



智慧运营 数据造未来

2016 · 永洪科技上海大数据峰会

大数据底层技术从此触手可及



薛铮 | 大快搜索
联合创始人&CMO

大快搜索解决了：

- 获取数据
- 高效处理海量数据
- 快速开发大数据应用



1 / 基于机器学习的自定义通用爬虫



2 / DKH企业级大数据通用计算平台



3 / FreeRCH大数据一体化开发框架

自定义爬虫

dkso.com | 数据管理

2505326688@qq.com

返回首页 / 我的任务 / 数据采集

设置网址 (URL)

编辑你的URL并设置规则, 设置完成后请保存 任务ID: 1604270012085019 任务名称: 青岛淘宝女装

`https://shopsearch.taobao.com/search?q=%E9%9D%92%E5%B2%9D%E5%A5%B3%E8%A3%85&app=shopsearch&spm=a217f.7278617.1000187.3&loc=%E9%9D%92%E5%B2%9D` 删除 设置

设置

起始URL: `https://shopsearch.taobao.com/search?q=%E9%9D%92%E5%B2%9D%E5%A5%B3%E8%A3%85&app=shopsearch&spm=a217f.7278617.1000187.3&loc=%E9%9D%92%E5%B2%9D`

向下抓取深度: 2

页面抓取间隔: 10000

是否是列表页: 否 是

起始页详细设置 第二层设置

保存 关闭 恢复默认设置

自定义搜索规则

第二层设置

起始URL: `https://shopsearch.taobao.com/search?q=%E9%9D%92%E5%B2%9D%E5%A5%B3%E8%A3%85&app=shopsearch&spm=a217f.7278617.1000187.3&loc=%E9%9D%92%E5%B2%9D`

当前层级: 第二层设置

本层是否保存: 保存 不保存

URL过滤

不过滤 关键字过滤 正则表达式过滤

内容抓取 整页抓取 精确抓取

不过滤 关键字过滤 正则表达式过滤

保存 关闭

- 首页
- 我的任务
- 任务列表
- 新建定制搜索任务
- 新建数据采集任务
- 导入任务
- 接口使用
- 会员说明
- 升级申请
- 设置
- 资料修改
- 密码修改
- 注销

设置网址 (URL)

编辑你的URL并设置规则，设置完成后请保存 任务ID: 1604270012085019 任务名称: 青岛淘宝女装

https://shopsearch.taobao.com/search?q=%E9%9D%92%E5%B2%9D%E5%A5%B3%E8%A3%85&app=shopsearch&spm=a217f.7278617.1000187.3&loc=%E9%9D%92%E5%B2%9D

删除

设置

设置

起始URL: https://shopsearch.taobao.com/search?q=%E9%9D%92%E5%B2%9D%E5%A5%B3%E8%A3%85&app=shopsearch&spm=a217f.7278617.1000187.3&loc=%E9%9D%92%E5%B2%9D

向下抓取深度: 2

页面抓取间隔: 10000

是否是列表页: 否 是

起始页详细设置

第二层设置

保存

关闭

恢复默认设置

自定义搜索规则

第二层设置

起始URL: https://shopsearch.taobao.com/search?q=%E9%9D%92%E5%B2%9D%E5%A5%B3%E8%A3%85&app=shopsearch&spm=a217f.7278617.1000187.3&loc=%E9%9D%92%E5%B2%9D

当前层级: 第二层设置

本层是否保存: 保存 不保存

URL过滤

不过滤 关键字过滤 正则表达式过滤

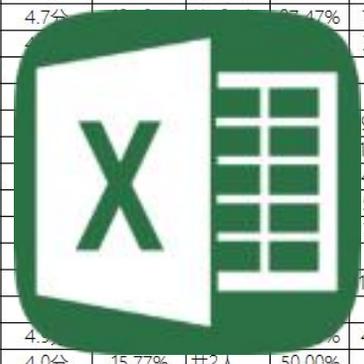
内容抓取 整页抓取 精确抓取

不过滤 关键字过滤 正则表达式过滤

保存

关闭

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	店铺名称	店铺地址	卖家名称	当前主营	所在地区	建店时间	卖家信用	买家信用	卖家好评率	主营行业 卖家信用	主营占比	宝贝与描述相符	比同行业水平高	评价人数	打5分人数比	打4分人数比	打3分人数比	打2分人数比	打1分人数比	卖家的服务态	比同行业水平高	评价人数
2	唯恩	http://store.taobao.com/shop/view_shop.htm	小舞舞小	服饰鞋包	青岛	2009-10-30	31692	641	99.64%	31652	99.87%	4.8分	53.66%	共1100人	92.36%	5.45%	0.91%	0.73%	0.55%	4.9分	61.10%	共1073人
3	三江雕刻	http://store.taobao.com/shop/view_shop-c674	三江雕刻	服饰鞋包	青岛	2011-1-9	6506	140	99.16%	6488	99.72%	4.8分	59.37%	共245人	92.65%	4.90%	2.04%	0.41%	暂无	4.9分	54.80%	共228人
4	美丽心情依衣阁	http://store.taobao.com/shop/view_shop-357	美丽心情依衣阁	服饰鞋包	青岛	2010-3-31	46187	331	98.66%	46140	99.90%	4.7分	9.73%	共944人	87.08%	7.73%	2.65%	0.53%	2.01%	4.8分	24.26%	共898人
5	bule128	http://store.taobao.com/shop/view_shop-0d4	bule128	服饰鞋包	青岛	2009-5-12	11702	55	99.34%	11527	98.50%	4.7分	9.88%	共1140人	86.14%	8.25%	3.68%	0.70%	1.23%	4.8分	14.82%	共980人
6	zhangyongli_2005	http://store.taobao.com/shop/view_shop-955	zhangyongli_2005	服饰鞋包	青岛	2008-12-7	2288	1141	99.74%	2206	96.42%	4.9分	67.64%	共283人	93.99%	4.59%	1.06%	暂无	0.35%	4.9分	74.42%	共282人
7	邵言2	http://store.taobao.com/shop/view_shop-97e	邵言2	服饰鞋包	青岛	2012-12-5	1150	60	96.91%	1141	99.22%	4.9分	71.56%	共14人	92.86%	7.14%	暂无	暂无	暂无	5.0分	100.00%	共14人
8	大码大号衣橱	http://store.taobao.com/shop/view_shop.htm	生活乐园_2008	服饰鞋包	青岛	2008-11-18	3326	195	98.55%	3317	99.73%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	4.8分	29.56%	共395人
9	铭衣尚	http://store.taobao.com/shop/view_shop.htm	时尚芭莎_1914	服饰鞋包	青岛	2012-12-18	2099	83	99.01%	2099	100%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	4.8分	36.38%	共81人
10	好衣居	http://store.taobao.com/shop/view_shop.htm	tb2477029_2012	服饰鞋包	青岛	2012-12-18	45	26	97.87%	43	95.56%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	4.8分	36.38%	共81人
11	中老年女装、外贸原单	http://store.taobao.com/shop/view_shop.htm	waj0088	服饰鞋包	青岛	2012-12-18	51	12	100.00%	849	99.76%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	5.0分	100.00%	共13人
12	毒之衣衣女装裤装	http://store.taobao.com/shop/view_shop.htm	衣度飘飘	服饰鞋包	青岛	2012-12-18	32	65	99.46%	142879	92.04%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	4.8分	2.81%	共896人
13	三生不弃	http://store.taobao.com/shop/view_shop-0e2	三生不弃	服饰鞋包	青岛	2012-12-18	66	66	96.21%	248	99.20%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	4.7分	1.08%	共130人
14	鑫源运动户外正品店	http://store.taobao.com/shop/view_shop.htm	鑫源商社	服饰鞋包	青岛	2012-12-18	225	225	99.15%	689	46.15%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	4.9分	33.67%	共133人
15	华宝针织服饰	http://store.taobao.com/shop/view_shop.htm	wanhua791114	服饰鞋包	青岛	2012-12-18	31	31	98.58%	128	92.09%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	4.8分	9.07%	共17人
16	秋雨红茶	http://store.taobao.com/shop/view_shop-a6e	秋雨红茶	服饰鞋包	青岛	2012-12-18	5	114	95.99%	1085	100%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	4.7分	0.86%	共204人
17	小毛外贸杂货铺	http://store.taobao.com/shop/view_shop.htm	maomao19850512	服饰鞋包	青岛	2012-12-18	71	754	100.00%	63	56.76%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	5.0分	100.00%	共5人
18	喜洋洋6520	http://store.taobao.com/shop/view_shop-738	喜洋洋6520	服饰鞋包	青岛	2012-12-18	47	124	95.92%	47	100%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	4.8分	42.75%	共9人
19	痞太郎	http://store.taobao.com/shop/view_shop-e4c	痞太郎	服饰鞋包	青岛	2012-12-18	305	267	99.67%	305	100%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	5.0分	100.00%	共1人
20	sm6428	http://store.taobao.com/shop/view_shop-c2c	sm6428	服饰鞋包	青岛	2012-12-18	66	381	100.00%	66	100%	4.7分	100%	共14人	95.47%	7.37%	3.69%	0.25%	1.23%	5.0分	100.00%	共44人
21	爱上小妮的店	http://store.taobao.com/shop/view_shop.htm	王秀艳朱元超	服饰鞋包	青岛	2015-3-27	4	249	100.00%	0	0.00%	4.0分	15.77%	共2人	50.00%	暂无	50.00%	暂无	暂无	4.0分	16.77%	共2人
22	zxy72999	http://store.taobao.com/shop/view_shop-890	zxy72999	服饰鞋包	青岛	2010-12-11	538	282	98.90%	91	16.91%	4.9分	93.25%	共59人	98.31%	1.69%	暂无	暂无	暂无	5.0分	100.00%	共57人
23	last_next	http://store.taobao.com/shop/view_shop-779	last_next	服饰鞋包	青岛	2009-1-25	81	938	98.81%	55	67.90%	5.0分	100.00%	共2人	100.00%	暂无	暂无	暂无	暂无	5.0分	100.00%	共2人
24	依依不舍外贸女装美衣	http://store.taobao.com/shop/view_shop.htm	mu944429935	服饰鞋包	青岛	2012-11-17	5082	341	99.24%	5076	99.88%	4.7分	17.86%	共500人	84.20%	9.20%	3.40%	0.20%	1.00%	4.8分	23.19%	共483人
25	172当铺	http://store.taobao.com/shop/view_shop.htm	wuguanshifei	服饰鞋包	青岛	2013-11-17	780	746	94.77%	780	100%	4.4分	6.88%	共32人	100.00%	7.45%	7.76%	3.11%	5.90%	4.6分	3.85%	共306人
26	小周同学2008	http://store.taobao.com/shop/view_shop-5fd	小周同学2008	服饰鞋包	青岛	2007-10-20	11637	301	99.79%	11272	96.86%	4.8分	41.32%	共604人	89.40%	8.11%	1.66%	暂无	0.83%	4.8分	48.38%	共589人
27	石头休闲女装平价外贸	http://store.taobao.com/shop/view_shop.htm	tt外贸1	服饰鞋包	青岛	2014-4-13	8	65	80.00%	8	100%	0.0分	100.00%	共0人	暂无	暂无	暂无	暂无	暂无	0.0分	100.00%	共0人
28	兮兮的美衣	http://store.taobao.com/shop/view_shop.htm	angel、小可	服饰鞋包	青岛	2014-7-13	17	46	100.00%	17	100%	4.9分	69.37%	共13人	92.31%	7.69%	暂无	暂无	暂无	5.0分	100.00%	共12人
29	海星0724_2009	http://store.taobao.com/shop/view_shop-c315	海星0724_2009	服饰鞋包	青岛	2010-5-11	158	221	100.00%	150	94.94%	4.9分	77.88%	共36人	94.44%	5.56%	暂无	暂无	暂无	4.9分	69.69%	共34人
30	小C淘淘窝	http://store.taobao.com/shop/view_shop.htm	scx0222	珠宝/配饰	青岛	2015-1-5	4	550	100.00%	0	0.00%	5.0分	100.00%	共1人	100.00%	暂无	暂无	暂无	5.0分	100.00%	共1人	
31	炫出幸福	http://store.taobao.com/shop/view_shop-f838	炫出幸福	服饰鞋包	青岛	2010-10-17	315795	593	98.53%	314223	99.50%	4.6分	1.66%	共77180人	82.37%	8.52%	5.14%	1.68%	2.29%	4.7分	1.27%	共67350人
32	丝女装、真丝丝巾真丝腰	http://store.taobao.com/shop/view_shop.htm	songxiaolang	服饰鞋包	青岛	2007-11-8	7204	816	99.60%	5175	71.84%	4.8分	32.77%	共77人	88.31%	9.09%	1.30%	暂无	1.30%	4.8分	26.39%	共77人
33	索圣伊曼	http://store.taobao.com/shop/view_shop.htm	qqweichao1987	服饰鞋包	青岛	2010-11-17	7692	609	99.27%	1137	14.78%	4.9分	89.06%	共1056人	98.77%	0.38%	0.47%	0.10%	0.28%	4.9分	87.64%	共1042人
34	靓衣风尚vip	http://store.taobao.com/shop/view_shop.htm	夺命小砖头6216	服饰鞋包	青岛	2013-7-14	24	180	100.00%	17	70.83%	4.9分	76.58%	共17人	94.12%	5.88%	暂无	暂无	4.9分	69.69%	共17人	



RESTful

EXCEL

自定义爬虫

技术优势

1

WEB设置，简单易用



2

基于机器学习，精准抓取



3

多任务并行抓取，无需本地部署



4

专利技术，即抓即用，无重复任务

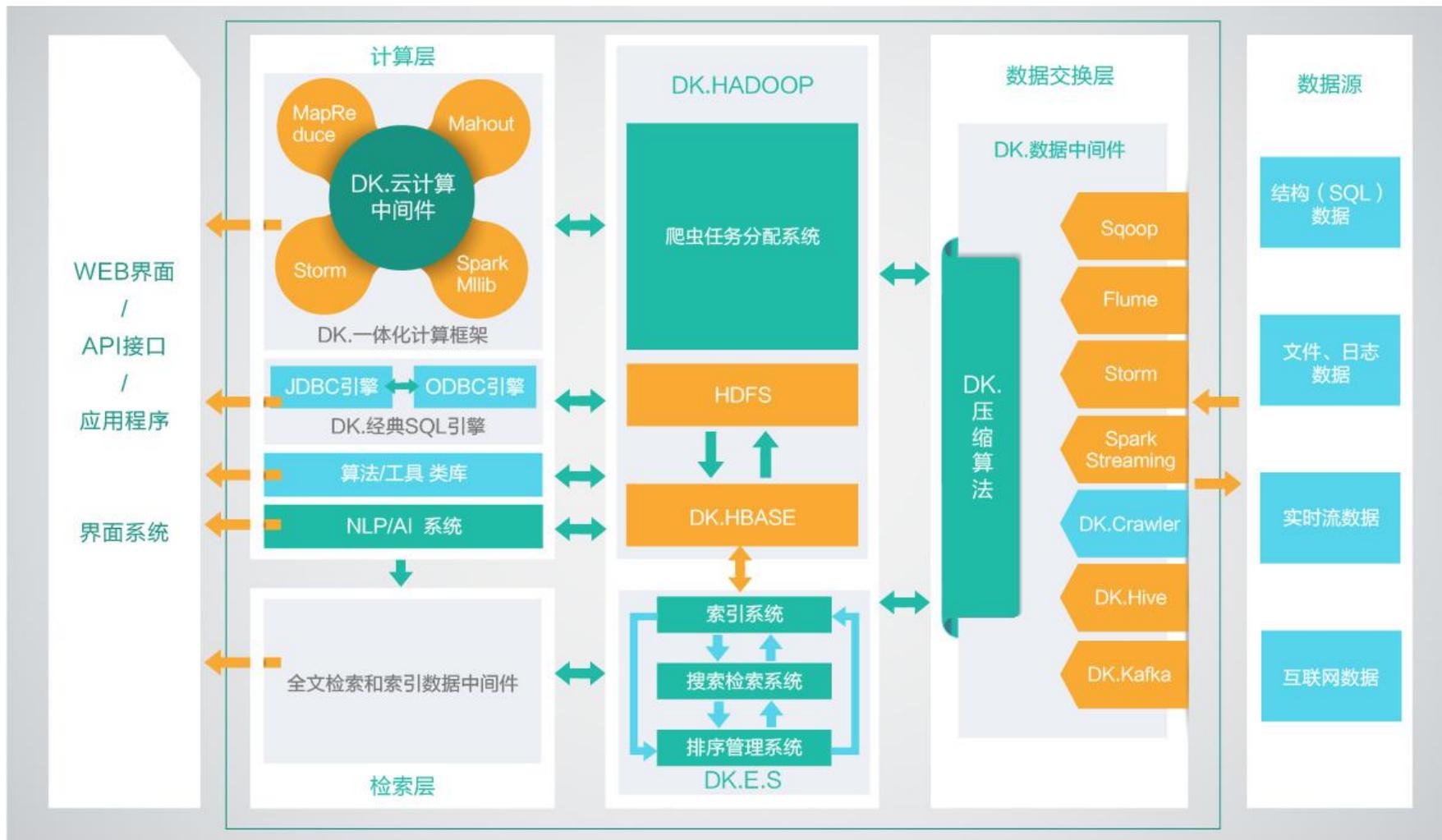


5

支持脚本编写，集成开发，功能强大



DKH企业级通用大数据计算平台



版本号：V1.2

通用大数据计算平台

技术优势

1

底层深度整合，一体化封装

2

各部件完美协同，大快开发框架集成，数据处理更高效

3

一体化部署，平台化安装，线性扩展，无需调试

大数据平台工程师的噩梦从此结束

通用大数据计算平台

技术优势



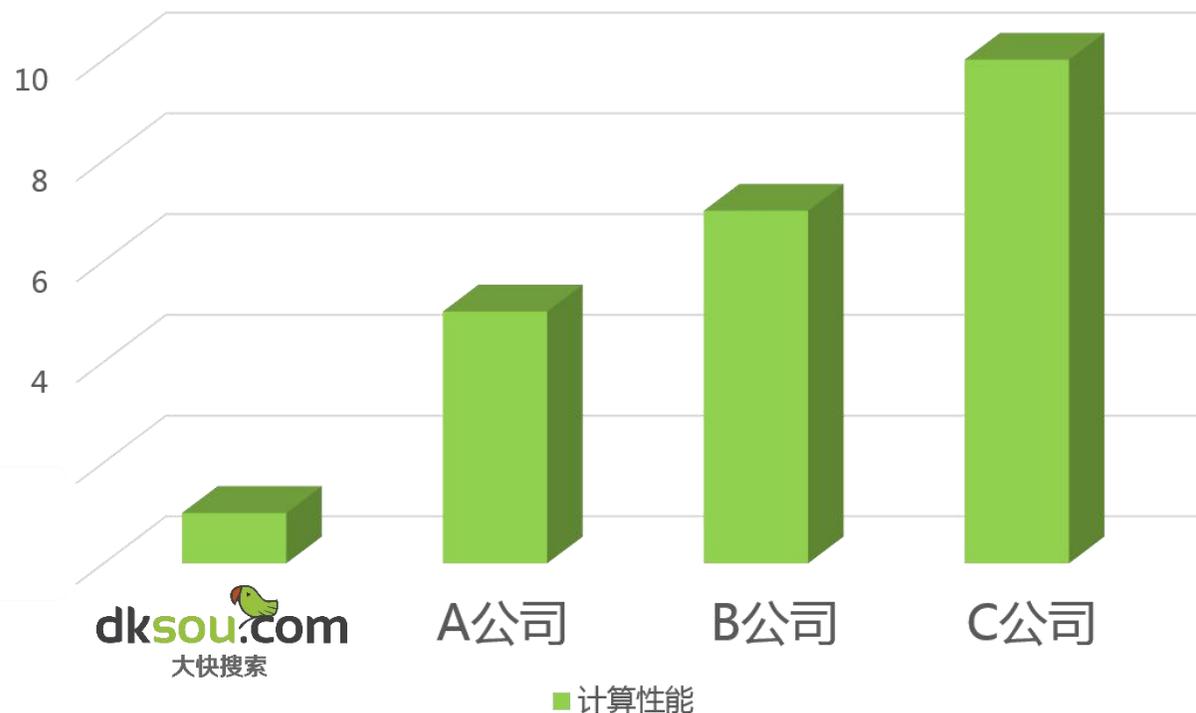
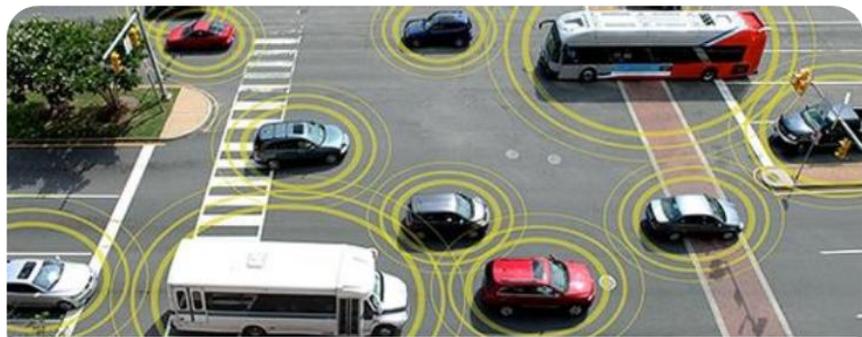
1 与开源平台100%兼容

2 主节点，管理节点，计算节点
集群节点简化至3种

3 维护更便捷

搜索引擎级大数据计算平台

100亿条交通流数据，按条件检索6000万条



Hisense 海信网络科技

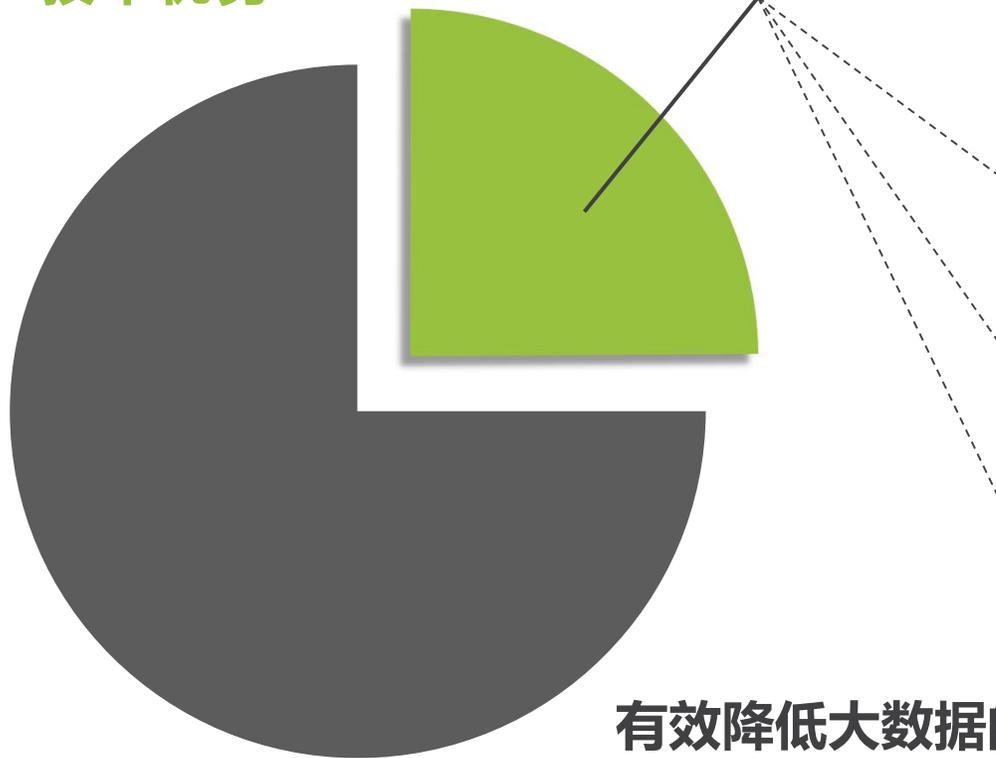


FreeRCH大数据一体化开发框架



大数据一体化开发框架

技术优势



1. 支持PHP\ .NET\ JAVA等主流开发语言
2. 集成度高，模块化编程，开发高效便捷
3. 与开源大数据系统100%兼容
4. 各框架组件可独立部署

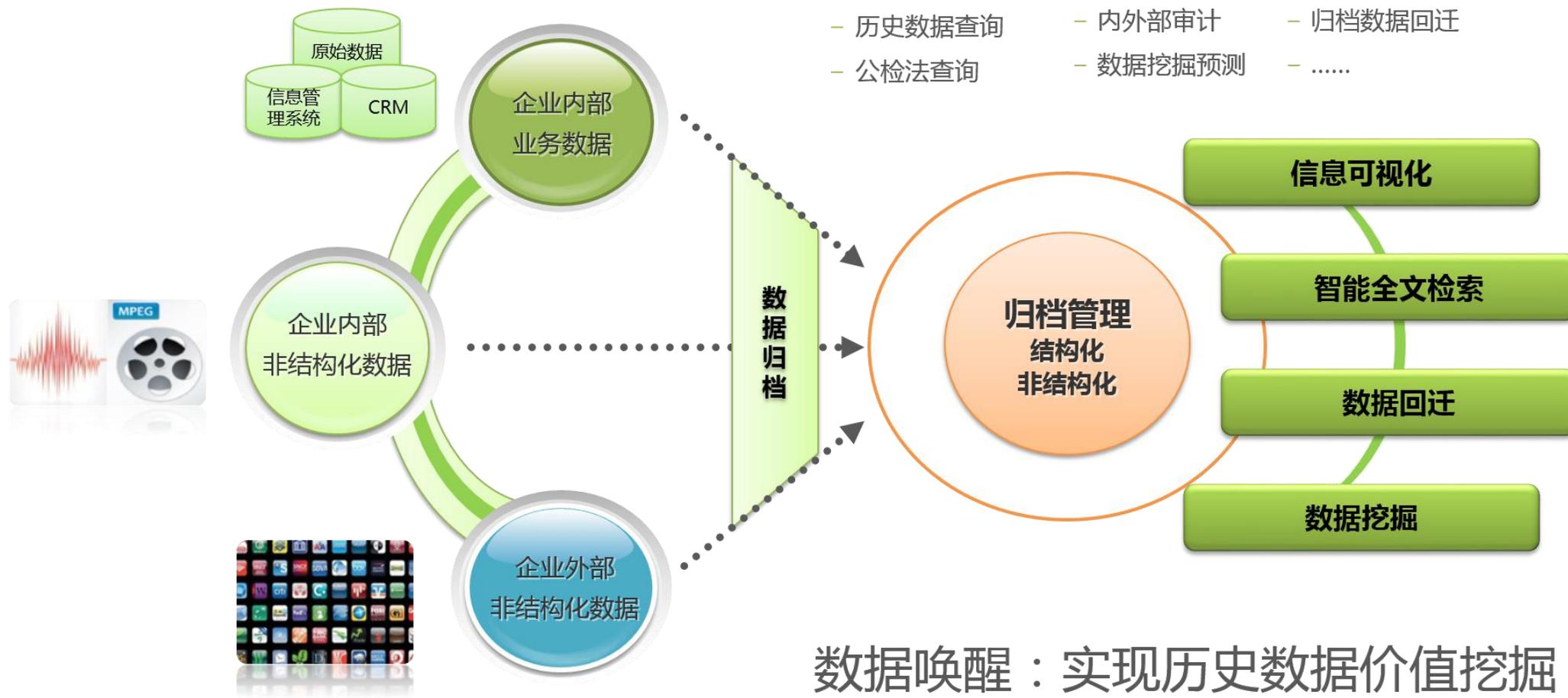
有效降低大数据的开发和学习难度：让每个程序员都能开发大数据应用，降低**10**倍开发成本，提高项目开发效率。

让应届毕业生也能开发出一流的大数据应用

永洪BI+大快搜索 无缝部署大数据项目

■ 应用场景：

- 历史数据查询
- 公检法查询
- 内外部审计
- 数据挖掘预测
- 归档数据回迁
-



数据唤醒：实现历史数据价值挖掘

■ 大快搜索 · 战略合作伙伴

Hisense 海信网络科技

Haier 海尔软件



中电瑞达
CERD



九次方大数据
JUSFOUN BIG DATA



HumanPool 汇思集团

大连华信

阿里云
aliyun.com



成功帮助合作伙伴高效部署和开发

dkSou.com
大快搜索

算法改变世界，数学打开未来之门





智慧运营 数造未来

大数据底层技术从此触手可及

