

旅游行业的数据应用

去哪儿网 搜索SI 涂强

Geekbang>

极客邦科技

全球领先的技术人学习和交流平台

扫我，码上开启新世界



Geekbang>

InfoQ | EGO | StuQ

InfoQ

专注中高端技术
人员的社区媒体

EGO

EXTRA GEEKS' ORGANIZATION
NETWORKS

高端技术人员
学习型社交网络

StuQ

实践驱动的IT职业
学习和服务平台

InfoQ new

促进软件开发领域知识与创新的传播

ArchSummit
全球架构师峰会

实践第一 案例为主

时间：2015年12月18-19日 / 地点：北京·国际会议中心

欢迎您参加ArchSummit北京2015，技术因你而不同



ArchSummit北京二维码

QCon
全球软件开发大会

【北京站】

2016年04月21日-23日



全球最大的中文旅游平台

成立时间

2005年成立至今，已有10年历史

上市时间

2013年 美国纳斯达克上市

用户情况

月活跃用户**4千万**，累计下载用户**10亿**！

预订率

网络预订100%；无线预订率85%

员工数量

8000+员工；60%产品开发人员

关于我

- 2008年 北京大学语言信息工程系硕士毕业
- 2013年 北京大学国际MBA

- 2007.6~2008.9 豆瓣
- 2008.9~2013.9 搜狗
- 2013.9~今
 - 无线商业战略投资总监
 - 无线创新研发总监
 - 无线智能旅游SI负责人

在线旅游的现状

- Online售卖比例越来越高
- 千万级SKU
- 旅游无线化
- 碎片化

意义

- 减少用户的操作路径
 - 发现规律
 - 效率
- 服务无处不在

我们的产品

- 全球第一个智能酒店预订系统 4006663223
- 无线推荐
- Qunar的无线智能搜索体系

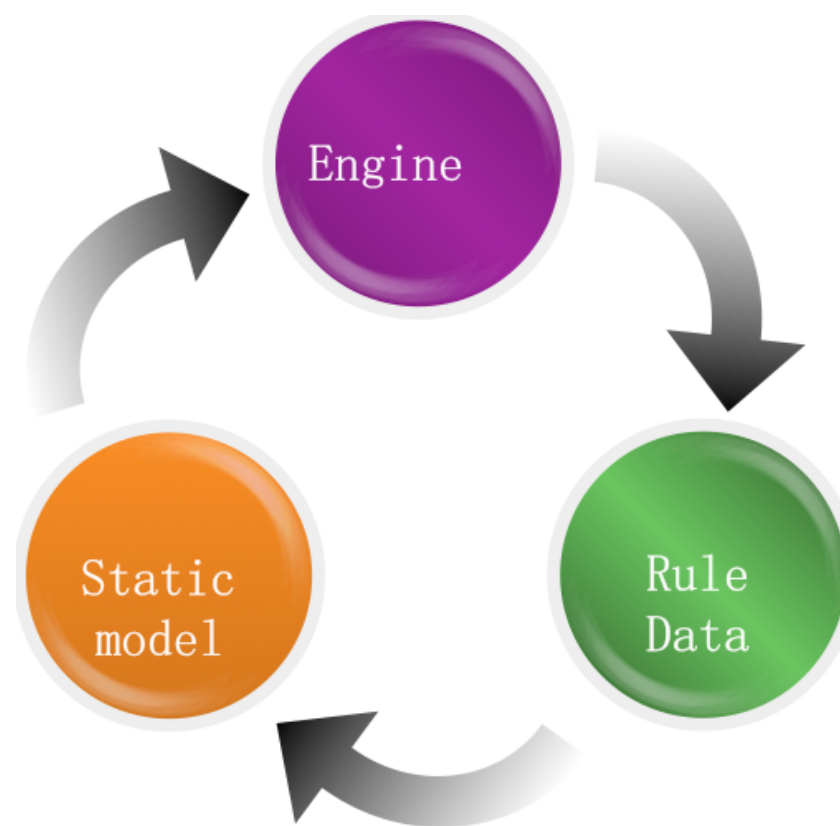


智能语音预定系统

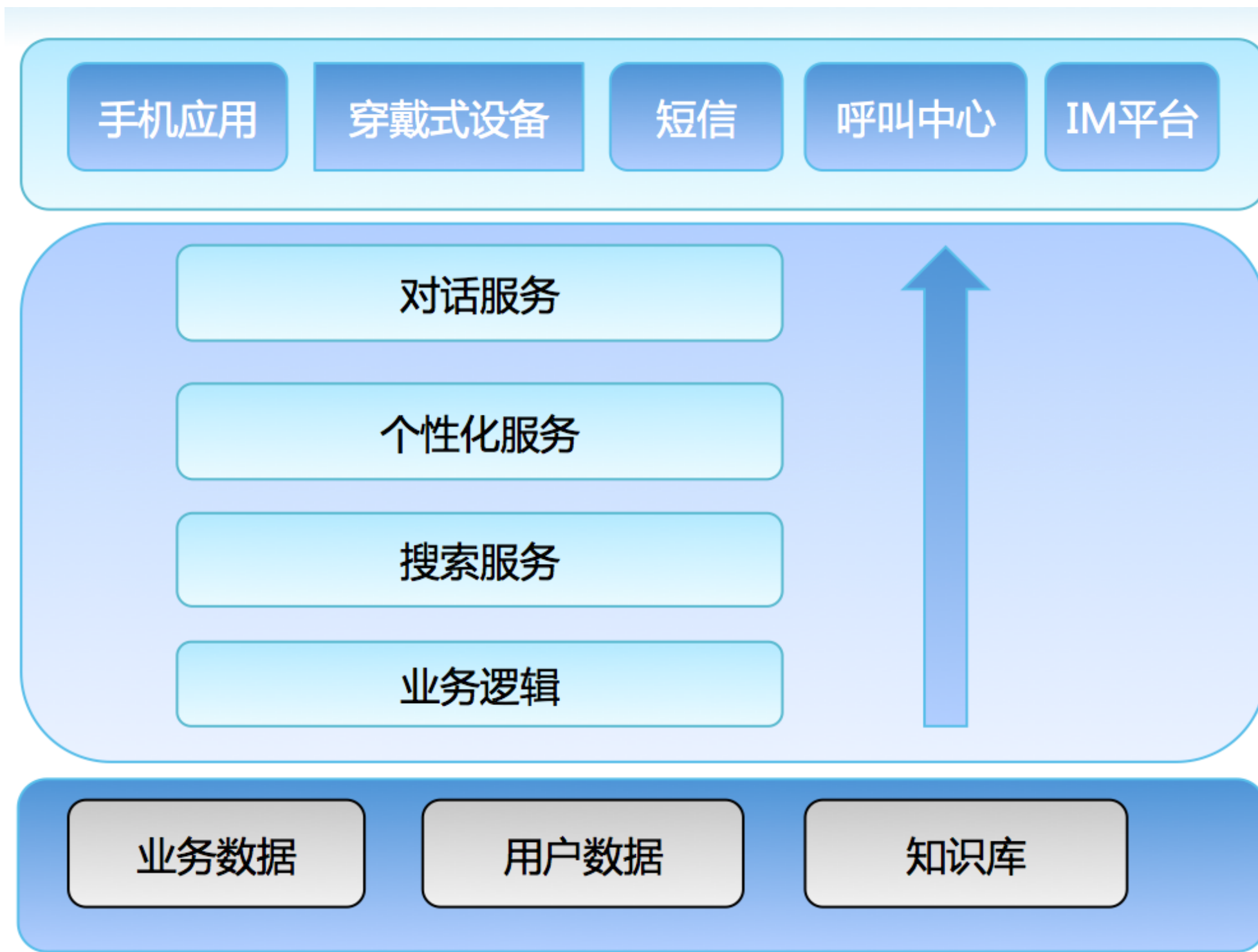
- 有目标和状态多轮对话系统
- ASR & TTS

语义服务

Engine	对话引擎，作为容器，支撑各种规则和模型数据
统计模型	实体识别模型，文本分类模型等
规则数据	pattern数据，entrance数据，flow数据，action脚本数据，各种词典数据



对话应用



人机对话核心指标

- 旅游领域语音识别句准确率：85+%
- 语义识别准确率：90%
- 关键业务词识别准确率：95%
- U20：20%

排序推荐演变

- 基于经验的score体系
- 加入点击反馈
- 机器学习的加入(LR)
- $P(\text{Order}|\text{Hotel})$ 演变到 $P(\text{Order}|\text{Hotel}, \text{User})$
- 特征工程

如何评估?

- 线下评估
- AB Test
- U2O提升:20%~40%

挑战

- 数据抽取与组织&业务理解
- 数据稀疏
- POI因素

为什么做搜索

- 旅游是一个方案
 - 基于供应链产品 转到 基于用户需求
 - 跨业务
 - 排序即推荐
 - 发现

搜索遇到的挑战

- 需求识别
- 基础数据&业务数据组织
- NLP
- 知识图谱
- 各种规则

未来的旅游

- “明天去上海出差，3天时间”
- “我下个月有五天年假”

智能旅游

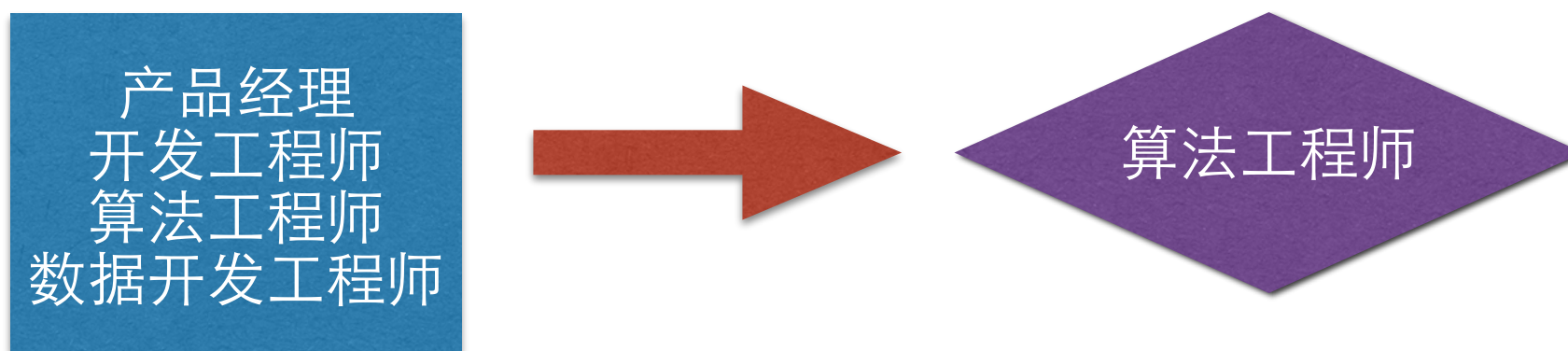
用户

搜索
推荐
行程规划
人机对话
智能客服
商业化
收益管理
营销平台

供应链

基础数据

我们的方式





邀请你的加入

- 资深产品经理
- 资深算法工程师



谢谢

微信:22016557