响应格式为JSON或ProtoBuf。
- b) 流式传输API：遵循gRPC-streaming，支持实时传感器数据流。

7.3.3.2 订阅规则

订阅规则包含以下内容：

- a) 事件驱动：数据更新时推送（环境变更>5%触发告警）。

- b) 周期推送：按固定间隔（每5分钟发送机器人状态快照）。
- c) 配额限制：单主体最大订阅流数 ≤ 100 路（防资源滥用）。

7.3.3.3 计费模型

计费模型包含以下内容：

- a) 按数据量计费：按时长（小时）计价；
- b) 按价值密度计费：高价值数据（如成功抓取轨迹）溢价系数1.2~1.5倍；
- c) 按数据标注精细度计费：高质量人工标注数据溢价系数1.2~1.5倍；
- d) 按调用频次计费：API请求次数。

7.3.4 质量监控

7.3.4.1 数据产品 SLA（服务等级协议）：每个数据产品应明确定义其 SLA，包括可用性（如 99.9%）、数据新鲜度、准确率下限等承诺。

7.3.4.2 异常告警：建立监控系统，对数据管道的中断、SLA 违约、异常访问行为等进行实时监测和告警。

7.4 接口与协议

为实现自动化数据交互，需规定标准接口与通信协议：

- a) 查询接口：HTTP协议
示例：`GET/api/v2/dataset?robot_id=xxx&scene=home`
- b) 订阅接口：WebSocket
示例：`npxwscat-c" wss://data.hri.org/stream"-H"Authorization:Bearer<YOUR_JWT>"`
- c) 协议栈：ROS2。

7.5 数据集全生命周期管理与角色权限

7.5.1 数据生命周期

数据集应遵循以下生命周期阶段进行管理：

- a) 采集：根据规范进行原始数据采集。
- b) 标注与质检：依据第5、6部分进行标注和质量检验。
- c) 发布：打上元数据标签，发布到数据资产目录和市场。
- d) 版本控制：任何变更必须生成新版本，保留历史版本以供追溯和模型复现。
- e) 退役：对过期或无效数据进行归档或安全删除。

7.5.2 角色权限矩阵

应对不同角色分配其在数据全生命周期中的最小必要权限：

- a) 数据工程师：拥有最高权限，负责数据管道的搭建、维护和生命周期管理。
- b) 标注员：权限仅限于被分配的任务数据的上传和标注，无法访问其他数据。
- c) 算法工程师：权限为查询数据目录、订阅和下载已授权的数据产品，无法修改原始数据。
- d) 质检员：权限为对标注结果进行审核和打回，无法修改原始数据。
- e) 项目经理：权限为查看项目整体数据进度和质量报告，无法下载原始数据。
- f) 外部用户：权限仅限于在数据市场上浏览公开的元数据信息，购买后方可获得下载权限。所有数据访问行为均需被审计日志记录。

8 数据质量评价方法

8.1 基础性能评价

基础性能评价包括完整性、一致性、可访问性、安全性评价，其应符合T/SAIAS 027—2025的规定。

8.2 多样性评价

8.2.1 环境多样性

8.2.1.1 评价指标

评价指标包含以下内容：

- a) 数据集实际覆盖的光照类型数量：应 ≥ 5 类。
- b) 数据集实际覆盖的背景场景类型数量：应 ≥ 5 类。

8.2.1.2 评价方法

评价方法包含以下内容：

a) 元数据初筛

从数据集元数据（需符合T/SAIAS 026—2025）中提取“光照类型标签”（如“晴天自然光”“夜间室内光”）、“场景类型标签”（如“家庭”“商场”），统计初步分类结果，判断是否存在明显缺失的光照/场景类别。

b) 分层抽样检验

按“光照类型-场景类型”交叉维度分层随机抽样，抽样比例不低于数据集总样本量的10%，每个光照/场景类别抽取样本数不少于50个以确保统计显著性。

8.2.2 对象多样性

8.2.2.1 评价指标

评价指标包含以下内容：

- a) 物体类别覆盖度：数据集实际覆盖的日常物品类别数量“ ≥ 200 类”的比例。
- b) 材质覆盖完整性：数据集对“刚性、柔性、透明、反光”四类材质的覆盖情况及标注准确性。
- c) 任务时序覆盖度：数据集对“静态操作、短时序动作、长时序复杂任务链”三类时序任务的覆盖比例。

8.2.2.2 评价方法

评价方法包含以下内容：

a) 物体类别核验

- 1) 元数据提取：从数据集标注文件中读取物体类别名称，统计初步类别数量。
- 2) 抽样验证：按“类别均匀抽样”原则，每类物体抽取20个实例，使用目标检测模型辅助人工核验物体类别。

b) 材质覆盖核验

- 1) 材质分类确认：从数据集标注中提取材质标签，确认是否覆盖“刚性、柔性、透明、反光”四类。
- 2) 实例数量核查：统计每类材质的实例数，确保每类不少于30个。

c) 任务时序核验

- 1) 任务类型划分：根据数据集任务描述，将样本划分为“静态操作”（如单一物体抓取定位）、“短时序动作”（如开门/拧瓶盖，时长 $< 30s$ ）、“长时序复杂任务链”（如泡茶/整理桌面，时长 $\geq 30s$ ）。
- 2) 覆盖比例统计：计算三类任务样本占总任务样本的比例，确保无类别缺失。
- 3) 时序完整性验证：随机抽取10%的长时序任务样本，核查任务步骤标注是否连续（如“泡茶”需包含“拿水壶-接水-加热-倒茶”等完整步骤），无关键步骤缺失。

8.3 数据来源要求评价

8.3.1 数据来源评价应从数据集采集系统评价及数据采集训练场评价两个维度进行考虑。

8.3.2 数据集采集系统评价宜符合《人形机器人数据采集系统通用要求》。

8.3.3 数据采集训练场评价宜符合《人形机器人数据采集训练场技术规范》。

9 数据标注评测方法

9.1 评测流程概述

数据标注评测应遵循一个系统化的流程，该流程包含以下关键步骤：

- a) 抽样：从已标注数据集中按预设比例（不低于10%）进行随机分层抽样，确保样本覆盖所有标注类别和复杂场景。
- b) 评测执行：由经过认证的评测员依据本规范，对抽样样本进行定性或定量评测。
- c) 结果计算与报告：计算各项评测指标，生成详细的质量评测报告。
- d) 整改与复验：对未达到质量要求的数据批次，退回标注方整改，并重新抽样评测直至合格。

9.1.1 评测环境与工具

评测应在受控的环境中进行，使用统一的评测软件或平台，该平台应能记录评测过程、支持标注对比、并自动计算核心指标，以确保评测的一致性和公正性。

9.1.2 评测员资质

参与最终质量评测的人员（评测员）应具备以下资质：

- a) 接受过本规范及评测工具的专业培训并通过考核。
- b) 理解标注任务的目标和所有细则。
- c) 与标注团队无利益关联，保证评测的独立性。
- d) 建议持有相关领域（如计算机视觉、数据质量管理）的专业认证。

9.2 评测方法

定性评测主要依赖于评测员的主观判断，但需通过明确的规则和标准化流程将其客观化。

9.2.1 准确性

采用对标注结果抽样的方式，抽样方法应符合GB/T 10111的要求，按照标注需求方提供的验证方法对抽样样本或全体样本进行计算，得出评价结果，评价结果应满足标注需求方给出的指标要求

9.2.2 一致性

9.2.2.1 检查文件中标注人员一致性内容的说明，应符合 6.3.2 的要求。

9.2.2.2 评测方法应符合 T/SAIAS 025—2025 的规定。

9.2.2.3 检查说明文件中原始数据一致性内容应符合 T/SAIAS 025—2025 的规定。

9.2.3 时效性

评测方法应符合T/SAIAS 025—2025的规定。

9.2.4 可访问性

检查文件中关于数据可访问性的说明，应符合T/SAIAS 025—2025的规定。

9.2.5 可移植性

通过实际操作验证不同操作环境下，应符合T/SAIAS 025—2025的规定。

9.2.6 保密性

检查文件中数据集保密性的内容说明，应符合T/SAIAS 025—2025的规定。

9.2.7 可追溯性

检查文件中数据集可追溯性的内容说明，应符合T/SAIAS 025—2025的规定。

9.3 评测报告与质量控制

9.3.1 标注质量控制与评价应符合 T/SAIAS 025—2025 的规定。

9.3.2 标注结果与相关交付内容应符合 T/SAIAS 025—2025 的规定。

10 数据互联互通评价方法

10.1 评价流程

评价流程应包括：

- a) 评价准备（明确评价范围、采集样本数据）。
- b) 符合性测试。
- c) 问题分析与整改。
- d) 综合评价报告编制。

10.2 评价测试

10.2.1 统一数据模型评价测试

测试多场景测试数据集（工业巡检、家庭服务等），检查五大实体（RobotState、SensorData等）是否覆盖全部必选字段。

10.2.2 元数据规范评价测试

测试方法参考T/SAIAS 026—2025。

10.2.3 数据经纬架构评价测试

数据经纬架构评价测试包含以下内容：

- a) 数据资产目录效能测试
测试数据集条数是否 ≥ 100 万。
- b) 权限中心安全测试
使用不同权限的账户对数据库进行访问，判断是否越权拦截率为100%。
- c) 数据市场稳定性测试
 - 1) 验证是否支持gRPC-streaming通信。
 - 2) 评价是否满足相应的订阅规则。
 - 3) 模拟计费异常（如溢价系数错误），检查审计告警响应时间。

10.2.4 接口与协议评价测试

接口与协议评价测试包含以下内容：

- a) 查询接口（REST）连通性
发送查询请求，验证是否收到服务端响应头（HTTP/1.12000K或401/403），判断网络是否连通及域名可达。
- b) 订阅接口（WebSocket）连通性
发送订阅请求，终端如显示connected/无报错即标识握手成功。

10.2.5 数据集生命周期与权限评价测试

该测试用于保障数据合规性与操作可审计，包含以下内容。

- a) 生命周期完整性测试
检查数据集是否满足从采集到退役全流程追踪。
- b) 角色权限测试
使用不同权限的账户对数据库进行访问，判断是否越权拦截率为100%。

参 考 文 献

- [1] GB/T 36344-2018 《信息技术 数据质量评价指标》
- [2] T/SAIAS 《人形机器人数据采集环境及作业流程规范》
- [3] T/SAIAS 025—2025 《人形机器人数据集数据标注规范》
- [4] T/SAIAS 026—2025 《人形机器人数据集格式要求》
- [5] T/SAIAS 027—2025 《人形机器人数据集质量评价》
- [6] T/CESA 1040-2019 信息技术 人工智能 面向机器学习的数据标注规程
- [7] ISO/IEC 25012:2008 《软件工程 软件产品质量要求和评估 数据质量模型》
- [8] ISO/IEC 27001:2022 《信息安全管理体系标准》
- [9] IEEE 1872-2015 《机器人和自动化标准本体》

TCAR