**数据采集功能设计说明书**

**编制部门：研发--软件部**

**编 制 人：粟德森 梁雪峰**

**审 核：**

**会 签：**

**批 准：**

[1引言 4](#_Toc1900)

[1.1编写目的 4](#_Toc28091)

[1.2背景 4](#_Toc26140)

[1.3定义 4](#_Toc29762)

[1.4参考资料 4](#_Toc14431)

[2总体设计 5](#_Toc25350)

[2.1系统模块 5](#_Toc30218)

[2.2平台结构 5](#_Toc23622)

[a. 平台架构 5](#_Toc811)

[b. 基本数据 6](#_Toc7003)

[c. 交互数据 6](#_Toc7964)

[2.3数据采集流程 7](#_Toc5658)

[d. APP端数据流程 7](#_Toc13519)

[b.Robot端数据流程 8](#_Toc28239)

[3接口设计 9](#_Toc24292)

[3.1[APP]数据采集 9](#_Toc23578)

[1) 记录按钮点击数据 9](#_Toc17672)

[2) 记录APP端基本信息 9](#_Toc16440)

[3) 记录新增注册用户信息 10](#_Toc26028)

[4) 记录用户在线状态（登录次数） 10](#_Toc17782)

[5) 记录用户行为信息 11](#_Toc13847)

[3.2[ROBOT]数据采集 12](#_Toc29412)

[1) 记录机器人基本信息 12](#_Toc5213)

[2) 记录机器人在线状态信息 12](#_Toc23686)

[3) 记录机器人第三方调用情况 13](#_Toc8644)

[4) 聊天[问答]记录数据采集[问答] 13](#_Toc20107)

[3.3数据分析报表 14](#_Toc4511)

[1) [APP]获取按钮点击报表 14](#_Toc28509)

[2) [APP]用户注册统计报表 15](#_Toc6051)

[3) [APP]获取用户登录统计报表 16](#_Toc30465)

[4) [APP]获取用户平均在线时长统计报表 16](#_Toc14914)

[5) [ROBOT]获取第三方接口调用统计报表 17](#_Toc20128)

[6) [ROBOT]获取日活机器数统计报表 18](#_Toc12438)

[7) [ROBOT]统计在线机器人报表 18](#_Toc16688)

[8) [ROBOT]统计聊天数据报表 20](#_Toc28893)

[9) [ROBOT]统计聊天数据意图报表 20](#_Toc28313)

[10) [ROBOT]统计聊天相应成功失败率 21](#_Toc31635)

[4运行设计 22](#_Toc31942)

[4.1运行模块组合 22](#_Toc4276)

[4.2运行控制 22](#_Toc7815)

[4.3运行时间 22](#_Toc10152)

[5系统数据结构设计 23](#_Toc1579)

[5.1逻辑结构设计要点 23](#_Toc2015)

[5.2数据库设计 23](#_Toc19000)

[1) [APP]用户登录日志表 t\_user\_login 23](#_Toc3900)

[2) [APP]手机基本信息表 t\_phone\_info 23](#_Toc13494)

[3) [APP]用户注册日志表 t\_user\_register 24](#_Toc23721)

[4) [APP]用户绑定机器人信息表 t\_user\_robot\_relation 24](#_Toc31194)

[5) [APP]按钮请求日志表 t\_function\_request 24](#_Toc2199)

[6) [APP]点赞日志表 t\_prise\_log 25](#_Toc31008)

[7) [APP]点赞日志表 t\_comment\_log 25](#_Toc8759)

[8) [APP]收藏日志表 t\_comment\_log 25](#_Toc1039)

[9) [APP]后台请求日志表 t\_function\_request 26](#_Toc8101)

[10) [ROBOT]机器人信息表 t\_robot\_device\_info 26](#_Toc16441)

[11) [ROBOT]闲聊请求记录表 t\_function\_request 27](#_Toc8298)

[12) [ROBOT]第三方请求日志表 third\_party\_call\_collect 27](#_Toc9791)

[5.3数据结构与程序的关系 27](#_Toc7414)

[6系统出错处理设计 28](#_Toc28121)

[6.1出错信息 28](#_Toc13061)

[6.2补救措施 28](#_Toc30514)

[6.3系统维护设计 28](#_Toc32266)

**数据采集分析平台设计说明书**

# 1引言

## 1.1编写目的

介绍云平台数据采集的功能模块，数据采集机制；数据简单分析功能，为后面分析出日常用户对机器人的使用场景、频率和粘性，以及统计出用户偏好的日常功能和闲聊平台，为机器人的日常功能开发扩展，平台的选型及数量采购提供了可参考的数据依据。

本设计说明书的预期读者为产品、开发、测试及市场人员。

## 1.2背景

公司机器人已进入市场，用户及机器使用数量已初具规模，产品提出要对用户对机器人的使用情况做一个全方位的数据统计，为以后的产品开发和升级提供数据参考。

## 1.3定义

## 1.4参考资料

# 2总体设计

## 2.1系统模块

数据统计采集分为三个模块：用户数据收集、机器人数据收集、数据分析平台。

## 2.2平台结构

数据采集是基于云平台整体框架设计，作为一个微服务运行在云平台服务器中。

### 平台架构



### 基本数据



### 交互数据



## 2.3数据采集流程

### APP端数据流程



1. 接口访问采集：记录用户后台接口调用情况，可以分析功能的使用。
2. 用户绑定机器人记录。（由各自对应的系统自行收集）
3. 在线时长记录：通过心跳记录，可以分析出用户的在线时长。
4. 功能点击记录：通过pagId buttonId 记录每个产品的界面与按钮ID。
5. 每次登录状态记录：通过定时发送心跳来分析出用户的登录状态。
6. Alpha1/jimu 在线时间（链接蓝牙开始记录）
7. 同步动作 固件传输成功率。（成功 失败）
8. 停留时间，通过按钮点击时间统计出。
9. 提供时间同步的接口。

### b.Robot端数据流程



1. 应用调用上报：闲聊，翻译，音乐，故事等。

2. 登录时长记录：通过心跳记录。

3. 应用信息：应用版本 应用包名 机器人序列号

4. 第三方调用情况上报：可以分析出第三方调用情况。

5. 机器人运行状态信息：内存大小，CPU运行状态，空间大小，报错日志等相关信息。

# 3接口设计

## 3.1[APP]数据采集

### **记录按钮点击数据**

* 接口地址：POST [http://localhost:8085/btnClick/a](http://localhost:8085/uFunctionRequest/stasticsByFunction?functionName=音乐&requestDate=2017-05-12)dd
* 应用范围：手机端APP,分析出用户的行为轨迹。
* 上报策略：通过文件格式上报，APP上有按钮点击，保存文件，等APP下次打开的时候上报文件
* 上传文件格式：通过#分割

clickPageId#clickButtonId#fromPageId#fromButtonId#createTime#deviceId#Token

|  |  |  |  |
| --- | --- | --- | --- |
| 参数名称 | 是否必须 | 参数类型 | 备注 |
| token | 否 | String | 用户使用userName登录后，后台通过userName生成的token，每次登录token值会变，但是后台可以通过token解析出userName，然后保存数据。 |
| clickPageId | 是 | int | 点击页面ID |
| clickButtonId | 是 | int | 点击按钮ID |
| deviceId | 是 | int | 设备唯一标识 |
| fromPageId | 是 |  | 上一个页面ID |
| fromButtonId | 否 | String | 上一个按钮的ID |
| clickTime | 是 | Long | 创建时间 |

* 返回信息：

{

{ }

}

* 参数说明：
* 产品编号定义：

参考APP deviceid 定义。

* 页面和按钮编号定义:

需要和每个产品定义。

### **记录用户在线状态**

* 接口地址：

POST [http://localhost:8085/online/a](http://localhost:8085/uFunctionRequest/stasticsByFunction?functionName=音乐&requestDate=2017-05-12)ddOrUpdate

* 应用范围：手机端APP（APP端维持心跳，定时上报每日在线信息，由后台判断是每日首次上报还是每日最后上报）
* 上报策略：打开APP后，定时上报心跳，后台如果10分钟没收到心跳，可以判断已经离线。
* 参数说明（Body）：

|  |  |  |  |
| --- | --- | --- | --- |
| 参数名称 | 是否必须 | 参数类型 | 备注 |
| deviceId | 是 | String | 设备唯一标识 |
| token | 是 | String | token |
| onlineTime | 是 | Long | 上报时间时间戳 |
| timeZone | 是 | String | 时区 |

* 返回信息：

{

{ }

}

### **APP启动次数上报**

* 接口地址：POST [http://localhost:8085/appStart/a](http://localhost:8085/uFunctionRequest/stasticsByFunction?functionName=音乐&requestDate=2017-05-12)dd
* 应用范围：手机端APP,统计用户的启动次数
* 上报策略：打开app的时候上报
* 参数说明（Body）：

|  |  |  |  |
| --- | --- | --- | --- |
| 参数名称 | 是否必须 | 参数类型 | 备注 |
| token | 否 | String | 如果登录了，就传token；token是用户使用userName登录后，后台通过userName生成的token，每次登录token值会变，但是后台可以通过token解析出userName，然后保存数据。 |
| deviceId | 是 | String | 设备ID |
| appVersion | 是 | String | App版本 |
| AppId | 是 | Int | AppId |
| Opentime | 是 | int | 打开app的时间戳 |

* 返回信息：

{

{ }

}

* 参数说明：
* 产品编号定义：

参考APP deviceid 定义。

* 页面和按钮编号定义:

需要和每个产品定义。

## 3.2[ROBOT]数据采集

### **记录机器人在线状态信息**

* 接口地址：

POST [http://localhost:8085/online/a](http://localhost:8085/uFunctionRequest/stasticsByFunction?functionName=音乐&requestDate=2017-05-12)ddOrUpdate

* 应用范围：手机端APP，当机器人联网 关机 （异常断网？）的时候调用
* 调用时间：定时调用
* 参数说明（Body）：

|  |  |  |  |
| --- | --- | --- | --- |
| 参数名称 | 是否必须 | 参数类型 | 备注 |
| userTypeId | 是 | String | 默认为2 |
| RobotSeq | 是 | String | 机器人序列号 |
| timeZone | 是 | String | 时区 |
| reportTime | 是 | Long | 上报时间 |

### **记录机器人第三方调用情况**

* 接口地址：

POST [http://localhost:8085/thridParty/add](http://localhost:8085/uFunctionRequest/stasticsByFunction?functionName=音乐&requestDate=2017-05-12)

* 应用范围：机器人端
* 上报策略：通过文件格式上报，机器人调用第三方记录保存文件，等APP下次打开的时候上报文件。
* 参数说明（Body）：

|  |  |  |  |
| --- | --- | --- | --- |
| 参数名称 | 是否必须 | 参数类型 | 备注 |
| deviceNo | 是 | String | 序列号： |
| callCostTime | 是 | Integer | 调用耗时单位毫秒 |
| callRequestParams | 否 | String | 调用参数 |
| callResult | 否 | String | 调用结果 |
| isSuccess | 是 | int | 调用是否成功 |
| callType | 是 | String | 调用的接口编号 详见：third\_part\_call\_type表 |
| appName | 否 | String | App名称 |
| appVersion | 否 | String | App 版本 |
| ipAddress | 是 | String | 调用方地址 |
| timeZone | 是 | String | 时区 |

* 返回信息：

{

{ }

}

* 参数说明：

### **聊天[问答]记录数据采集**

* 接口地址：

POST [http://localhost:8085/thridParty/add](http://localhost:8085/uFunctionRequest/stasticsByFunction?functionName=音乐&requestDate=2017-05-12)

* 应用范围：机器人端
* 上传时间：语音交互的问题和答案记录
* 参数说明（Body）：

|  |  |  |  |
| --- | --- | --- | --- |
| 参数名称 | 是否必须 | 参数类型 | 备注 |
| deviceNo | 是 | String | 序列号： |
| questionInfo | 是 | Integer | 问的信息 |
| answerInfo | 是 | String | 回答信息 |
| callStatus | 是 | String | 回答状态 1 表示成功 0 表示没有正确回答 -1 表示错误 |
| answerUrl | 否 | String | 问题扩展，如果与url在这里返回，比如问音乐返回对应链接。 |
| callCostTime | 是 | int | 相应时间 |
| requestTime | 是 | Time-stamp | 发起聊天时间 |
| requestType | 是 | int | App 版本 |
| [intent](http://www.baidu.com/link?url=_vZCHwu12HyBAuZCwRRBtcoN1B6zshPuOE2WhKKLIxmu2bIDKct0P9OVUWdja4m873QgdhmMWhLSTFTiRtNthZFmfFbZr5FEN7J0-WhxeIi" \t "https://www.baidu.com/_blank) | 是 | String | 意图：闲聊 翻译 问答 等 |
| prodType | 是 | int | Alpha1 ：1 alpha2 2 jimu：3  Cruzr：5 |
| timeZone | 是 | String | 时区 |

* 返回信息：

{

{ }

}

* 参数说明：

### **机器人运行状态采集**

* 接口地址：

POST [http://localhost:8085/runStatus/add](http://localhost:8085/uFunctionRequest/stasticsByFunction?functionName=音乐&requestDate=2017-05-12)

* 应用范围：机器人端
* 上传时间：定时收集机器人运行状态（内存、空间、CPU、崩溃日志）记录在文件，上传文件
* 参数说明（Body）：

|  |  |  |  |
| --- | --- | --- | --- |
| 参数名称 | 是否必须 | 参数类型 | 备注 |
| deviceId | 是 | String | 设备唯一标识 |
| Cpu | 是 | Integer | Cpu使用率 |
| Memory | 是 | String | 内存信息 |
| Disk | 是 | String | 磁盘信息 |
| Errormessage | 否 | String | 错误信息 |
| Runtime | 是 | int | 运行时长 |
| markTime | 是 | Time-stamp | 记录时间 |

* 返回信息：

{

{ }

}

* 参数说明：

# 4运行设计

## 4.1运行模块组合

## 4.2运行控制

定时任务生成报表。

## 4.3运行时间

# 5系统数据结构设计

## 5.1逻辑结构设计要点

## 5.2数据库设计

### [APP]用户登录日志表 t\_user\_login

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 注释 | 类型 | 默认值 | 是否允许为null |
| id | 主键ID | Int(11) | auto\_increment | Not null |
| user\_id | 用户id | int(11) | Null | Not null |
| app\_id | 应用ID | Int(11) | 0 | Not null |
| action\_date | 登录、登出时间 | Time-stamp | Null | Not null |
| new\_added | 是否为首次登录 | Int(8) | 0 | Not Null |
| phone\_mac | 手机mac地址 | varchar(128) | Null |  |
| ip | 登录时的ip地址 | varchar(128) | Null |  |

可以分析出：用户每天在线时长，用户在线状态

[APP]用户注册日志表 t\_user\_registe

### [APP]用户绑定机器人信息表 t\_user\_robot\_relation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 注释 | 类型 | 默认值 | 是否允许为null |
| id | 主键ID | Int(11) | auto\_increment | Not null |
| user\_id | 用户id | int(11) | Null | Not null |
| app\_id | 应用ID | Int(11) | 0 | Not null |
| bind\_date | 绑定时间 | datetime | Null | Not null |
| robot\_seq | 机器人序列号 | Varchar(25) |  |  |

可以分析出：有效用户数

### [APP]按钮请求日志表 t\_function\_request

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 注释 | 类型 | 默认值 | 是否允许为null |
| id | 主键ID | Int(11) | auto\_increment | Not null |
| user\_id | 用户id | int(11) | Null | YES |
| function\_name | 按钮名称 | varchar(64) | Null | Not null |
| create\_time | 创建时间 | Time-stamp | null | Not null |
| app\_id | App应用ID | int(2) | null | Not null |

可以分析出：APP使用功能的频率，喜好。

### [APP]点赞日志表 t\_prise\_log

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 注释 | 类型 | 是否允许为null | 默认值 |
| id | 主键 | Int(11) | Not null | auto\_increment |
| user\_id | 用户ID | Int(11) | Not null |  |
| app\_id | 应用ID | int(8) | Not null |  |
| prise\_id | 被赞对象id | int(11) | Not null |  |
| prise\_time | 点赞时间 | long | Null |  |

可以分析出：动作、应用、评论被点赞的次数

### [APP]评论日志表 t\_comment\_log

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 注释 | 类型 | 是否允许为null | 默认值 |
| id | 主键 | Int(11) | Not null | auto\_increment |
| user\_id | 用户ID | Int(11) | Not null |  |
| app\_id | 应用ID | Int(8) | Not null |  |
| comment\_type | 评论类型（1、动作 2、应用 3、评论 0：其他） | int(8) | Not null |  |
| comment\_id | 被评论对象id | int(11) | Not null |  |
| comment\_time | 点赞时间 | long | Null |  |

可以分析出：动作、应用被评论的次数

### [APP]收藏日志表 t\_collect\_log

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 注释 | 类型 | 是否允许为null | 默认值 |
| id | 主键 | Int(11) | Not null | auto\_increment |
| user\_id | 用户ID | Int(11) | Not null |  |
| app\_id | 应用ID | Int（8） | Not null |  |
| collect\_type | 收藏类型（1、动作 2、应用 3、评论 0：其他） | int(8) | Not null |  |
| collect\_id | 被收藏对象id | int(11) | Not null |  |
| collect\_time | 收藏时间 | long | Null |  |

可以分析出：动作、应用被评论的次数。

### [APP]后台请求日志表 t\_function\_request

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 注释 | 类型 | 是否允许为null | 默认值 |
| id | 主键 | Int(11) | Not null |  |
| user\_id | 用户ID | Int(11) | Not null |  |
| content | 内容日志 | varchar(1024) |  |  |
| App\_id | 应用ID | Int(8) |  |  |
| ip\_address | IP地址 | varchar(32) |  |  |
| operation | 具体操作 | Varchar(50) |  |  |
| operation\_type | 操作类型 | Varchar(50) |  |  |
| cost\_time | 耗时 | double |  |  |
| create\_time | 创建时间 | timestamp |  |  |

可以分析出：后台接口请求的频率，系统耗时，功能使用的频率。

### [ROBOT]闲聊请求记录表 t\_function\_request

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名 | 注释 | 类型 | 默认值 | 是否允许为null |
| id | 主键ID | Int(11) | auto\_increment | Not null |
| robot\_seq | 机器人id | int(11) |  | Not null |
| robot\_type | 机器人类型id:alpha1:1 alpha2:2,jimu:3,cruzr:4 | Int(11) |  | Not null |
| skill\_id | 技能分类（如聊天，具体见skill\_type 表） | varchar(16) |  | Not null |
| user\_id | 使用机器人的用户ID | int(11) |  | Not null |
| App\_id | 应用ID | Int(8) |  |  |
| user\_msg | 用户消息 | text |  | Not null |
| robot\_msg | 聊天时机器人回复的对话 | text |  |  |
| push\_content | 机器人推送的内容 | text |  |  |
| push\_link | 机器人推送的链接 | varchar(256) |  |  |
| createtime | 创建时间 | long |  | Not null |

### [ROBOT]第三方请求日志表 third\_party\_call\_collect

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | | 注释 | 类型 | 默认值 | 是否允许为null |
| id | 主键ID | | Int(11) | auto\_increment | Not null |
| sys\_name | 系统名称 | | varchar(64) | null | Not null |
| sys\_version | 系统版本 | | varchar(64) | null | not null |
| device\_type | 设备类型 | | varchar(64) | null | not null |
| device\_no | 设备编号 | | varchar(64) | null | not\_null |
| call\_cost\_time | 调用耗时 | | int | 0 |  |
| call\_request\_params | 调用参数 | | text | null |  |
| call\_result | 请求结果(1:成功 0：失败) | | Int(8) | 0 | not null |
| call\_type | 第三方服务编号 | | varchar(64) | null |  |
| ip\_address | 请求终端ip地址 | | varchar(64) | null |  |
| app\_id | 应用ID | | varchar(64) | null |  |

## 5.3数据结构与程序的关系

# 6系统出错处理设计

## 6.1出错信息

## 6.2补救措施

## 6.3系统维护设计