

AI筑巢：机器学习在 百度凤巢的深度应用

刘斌新

百度 商务搜索智能交互部负责人

2017.7



CNUTCon 2017

全球运维技术大会

上海·光大会展中心大酒店 | 2017.9.10-11

智能时代的新运维

大数据运维
安全
SRE
DevOps
Kubernetes
Serverless
游戏运维
AI Ops
智能化运维
基础架构
监控
互联网金融



StuQ

斯达克学院

实践驱动的IT教育



斯达克学院(StuQ)，极客邦旗下实践驱动的IT教育平台。通过线下和线上多种形式的综合学习解决方案，帮助IT从业者和研发团队提升技能水平。



10大职业技术领域课程

<http://www.stuq.org>

SPEAKER INTRODUCE

刘斌新

- 现负责凤巢核心策略、产品团队，包括CTR预估、动态内容、富媒体广告等方向。
- 2010年正式加入百度，作为负责人先后完成移动凤巢、应用商店变现、联盟流量、信息流变现等关键业务的攻坚；
- 带领团队连续4年6个项目入围最高奖TOP 10并两获最高奖（百万美金大奖）。



TABLE OF
CONTENTS 大纲

- 搜索广告业务简介
- CTR预估的前世今生
- 生成式广告触发
- 机器学习在内容生态的创新
- 未来的探索

搜索广告业务简介



关键词购买

按点击计费

GSP二价计费

搜索广告业务简介

技术围绕业务场景

CTR预估、触发、文案生成、样式选择、弹性计算、风控.....
针对业务优化点，机器学习在凤巢几乎无处不在

例如：

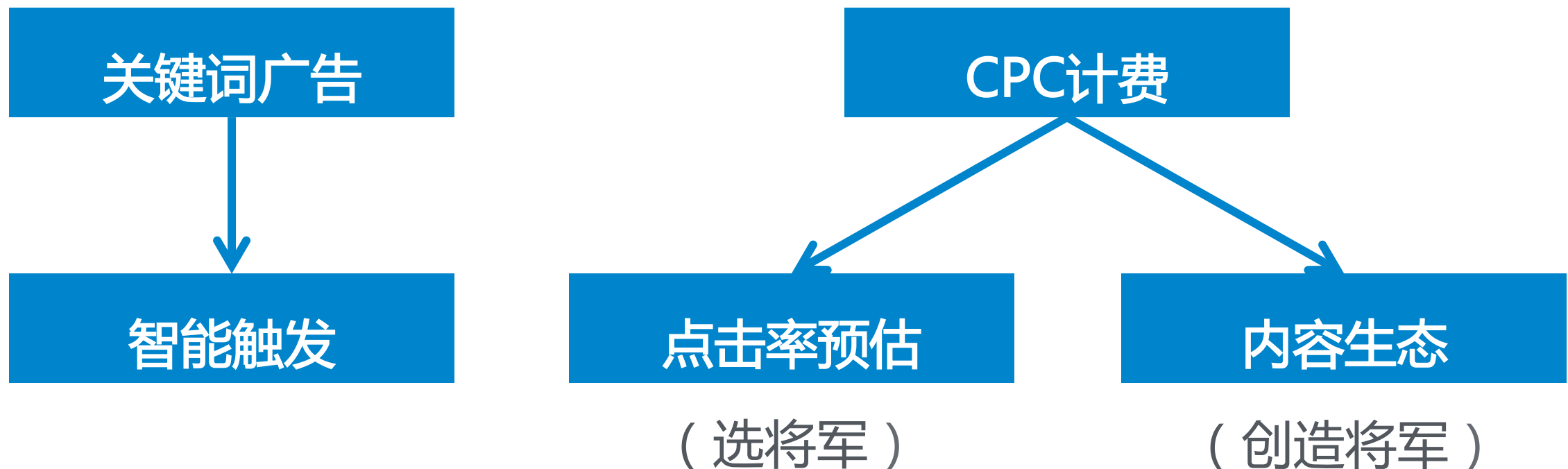
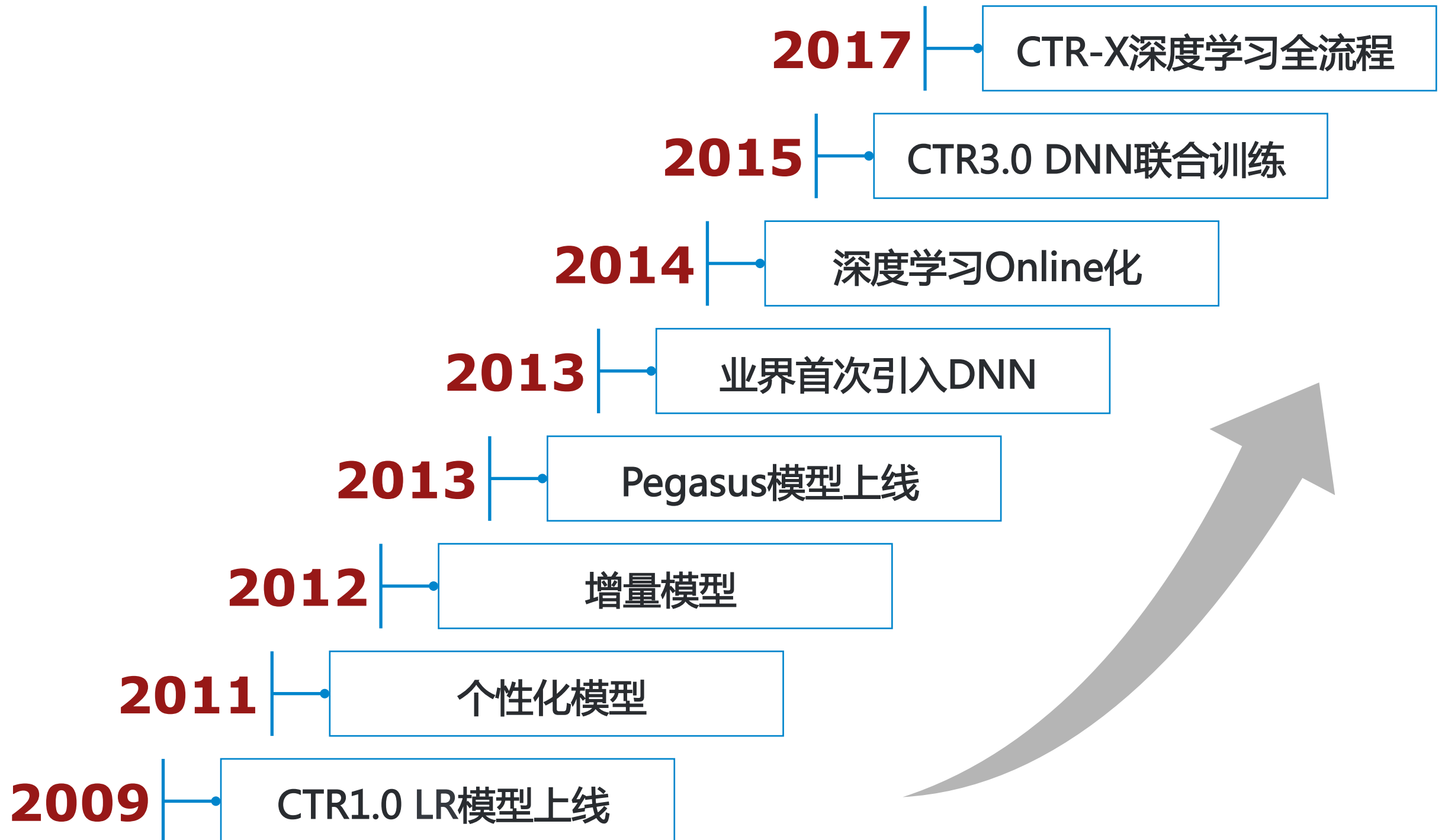


TABLE OF CONTENTS 大纲

- 搜索广告业务简介
- CTR预估的前世今生
- 生成式广告触发
- 机器学习在内容生态的创新
- 未来的探索

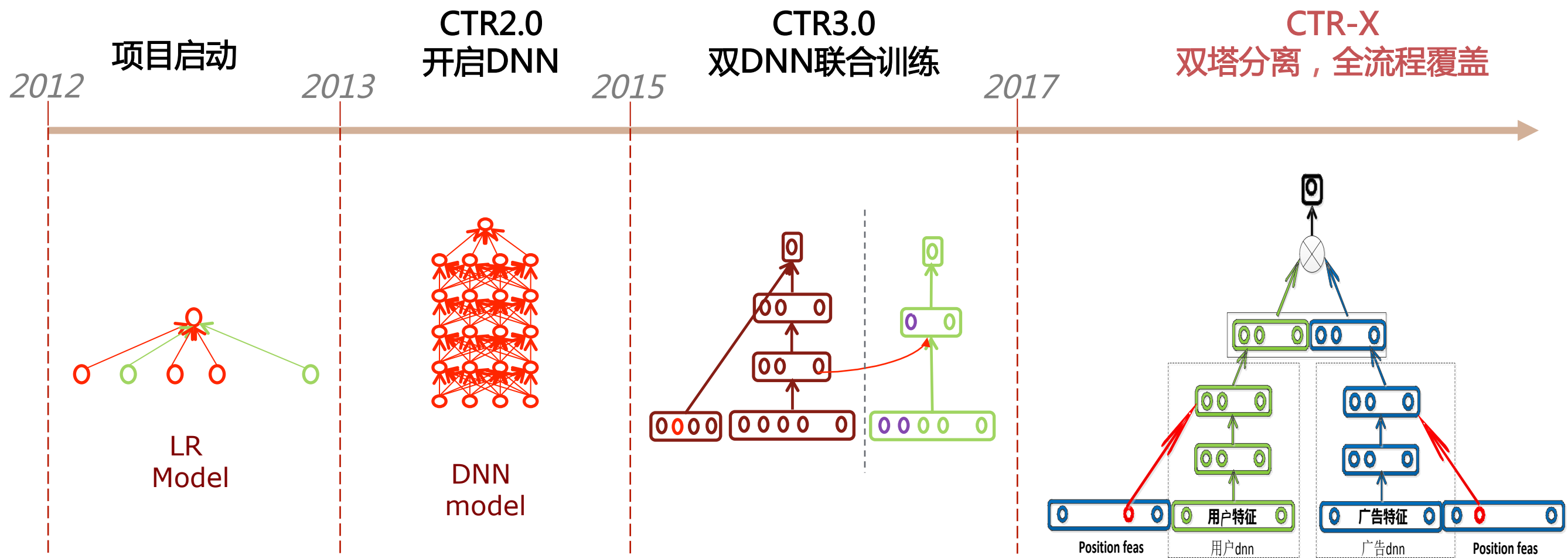
CTR预估发展回顾(2009~2017)



> 5000

experiments

深度学习引入CTR进展回顾

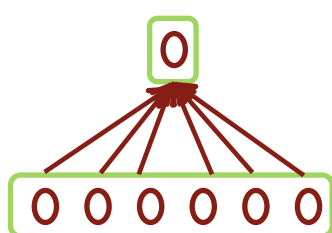


CTR-X：深度学习覆盖基础排序全流程

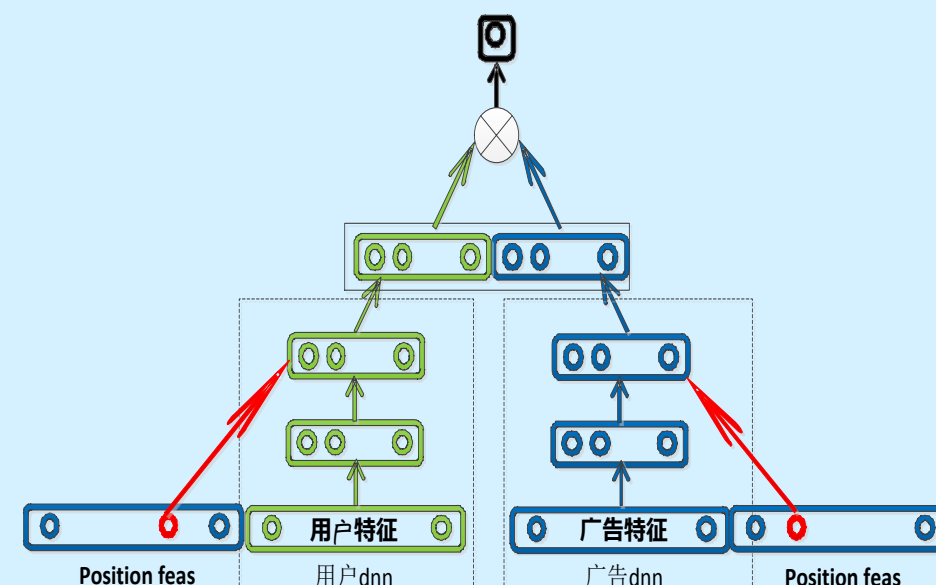
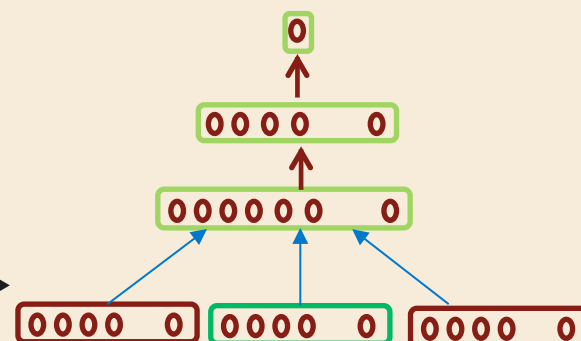
- 广告初筛阶段，计算量巨大，使用DNN有挑战。
- 双塔分离DNN模型架构，优化计算逻辑和模型泛化能力，提升变现能力。

Before：LR模型

After：双塔分离DNN模型



普通DNN模型



✗ 长耗时/资源损耗，无法由LR直接升级到DNN

✓ 用户&广告两个维度的模型分开训练
由LR升级为双塔大规模离散DNN

耗时
优化

特征抽取：用户侧、AD侧分别抽取

分布式字典：网络查询 → 内存查询

计算方面：10万级别的float相乘 → 32个float相乘

自有大规模机器学习平台

Abacus | 超大规模

复杂性

可支持**千亿**级别特征和**万亿**模型参数

- ✓ 允许多机多线程异步实时访问和更新
- ✓ 用于支持大规模离散模型（如离散DNN）的异步sgd训练

扩展性

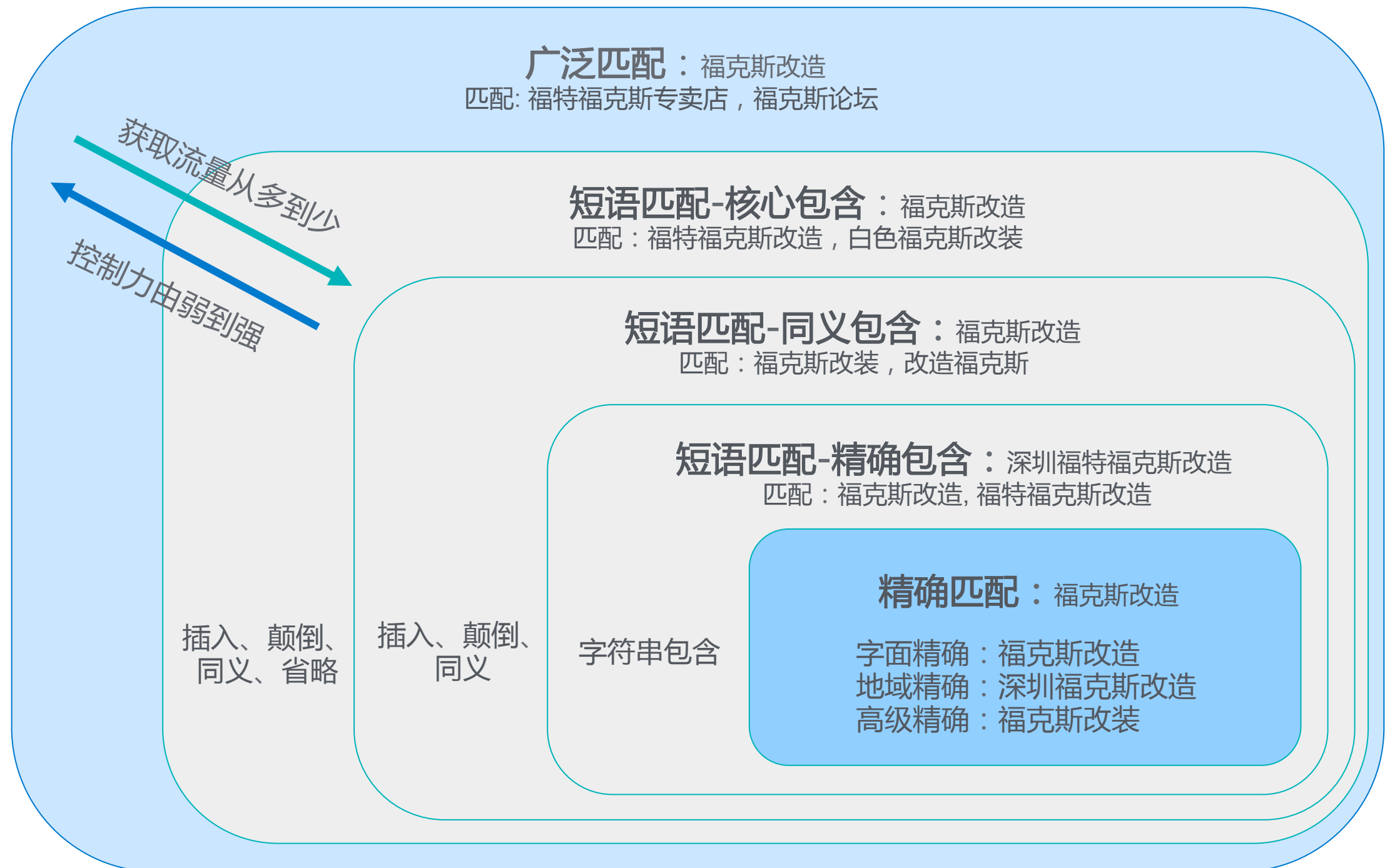
自主的**Anti Over-Fitting**机制，可随意使用扩展

- ✓ 应用于凤巢CTR模型
- ✓ 应用于手百Feed推荐模型

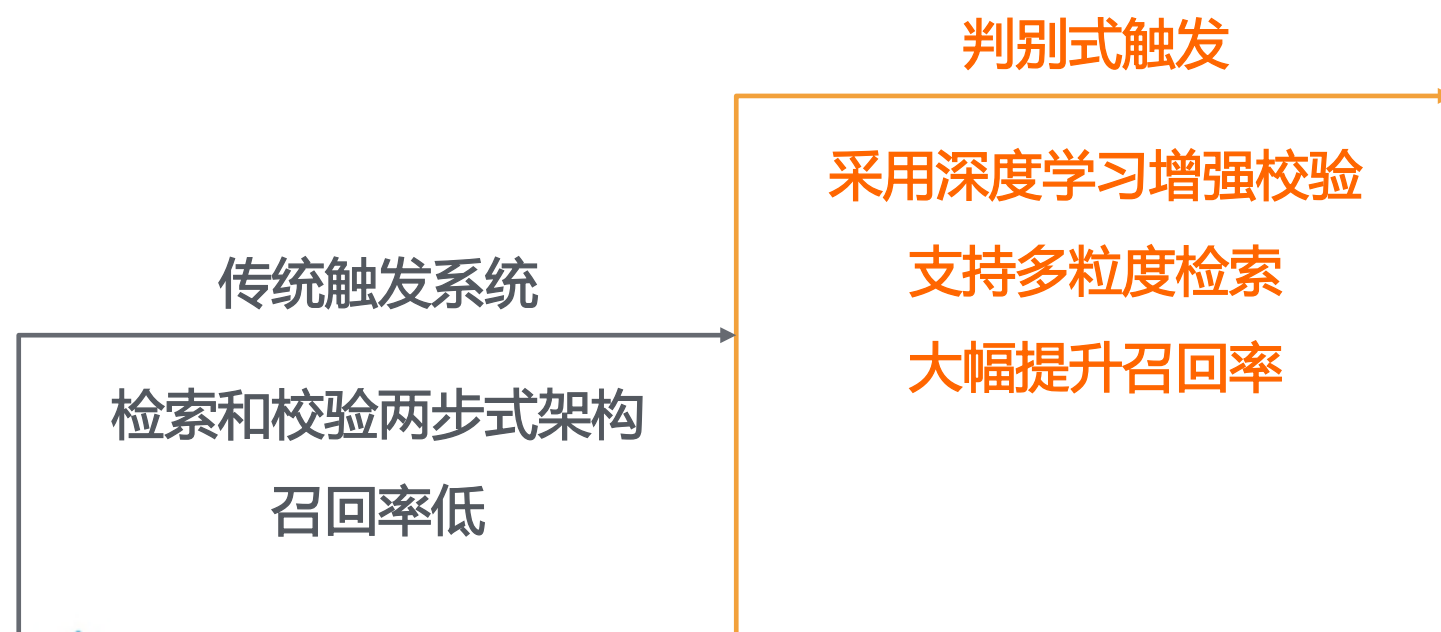
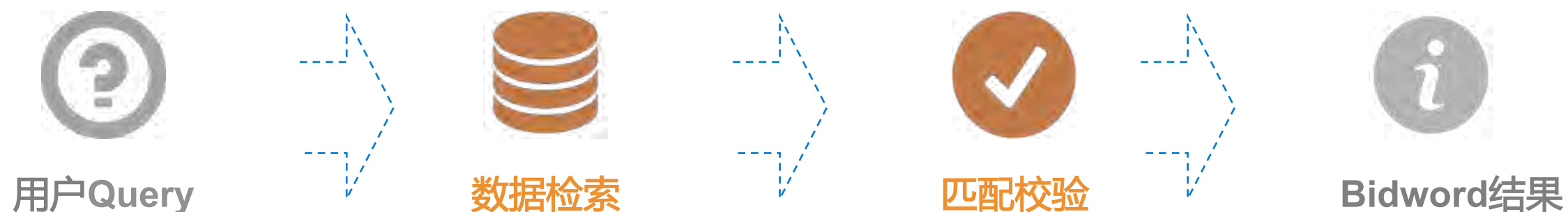
TABLE OF CONTENTS 大纲

- 搜索广告业务简介
- CTR预估的前世今生
- 生成式广告触发
- 机器学习在内容生态的创新
- 未来的探索

Keyword Targeting—传统匹配模式



从传统触发模式到生成式触发， 构建高度智能化的触发引擎



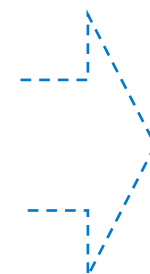
从传统触发模式到生成式触发， 构建高度智能化的触发引擎



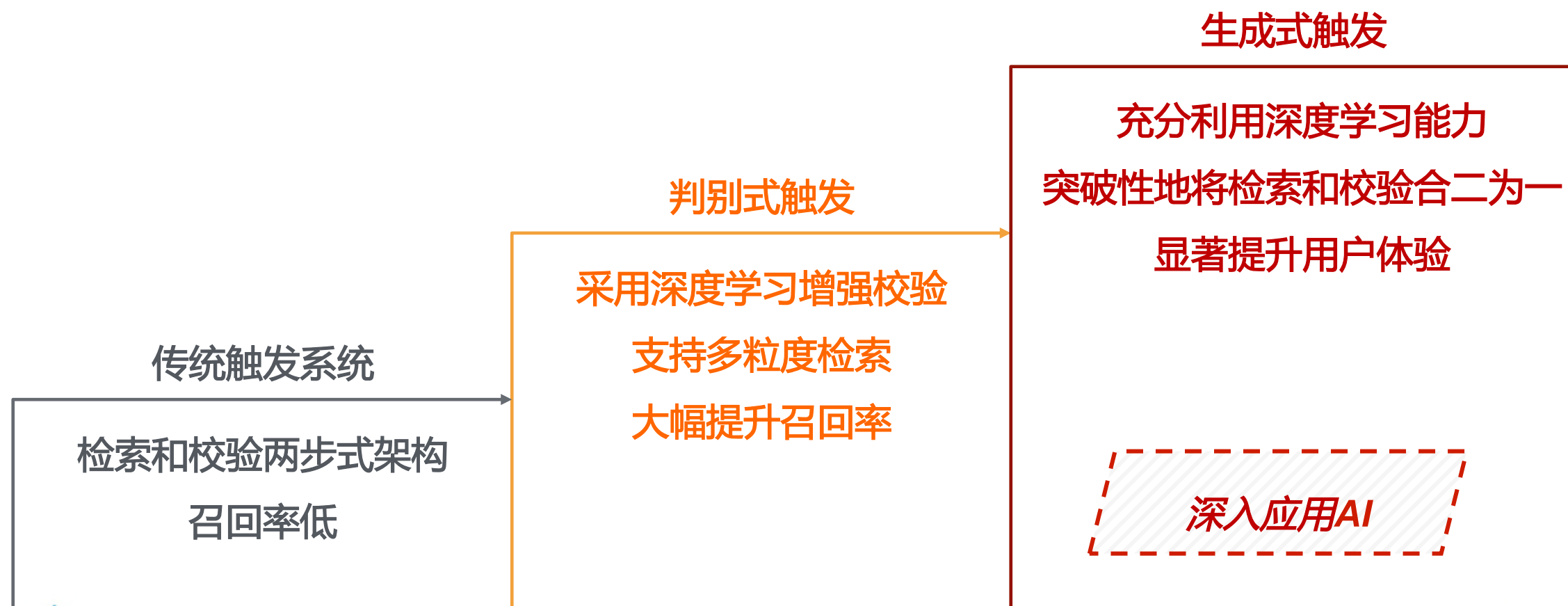
用户Query



生成式触发

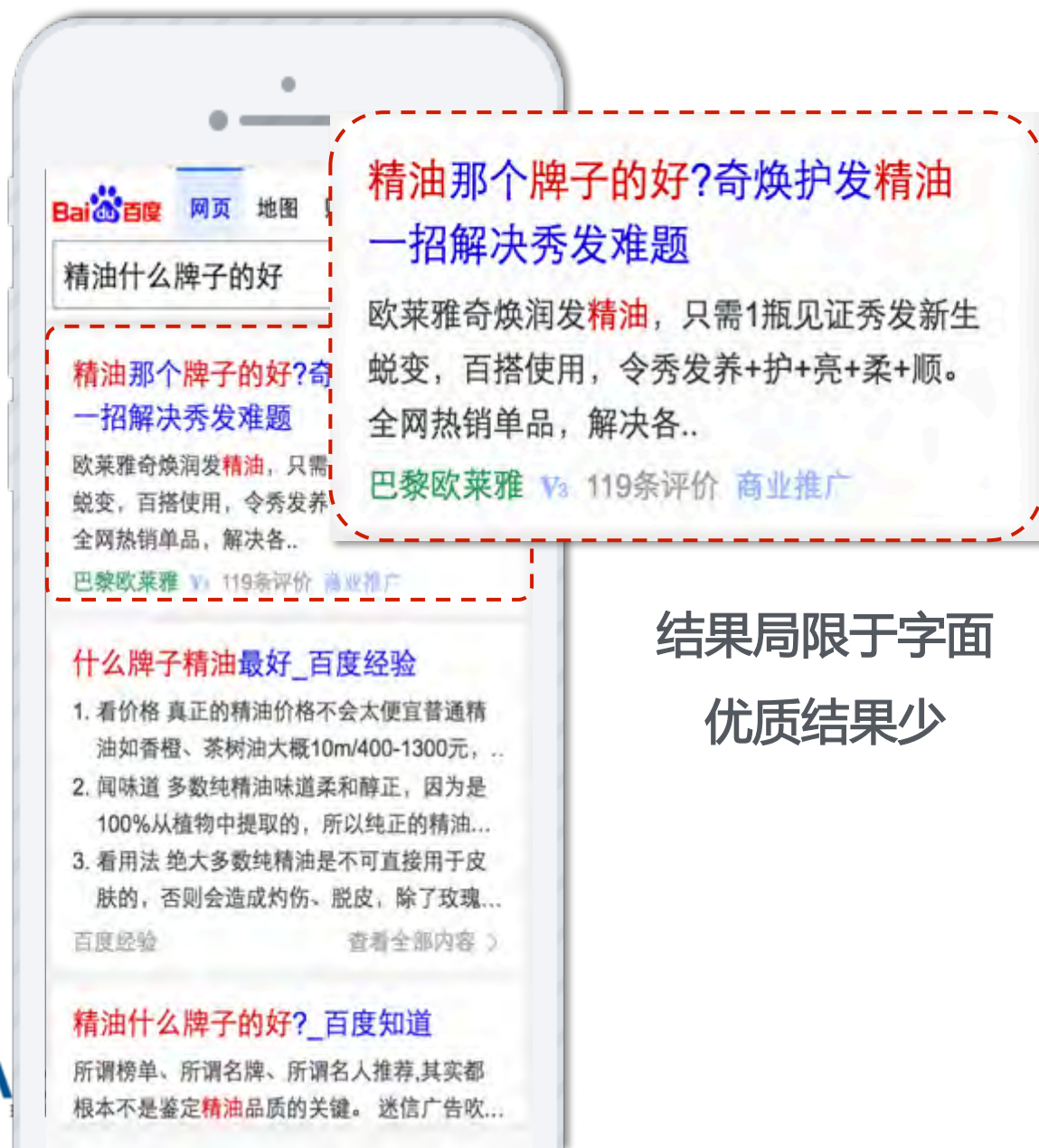


Bidword结果



深刻理解用户意图，提升用户体验

Before



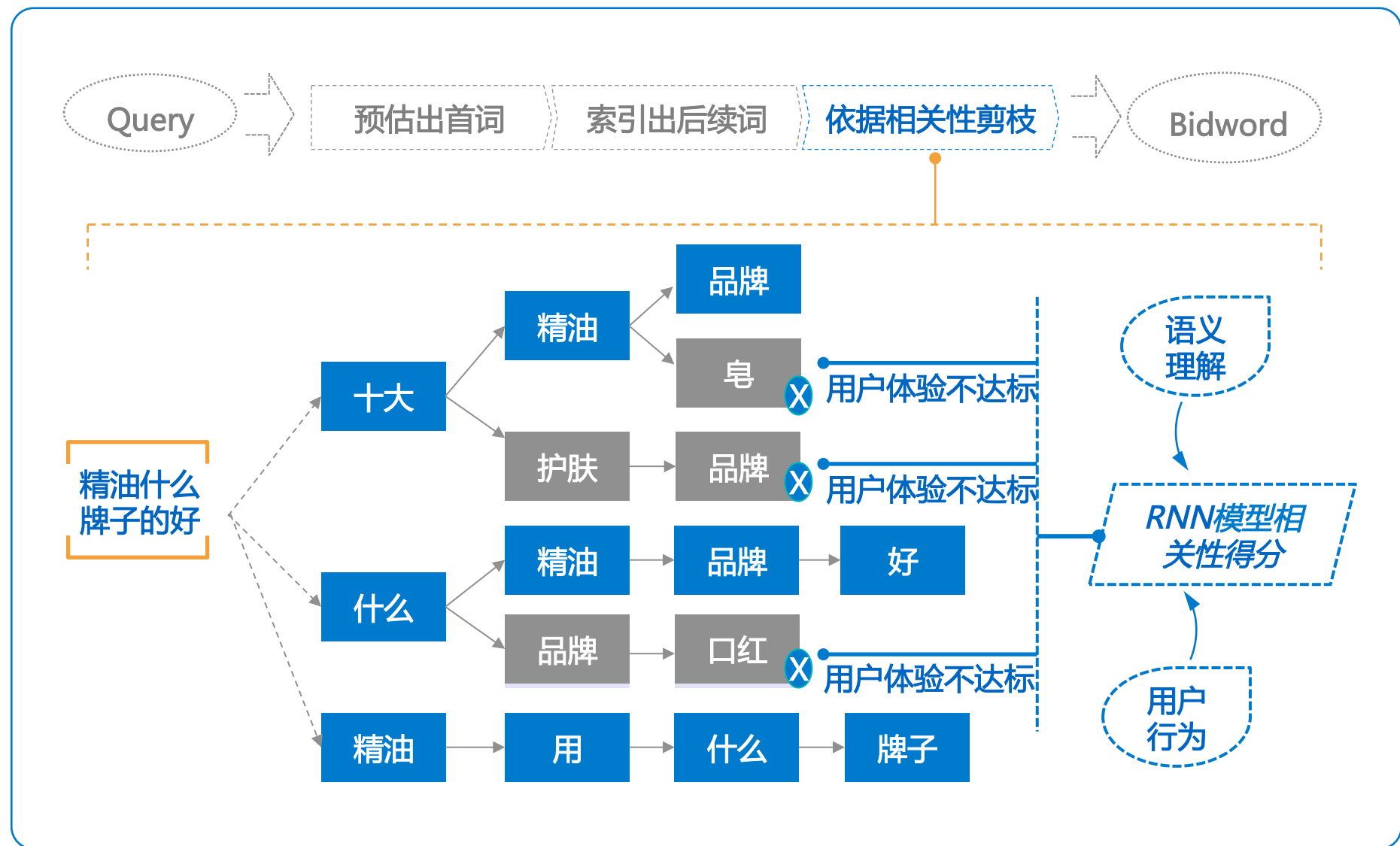
结果局限于字面
优质结果少

After



具备AI“思考”能力
深度理解用户意图
呈现更多优质结果

生成式触发模型， 以AI技术深度理解用户需求



生成式触发模型， 以AI技术深度理解用户需求

编码器: Bi-RNN

注意力模型

解码器: RNN

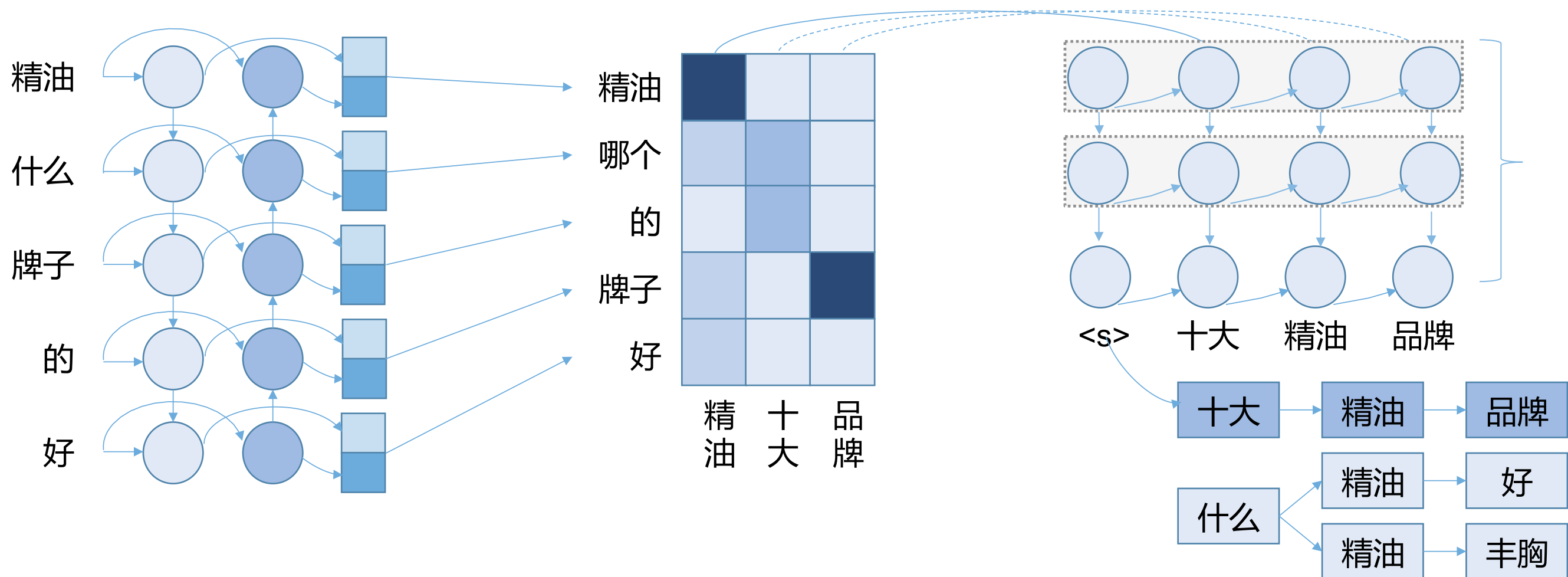


TABLE OF CONTENTS 大纲

- 搜索广告业务简介
- CTR预估的前世今生
- 生成式广告触发
- 机器学习在内容生态的创新
- 未来的探索

自动生成、校验商业文案，提升可信度

去八达岭长城 只需59元起 不满意退款

去八达岭长城-北京国际旅行总社官方网，市区免费接送，给您完美旅行体验，多款京城游玩线路，供您选择，价格优惠，欢迎咨询:010-59422266

m.bjglwj.com 评价 商业推广

去八达岭长城 仅售100元 无购物 多景点

去八达岭长城-北京国际旅行总社官方网，市区免费接送，给您完美旅行体验，多款京城游玩线路，供您选择

长城一日游[豪华团] 2266

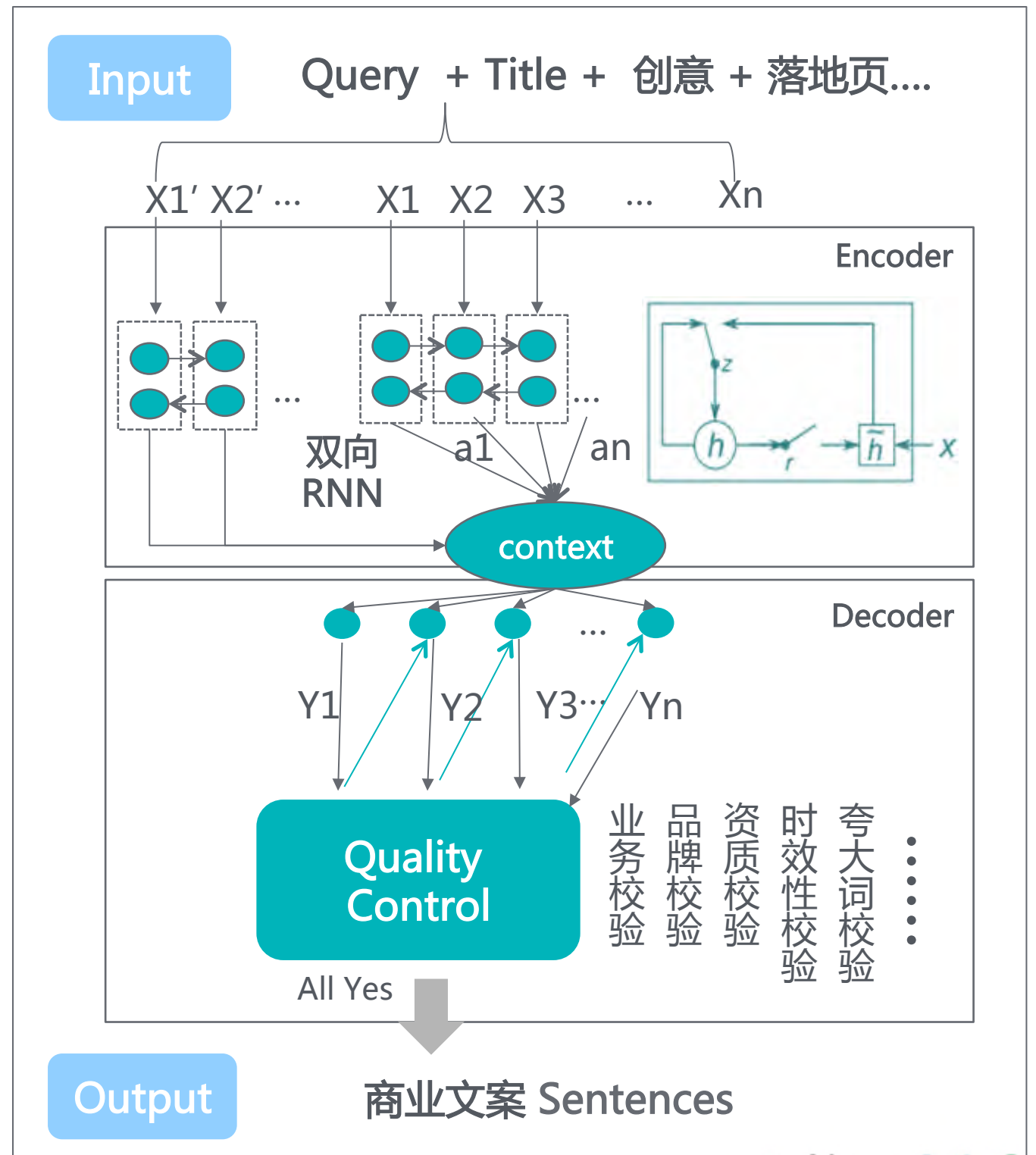
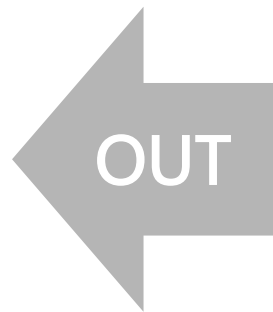
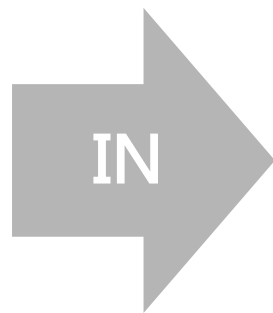
仅售100元! 享受中国国旅提供价值200元 无购物: 无自费长城一日游[豪华团] <天天发团>

在线咨询

线路推荐理由

落地页

ArchSummit 全球架构师峰会 2017



商业文案从人工撰写迈向AI自动生成

Before

雅思口语考试 雅思7分通过率90%

雅思口语考试签约直达，不达标免费重读雅思名师教学，致力于全程一对一个性化课时辅导，致力于打造高分之路！

m.beijing.gedu.org V1 评价 商业推广

- 标题描述夸大不一致
- 创意空洞冗长，无重点

“ 这些优势帮你高分通关 ”

品牌

全球最大教育集团—雅思领军品牌

雅思7分通过率达70%。全球600多所学习中心每年培训学员达60万人。

7分通过率70%

After : 智能创意

雅思口语考试 环球英语培训学校

【常见问题】雅思考试口语水平如何提高?

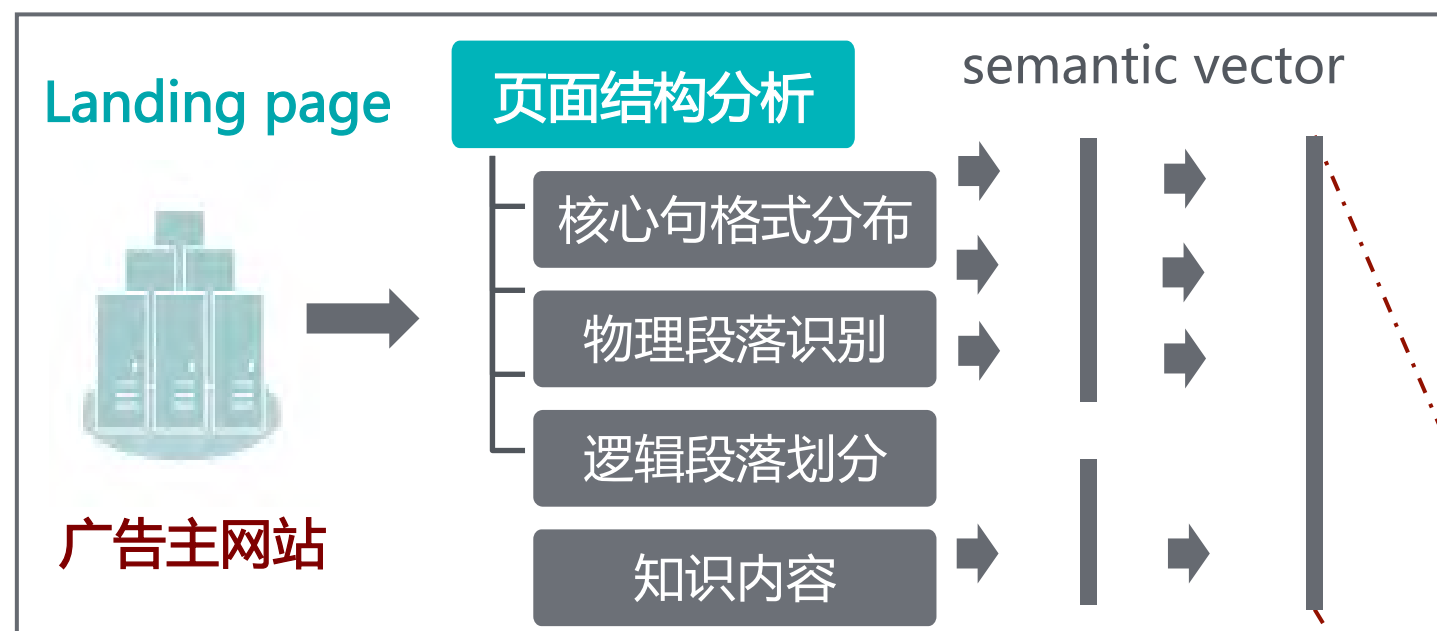
【推荐答案】1.模仿磁带或广播。2.大声朗读。3.用自己的话复述。4.选择语言地道、内容新颖的教材。详情...

m.beijing.gedu.org 评价 商业推广 电话咨询

- AI知识和内容理解
- 结构化核心内容前置
- 知识直达最相关页面



隐式知识理解与凝练，提高用户获取信息效率



Query：雅思口语考试

雅思口语考试 环球英语培训学校

[常见问题] 雅思考试口语水平如何提高?

[推荐答案] 1.模仿磁带或广播。2.大声朗读。3.用自己的话复述。4.选择语言地道、内容新颖的教材。详情..

m.beijing.gedu.org 评价 商业推广 电话咨询

Knowledge Category

Summary

Detail

Irrelevant

基于用户意图智能出图，快速Catch用户

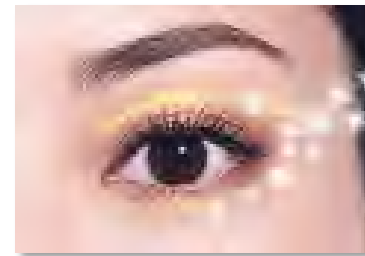


query
双眼皮整形 → 意图识别

cross-domain deep
metric learning

双眼皮价格?华韩双眼皮「3.6折」
双眼皮什么价位?数十位中韩双眼皮整形医生,
双眼皮超100,000成功案例.双眼皮价格?精心
缔造只属您的双眼皮!
北京华韩医疗美容医院

智能出图



基于图片内容的主体识别和智能裁剪



低质量图片过滤

机器学习打造智能落地页

Before

餐饮油烟净化器 油烟净化设备 全新厨..
全新厨具设备, 油烟净化器产品, 好品牌就选东莞市
科怡环保油烟净化器
www.keyisb.com 价格 厂家直销



- 内容空洞千篇一律
- 专业化、结构化不足
- 移动用户体验质量堪忧

After : 智能落地页

餐饮油烟净化器 节能耐用 超强静音
相关产品: 约11款 优势: 环保|节能|可定做|质量领先
专业油烟净化器生产商, 提供新研制的低空排放油烟..
www.keyisb.com 价格 厂家直销

- 重构高质量落地页
- 智能化、专业结构化
- 实现内容体验闭环



机器学习打造智能落地页

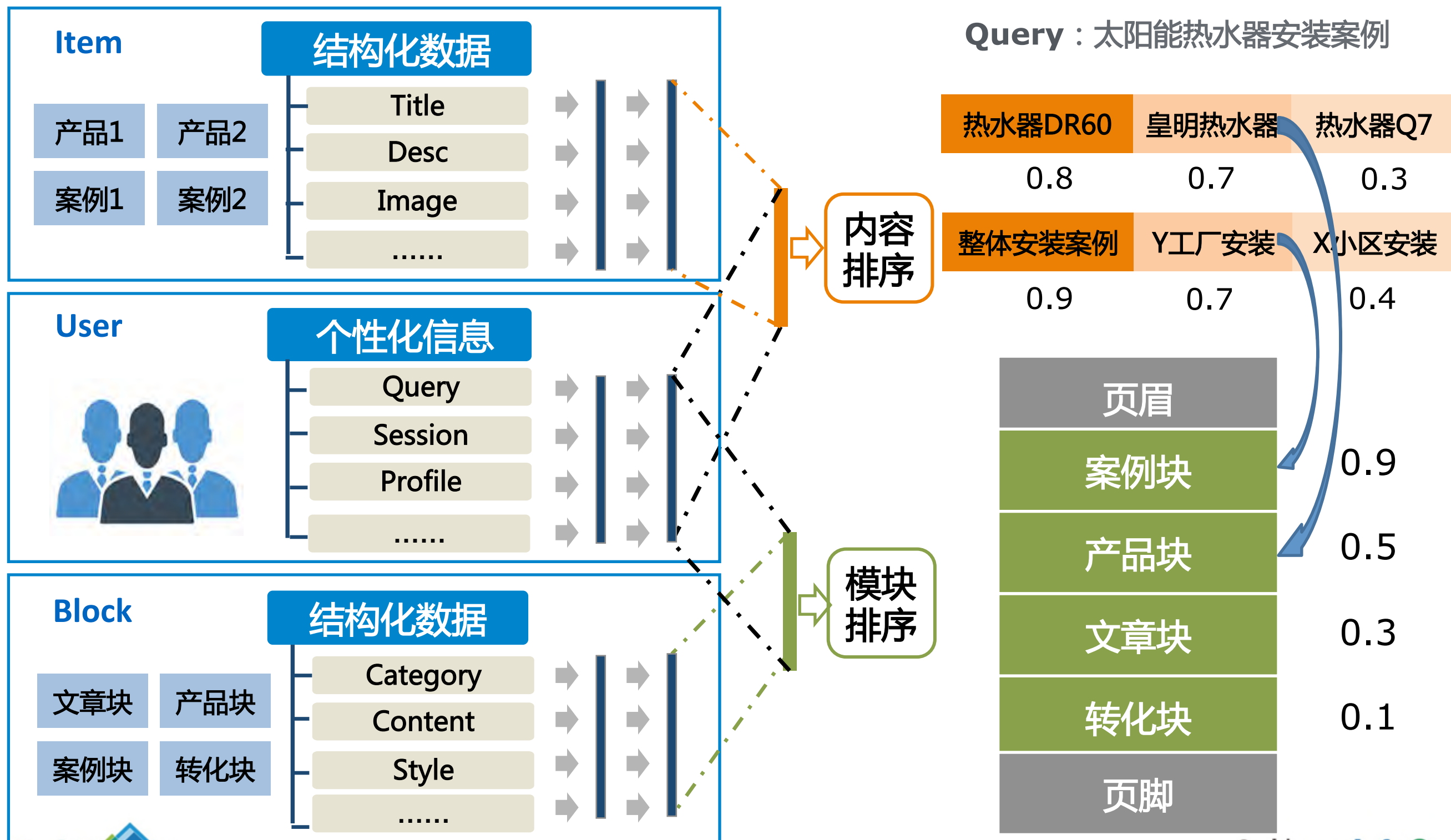


TABLE OF CONTENTS 大纲

- 搜索广告业务简介
- CTR预估的前世今生
- 生成式广告触发
- 机器学习在内容生态的创新
- 未来的探索

未来的探索

时代在变，内容的形式也在变



文本



图片



视频

.....

以图片化信息集成快速吸引用户并满足需求

实现图片创意的“一键美颜”



实现图片创意信息的“一览无遗”



以图片化信息集成快速吸引用户并满足需求

实现图片创意的 “一键美颜”

实现图片创意信息的“一览无遗”



加强端到端的富媒体化程度...

Before



After



营销信息提炼成短视频呈现，让用户一秒读懂



视频合成

订婚宴 就选花嫁丽舍

未来的探索

Video-Out视频广告：视频处理、图像理解、检测技术



基于AI的角标广告

未来的探索

Video-In视频广告：图像前景背景检测与分割、AR虚拟植入



视频植入广告

未来的探索

Video-In视频广告：图像前景背景检测与分割、AR虚拟植入



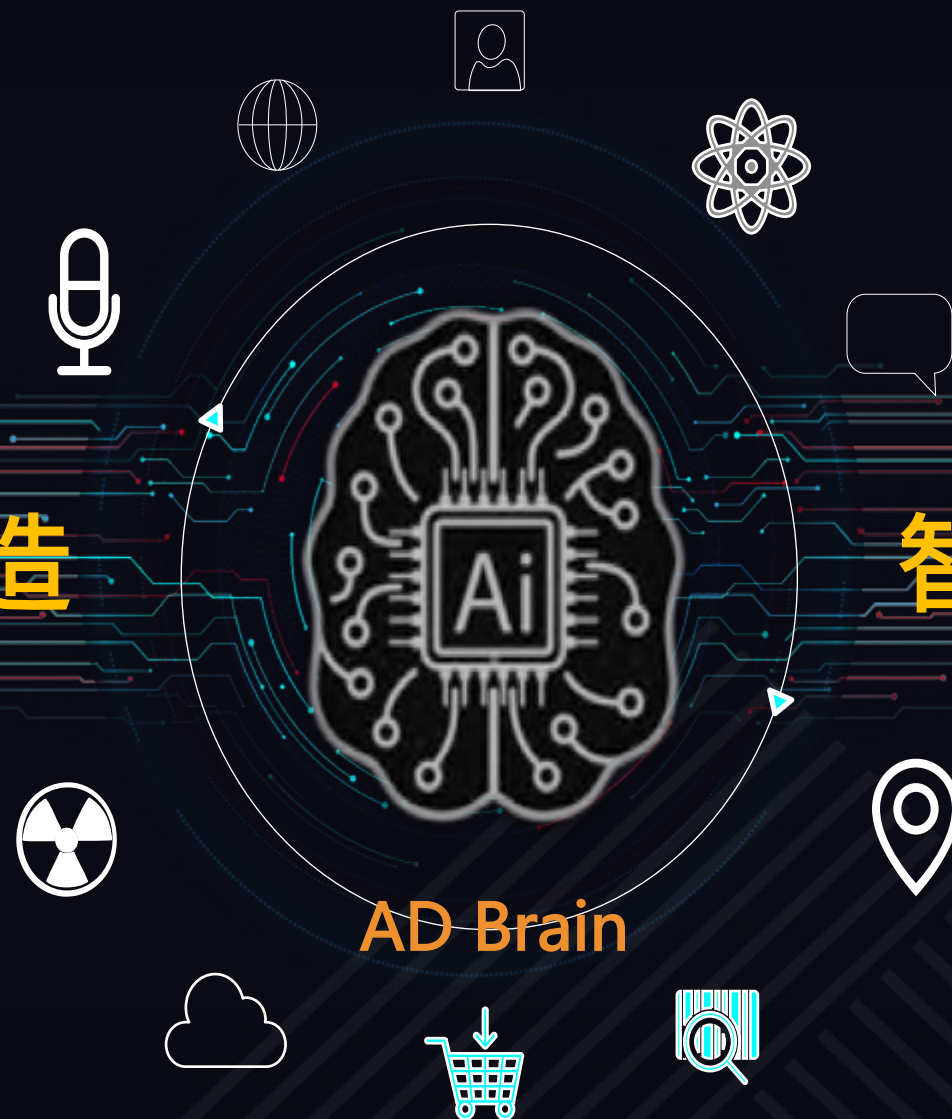
视频植入广告

新媒体的产生渴求新技术，
在AI时代这波浪潮中，欢迎各位大咖加入

欢迎联系：fc-ai@baidu.com

让我们一起打造

智能AI商业体系



THANKS!

让创新技术推动社会进步

HELP TO BUILD A BETTER SOCIETY WITH
INNOVATIVE TECHNOLOGIES

Geekbang>

极客邦科技

InfoQ

专注中高端技术人员的技术媒体



EGO EXTRA GEEKS' ORGANIZATION
NETWORKS

高端技术人员学习型社交平台



StuQ
斯达克学院

实践驱动的 IT 教育平台

