

T/SAIAS

上海市人工智能行业协会团体标准

T/SAIAS XXXX—2025

人形机器人数据集应用 第1部分：家政场景数据集分类分级评估规范

Humanoid robot dataset application - part 1: Specification for domestic service scene dataset classification and grading evaluation

(征求意见稿)

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

2025 - XX - XX 发布

2025 - XX - XX 实施

上海市人工智能行业协会 发布

目 次

前言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 基本要求	1
5 分级维度和等级划分	1
6 家政场景任务分类分级	4
7 测评组织与流程	14
7.1 评价组织	14
7.2 评价流程	14
8 测评结果要求	15
8.1 基本要求	15
8.2 结果报告	15
参 考 文 献	16

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。

本文件由上海市人工智能行业协会提出。

本文件由上海市人工智能行业协会归口。

本文件起草单位：上海机器人产业技术研究院有限公司。

本文件主要起草人：

本文件为首次发布。

上海机器人产业技术研究院有限公司

人形机器人数据集应用 第1部分：家政场景数据集分类分级评估规范

1 范围

本文件规定了家政场景数据集分级维度和等级划分、家政场景任务分类分级、测评组织与流程以及测评结果要求。

本文件适用于家政场景人形机器人训练用数据集（以下简称“数据集”）。

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。

- SB/T 10847—2012 家政服务业通用术语
- T/CIE 298—2025 人形机器人智能化分级
- T/SAIAS 024—2025 人形机器人数据集分类与编码
- T/SAIAS 027—2025 人形机器人数据集质量评价

3 术语和定义

T/SAIAS 024—2025界定的以及下列术语和定义适用于本文件。

3.1

家政 homemaking

以改善或提高家庭生活质量为目标，按照服务提供方与服务接受方约定的事项，综合运用相关知识和业务技能对家庭生活进行科学的规划与管理，并采取相应的措施与方法保证目标的实现。

[来源：SB/T 10847—2012，2.1，有修改]

3.2

训练 training

使用数据集使人形机器人获得通用知识和技能的过程。

4 基本要求

数据集应符合T/SAIAS 027—2025《人形机器人数据集质量评价》的要求。

5 分级维度和等级划分

数据集分级维度为6个维度，分别为场景复杂程度、交互精细程度、任务链复杂程度、认知理解程度、数据集获取的难度和安全可靠程度。每个维度按照难易程度从低到高划分为5个等级（L1-L5级）。分级维度和等级划分为家政场景任务分类分级（详见第6章）提供依据，见表1。

表1 数据集分级维度和等级划分

维度	等级划分	等级描述	等级描述定义与举例
场景复杂程度	L1	结构化场景	数据集中环境固定、物体固定。 如固定的卧室环境，其中的床、衣柜、床头柜等家具位置固定。

表 1 数据集分级维度和等级划分（续）

维度	等级划分	等级描述	等级描述定义与举例
场景复杂程度	L2	半结构化场景	数据集中环境固定、部分物体存在变动。 如普通客厅，沙发、茶几位置相对稳定，但玩具、书籍等物品摆放可能因日常使用而有随机性。
	L3	非结构化场景	数据集中环境固定、绝大部分物体存在变动且无固定模式规律。 如经过重新装修的房间，家具、杂物杂乱摆放，还未整理归位。
	L4	动态变化场景	数据集中环境的场景、物体等会随时间发生动态改变。 如正在举办家庭聚会的餐厅，桌椅频繁移动，桌上餐具、食物位置不断变化。
	L5	复杂动态混合场景	数据集中融合多种复杂因素，既包含非结构化环境特征，又存在动态变化物体，且物体之间相互影响。 如正在进行大规模家庭空间改造的场景，既有杂乱堆放的建筑材料，又有工人不断施工、搬运物品。
交互精细程度	L1	与刚性物体的基础精细程度交互	构建数据集，能够覆盖人形机器人对简单刚性物体进行基础精细程度交互的能力。 如家居保洁的清洗餐具任务中，人形机器人在宽大的不锈钢水槽中简单冲洗厚底不锈钢煎锅。
	L2	与半刚性物体的中等精细程度交互	构建数据集，能够覆盖人形机器人对半刚性物体进行中等精细程度交互的能力。 如在家庭教育的儿童家教任务中，人形机器人将塑料直尺摆放至书桌的角落。
	L3	与柔性规则物体的较高精细程度交互	构建数据集，能够覆盖人形机器人对柔性规则物体进行较高精细程度交互的能力。 如在家居保洁的整理床铺任务中，人形机器人将衣物折叠后垂直码放。
	L4	与柔性不规则物体的高精度程度交互	构建数据集，能够覆盖人形机器人对柔性不规则物体进行高精度程度交互的能力。 如在家居保洁的清扫地面任务中，人形机器人将缠绕的各类电线理顺并分类。
	L5	与特殊属性物体的极致精细程度交互	构建数据集，能够覆盖人形机器人对附带特殊属性的柔性物体进行极致精细程度交互的能力。 如在家居保洁清洗餐具场景中，人形机器人抓取、清洗、擦干透明易碎玻璃材质的高脚红酒杯。
任务链复杂程度	L1	单链、无并行任务	构建数据集，能够覆盖人形机器人完成单一、无并行任务的能力。 如在家庭烹饪的灶台普通饮食烹饪任务中，人形机器人从已打开门的橱柜中取出一个餐盘并放置在灶台上，过程中无法同时进行其他任何操作。
	L2	短链、少量并行任务	构建数据集，能够覆盖人形机器人完成流程相对较短、同时执行少量并行任务的能力。 如在家居保洁的整理书桌场景中，人形机器人将桌上的书本整理归位至书架，同时擦拭桌面灰尘。
	L3	中等链、适量并行任务	构建数据集，能够覆盖人形机器人完成具有一定步骤数量、同时并行执行适量任务的能力。 如在家庭烹饪的灶台普通饮食烹饪任务和其他场景特殊饮食烹饪中，人形机器人完成清洗蔬菜、切菜、炒菜并装盘这一系列操作，同时可并行进行烧水、预热烤箱等任务。
	L4	较长链、较多并行任务	构建数据集，能够覆盖人形机器人完成步骤较多、同时并行处理较多关联复杂任务的能力。 如在家居保洁场景中，人形机器人依次完成卧室的整理床铺、清理地面、擦拭家具以及整理衣柜等多项任务，同时还能兼顾客厅的简单除尘、餐厅的桌面整理等任务。

表 1 数据集分级维度和等级划分（续）

维度	等级划分	等级描述	等级描述定义与举例
任务链复杂程度	L5	长链、大量并行任务	构建数据集，能够覆盖人形机器人完成多环节连贯性强、同时并行执行大量复杂决策任务的能力。 如在家庭护理的母婴护理场景中，人形机器人从判断婴儿哭闹原因、采取相应安抚措施，到测量体温、准备应急药品，再到联系监护人等一整套流程，同时还能处理婴儿的日常护理、环境安全监测、设备维护等任务。
认知理解程度	L1	简单指令理解	构建数据集，能够覆盖人形机器人准确理解并执行简单、明确指令的能力。 如人形机器人执行“拿起遥控器”、“关上窗户”等指令。
	L2	常见场景需求理解	构建数据集，能够覆盖人形机器人在常见场景中根据环境信息和一般常识理解需求并做出反应的能力。 如人形机器人在厨房看到水槽里的脏碗碟，能理解清晰需求将其清洗。
	L3	简单推理理解	构建数据集，能够覆盖人形机器人可根据已有的信息进行简单推理来完成的任务的能力。 如人形机器人看到餐桌上的剩菜和打开的冰箱，能推理出需要将剩菜收纳进冰箱。
	L4	模糊指令理解	构建数据集，能够覆盖人形机器人处理语义模糊、不明确的指令，并通过一定方式（如询问澄清、结合环境分析等）来执行任务的能力。 如指令“把那边的东西收拾下”，人形机器人能通过进一步交互确定“东西”具体所指并完成收拾。
	L5	深度意图理解与创造	构建数据集，能够覆盖人形机器人深刻理解人类的潜在意图，并在一定程度上发挥创造性来完成的任务的能力。 如主人说想要一个温馨舒适的家庭氛围，人形机器人能综合考虑灯光调节、音乐播放、室内布置等因素，自主创造出适宜的环境。
数据集获取的难度	L1	简易获取	数据集采集环境易于搭建。数据集采集过程简单直接。 如在一个布置简单的单人出租屋，对人形机器人简单动作进行采集。
	L2	普通难度获取	数据集采集环境需要一定的设备投入和布置。数据集采集过程存在一定难度的操作流程。 如搭建一个常见的两居室家庭模拟场景，采集人形机器人日常家居动作数据。
	L3	中等难度获取	数据集采集环境较为复杂，需要专业设备和一定技术支持。数据集采集过程需要应对一些技术问题和变化因素。 如在一个布置较为奢华、家具复杂的客厅场景中，采集人形机器人相关数据。
	L4	高难度获取	数据集采集环境的真实性、复杂性高，需要高端专业设备和大量专业人员参与。数据集采集过程面临诸多技术挑战和不确定性。 如在真实的多人口大家庭环境中，采集人形机器人日常生活的数据。
	L5	极难获取	数据集采集环境极难复现，且极端复杂、场景动态多变。数据集采集过程面临巨大技术难题和伦理等多方面限制。 如在一个人员众多、情况复杂的大型家庭活动场景中，采集人形机器人各种任务的数据。
安全可靠程度	L1	需在隔离环境下工作	构建数据集，能够覆盖人形机器人在与人隔离的环境下工作的能力。 如在进行厨房清洁时，需提前将人形机器人放置在划定好的操作区域，并用隔离栏与家人活动区域隔开，防止意外碰撞。
	L2	初步的人员识别与基础安全防护	构建数据集，能够覆盖人形机器人识别近距离人员并在发生意外碰撞时停止动作的能力。 如人形机器人在整理床铺时，能识别突然靠近的小孩，并在发生碰撞时停止整理动作。

表1 数据集分级维度和等级划分（续）

维度	等级划分	等级描述	等级描述定义与举例
安全可靠程度	L3	主动预判风险与全面安全防护	构建数据集，能够覆盖人形机器人主动预判安全风险并采取安全防护的能力。 如在灶台普通饮食烹饪场景中，人形机器人能主动预判炉灶可能发生火灾风险（油锅温度可能快速升高、油烟浓度可能超出正常范围等），并在火灾发生时能立即自动关闭燃气阀门，并启动厨房内的灭火装置，同时发出警报通知家人。
	L4	复杂环境的风险自适应防护	构建数据集，能够覆盖人形机器人在复杂多变的家政场景中高度自主应对安全风险的能力。 如在人群密集且环境复杂的家庭聚会场景中，人形机器人能灵活避让行人与移动的物体，并在发生意外时迅速调整自身行动轨迹与动作模式。
	L5	全方位协同的极致安全保障	构建数据集，能够覆盖人形机器人在各种家政场景中与人类、环境及其他设备完全协同实现极致安全保障的能力。 如在家庭护理的老人陪护场景，人形机器人能依据老人行为、健康数据，提前预判如卫生间滑倒等风险，提前铺防滑垫、调水温。并在老人突发疾病时迅速拨打急救电话并配合急救，全方位保障安全。

6 家政场景任务分类分级

结合家政场景的实际需求，将家政场景任务分为5个一类任务，每个一类任务包含若干二类任务。一类任务分别为家居保洁（整理床铺、整理衣橱、清扫地面和清洗餐具）、家庭烹饪（清洗食材、灶台普通饮食烹饪和其他场景特殊饮食烹饪）、家庭护理（母婴护理、老人陪护）、家庭教育（儿童家教、成人教育）和动物植物养护（动物养护、植物养护）。

依据分级维度和等级划分（详见第5章）对家政场景任务分类进行分级，并提供人形机器人能够覆盖的技能示例，见表2。

表2 家政场景任务分类分级

一类任务	二类任务	维度	等级划分	人形机器人能够覆盖的技能示例
家居保洁	整理床铺	场景复杂程度	L1	单人床，床铺规整无杂物。
			L2	双人床，床上有少量衣物散落。
			L3	儿童床，床上玩具、衣物杂乱。
			L4	床铺东西多，床铺随时被翻动。
			L5	卧室混乱，床铺用品散落一地。
		交互精细程度	L1	平整硬质床单简单平铺。
			L2	将棉质薄被折叠放置床头。
			L3	把柔软羽绒被叠成豆腐块。
			L4	理顺并整理复杂的床幔。
			L5	布置有特殊装饰的床铺，精细摆放。
		任务链复杂程度	L1	只将床单铺平。
			L2	铺床单同时整理一个枕头。
			L3	铺床单、套枕套、整理被子。
			L4	完成床铺整理，同时清理床边地面杂物。
			L5	整理床铺，调整卧室灯光、香薰从而营造氛围。
认知理解程度	L1	执行“铺好床单”指令。		
	L2	看到皱床单主动整理。		

表2 家政场景任务分类分级（续）

一类任务	二类任务	维度	等级划分	人形机器人能够覆盖的技能示例	
家居保洁	整理床铺	认知理解程度	L3	依据季节更换合适床品。	
			L4	按主人模糊要求“将床铺整理得舒服点”整理。	
			L5	根据主人心情打造专属睡眠环境。	
		数据集获取的难度	L1	在简单单人床房间采集。	
			L2	在普通家庭卧室采集。	
			L3	在布置复杂的卧室采集。	
			L4	在多人居住且常变动的卧室采集。	
			L5	在不断变化主题的特色卧室采集。	
		安全可靠程度	L1	在无人卧室独自整理。	
			L2	有人靠近时停止动作。	
			L3	预判整理时床架损坏风险并防护。	
			L4	在卧室人员走动频繁时灵活操作。	
			L5	与睡眠中的人协同，不打扰休息。	
		整理衣橱	场景复杂程度	L1	小型衣柜，衣物分类明确。
				L2	普通衣柜，衣物稍有凌乱。
	L3			多功能衣柜，杂物多。	
	L4			换季时衣柜，衣物频繁翻动。	
	L5			经历衣物大采购后，衣柜杂乱。	
	交互精细程度		L1	将硬纸盒衣物整齐放入衣柜。	
			L2	折叠普通棉质衣物并放置。	
			L3	折叠丝绸衣物并分类摆放。	
			L4	整理缠绕的围巾、领带。	
			L5	精细调整名贵衣物位置与展示。	
	任务链复杂程度		L1	把一件衣服挂进衣柜。	
			L2	挂衣服同时整理一层衣柜。	
			L3	分类整理当季衣物入衣柜。	
			L4	整理衣柜同时清洁衣柜表面。	
			L5	重新规划衣柜布局并整理。	
	认知理解程度		L1	按指令挂放指定衣物。	
			L2	看到散落衣物主动收拾。	
			L3	根据季节换装整理衣柜。	
			L4	按主人描述找衣物并整理。	
			L5	按主人风格偏好重新搭配整理。	
	数据集获取的难度		L1	在简易衣柜房间采集。	
			L2	在普通家庭衣柜处采集。	
			L3	在大型复杂衣柜房间采集。	
			L4	在频繁使用变动的衣柜处采集。	
			L5	在经常变换风格的衣柜处采集。	
	安全可靠程度		L1	在周围无人的环境整理衣柜。	
			L2	有人靠近衣柜时暂停。	
L3			预判衣柜门损坏风险并预防。		
L4			在家人拿取衣物时协作整理。		

表2 家政场景任务分类分级（续）

一类任务	二类任务	维度	等级划分	人形机器人能够覆盖的技能示例
家居保洁	整理衣橱	安全可靠程度	L5	与主人协同，保障隐私安全。
	清扫地面	场景复杂程度	L1	小房间，地面无杂物。
			L2	客厅，有少量家具和小物件。
			L3	书房，书籍纸张散落地面。
			L4	儿童游乐区，玩具满地。
			L5	举办派对的房间，垃圾杂物多。
		交互精细程度	L1	清扫硬质地面大块垃圾。
			L2	清扫普通灰尘。
			L3	清扫地毯上的毛发。
			L4	清理地面上的粘稠污渍。
			L5	小心清扫易碎物品周围垃圾。
		任务链复杂程度	L1	清扫一个小区地面。
			L2	清扫地面同时清理小型垃圾桶。
			L3	清扫全屋地面并擦拭踢脚线。
			L4	完成地面清扫，同时整理地面物品。
			L5	清扫地面，根据房间用途调整清洁方式。
		认知理解程度	L1	执行“清扫这块地面”指令。
			L2	看到地面脏乱主动清扫。
			L3	根据地面材质选择合适工具。
			L4	理解主人模糊的清扫要求。
			L5	根据不同时段调整清扫方式。
		数据集获取的难度	L1	在简易空旷房间采集。
			L2	在普通客厅采集。
			L3	在杂物多的房间采集。
			L4	在经常变动布局的房间采集。
			L5	在复杂活动频繁的房间采集。
	安全可靠程度	L1	在无人房间清扫。	
		L2	有人经过时暂停清扫。	
		L3	预判清扫工具碰撞风险。	
		L4	在人员走动多的地方安全清扫。	
		L5	与宠物协同，避免打扰。	
	清洗餐具	场景复杂程度	L1	简易厨房，少量餐具。
			L2	普通厨房，餐具稍有杂乱。
			L3	聚餐后厨房，餐具堆满水槽。
			L4	餐厅厨房，餐具随时增加。
			L5	举办大型宴会后的厨房，餐具杂乱且有特殊餐具。
		交互精细程度	L1	清洗厚实不锈钢餐具。
			L2	清洗普通塑料餐具。
			L3	清洗精致玻璃酒杯。
			L4	清洗形状不规则餐具。
			L5	清洗有特殊工艺的餐具。
		任务链复杂程度	L1	清洗一个餐具。

表2 家政场景任务分类分级（续）

一类任务	二类任务	维度	等级划分	人形机器人能够覆盖的技能示例
家居保洁	清洗餐具	任务链复杂程度	L2	清洗餐具同时擦拭台面。
			L3	清洗、分类、摆放餐具。
			L4	清洗餐具，同时清理厨房水槽。
			L5	清洗不同材质餐具，按餐厅要求摆放。
		认知理解程度	L1	按指令清洗指定餐具。
			L2	看到脏餐具主动清洗。
			L3	根据餐具材质选择清洁剂。
			L4	理解主人特殊清洗要求。
			L5	根据用餐人数、风格调整餐具准备。
		数据集获取的难度	L1	在简易厨房水槽采集。
			L2	在普通家庭厨房采集。
			L3	在复杂厨房环境采集。
			L4	在餐厅厨房实时采集。
			L5	在大型宴会厨房采集。
		安全可靠程度	L1	在无人的厨房清洗。
			L2	有人靠近时注意安全操作。
			L3	预判热水烫伤、餐具滑落风险。
			L4	在厨房人员流动大时安全操作。
			L5	与厨师协同，保障食品安全。
		家庭烹饪	清洗食材	场景复杂程度
L2	普通厨房，几种常见食材。			
L3	准备家庭宴会时，多种食材。			
L4	餐厅厨房，食材随着订单增加。			
L5	大型宴席厨房，大量复杂食材。			
交互精细程度	L1			清洗大块硬质根茎食材。
	L2			清洗普通叶菜。
	L3			清洗需要保持外表完整的菌类食材。
	L4			清洗带壳或有刺的食材。
	L5			清洗有特殊处理要求的食材。
任务链复杂程度	L1			清洗一种食材。
	L2			清洗食材的同时整理菜板。
	L3			清洗、切配多种食材。
	L4			清洗食材，同时清理厨房垃圾。
	L5			根据菜品需求预处理食材。
认知理解程度	L1			按指令清洗某食材。
	L2			看到待洗食材主动清洗。
	L3			根据食材用途选择清洗方式。
	L4			理解主人对于特殊食材的处理要求。
	L5			根据客人饮食禁忌处理食材。
数据集获取的难度	L1	在简易厨房采集。		
	L2	在普通家庭厨房采集。		
	L3	在准备丰盛菜肴的厨房采集。		

表2 家政场景任务分类分级（续）

一类任务	二类任务	维度	等级划分	人形机器人能够覆盖的技能示例	
家庭烹饪	清洗食材	数据集获取的难度	L4	在餐厅厨房实时采集。	
			L5	在大型宴会厨房采集。	
		安全可靠程度	L1	在无人厨房清洗。	
			L2	有人靠近时规范操作。	
			L3	预判食材可能产生的交叉污染风险。	
			L4	在厨房人员忙碌时安全操作。	
			L5	与厨师配合，保障食材的安全。	
		灶台普通 饮食烹饪	场景复杂程度	L1	简易厨房，设备少且固定。
				L2	普通厨房，有常见厨具。
				L3	厨房设备较多，布局稍乱。
	L4			多人同时使用厨房，环境嘈杂。	
	L5			餐厅厨房，高峰时段忙碌。	
	交互精细程度		L1	翻炒大锅菜，动作粗放。	
			L2	煎普通鸡蛋。	
			L3	烹饪需要保持表面完整的鱼片。	
			L4	制作复杂造型的点心。	
			L5	制作有精细工艺的菜品种。	
	任务链复杂程度		L1	加热一个菜品。	
			L2	炒一个菜同时烧一壶水。	
			L3	完成一道菜的多步骤制作。	
			L4	同时烹饪多个菜品。	
			L5	按餐厅订单节奏烹饪多道菜。	
	认知理解程度		L1	按指令烹饪简单菜品。	
			L2	根据常见菜谱做菜。	
			L3	根据现有的食材搭配菜品。	
			L4	理解主人模糊的口味要求。	
			L5	根据客人喜好和营养需求定制菜品。	
	数据集获取的难度	数据集获取的难度	L1	在简易厨房灶台采集。	
			L2	在普通家庭厨房采集。	
			L3	在复杂厨房环境采集。	
			L4	在多人频繁使用的厨房采集。	
			L5	在餐厅高峰时段厨房采集。	
		安全可靠程度	L1	在无人厨房烹饪。	
			L2	有人靠近时注意安全。	
			L3	预判火灾、燃气泄漏风险。	
			L4	在厨房环境复杂时安全烹饪。	
			L5	与餐厅服务员协同，保障上菜安全。	
	其他场景 特殊饮食 烹饪	场景复杂程度	L1	简易厨房，设备少（如微波炉）且固定。	
			L2	普通厨房，有常见烹饪设备（如空气炸锅）。	
			L3	稍复杂厨房，需操作电烤箱等制作特殊餐点，物品摆放较乱。	
L4			多人同时使用厨房，环境嘈杂。		

表2 家政场景任务分类分级（续）

一类任务	二类任务	维度	等级划分	人形机器人能够覆盖的技能示例
家庭烹饪	其他场景 特殊饮食 烹饪	场景复杂程度	L5	专业特殊饮食厨房，设备多样且特殊，需严格按照流程操作多种设备制作复杂特殊饮食，有他人频繁走动干扰。
		交互精细程度	L1	简单操作微波炉，设定时间与温度。
			L2	熟练操作面包机，放入食材、选择程序。
			L3	精准操作电烤箱，调整火候、放置食材位置。
			L4	精细操作多种设备，制作复杂特殊饮食，关注设备运行状态。
			L5	高难度操作专业特殊饮食设备，与助手配合，把控饮食制作的每个细节。
		任务链复杂程度	L1	用微波炉完成一次特殊食物加热。
			L2	用面包机完成特殊面包制作流程。
			L3	操作空气炸锅和电烤箱，完成一道特殊菜品制作。
			L4	操作多种设备，完成多道特殊饮食菜品制作。
			L5	操作多种专业设备，完成一整套特殊饮食套餐制作。
		认知理解程度	L1	按常规步骤操作单一特殊烹饪设备。
			L2	根据特殊食物调整面包机设置。
			L3	理解不同特殊食材对空气炸锅和电烤箱操作的不同要求。
			L4	理解多种特殊饮食需求，合理安排设备操作顺序。
			L5	深入理解特殊饮食原理，灵活调整设备和制作流程。
		数据集获取的难度	L1	在简单厨房操作微波炉场景采集。
			L2	在普通厨房操作面包机场景采集。
			L3	在稍复杂厨房操作多种设备场景采集。
			L4	在开放式复杂厨房操作多种设备场景采集。
			L5	在专业特殊饮食厨房操作多种设备场景采集。
		安全可靠程度	L1	在无人厨房操作微波炉，无安全隐患。
			L2	操作面包机时，有人靠近不影响操作，注意基本安全。
			L3	操作空气炸锅和电烤箱，预判设备使用风险并预防。
			L4	在人员走动多的厨房安全操作多种设备。
L5	与专业人员协同，保障特殊饮食制作过程安全。			
家庭护理	母婴护理	场景复杂程度	L1	婴儿房，布局简单设备少。
			L2	普通婴儿房，有常见婴儿用品。
			L3	婴儿房物品较多，稍显杂乱。
			L4	婴儿活动区域，物品常变动。
			L5	家庭聚会时，婴儿在复杂环境。
		交互精细程度	L1	给婴儿换纸尿裤。
			L2	给婴儿穿棉质衣服。
			L3	轻柔安抚哭闹婴儿。
			L4	给婴儿洗澡，注意水温。
			L5	给婴儿做精细的抚触按摩。
任务链复杂程度	L1	给婴儿喂一次奶。		
	L2	喂奶同时拍嗝。		

表2 家政场景任务分类分级（续）

一类任务	二类任务	维度	等级划分	人形机器人能够覆盖的技能示例
家庭护理	母婴护理	任务链复杂程度	L3	完成婴儿日常护理流程。
			L4	照顾婴儿同时处理一些家务。
			L5	全面照顾婴儿，协调家庭活动。
		认知理解程度	L1	按指令给婴儿做简单护理。
			L2	根据婴儿常见需求照顾。
			L3	根据婴儿表情动作判断需求。
			L4	理解妈妈模糊的婴儿护理要求。
			L5	根据婴儿成长阶段调整护理。
		数据集获取的难度	L1	在简易婴儿房采集。
			L2	在普通家庭婴儿房采集。
			L3	在物品多的婴儿房采集。
			L4	在婴儿活动频繁区域采集。
			L5	在家庭复杂场景中采集。
		安全可靠程度	L1	在无人婴儿房护理。
			L2	有人靠近婴儿时注意安全。
	L3		预判婴儿坠床、烫伤风险。	
	L4		在家庭人员走动多的环境安全护理。	
	L5		与家人协同，保障婴儿全方位安全。	
	老人陪护	场景复杂程度	L1	老人房间，布局简单。
			L2	普通老人居住环境，有日常用品。
			L3	老人房间杂物较多。
			L4	老人活动区域，环境常变动。
			L5	养老院公共区域，人员复杂。
		交互精细程度	L1	扶老人在平地行走。
			L2	给老人递送普通物品。
			L3	帮老人修剪指甲。
			L4	给老人做简单康复按摩。
			L5	协助老人进行精细手工活动。
		任务链复杂程度	L1	陪老人坐一会儿。
			L2	陪老人聊天同时递水。
			L3	完成老人日常照料流程。
			L4	照顾老人同时处理一些家务。
			L5	全面照顾老人，协调医疗等事务。
	认知理解程度	L1	按指令陪老人做简单事。	
		L2	根据老人常见需求照顾。	
		L3	根据老人表情判断需求。	
		L4	理解老人模糊的需求。	
		L5	根据老人健康变化调整陪护。	
	数据集获取的难度	L1	在简易老人房间采集。	
		L2	在普通家庭老人居住处采集。	
L3		在杂物多的老人房间采集。		
L4		在老人经常活动的区域采集。		

表2 家政场景任务分类分级（续）

一类任务	二类任务	维度	等级划分	人形机器人能够覆盖的技能示例
家庭护理	老人陪护	数据集获取的难度	L5	在养老院复杂环境采集。
		安全可靠程度	L1	在无人的老人房间陪护。
			L2	有人靠近老人时注意安全。
			L3	预判老人摔倒、疾病突发风险。
			L4	在养老院人员多的地方安全陪护。
L5	与医护人员、家属协同，保障老人安全。			
家庭教育	儿童家教	场景复杂程度	L1	安静书房，教学用品规整，仅辅导基础科目。
			L2	普通儿童房，有少量玩具，辅导日常作业。
			L3	儿童学习娱乐一体空间，物品杂乱，辅导多科目。
			L4	兴趣班教室，环境嘈杂，需灵活调整教学。
			L5	户外教学场景，干扰因素多，教学内容复杂。
		交互精细程度	L1	讲解简单数学运算，动作常规。
			L2	教写规范的汉字，纠正书写姿势。
			L3	生动讲解故事，模仿角色互动。
			L4	指导复杂手工制作，精准协助操作。
			L5	进行情感沟通引导，细腻把握孩子情绪。
		任务链复杂程度	L1	完成一道数学题辅导。
			L2	辅导作业同时解答简单疑问。
			L3	完成多科目作业辅导并总结知识。
			L4	辅导学习，组织简单学习活动。
			L5	制定全面学习计划并跟踪执行。
	认知理解程度	L1	按教材讲解知识点。	
		L2	根据孩子错误针对性辅导。	
		L3	结合生活实例拓展知识。	
		L4	理解孩子模糊问题并解答。	
		L5	根据孩子学习风格调整教学策略。	
	数据集获取的难度	L1	在简易书房环境采集。	
		L2	在普通家庭儿童房采集。	
		L3	在布置复杂的儿童学习区采集。	
		L4	在兴趣班教室实时采集。	
		L5	在多样化户外教学点采集。	
	安全可靠程度	L1	在仅有被辅导者的书房教学。	
		L2	有他人靠近时正常教学不受干扰。	
		L3	预判教学用品使用风险并防范。	
		L4	在人员走动多且嘈杂的场所安全教学。	
		L5	与家长协同，保障孩子身心安全。	
	成人教育	场景复杂程度	L1	安静自习室，设施简单，学习内容单一。
			L2	普通教室，有常见教学设备，课程常规。
L3			培训中心，设备多样，人员流动大，课程复杂。	
L4			企业内部培训室，结合工作场景，需求多变。	
L5			线上线下混合教学，环境复杂，协调难度大。	
交互精细程度		L1	讲解理论知识，互动形式简单。	

表2 家政场景任务分类分级（续）

一类任务	二类任务	维度	等级划分	人形机器人能够覆盖的技能示例
家庭教育	成人教育	交互精细程度	L2	组织小组讨论，引导交流观点。
			L3	案例分析，深度互动答疑。
			L4	模拟工作场景，精准指导操作。
			L5	个性化职业规划咨询，深入沟通建议。
		任务链复杂程度	L1	完成一次课程讲解。
			L2	讲课同时批改简单作业。
			L3	完成多环节课程教学与评估。
			L4	教学同时协助学员解决工作问题。
			L5	整合多方资源，推动学员职业发展。
		认知理解程度	L1	按课件传授知识。
			L2	根据学员疑问针对性讲解。
			L3	结合行业动态拓展知识。
			L4	理解学员职业困惑并提供方案。
			L5	根据学员职业目标调整教学重点。
		数据集获取的难度	L1	在简易自习室采集。
			L2	在普通教室环境采集。
			L3	在复杂培训中心采集。
			L4	在企业工作场景采集。
			L5	在多元混合教学环境采集。
		安全可靠程度	L1	在无人自习室授课。
L2	有人靠近教室时正常教学。			
L3	预判教学设备使用风险并预防。			
L4	在人员复杂场所安全教学。			
L5	与企业等协同，保障学员权益安全。			
动物植物养护	动物养护	场景复杂程度	L1	小型宠物笼，宠物单一，环境简单。
			L2	普通宠物房，有常见宠物用品，一只宠物。
			L3	多只宠物的居住空间，物品较多稍乱。
			L4	宠物活动区，设施常变动，宠物数量多。
			L5	动物医院病房，情况复杂，多种伤病宠物。
		交互精细程度	L1	给宠物喂食普通干粮。
			L2	为宠物梳理短毛。
			L3	给宠物洗澡，动作轻柔。
			L4	给宠物包扎简单伤口。
			L5	为宠物进行精细康复护理。
		任务链复杂程度	L1	清理一次宠物粪便。
			L2	喂食同时更换饮用水。
			L3	完成宠物短期日常养护流程。
			L4	养护宠物同时处理周边环境。
			L5	协调宠物医疗、训练等多项长期任务。
认知理解程度	L1	按指令照顾宠物基本需求。		
	L2	根据宠物常见行为判断需求。		
	L3	依据宠物健康状况调整养护。		

表2 家政场景任务分类分级（续）

一类任务	二类任务	维度	等级划分	人形机器人能够覆盖的技能示例	
动物植物 养护	动物养护	认知理解程度	L4	理解主人模糊的养护要求。	
			L5	根据宠物习性和主人生活调整方案。	
		数据集获取的难度	L1	在简易宠物笼处采集。	
			L2	在普通家庭宠物房采集。	
			L3	在多宠物居住空间采集。	
			L4	在宠物活动频繁区域采集。	
			L5	在动物医院复杂环境采集。	
		安全可靠程度	L1	在无人宠物房养护。	
			L2	有人靠近宠物时注意安全。	
			L3	预判宠物攻击、受伤风险。	
			L4	在宠物活动多的地方安全操作。	
			L5	与兽医、主人协同，保障宠物安全。	
		植物养护	场景复杂程度	L1	简易阳台，少量绿植，环境简单。
				L2	普通室内植物角，几种常见植物。
				L3	小型温室，植物种类多，布局稍乱。
	L4			大型花园，植物生长状态不一，常变动。	
	L5			植物园特殊展区，珍稀植物要求高。	
	交互精细程度		L1	给植物浇适量水。	
			L2	修剪植物的枝叶。	
			L3	给娇嫩花卉施肥，控制用量。	
			L4	为攀爬植物搭建支架，精细调整。	
			L5	对濒危植物进行特殊养护操作。	
	任务链复杂程度		L1	清理一盆植物的枯叶。	
			L2	浇水同时检查植物状态。	
			L3	完成多种植物短期的日常养护工作。	
			L4	养护植物同时整理周边环境。	
			L5	制定植物长期的养护规划并执行。	
	认知理解程度		L1	按指令照顾单一植物。	
			L2	根据植物常见症状判断需求。	
		L3	依据季节和植物习性调整养护。		
		L4	理解主人模糊的植物养护期望。		
		L5	根据植物生长变化优化养护方案。		
数据集获取的难度	L1	在简易阳台植物区采集。			
	L2	在普通室内植物角采集。			
	L3	在小型温室环境采集。			
	L4	在大型花园场景采集。			
	L5	在植物园、湿地等特殊展区采集。			
安全可靠程度	L1	在无人的阳台养护植物。			
	L2	有人靠近时正常养护操作。			
	L3	预判养护工具和植物损伤风险。			
	L4	在人员走动多的花园安全养护。			
	L5	与园林专家协同，保障植物安全。			

7 测评组织与流程

7.1 评价组织

家政场景数据集分类分级评估规范应按以下方式组织开展：

- a) 评价团队由 3 人及以上单数构成；
- b) 评价人员应掌握家政场景机器人数据集建设工作的相关专业知 识，对家政场景、任务、用品等有清晰的认知；
- c) 评价人员应通过能力测评获取相应证书，熟悉相关文件与标准要求，掌握机器人数据集建设工作的相关专业知 识，遵纪守法、诚实正直、坚持原则、实事求是、科学公正，定期开展技能培训；
- d) 具备保障评价活动的规范性文件。

7.2 评价流程

7.2.1 概述

机器人训练用数据集测试评价流程划分为3个阶段，见图1：

- a) 评价准备；
- b) 评价实施；
- c) 结果评估。

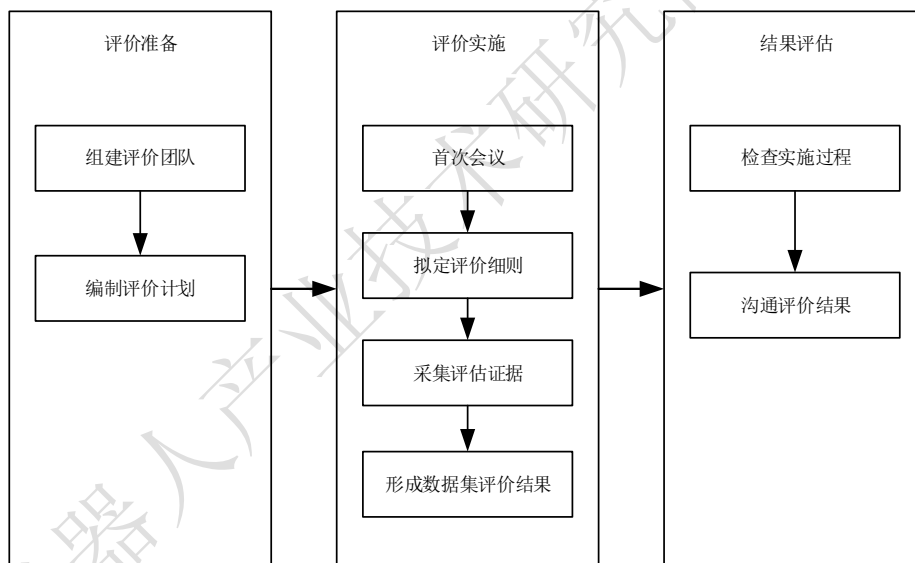


图1 评价流程图

7.2.2 评价准备

7.2.2.1 组建评价团队包括：

- a) 组建 1 个具备评估能力的评估团队实施现场评估活动，见 7.1；
- b) 确认 1 名评估组长及多名评估组员；
- c) 提前将评审时间告知评审团队。

7.2.2.2 编制评价计划包括：

- a) 评价团队讨论确定分级维度和等级划分、家政场景任务分类分级要求，见第 5 章和第 6 章；
- b) 评价团队编制评估计划，并与受评估方确认。评估计划至少包括评价目的、评价任务、评价时间、评价人员、评价日程安排等。

7.2.3 评价实施

- 7.2.3.1 首次会议要求包括：
- 确认相关方对评价计划的安排达成一致；
 - 介绍评价团队；
 - 确保策划的质量评价活动可执行；
 - 会议说明评价目的、介绍评价方法、确定评价日程以及明确其他需要提前沟通的事项。
- 7.2.3.2 在实际评价中，应根据本文件第5章和第6章制定评价细则。评价细则包括：
- 人工审查：
 - 待评价的数据集应至少包含一个二类任务；
 - 数据集应首先满足第4章的基本要求，再依据第5章和第6章进行分类分级评价。
 - 可视化工具或平台测评：通过算法/模型/系统对数据集进行分级和评价。
- 7.2.3.3 采集评估证据包括：
- 实施评价过程中，通过适当的方法收集并验证与评价目标、评价指标有关的证据；
 - 采集的证据予以记录，采集方式主要为包括访谈、文件评审等。
- 7.2.3.4 形成数据集评价结果包括：
- 对照评价细则，将采集的证据与其进行对比形成评价结果，包括具有证据支持的符合事项、弱项；
 - 完成测评后出具结果报告，结果报告包括但不限于测评报告；
 - 评价团队对评价结果达成一致意见，必要时进行组内评审。

7.2.4 结果评估

7.2.4.1 检查实施过程包括：

- 核查采集证据；
- 核查评价过程记录；
- 核查评价结果。

7.2.4.2 沟通评价结果包括：评价团队将评价结果与受评估方代表进行通报，并由评价团队确定最终结果。

8 测评结果要求

8.1 基本要求

完成测评后应出具测试结果报告。

8.2 结果报告

测试结果报告内容应支持试验可重复性，包括但不限于：

- 被测数据集的基本描述；
- 辅助设备的基本描述，如商标名称、产品型号和序列号；
- 测评设备和配置：如商标名称、产品型号和序列号以及计量/校准有效期；
- 测评环境；
- 任何进行测评所需的特殊环境条件（如有）；
- 测试数据和评价结果。

参 考 文 献

- [1] GB/T 12643—2013 机器人与机器人装备 词汇
- [2] GB/T 18391.1—2009 信息技术 实现元数据注册系统（MDR）内容一致性的规程 第1部分：
数据元
- [3] GB/T 36344—2018 信息技术 数据质量评价指标
- [4] GB/T 38667—2020 信息技术 大数据 数据分类指南
- [5] GB/T 41352—2022 再制造 机械产品质量评价通则
- [6] GB/Z 42007—2022 跨境电子商务交易服务质量评价
- [7] GB/T 45440—2025 电子商务家政 家政服务人员能力信息描述
- [8] CY/T 103—2014 数据库出版物质量评价规范
- [9] DB35/T 1952—2020 公共信息资源开发 数据质量评价规范
- [10] DB3502/T 186—2025 海峡两岸家政服务术语对照
- [11] ISO 19115-1: 2014 Geographic information — Metadata Part 1: Fundamentals
- [12] YY/T 1833.2—2022 人工智能医疗器械 质量要求和评价第2部分:数据集通用要求