

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		小蜜蜂采集系统产品介绍
5	数据采集	小蜜蜂采集系统技术优势
6		小蜜蜂采集系统操作界面
8		小蜜蜂采集系统落地案例
12		大数据需求分析
		比一比大数据技术优势
	大数据平台	比一比大数据落地案例

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

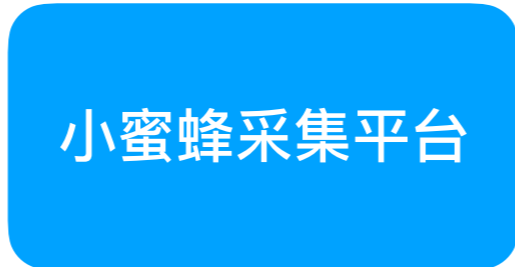
技术优势-小蜜蜂

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

小蜜蜂采集平台-配置页

设置配置页

采集配置 基础属性

浏览器

BrowserType:
WebClient

采集

允许的最大采集时长 (秒):
每个CrawlStep的采集间隔时间 (秒):
允许的最大采集次数:
允许的最大采集深度:
允许的最大连续采集失败次数:
最大采集线程数:
起始验证码抽取表达式:
弹出验证码抽取表达式:
刷新验证码js:
提交按钮抽取表达式:
起始验证码url:

HTTP

HTTP请求的超时时长 (秒):
当HTTP失败时, 重试前的等待时间 (秒):
当HTTP失败时, 允许的重试次数:
可接受的最大HTTP Size:
 使用公共Cookie 不使用公共Cookie
用户代理 (UserAgent):
Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/50.0.2661.102 Safari/537.36
主机Host:
请求Referer:
ContentType:

模版配置页

10086·优惠专区

注: 您在此所做的任何修改都不会影响目前正式环境中运行中的内容, 只有当您手动“发布”时, 才会将这些修改更新到正式环境。 退出以进行编辑

模型属性 发布 基础配置 历史版本 FullStack浏览器设置

http://www.10086.cn/support/service/events/zj/index_571_571.html

测试

活动: .llyh

活动主题: p.text[1]/活动主题: (.+) \$->\$1

活动时间: p.text[2]/活动时间: (.+) \$->\$1

活动范围: p.text[3]/活动范围: (.+) \$->\$1

参与对象: p.text[4]/参与对象: (.+) \$->\$1

图片: img.src

活动地址: .right a.href

省市: .topcity.text

+属性 / 属性组

案例分析-小蜜蜂

小蜜蜂数据中心



案例分析-小蜜蜂

小蜜蜂数据中心-详情页

BDatahub 数据管理系统 658,873.00

Search

Dashboard 查看索引 MongoDB-Query

其他网站 中国证监会 新浪微博 无讼 微信搜索 天眼查 天猫 天津社保 搜狗 企查查 美团 链家网 京东 金十数据 吉林省建设信息网 和讯网 东方财富网 成都社保 百度 安居客 10086

api.cportal.cctv.com

PageKey: 页面模型: UrlRegex:

Query:

Showing 1 to 100 of 41255 entries

第一页 上一页 1 2 3 ... 412 413 下一页 最后一页

Search

央视新闻-内容 <http://api.cportal.cctv.com/api/rest/articleInfo?id=Arti2gqCXjk24ITnHNNH4lkdj1904...> (👤4分钟前 / 🗨️无) 查看 查看表格

- 标题: 习近平出席2019年中国北京世界园艺博览会开幕式并发表重要讲话 宣布北京世界园艺博览会开幕 强调各方要共同建设美丽地球家园
- 内容:

国家主席习近平28日在北京延庆出席2019年中国北京世界园艺博览会开幕式,并发表题为《共谋绿色生活,共建美丽家园》的重要讲话,强调顺应自然、保护生态的绿色发展昭示着未来。地球是全人类赖以生存的唯一家园。中国愿同各国一道,共同建设美丽地球家园,共同构建人类命运共同体。

4月底的北京,万物复苏,芳菲满目。雄伟的长城脚下,美丽的妫水河畔,作为开幕式主会场的妫汭剧场,宛若一只展翅的彩蝶。

晚8时许,习近平和夫人彭丽媛同刚刚参加完第二届“一带一路”国际合作高峰论坛的捷克总统泽曼、吉布提总统盖莱、吉尔吉斯斯坦总统热恩别科夫、尼泊尔总统班达里、塔吉克斯坦总统拉赫蒙、柬埔寨首相洪森、缅甸国务资政昂山素季、巴基斯坦总理伊姆兰·汗、新加坡总理李显龙、日本首相特别代表二阶俊博等外方领导人夫妇共同步入会场。

现场奏唱中华人民共和国国歌,五星红旗冉冉升起。伴随雄壮的乐曲声,国际展览局局旗、国际园艺生产者协会会旗、北京世园会会旗依次升起。

在热烈的掌声中,习近平发表重要讲话,代表中国政府和中国人民,热烈欢迎远道而来的各位嘉宾。习近平指出,锦绣中华大地,是中华民族赖以生存和发展的家园,孕育了中华民族5000多年的灿烂文明,造就了中华民族天人合一的崇高追求。现在,生态文明建设已经纳入中国国家发展总体布局,建设美丽中国已经成为中国人民心向往之的奋斗目标。中国生态文明建设进入了快车道,天更蓝、山更绿、水更清将不断展现在世人面前。

习近平强调,纵观人类文明发展史,生态兴则文明兴,生态衰则文明衰。杀鸡取卵、竭泽而渔的发展方式走到了尽头,顺应自然、保护生态的绿色发展昭示着未来。地球是全人类赖以生存的唯一家园。我们要像保护自己的眼睛一样保护生态环境,像对待生命一样对待生态环境,同筑生态文明之基,同走绿色发展之路。

第一,我们应该追求人与自然和谐。无序开发、粗暴掠夺,人类定会遭到大自然的无情报复;合理利用、友好保护,人类必将获得大自然的慷慨回报。我们要维持地球生态整体平衡,让子孙后代既能享有丰富的物质财富,又能遥望星空、看见青山、闻到花香。

第二,我们应该追求绿色发展繁荣。绿水青山就是金山银山,改善生态环境就是发展生产力。良好生态本身能够源源不断创造综合效益,实现经济社会可持续发展。

第三,我们应该追求热爱自然情怀。我们要倡导简约适度、绿色低碳的生活方式,形成文明健康的生活风尚;倡导环保意识、生态意识,构建全社会共同参与的环境治理体系;倡导尊重自然、爱护自然的绿色价值观念,形成深刻的人文情怀。

第四,我们应该追求科学治理精神。生态治理必须遵循规律,科学规划,因地制宜,统筹兼顾,打造多元共生的生态系统。生态治理,道阻且长,行则将至。我们既要有只争朝夕的精神,更要有持之以恒的坚守。

第五,我们应该追求携手合作应对。面对生态环境挑战,人类是一荣俱荣、一损俱损的命运共同体。唯有携手合作,我们才能有效应对气候变化、海洋污染、生物保护等全球性环境问题,实现联合国2030年可持续发展目标,让绿色发展理念深入人心、全球生态文明之路行稳致远。

习近平强调,昨天,第二届“一带一路”国际合作高峰论坛成功闭幕。共建“一带一路”就是要建设一条开放发展之路,同时也必须是一条绿色发展之路。中国愿同各国一道,共同建设美丽地球家园,共同构建人类命运共同体。建设生态文明,功在当代,利在千秋。让我们从自己、从现在做起,把接力棒一棒一棒传下去。

习近平宣布,2019年中国北京世界园艺博览会开幕。

剧场内响起热烈的掌声和欢呼声。舞台上,志愿者们挥舞参展国国旗和参展组织旗帜,将舞台装扮成色彩绚烂的海洋。

开幕式上,国际展览局秘书长洛塞泰斯、国际园艺生产者协会主席欧斯特罗姆分别致辞。他们盛赞中国是绿色发展的典范,高度评价中国政府重视生态文明和美丽中国建设。他们感谢中国政府举办了有史以来规模最大的世园会,表示相信北京世园会将极大地弘扬可持续发展理念,为子孙后代留下一份丰厚的绿色遗产。

随后,习近平和彭丽媛同各国嘉宾共同观看了以“美丽家园”为主题的文艺晚会。

演出在童声合唱《心底的天籁》中拉开序幕,孩子们动听的歌声和深情的朗诵,邀约大家携手播撒爱与希望。动听的钢琴演奏和灵动的芭蕾舞表演,共同演绎了春夏秋冬绮丽的田园。《东方的墨韵》以中国民族乐器演奏,呈现了一幅雅致唯美的月夜画卷。探戈、踢踏、非洲鼓舞和中国扇子舞《彩翼的国度》,跳出世界各国人民的欢乐和喜悦。歌曲《月影的深情》演绎了一幅镜花水月、空灵梦幻的唯美画境。中国古典舞《彩蝶的虹桥》,以翩翩起舞的蝴蝶展现了大自然的五彩斑斓。各国模特身着参展国名花图案服饰,用花语共叙友谊。主题曲《美丽的家园》唱出对家园的赞美、对自然的热爱、对未来的希冀、对绿色的承诺。精彩的演出,赢得观众阵阵掌声。演出在绚丽的焰火表演中落下帷幕。

丁薛祥、杨洁篪、黄坤明、蔡奇、王毅、赵克志等出席上述活动。

胡春华主持开幕式。

2019年中国北京世界园艺博览会将于2019年4月29日至10月7日在北京延庆举行,以“绿色生活,美丽家园”为主题。共有约110个国家和国际组织参展。
- id: Arti2gqCXjk24ITnHNNH4lkdj190428
- 来源: 央视新闻客户端
- 时间: 2019-04-28 23:13:03

落地案例-小蜜蜂

中信数据采集平台

通过高并发，大吞吐、稳定的自动化采集技术，帮助中信银行自动化的收集用户社保以及公积金信息，辅助风控部门决策。并预知信用卡部门相关的决策风险。为银行提升40%的审核效率以及300%的核卡速度。从而积累了良好的口碑。

我们的主要实现了以下功能：

(1) 品牌舆情监控和分析

通过采集微信公众号、微博、四大主流资讯网站、政务网站等互联网新闻数据渠道，监控中信银行、中信信用卡、中信银行高管和中信银行重点产品的网络舆情新闻事件。自动化对舆情进行NLP技术清洗。每日自动化生成并返回相关舆情分析报告。

(2) 智能抓取社保信息

为优化销售端分中心社保专案进件流程，提升客户体验，同时减少审批端审核环节人工重复操作，提升审批自动化率，在销售前端充分告知客户并获得客户授权的情况下，智能采集申请人社保信息（涉及全国各地社保网站）。

(3) 定期清洗企业工商信息

目前，单位信息完整性及有效性在风险防控及效率提升中起着至关重要的作用。因此，信贷审批部在9月底上线了单位信息标准化需求，初步实现了单位信息标准化功能，为能保证单位信息数据的及时性和准确性，精准识别风险，需定期采集、清洗、更新、完善网络上单位相关信息（涉及一些企业信息的网站）。

(4) 业务环节网查信息

当前审批流程中需各人工岗采集、处理大量的网查信息，如网查单位地址及电话等（涉及招聘类、地产类和团购类等网站）。

随着信息技术的飞速发展，在当前人工智能背景下，借助智能抓取技术，通过帮助各业务部门抓取所需要的网络数据信息，处理相关的业务需求，减少人工操作，降低人工成本，提升卡中心各业务部门的自动化率，提升客户体验，助力中信信用卡卡中心“三年三倍增”的战略目标。

落地案例-小蜜蜂

郑州银行外部数据管理平台

这个找方楠要!!!

需求分析

痛点分析

- 1.工业数据来源多样，结构不统一，管理和处理这些数据成为瓶颈。
- 2.对于日益高速增长的工业数据（其中不乏大量的非结构化数据），海量的数据场景下，传统的数据处理技术无法满足业务查询和分析需要。
- 3.工艺优化、质量提升、故障预测等需求，需要更智能先进的分析和技术手段得以实现，来取代人工分析。
- 5.大数据平台往往涉及众多服务器资源，运维人员如何高效管理这些资源成为难题。

技术优势

1.多样性

大数据平台支持多种数据存储，结构化，半结构化，非结构化数据

2.高吞吐

大数据平台每天可处理上PB级数据量可根据需求调配资源理论上处理量没有上限

3.低延时

大数据平台更新数据可达秒级

4.高可用

大数据平台可靠性高达99.99%

5.安全性

大数据平台可有效杜绝绝大部分的安全威胁，包括人为的、自然的因素

6.易管理

可视化的大数据平台极大提升企业对数据管理的效率

7.可扩展

大数据平台可方便扩充存储容量，理论上扩容量没有上限

8.高兼容

大数据平台具备高通用性的优点，支持多种行业应用的对接

金道BI

金道BI项目

金道BI项目是一个外包项目，我司通过大数据分析技术，为企业客户打造了一款商业智能软件，该项目主要针对海量的用户行为和公司各部门长期业务数据提供处理、计算和分析功能，帮助企业客户的战略决策提供高价值的参考。

比一比大数据中心

有问题问全艺

比一比大数据中心是自主研发的包含了数据存储、数据处理、数据分析、搜索引擎、数据挖掘等功能的综合性大平台，为金融、银行、保险、法律、新零售、旅游等多个行业的大数据项目提供厚实的底层技术和资源保障。