

电子商务语义库

阿里巴巴一淘及搜索事业部 陈维(随喜)

Suixi.cw@taobao.com



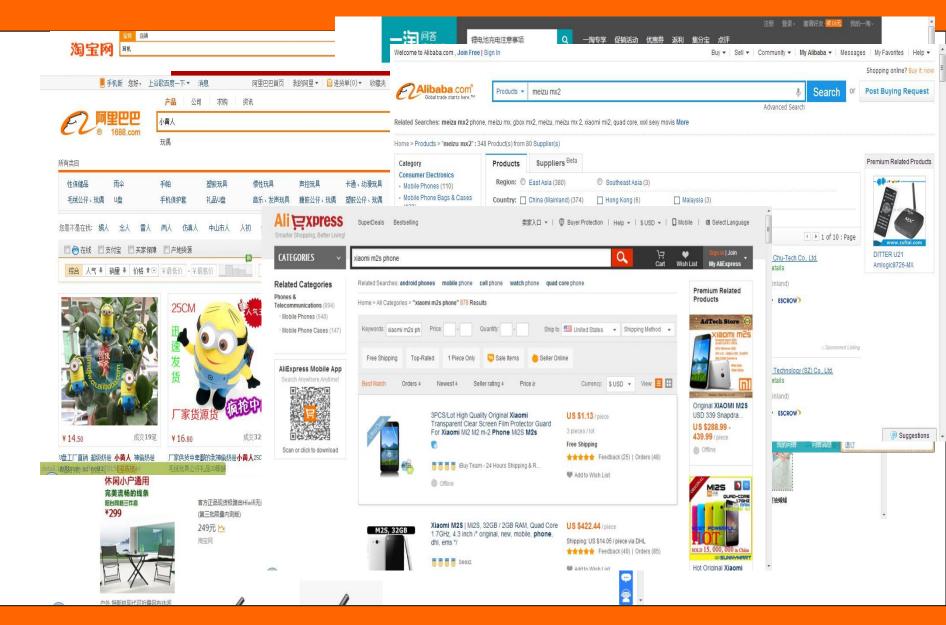
目录



- 关于阿里巴巴一淘及搜索事业部
- 阿里巴巴知识库
- 阿里巴巴产品库
- 电子商务语义搜索
- 挑战与未来

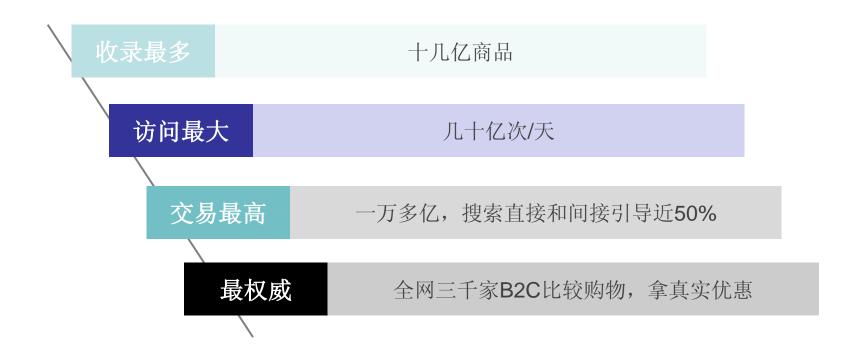
一淘及搜索事业部





四个"最"





目录



- 关于阿里巴巴一淘及搜索事业部
- 阿里巴巴知识库
- 阿里巴巴产品库
- 电子商务语义搜索
- 挑战与未来

2013-10-13 5

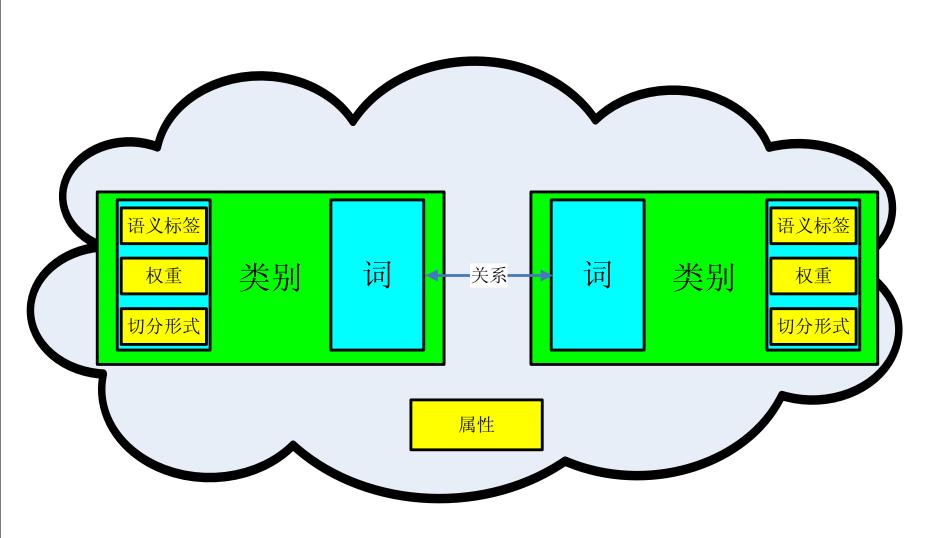
阿里巴巴知识库



基础库			电子	商务库	Ē				
基本词		人名	机	构名		属性词	ij	服务	
熟语		区划	产	品类型	修	饰词		促销词	
专有名词		色情词	品	牌	型号	#	规格	交易i	词
娱乐库			书文	库					
歌曲	赛事	体育	书	籍					
FLASH	i	电影电视	报	纸					
网站名	软件	游戏	杂	志					
生活库			杂类	类库					
房地产	公园		缩	写词					
餐馆酒店			邮	编					
财经									

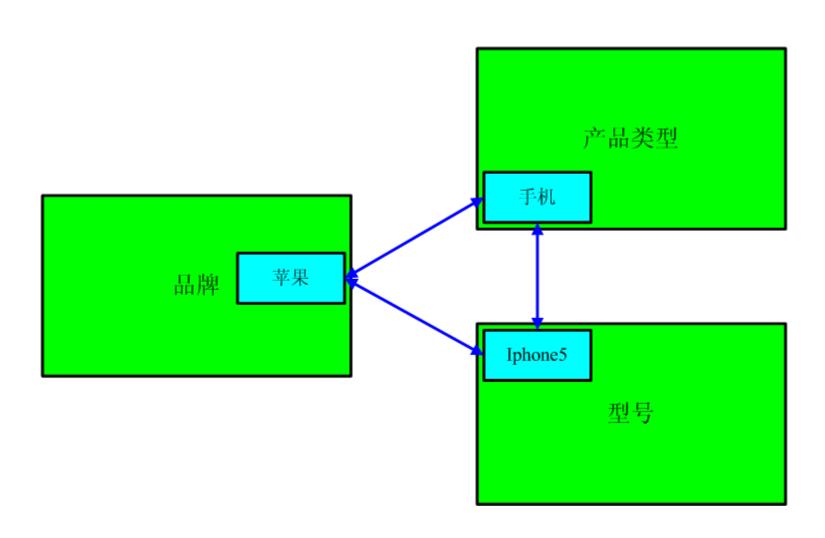
知识库----词条和关系





知识库----关系举例





基于知识库的文本分析



please input the title GO!						
标题 (title)	五折 包邮 卓欧2013夏三件套吊带雪纺连衣裙 F105002					
品牌 (Brand)	卓欧					
货号 (ItemNumber)	f105002					
产品类型 (ProductType)	吊带 连衣裙					
修饰词 (Modifiers)	雪纺					
数量/规格 (Quantifier)	三件套					
价格 (Price)						
特殊规 格 (InregularityQuantifier)						
过滤数量 (UselessAmount)						
颜色 (Color)						
折扣 (Discount)	五折					
赠品 (Giveaway)						
特殊服务 (SpecialService)	包邮					
时间词 (TimeWord)	2013夏					
促销词 (PromotionWord)						

目录



- 关于阿里巴巴一淘及搜索事业部
- 阿里巴巴知识库
- 阿里巴巴产品库
- 电子商务语义搜索
- 挑战与未来

商品与产品





产品库----行业





产品库---产品节点



APPLE/苹果 IPHONE 5S(16GB) iOS 7.0 手机 APPLE/苹果 IPHONE 5S 16G

▲ 纠错就得集分宝 〈 分享到

KIEHL'S/科颜氏小黄瓜水 KIEHL'S/科颜氏小黄瓜水

▲ 纠错就得集分宝 〈 分享到

小时代3.0 刺金时代

▲ 纠错就得集分宝 〈 分享到

DR BROWN/布朗博士 203 奶瓶

▲ 纠错就得集分宝 〈 分享到

₽ 降价提醒



参考价: 85.00元 (10个商家)

- 品牌: DR BROWN/布朗博士

■□径:标准□径

• 容量: 250ml/125ml

■形状:首身

直看更多参数:



• 材质: 玻璃

■是否带柄:否



一个月 150 100 8.24 9.1 9.9 9.17 9.25 10.3 10.11

与之前价格相比(2013年8月24日)-32.00元

价格走势

此产品目前处于降价中

网上商家(10)

产品参数

点评(55)

产品库----属性



APPLE/苹果 IPHONE 5S 16G 参数

基础参数	储存参数	摄像功能	数据传输	■WiFi: WIFI, IEEE 802.11 a/n/b/g ■蓝牙:支持蓝牙4.0 ■数据线接口:Lightning 接口
			Th 3th 42 .45	中外类型,不可控制等中外
	屏幕参数		电池参数	电池类型:不可拆卸式电池电池容量:1560mAh通话时间:600分钟待机时间:250小时
			其他参数	·颜色分类:黑色,金色,银色
			头心多数	
		个人助理		- 传感器类型:CMOS
	娱乐功能			- 重力感应:支持
				- 光线传感器:支持
				- 距离感应: 支持
			-GPS:内置GPS	, 支持GLONASS

- 通讯类型:单卡多模

•NFC功能:不支持

产品库----难点

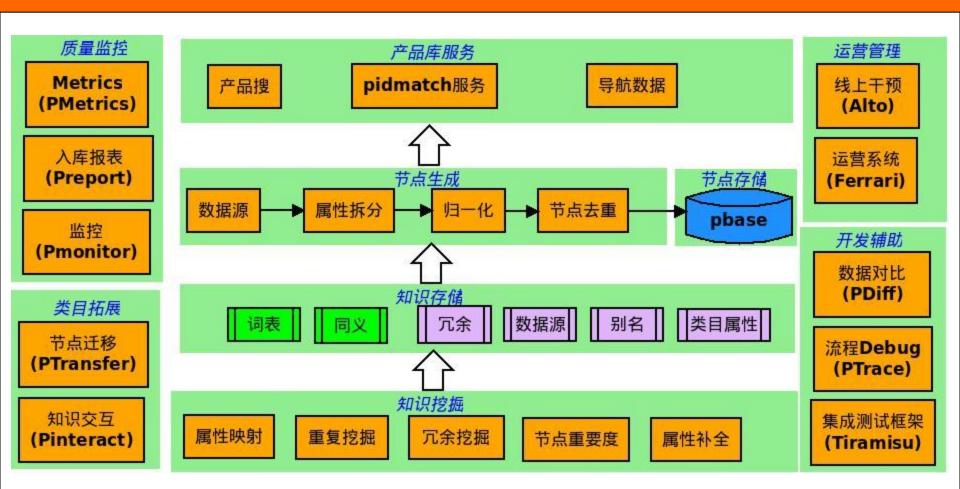


- 行业复杂性
 - 不同类目的属性不
 - 行业知识很多
- 商家多样化
 - 模板多样化
 - 表述多样化
- 重复率 vs 覆盖率
- 更新率



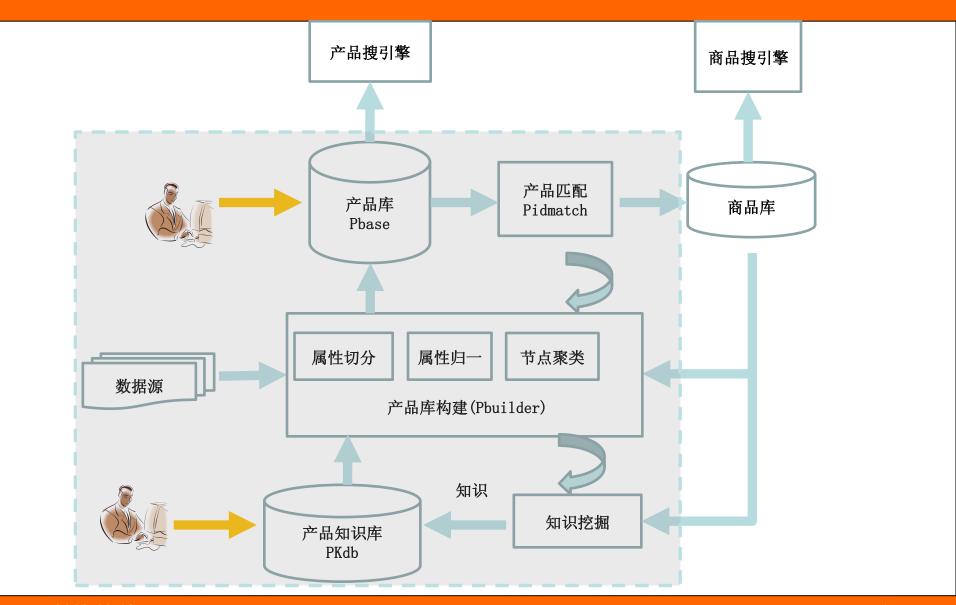
产品库----系统模块





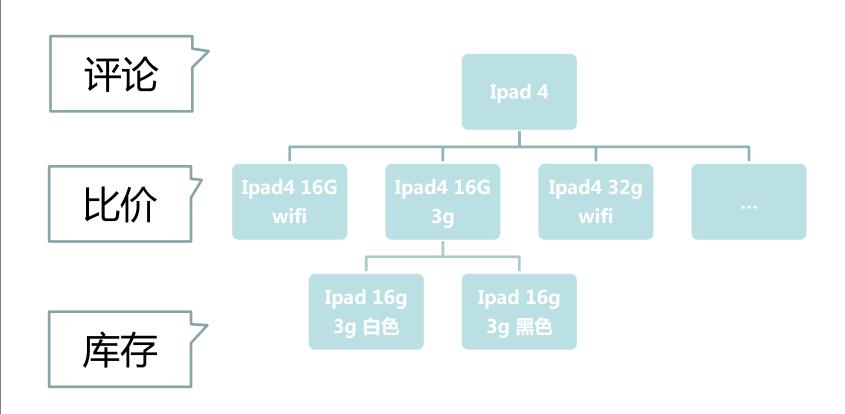
产品库----创建流程





产品库----关系





节点上下位关系

产品库----质量评估



- 按行业区分
- 按产品节点
 - 节点覆盖率
 - 节点重复率
 - 节点质量
 - 属性

	—— 1. L. ———————————————————————————————	
lacktriangle	ト社	

• 价格

.		覆盖率	重复率	属性 准确率	主图 质量	商品下挂 准确率	参考价 准确性
	3c	96.00%	0.87%	94.59%	98.68%	98.77%	99.94%
	化妆品	91.68%	0.91%	96.34%	98.60%	98.44%	86.97%

产品库----点评和资讯



APPLE/苹果 IPHONE 5S 16G 测评

优质皮革手感极佳 iPhone5s Case评测

每当我们购买了一款新手机之后,大多数用户第一件要做的事情是什么呢?相信很多人的回答都会是为心爱的手机配备一个好看又安全的手机壳,而这一点对于相对脆弱的Phone来说就更加重要了。... 查看全文

来源:中关村在线

两极化设计 iPhone5s外观对比iPhone5c

今天是中秋三天小长假后的第一天工作日,大家都过得如何呢?当然并不是所有人都能够第一时间回家和家人过节团聚,而我们ZOL报道团队也特此为大家带来了9月20日iPhone5s/5c的首销报道,算是为大家带来的... 查看全文

宋源:中关村在线

800W相差几何 iPhone5s拍照对比iPhone5

在苹果Phone5s的新卖点中,我们惊喜的发现其实它更多的是一种内在创新的体现,例如指纹识别、A7处理器的加入等等。不过摄像头方面的升级倒是引起了人们的猜疑,因为iPhone5s拍摄能力的提升从表面上来… **查看全文**

来源:中关村在线

不同时代之作 苹果iPhone5c对比iPhone5

苹果在今年首次在一年内推出两款Phone新机——在Phone5s之外的Phone5c在发布前便已经大量曝光。发布后,大部分人对这款手机的普遍印象便是"上代Phone5的塑料版"。的确,苹果在发布会当中也并未否… <mark>查看全文</mark>

来源:中关村在线

查看更多测评

目录



- 关于阿里巴巴一淘及搜索事业部
- 阿里巴巴知识库
- 阿里巴巴产品库
- 电子商务语义搜索
- 挑战与未来

电子商务语义搜索



- 产品库 + 知识库 = 语义库
- 商品分析 + 文本分析 = 语义分析
- 语义搜索 = 结构化搜索 + graph关联
 - 文本结构化
 - 商品结构化
 - Query结构化

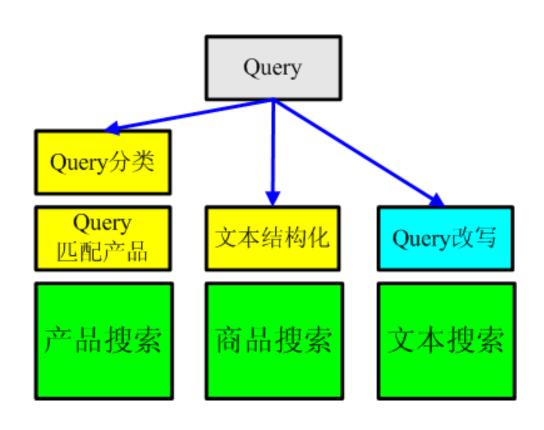
搜索offline



商品搜索 产品搜索 文本搜索 参考价计算 Sku抽取 商品聚合 关键词提取 商品结构化 匹配产品 文本分类 文本结构化 商品分类 分词 词表 属性

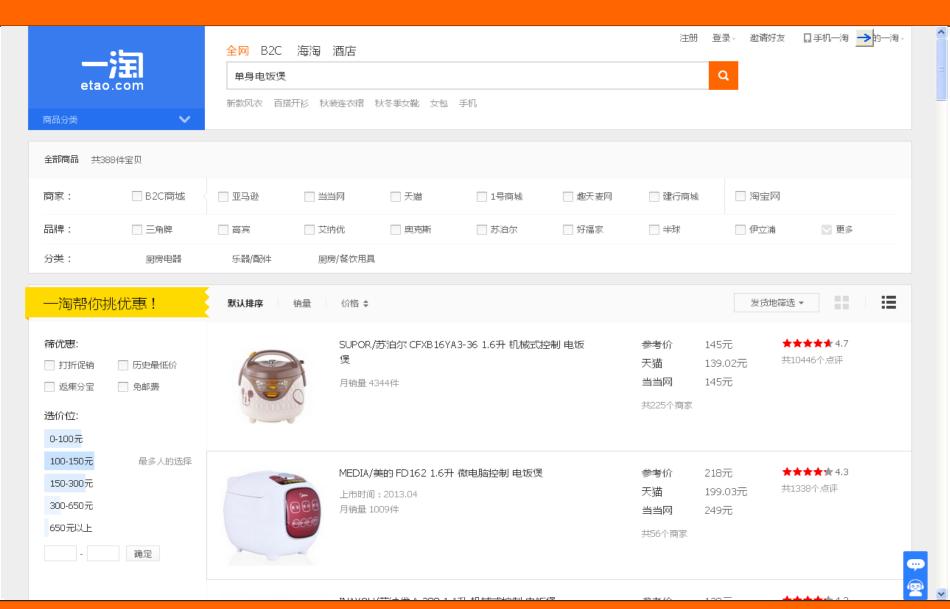
搜索online





搜索示例





目录



- 关于阿里巴巴一淘及搜索事业部
- 阿里巴巴知识库
- 阿里巴巴产品库
- 电子商务语义搜索
- 挑战与未来

挑战和未来



- 难以定义的节点粒度
- 上下位关系 vs 网状关系



- 产品库改进
 - -丰富关系
 - 节点与特色标签的关联
- 基于淘宝详情页的商品百科

Thank You!



