数据平台

种子服务解决方案

版本号：V1.0

|  |  |  |  |
| --- | --- | --- | --- |
| 拟制: | 王尧 | 日期： | 2016/5/10 |
| 审核: |  | 日期： |  |
| 批准: |  | 日期： |  |

**文档修订记录**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **修订版本** | 日期 | **作者** | 修改内容描述 | **备注** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

目录

1 引言 4

1.1 目的 4

1.2 预期读者 4

2 现有问题和解决思路 4

2.1 种子服务 4

2.1.1 概述 4

2.1.2 技术解决思路及措施 4

# 引言

## 目的

解决现有Log 混乱，没有标准的情况指定统一标准个格式，便于跟踪和分析。

## 预期读者

(1)爬虫开发人员、运维开发人员、产品经理、项目管理人员

# 现有问题和解决思路

## Log 格式

[%(asctime)s-%(process)d-%(filename)s-%(module)s-%(funcName)s-%(lineno)d-%(levelname)s]: %(msg)s'

e.g.:

[2016-05-11 15:18:41,204-11876-Logging.py-Logging-<module>-122-ERROR]: eg msg

**2.2 技术解决思路及措施**

1. Logger 模块基于logging 模块封装

1. 爬虫进程持有logger实例，以（type，PID）初始化以隔离不同的log
2. Log 提供关键性信息供分析
3. Log 分为基础log和扩展log

1>基础log 主要包含方法调用的进出信息

2> 扩展log主要包含爬取或则解析过程中关键词分析

5. Log信息提供本地文件，标准输出和第三方DB存储，可以同时输出到三个方向，也可以通过配置选择输出目的

**2.3 log 使用注意事项**

1. Import metaclass 的包

**class X**(object):  
 \_\_metaclass\_\_ = LogMetaclass

Logger=logger(type,pid)

要求客户端必须持有logger 类实例作为成员，不能放在\_\_init\_\_()里面

1. Logger 提供 5种级别的方法, 不要随意使用 error，critical等严重等级

logging.debug('debug message')

logging.info('info message')

logging.warning('warning message')

logging.error('error message')

logging.critical('critical message')