



美柚

美团点评  
技术团队

IT大咖说  
知识分享平台

# 个性化信息流构建实践

陈伟平

# 目 录

01

产品背景

02

策略算法

03

架构演进

04

效果评估

# 1

## 产品背景



美柚：近千万级dau 工具型APP

痛点：用完就走 需要才来 留存时间短

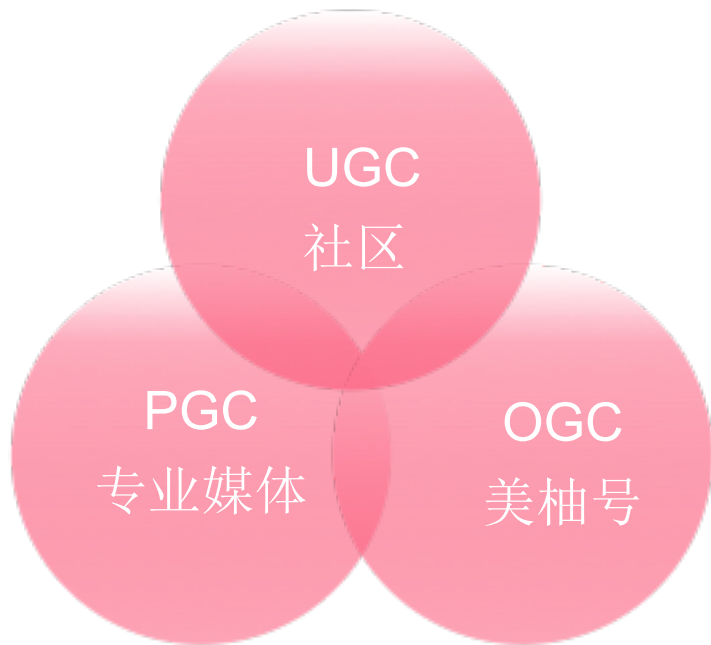
愿景：爱我别走



社区：有效提升留存

痛点：数量少 内容差 满意度低

愿景：提供有价值内容



UGC:美柚社区

PGC:网易 腾讯等专业媒体

OGC:美柚自媒体平台

内容池：海量内容 数十万+

愿景：千人千面 各取所需

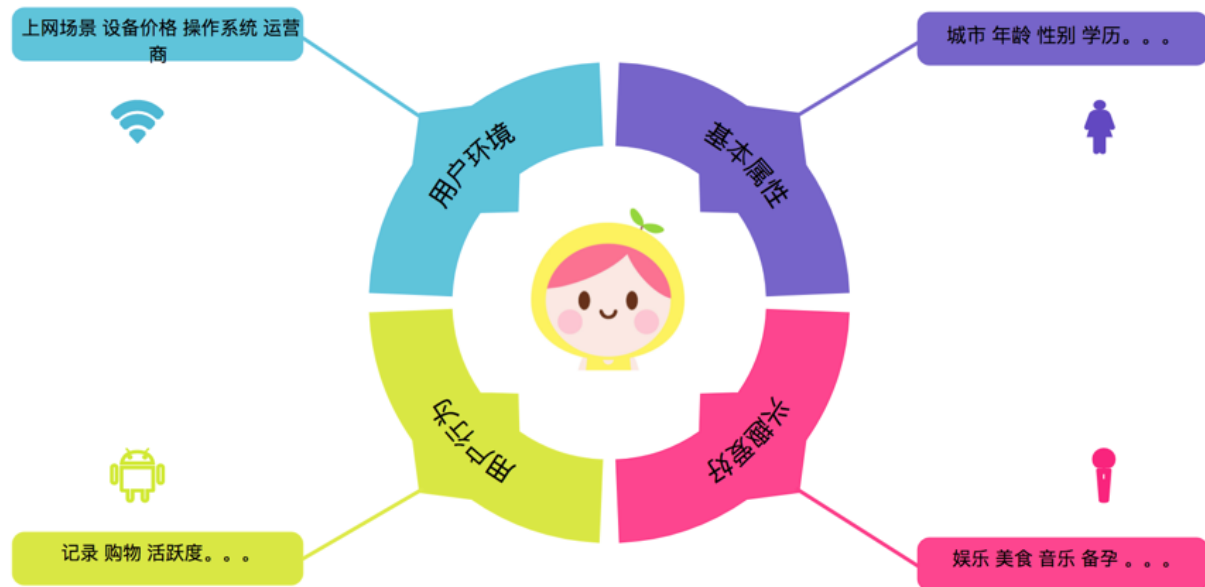
## 产品形态：个性化信息流



# 2

## 策略算法

## 推荐基础：用户画像





## Item画像：标签提取

### 1 基于K-means提取

步骤：词语自聚类 整理标签 由词标签向量计算文本标签

问题：聚类时间长 数目不好控 难以概括每一类词语 新增困难

### 2 基于Bayes提取

问题：无法解决语料中一词多义和多词一义的问题 它更像是词法分析，而非语义分析

### 3 基于改进式Bayes提取

## 标签体系

行业分类 23

娱乐八卦

体育休闲

新闻时事

孕育指南

...

类型标签 74

明星

居家

星座

情感

...

内容标签 10万+

周杰伦

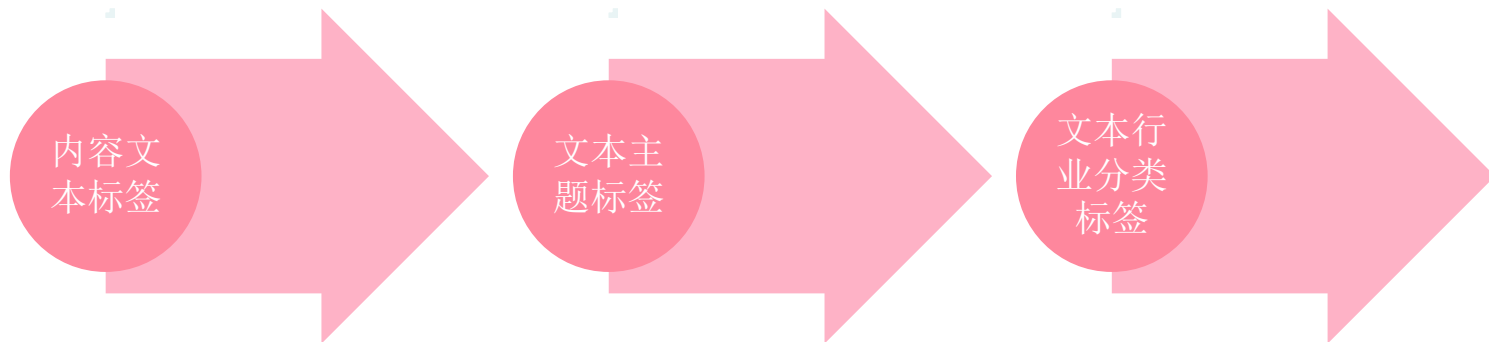
跑男

婆婆

宝宝

...

## 文本内容识别步骤



## 用户兴趣画像更新





## 其他策略

运营要闻推荐  
新内容定向投放  
关注用户  
活动专题  
多样性  
小说

女性身份信息  
设备信息  
地域

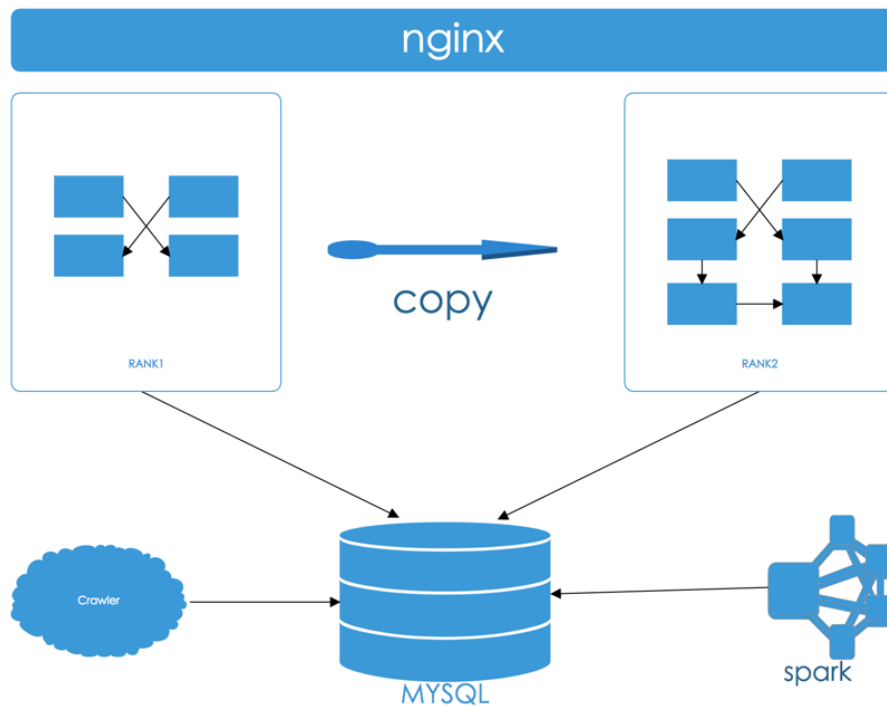
- 
- 
- 



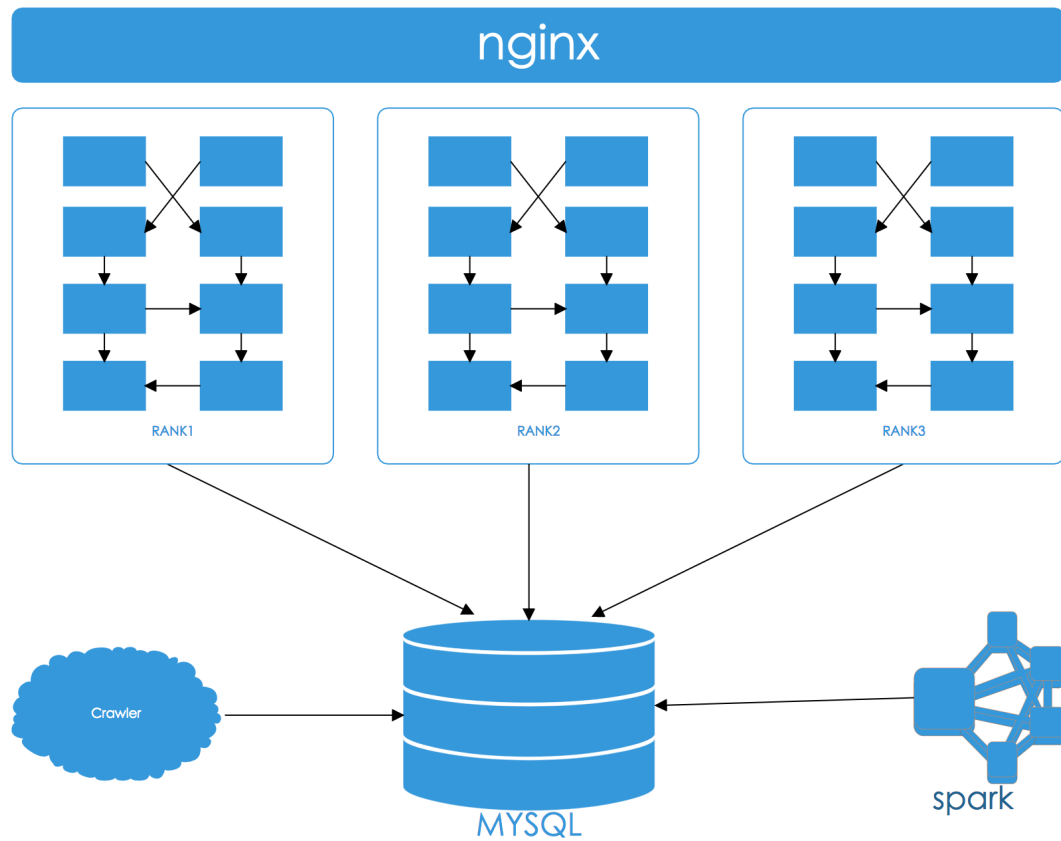
3

# 架构演进

## 架构1.0: 快



业务增长 需求复杂



## 架构1.0



模块多  
依赖大  
环境复杂  
难以升级

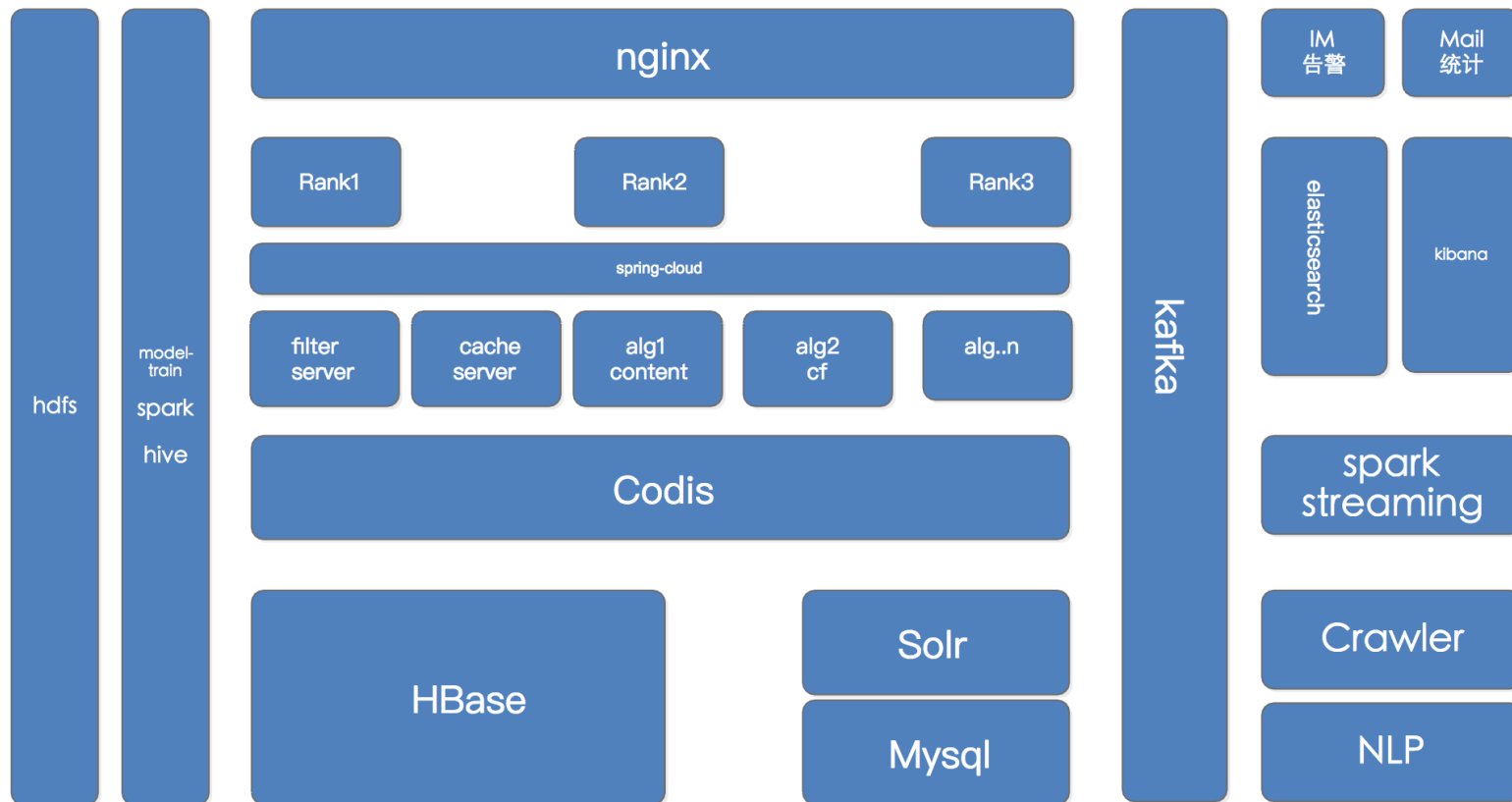


## 架构2.0



模块解耦  
调用灵活  
横向扩展  
高可用容错

# 架构 2.0



|          | 架构1.0（单体服务）                         | 架构1.1（共用库）                      | 架构2.0（微服务）                                                             |
|----------|-------------------------------------|---------------------------------|------------------------------------------------------------------------|
| 易用性      | 本地调用（胜）                             | 共享库引用 + 本地调用                    | API 调用                                                                 |
| 系统设计     | 高手能实现低耦合<br>数据最终还是耦合一起<br>一荣俱荣 一损俱损 | 在使用方的源代码中引用共<br>享库的类和方法<br>数据耦合 | 业务包装成服务，数<br>据代码物理上隔离，<br>实现高内聚低耦合相<br>对容易<br>（胜）                      |
| 升级，Bug修复 | 整体变动，容易导致不<br>相关模块的意外失败             | 改源代码，使用方重新打包                    | 单独升级，其它模块<br>无感知，容易实施敏<br>捷开发方法（胜）                                     |
| 运维效率     | 简单直接（胜）                             | 相对简单                            | 大系统被拆分为多个<br>小系统，部署和运维<br>难度加大，但可以利<br>用 <b>DevOps</b> 等方式将<br>运维工作自动化 |

总结：快速迭代 定期重构 结合业务

# 3

## 效果评估



用户轨迹查询  
异常排查

分析与评估

首页推荐

用户推荐列表

资讯相关推荐

用户标签查询

用户浏览记录查询

用户电商浏览记录查询

用户算法排序查询

相似资讯排查

首页新版指标统计

新版搜索

搜索数据分析

查询

UserID:

5

标签版本:

标签v2新模型

查询

用户标签信息

显示 5 项结果 搜索:

| TagID | 名称   | 权重      |
|-------|------|---------|
| 5     | 明星   | 0.0733  |
| 6     | 搞笑段子 | 0.05297 |
| 25    | 情感   | 0.04532 |
| 32    | 怀孕   | 0.03888 |
| 50    | 职场   | 0.03813 |

显示第 1 至 5 项结果, 共 60 项

首页 上一页 1 2 3 4 5 ... 12 下一页 末页

用户推荐列表

显示 10 项结果 搜索:

| ID       | 标题                               | 内容                                                                                                               | 类型 | 推荐时间                | 时间                    | 标签                      | 关键词                                                                  | 来源    | 点击率                 | 阅读时长  |
|----------|----------------------------------|------------------------------------------------------------------------------------------------------------------|----|---------------------|-----------------------|-------------------------|----------------------------------------------------------------------|-------|---------------------|-------|
| 9964625  | 杨幂现身, 大家却被她的下装吸引, 网友: 这裙子要火, 快去买 | 如今的杨幂可谓是事业红火, 人气爆棚, 每一次的出现都能给无数人带来不一样的惊喜。每次穿衣都能引领潮流, 让粉丝们纷纷求同款。前不久, 杨幂亮相《明日之子》录制灰色毛衣搭配牛仔短裙, 兔子"短裙", 大秀美腿, 同时搭... | 资讯 | 2017-10-16 14:19:06 | 2017-10-10 08:22:43.0 | 明星:0.23857; 服饰:0.50483; | 杨幂:2.0033333333333334; 要火:0.8999999999999999; 裙装:1.6833333333333331; | 关键字推荐 | 778324/2387520=0.33 | 21.75 |
| 10163877 | 1岁宝宝醒来后发现没有大人, 监控拍下的一幕惊呆了        | 你可能不会想像的到, 1岁的宝宝虽然还不会说话, 但他们心中其实什么都明白, 他们的聪明超乎你的想像。最近这组照片在网络上走红, 这个刚好1岁的小男孩的                                     | 资讯 | 2017-10-16 14:19:06 | 2017-10-15 17:00:05.0 | 社会民生:0.66368;           | 宝宝:1.3479999999999999; 妈妈:0.8119999999999999;                        | 要闻    | 143484/611534=0.23  | 30.00 |



T h a n k   y o u