

比一比-钢铁大数据应答PPT

应答目录

页码	应答目标	应答内容
3		内容采集需求分析
4		小蜜蜂产品介绍
5	数据采集	小蜜蜂技术优势
6		小蜜蜂操作界面
8		小蜜蜂落地案例
	大数据平台	

需求分析-采集

高时效数据

多维度数据

竞品数据

工业大数据特点

供应商数据	机器数据	控制数据	人员数据	物料数据	质量数据	客户数据	物流数据
<ul style="list-style-type: none">• 产品质量• 服务信息• 信用数据• 位置数据• 渠道依赖• 原料来源• Web信息• 业务信息• 行为信息	<ul style="list-style-type: none">• 多种类型• 时间序列• 数据真实• 数据海量• 并发较高	<ul style="list-style-type: none">• 数据多样• 时间戳• 程序数据• 结果数据	<ul style="list-style-type: none">• 基本信息• 行为信息	<ul style="list-style-type: none">• 基本信息• 计量信息• 位置信息• 物流信息• 加工信息• 装配信息• 追踪信息	<ul style="list-style-type: none">• 检验数据• 随机性• 概率特征• 相关性	<ul style="list-style-type: none">• 需求数据• 产品数据• 位置数据• 竞争对手• 信用数据• 业务数据• Web信息• 行为信息	<ul style="list-style-type: none">• 位置数据• 计量数据• 时间数据

高并发

分布式

APP数据

WEB数据

产品介绍-小蜜蜂

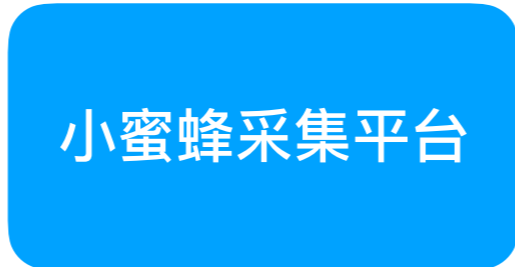
技术优势-小蜜蜂

成熟快速落地

无缝接入各种类型数据库
关系/非关系/ES/Hadoop

单渠道十分钟快速配置

在多个项目上并发300+



支持私有化部署或云部署

数据仓库集成管理

热下载节点拓展

积累80+渠道采集模版

案例分析-小蜜蜂

小蜜蜂采集平台

Bee · DataHub · Controller fangnan

Default

MAIN

- 采集种子
- 页面模型**
- 采集配置
- 用户列表
- 1 最先采集队列
- 2 优先采集队列
- 3 采集队列

FRUIT

- Fruit

LOG

- Step Log
- Step Log数据报表
- Origin Log

BEE

- 最后活动时间

开发人员工具

- 快速采集
- 选择器测试
- 模型匹配工具
- 采集调度
- 页面模型结构

+ 添加网站

10086 Apple balmain boots BOSS直聘 digitimes eciaauthorized IT桔子 kaoshi100 The Conran Shop

YOKA 爱彼迎 安居客 百度 百事通 北京市国土资源 北京市社保 比一比 成都社保 大众点评

东方财富网 饿了么 房天下 瓜子二手车 国家企业信息 和讯网 华润万家 吉林省建设信息网 加州大学旧金山分校

金十数据 京东 卡宝宝 看准网 科学技术部 拉勾网 链家网 美团 南方都市报 企查查 启信宝

人民日报 融360 厦门社保个人资料 上海证券交易所 深圳市国土资源 搜狗 天津社保 天猫 天眼查 同花顺

沱沱工社 微迷 微信搜索 无讼 西域 新浪微博 雪球 央视新闻 一号店 中关村 中国裁判文书网

中国国土资源 中国理财网 中国商标网 中国土地市场网 中国证监会 中国知识产权 重庆市国土资源 专利审查信息查询

+ 添加模型 刷新

模型名称	启用状态	发布状态	签出状态	自动翻页	Content-Type	Browser	删除
10086-流量充值	✔	未发布	未签出	未启用	Json	WebClient	✘
10086-优惠专区	✔	已发布	未签出	未启用	Html	Selenium_PhantomJs	✘
a	⊖	已发布	他签出	未启用	Html	WebClient	✘
a+	⊖	未发布	未签出	未启用	Html	WebClient	✘
test01	⊖	未发布	未签出	未启用	Html	WebClient	✘
test02	⊖	未发布	他签出	未启用	Html	WebClient	✘
北京国土资源管理	⊖	已发布	未签出	未启用	Html	WebClient	✘
房地产信息	⊖	未发布	未签出	未启用	Html	WebClient	✘
飞客茶馆	✔	已发布	他签出	未启用	Html	Selenium_Chrome	✘

案例分析-小蜜蜂

小蜜蜂数据中心

