

# DB4403

深 圳 市 地 方 标 准

DB4403/T XXX—XXXX

## 人形机器人应用环境数据采集规范

Humanoid robot—Application environment data collection criteria

(送审稿)

XXXX-XX-XX 发布

XXXX-XX-XX 实施

深圳市市场监督管理局 发 布



目 次

前 言 ..... II

1 范围 ..... 1

2 规范性引用文件 ..... 1

3 术语和定义 ..... 1

4 数据采集模型 ..... 2

5 数据类型 ..... 3

6 数据要求 ..... 3

7 数据接口要求 ..... 5

8 数据采集质量控制与安全控制要求 ..... 5

参 考 文 献 ..... 7

# 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。

本文件由深圳市工业和信息化局提出并归口。

本文件起草单位：

本文件主要起草人：

# 人形机器人应用环境数据采集规范

## 1 范围

本文件规定了人形机器人应用环境数据采集模型、数据类型、数据要求、数据接口要求和数据采集质量控制与安全控制要求。

本文件适用于人形机器人系统研发实际应用环境数据的体系规划、设计和建设。

## 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 30269.701 信息技术 传感器网络 第701部分：传感器接口：信号接口

GB/T 36344 信息技术 数据质量评价指标

## 3 术语和定义

以下术语和定义适用于本文件。

### 3.1

**人形机器人 humanoid robot**

具有躯干、头部和部分或全部肢体，外观和动作与人类相似，实现感知、认知、决策、执行等能力的机器人。

[来源：GB/T 12643—2025，4.15.5，有修改]

### 3.2

**内部参数 intrinsic parameter**

某硬件内部属性的参数。

注：内部参数与该硬件的物理属性有关。

### 3.3

**外部参数 extrinsic parameter**

某硬件在世界坐标系位置和方向的参数。

注：外部参数用于确定该硬件坐标系与世界坐标系之间的关系。

### 3.4

**数据采集 data acquisition**

从数据源中得到原始数据，通过标准化处理并转化为满足数据共享与利用需求的过程。

[来源：GB/T 36625.3—2021，3.2]

### 3.5

**字符串 string**

由两个（含）以上字符组成的序列。

注1：字符是字母、数字、标点符号、空格、控制字符（如换行符、制表符等）以及任何通过键盘或其他输入设备输入到计算机中的符号。

注2：字符串例如“Hello, 你好!”“把苹果放进冰箱A”等。

3.6

结构化格式 structured format

以一种预定义的、有组织的方式存储和展示的数据表示格式。

注：常用的数据结构化格式有数据库表、CSV文件（逗号分隔值）、JSON（JavaScript Object Notation）、XML（可扩展标记语言）等。

3.7

执行器 actuator

接收来自控制算法的指令，将控制系统的电信号、液压信号、气压信号等转化为机械运动，驱动机器人完成相应的机械动作（如行走、抓取、弯腰等）的设备。

3.8

传感器 sensor

依照一定的规则，对物理世界中的客观现象、物理属性进行监测，并将监测结果转化为可以进一步处理的信号的设备。

注1：电子、化学或者其他形式的传感器响应是信号的一种表现形式。

注2：1维、2维、3维或更高维度的数据是信号的一种表现形式。

[来源：GB/T 30269.2—2013，3.1.2]

4 数据采集模型

人形机器人的数据采集发生在其应用环境中，本体数据源包括传感器设备、执行器设备和任务程序，数据主要应用于人形机器人模型训练、系统开发、系统调试与验证等方面。人形机器人数据采集模型作为连接本体数据源与数据应用的中间层，包括采集数据类型和采集数据接口类型。人形机器人数据采集模型如图1所示。

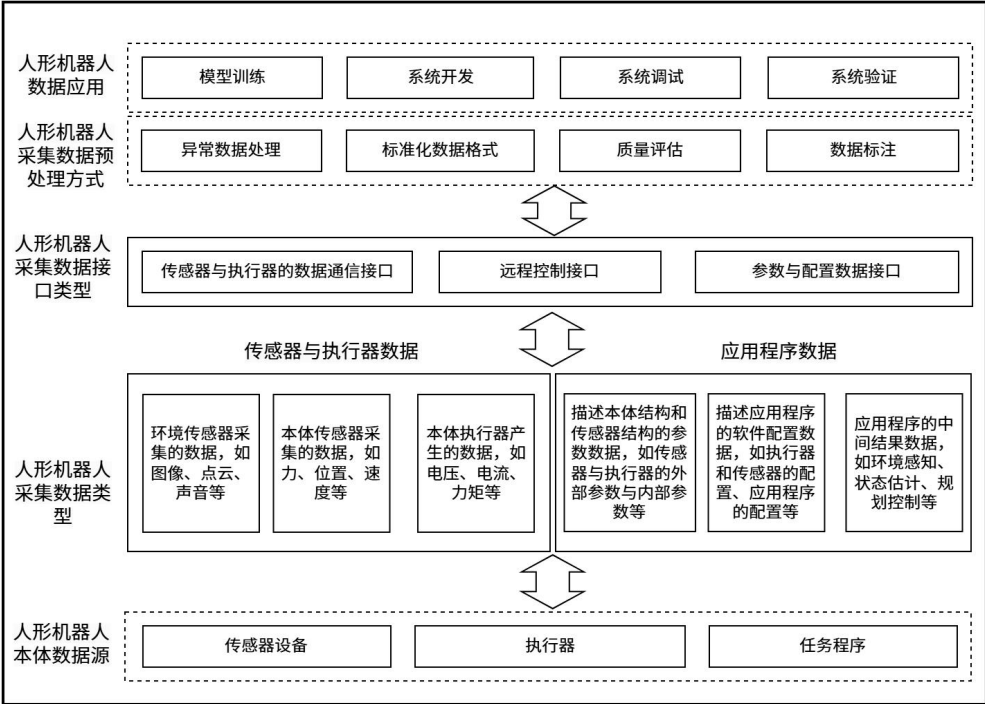


图 1 人形机器人数据采集模型

## 5 数据类型

按照数据来源划分，人形机器人采集数据可分为以下类型：

- a) 通过环境传感器和本体传感器采集和产生的数据：
  - 1) 环境传感器指人形机器人用于感知环境状态和接收外部环境信息的传感器，如摄像头、激光雷达、麦克风等；
  - 2) 本体传感器指人形机器人用于感知自身状态的传感器，如力、位置和速度传感器等。
- b) 驱动人形机器人执行器工作和反馈的数据：人形机器人执行器指控制关节运动的零部件，如电机、减速器等；
- c) 由参数数据、配置数据和应用程序中间结果数据组成的应用程序数据：
  - 1) 参数数据是描述人形机器人本体结构、传感器内外结构和执行器结构的参数，例如人形机器人机械臂、机械腿、灵巧手的结构参数、传感器和执行器的内部参数和外部参数；
  - 2) 配置数据是描述机器人硬件及软件的工作模式数据，例如人形机器人执行器和传感器的配置、应用程序的配置等；
  - 3) 应用程序中间结果数据是指机器人感知、定位、决策和控制等程序模块的输入和输出数据。

## 6 数据要求

### 6.1 数据内容要求

#### 6.1.1 传感器数据

传感器数据应包括但不限于以下内容：

- a) 时间戳；
- b) 传感器设备类型；
- c) 传感器设备在人形机器人本体的标识；
- d) 传感器数据格式；
- e) 传感器数据。

#### 6.1.2 执行器数据

执行器数据应包括但不限于以下内容：

- a) 时间戳；
- b) 执行器设备类型；
- c) 执行器设备在人形机器人本体的标识；
- d) 执行器数据格式；
- e) 执行器数值。

#### 6.1.3 应用程序中间结果数据

应用程序中间结果数据应包括但不限于以下内容：

- a) 时间戳；
- b) 软件模块标识；
- c) 软件数据格式；
- d) 软件模块产生的数据。

#### 6.1.4 参数数据

参数数据应包括但不限于以下内容：

- a) 数据版本；
- b) 更新时间；
- c) 硬件类型；
- d) 硬件标识；
- e) 参数类型；
- f) 参数值。

#### 6.1.5 配置数据

配置数据应包括但不限于以下内容：

- a) 数据版本；
- b) 更新时间；
- c) 硬件类型；
- d) 硬件标识；
- e) 软件版本；
- f) 软件编号；
- g) 配置参数。

### 6.2 数据格式要求

#### 6.2.1 概述

人形机器人数据采集应不限于单一数据类型。同类型多笔数据的保存格式应保持一致，且符合使用场景，优先选择XML、JSON、HDF5等常见格式。数据格式中的字符编码格式应统一形式化描述，宜遵循GB/T 30269.701第6章要求，确保数据在不同系统间交互的质量。

#### 6.2.2 传感器数据格式

传感器数据宜遵循机器人环境传感器（如RGB、深度、红外、点云、超声等）的数据格式要求。传感器数据的格式描述宜包括但不限于以下内容：

- a) 使用字符串或整数格式描述时间戳；
- b) 使用字符串或结构化格式描述传感器类型；
- c) 使用字符串或结构化格式描述传感器标识；
- d) 使用字符串或结构化格式描述传感器采集的数据类型。

#### 6.2.3 执行器数据格式

执行器数据宜遵循机器人执行器（如机器人手、臂、腿等）原始数据的格式要求。执行器数据的格式描述宜包括但不限于以下内容：

- a) 使用字符串或整数格式描述时间戳；
- b) 使用字符串或结构化格式描述执行器类型；
- c) 使用字符串或结构化格式描述执行器标识；
- d) 使用字符串或结构化格式描述执行器产生的数据类型。

#### 6.2.4 应用程序中间结果数据格式



应用程序中间结果数据格式描述宜包括但不限于以下内容：

- a) 使用字符串或整数格式描述时间戳；
- b) 使用字符串或结构化格式描述软件模块标识；
- c) 使用字符串或结构化格式描述软件数据格式；
- d) 使用结构化格式描述机器人软件的中间结果。

### 6.2.5 参数数据格式

参数数据的组成子项以及子项数据格式描述宜包括但不限于以下内容：

- a) 使用字符串或结构化格式描述数据版本；
- b) 使用字符串或整数的格式描述时间戳；
- c) 使用字符串或结构化格式描述硬件类型；
- d) 使用字符串或结构化格式描述硬件编号；
- e) 使用字符串或结构化格式描述参数类型；
- f) 使用结构化格式描述参数值。

### 6.2.6 配置数据格式

配置数据的组成子项以及子项数据格式描述宜包括但不限于以下内容：

- a) 使用字符串或结构化格式描述数据版本；
- b) 使用字符串或整数描述时间戳；
- c) 使用字符串或结构化格式描述硬件类型；
- d) 使用字符串或结构化格式描述硬件编号；
- e) 使用字符串或结构化格式描述软件类型；
- f) 使用字符串或结构化格式描述软件编号；
- g) 使用结构化格式描述配置数值。

## 7 数据接口要求

人形机器人采集数据接口应明确采集传感器数据、执行器数据、软件中间结果数据和硬件状态数据的通信协议、传输带宽及可靠性要求。这些要求应与通用数据传输标准保持一致，包括有线网络传输（如以太网、USB）、无线网络传输（如Wi-Fi）以及高速串口传输（如CAN）等。

## 8 数据采集质量控制与安全控制要求

### 8.1 数据采集质量控制要求

数据采集质量控制应贯穿采集的全过程，数据质量评价指标可参考GB/T 36344第5章内容。数据采集质量控制应遵循以下原则：

- a) 完整性：包含数据规则要求的数据的必要元素，识别并处理原始数据中的离群点或奇异值；
- b) 准确性：真实反映数据描述实体，符合模型训练学习的数据标准；
- c) 一致性：保证数据与其他特定上下文的数据无矛盾；
- d) 时效性：保证数据发生变化后及时更新；
- e) 可访问性：保证安全访问数据；
- f) 可追溯性：定位溯源数据；

- g) 信息标注：在原始数据上增加标注信息，转换为可供机器学习模型训练和评估的结构化数据，直接用于模型的训练和评估。

## 8.2 数据安全控制要求

### 8.2.1 数据安全控制对象

数据安全控制对象是唯一识别个人生理或行为特征的生物识别数据，包括但不限于含有人的图像、声音、指纹等数据。

### 8.2.2 数据安全控制原则

人形机器人数据安全控制应遵循以下原则：

- a) 数据保密性：传感器采集的个人数据（包括但不限于视频、音频等）采用适当的技术或组织措施，防止未经授权或非法访问，同时避免意外丢失、破坏或损坏；
- b) 目的限定性：个人数据的收集和处理明确合法用途，不得将数据用于其他目的；
- c) 数据最小化：个人数据的收集遵循与处理目的相关且必要的原则，不得过度收集无关数据；
- d) 存储限制：个人数据的存储期限满足处理目的所需最短时间要求，超过必要期限的数据予以删除或匿名化处理。

## 参 考 文 献

- [1] GB/T 39401—2020 工业机器人云服务平台数据交换
  - [2] GB/T 36625.3—2021 智慧城市 数据融合 第3部分：数据采集规范
  - [3] GB/T 41867—2022 信息技术 人工智能 术语
  - [4] GB/T 42127—2022 智能制造 工业数据 采集规范
  - [5] ISO 8373—2021 Robotics—Vocabulary
-