# 深远海船舶作业智能体数据规范（ Data Specification）

## — 、数据总体要求

### 数据来源

来⾃船舶甲板机械设备、推进系统、动⼒定位系统、发电机组等关键⼦系统。

所有数据需以时间序列结构化格式（CSV 或 JSON）存储。

⽂件命名需遵循统—映射表（详⻅第六部分）。

### 2. ⽂件编码与格式

⽂件编码：UTF-8（ ⽆ BOM）

时间字段统—为字符串格式 " YYYY-MM-DD HH: MM:SS" 示例：2025-07-07 12:00:00 . 所有数值字段均采⽤浮点数（float）或布尔型（ bool）。

### 3. 数据精度与采样频率

. 时间戳最⼩间隔建议 ≤ 1 秒；

. 电压、电流、功率类数据保留 1~2 位⼩数；

 各设备采样频率不同，需在元数据⽂件中注明（详⻅第七部分）。

## ⼆ 、数据结构与字段定义

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模块 | 关键字段 | 字段类型 | 单位/格式 | 描述 |
| A架设备 | csvTime | string | YYYY-MM-DD HH:MM:SS | 时间戳 |
|  | Ajia-0\_v | float |  | A架右舷⻆度 |
|  | Ajia-3\_v | float | A | 电流通道3数值 |
|  | Ajia-5\_v | float | A | 电流通道5数值 |
|  | key\_action | string | 枚举 | 标注关键动作 |
|  | running\_status | bool | True/False | 当前运⾏状态 |
| 动⼒定位系统 | csvTime | string | YYYY-MM-DD HH:MM:SS | 时间戳 |
|  | P3\_33 | float | kW | 推进变频器功率 |
|  | dp\_status | string | 枚举 | " ON DP " / " OFF DP " |
| 折臂吊车 | csvTime | string | YYYY-MM-DD HH:MM:SS | 时间戳 |
|  | 13-11-6\_v | float | kW | 功率 |
|  | stage\_field | string | 枚举 | " 待机" / " ⼯作" 阶段 |
| 发电机系统 | csvTime | string | YYYY-MM-DD HH:MM:SS | 时间戳 |
|  | P1\_66 | float | kW | 有功功率 |
|  | P1\_88.14 | bool | True/False | 是否额定转速 |
|  | P3\_15 | float | kW | 推进功率 |
| 航⾏状态表 | dp\_status | string | 枚举 | 当前DP状态 |

## 三、数据清洗与预处理规则

1. 时间格式规范化：

df[IcsvTime I] = pd.to\_datetime(df[IcsvTime I],format=I%Y-%m-%d % H:%M:%S I)

2. 异常值与缺失值处理

. 缺失值（ null/error）跳过；

. 连续零值超过10个时过滤；

. 异常负值需检查是否设备休眠或信号反转。

3. 单位换算

. 电流、电压除以10；频率除以100；1MJ = 1kWh / 3.6。

4. 多源数据对⻬

. 采⽤时间戳匹配，采样频率不同使⽤线性插值。

## 四、数据标注规范

|  |  |
| --- | --- |
| 标注对象 | 规则说明 |
| A架动作识别 | 电流 > 0 表示“有电流”；= 0 表示“回到位”。 |
| 动⼒定位阶段判定 | dp\_status 为 “ON DP” 表示开始，“OFF DP” 表示结束。 |
| 深海作业阶段 | 通过 A架 stage\_field 区分 “布放 / 回收”。 |
| 报警机制 | 按阈值规则触发报警。 |

## **五、报警与安全阈值参考**

|  |  |  |  |
| --- | --- | --- | --- |
| 设备 | 参数 | 阈值 | 说明 |
| 发电机 | 滑油压⼒ | < 210 kPa | 预警 |
| 发电机 | 滑油压⼒ | < 180 kPa | 停机 |
| A架 | 电流状态 | = 0 | “ 摆回到位” |
| 推进器 | 功率 | ≥ 1000 kW | 航渡状态 |

## **六、⽂件组织与命名规范**

|  |  |  |
| --- | --- | --- |
| 内容 | 示例 | 说明 |
| 输⼊数据⽂件夹 | /data/input\_data/ | 原始设备数据 |
| 输出数据⽂件夹 | /data/output\_data/ | 清洗 / 分析结果 |
| 命名格式 | {设备名}\_{编号}.csv | 例如 Ajia\_plc\_1.csv |
| 字段完整性 |  | 必须包含 csvTime + 关键字段 |

## **元数据与描述⽂件**

示例：

{

"device " : "A架吊机",

" sampling\_rate " : "1s ", "data\_format " : " CSV",

" columns " : [" csvTime ", "Ajia-0\_v ", "Ajia-3\_v ", "Ajia-5\_v ", "key\_action"], "description " : " 记录A架作业全过程电流与⻆度变化",

" version " : " v1.0"

}

## **⼋ 、数据版本与追踪**

- 每次数据更新须在 metadata.json 中更新 version；

. 建议使⽤ Git LFS 管理⼤型CSV⽂件；

 所有数据操作过程需⽇志记录（ 清洗、合并、插值步骤）。