



通用Query推荐平台 (金门)

IMPL 甘璐



Outline

- 金门概述
- 金门背景
- 架构设计
- **Feed落地**

金门概述 | 简介

金门是什么

通用Query推荐平台

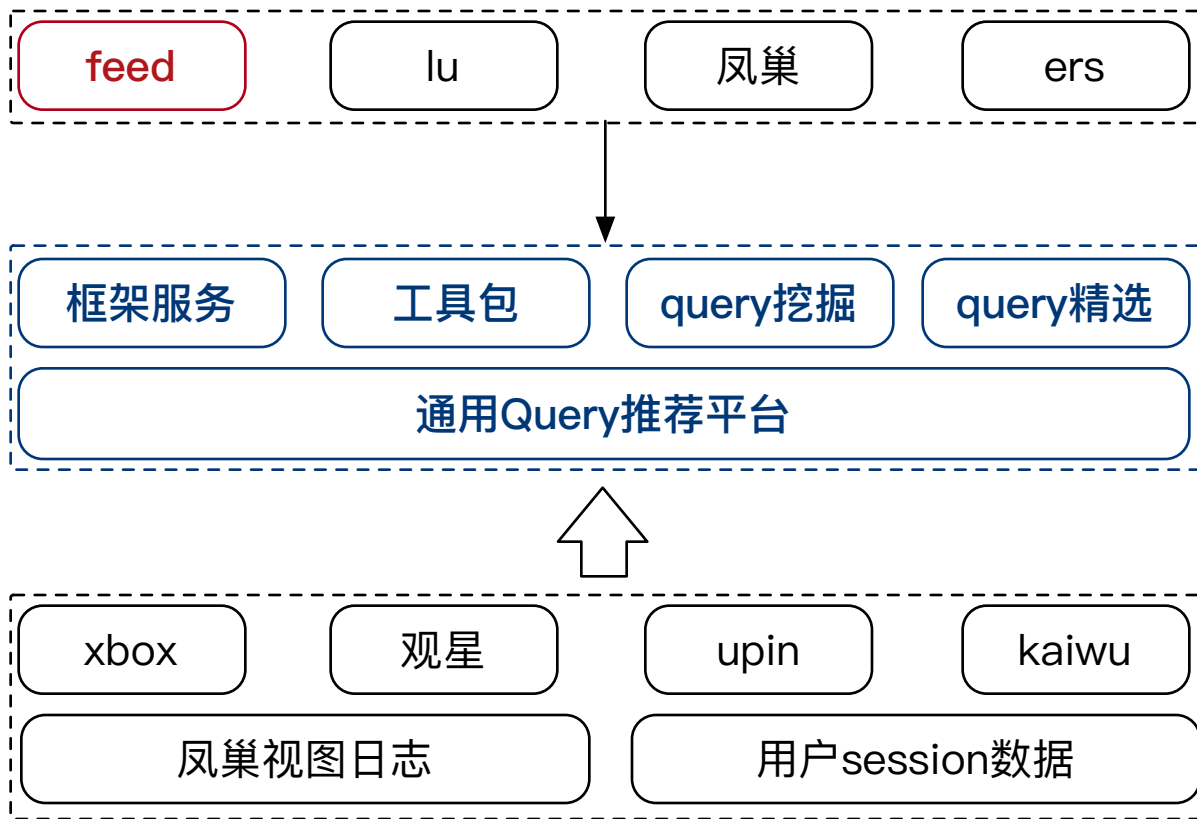
金门做什么

基于用户的意图分析
推荐用户意图query
用于广告检索

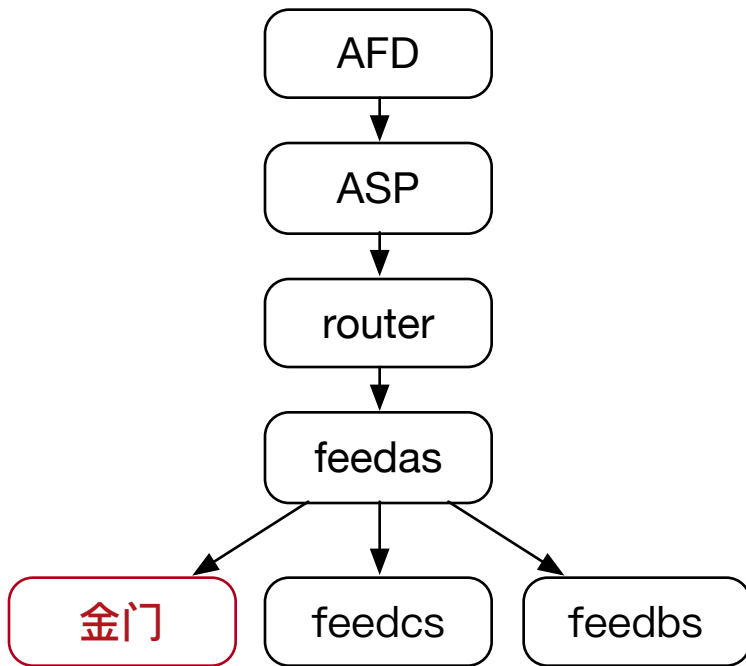
为什么叫金门

推荐类流量变现入口
影响千万收入

金门概述 | 定位



金门概述 | feed中的定位



金门背景 | 商业产品形态



Find the **best match** between a given **user** in a given context and a suitable **advertisement**

凤巢

- ✓ 用户意图显性: query
- ✓ 匹配: 以query为关键字的广告**检索**
- ✓ 可基于用户意图分析进行query改写

feed

- ✓ 用户意图隐性: 无输入
- ✓ 匹配: 基于用户意图分析的广告**推荐**

金門背景 | 意图表达形式

query

- ✓ 商业场景下意图表达通用
- ✓ 场景：凤巢，feed，商业导流

结构化意图

- ✓ 兴趣、年龄、性别、地域、教育等
- ✓ 场景：feed

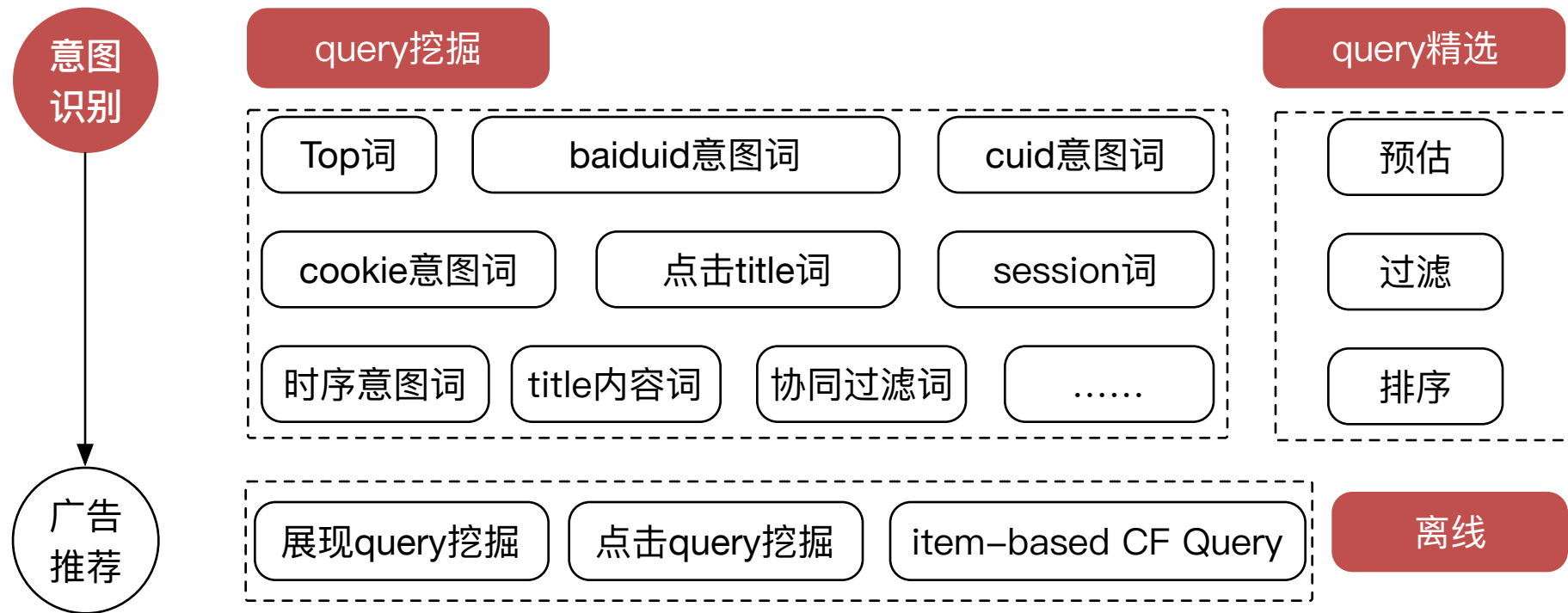
用户向量

- ✓ 用户意图向量化，ML方法匹配
- ✓ 场景：凤巢罗森桥

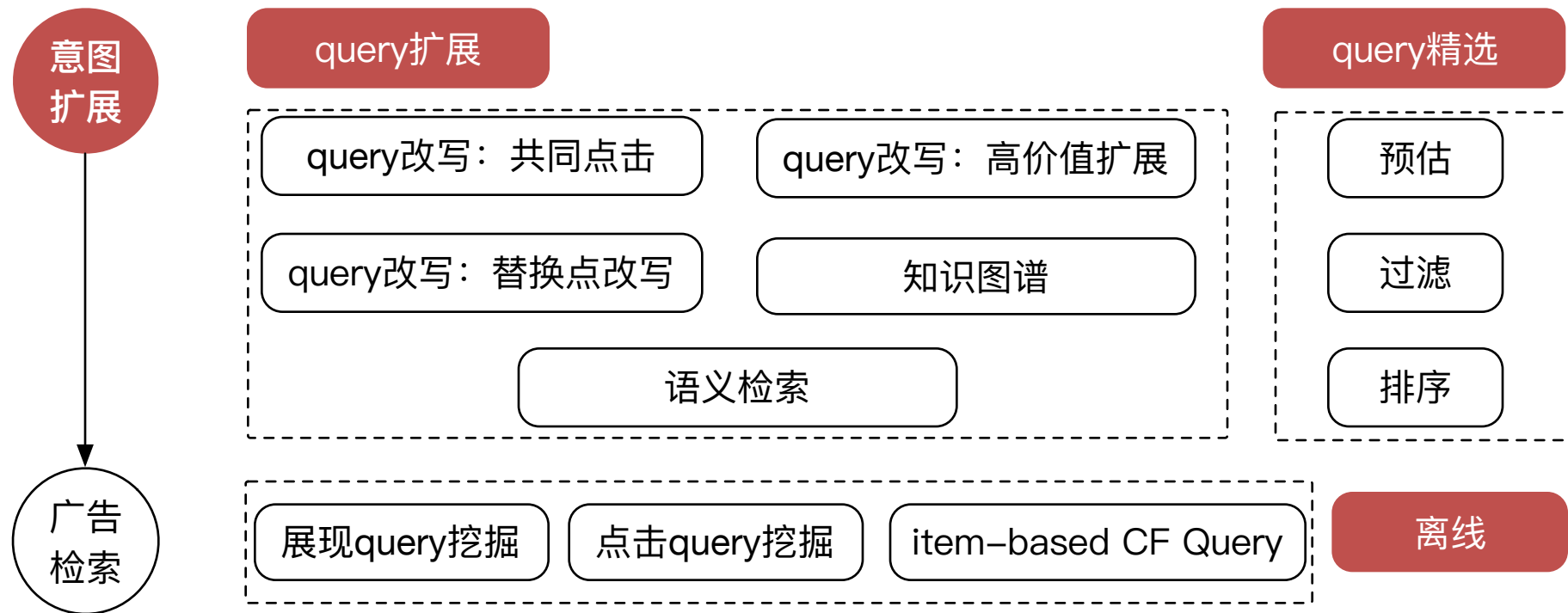
结论

- ✓ query更为通用，场景更多
- ✓ 以此为切入点

金門背景 | 推荐类场景分析



金門背景 | 搜索类场景分析



金门背景 | 场景分析

业务名称	业务场景	query场景	query处理阶段	部署模块
feed	信息流广告	无 → 有	① 生成query ② 筛选query	feedas
lu导流	商业导流	无 → 有	① 生成query ② 筛选query	lu-bs
凤巢	搜索广告	一个 → 多个	① 扩展query ② 筛选query	mixer
ers/esug 商业词推荐	商业导流	一个 → 多个	① 扩展query ② 筛选query	ers/esug

金門背景 | 场景分析

有query场景

- ① 扩展、改写query
- ② query筛选

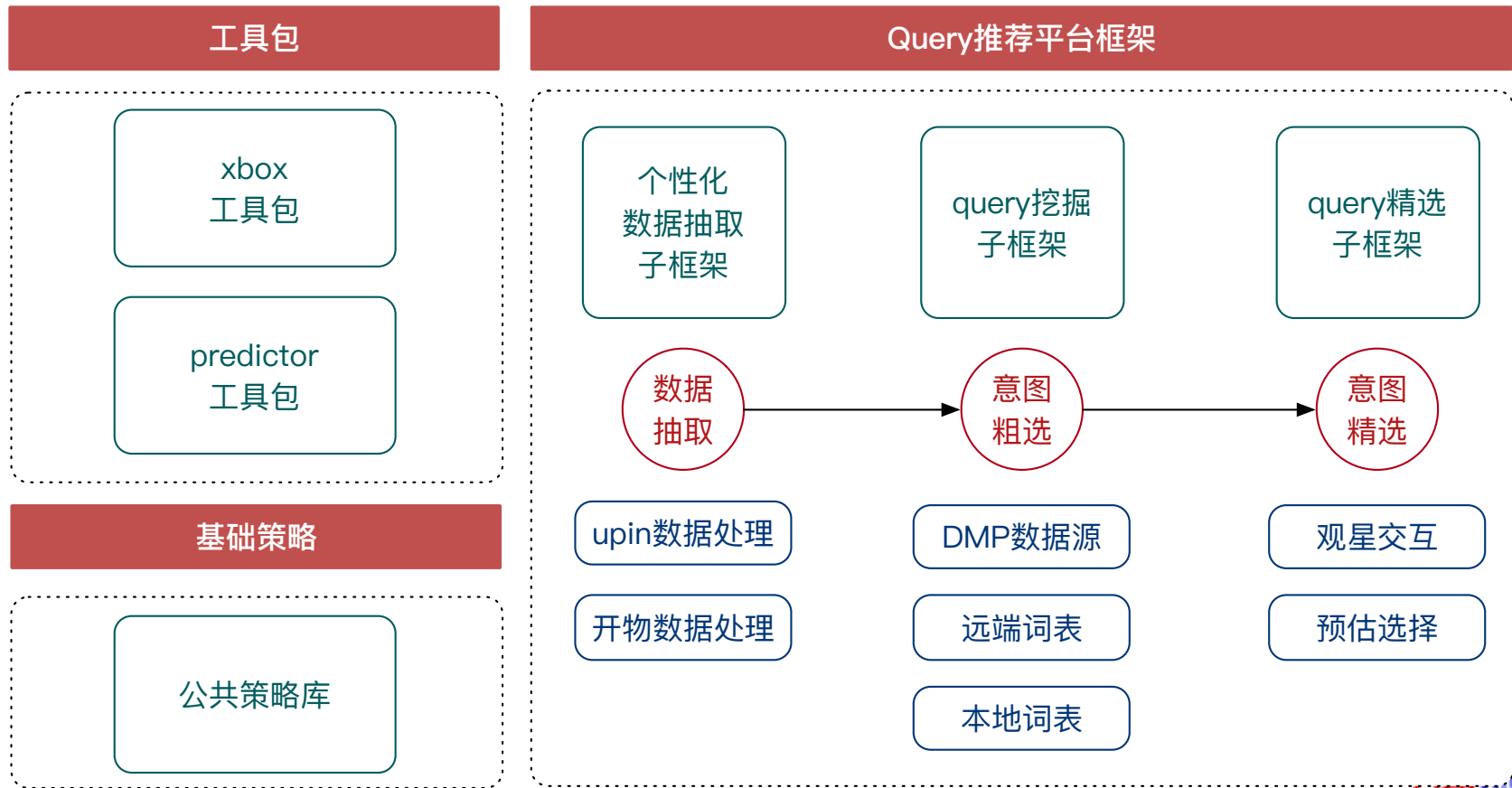
无query场景

- ① 挖掘query
- ② query筛选

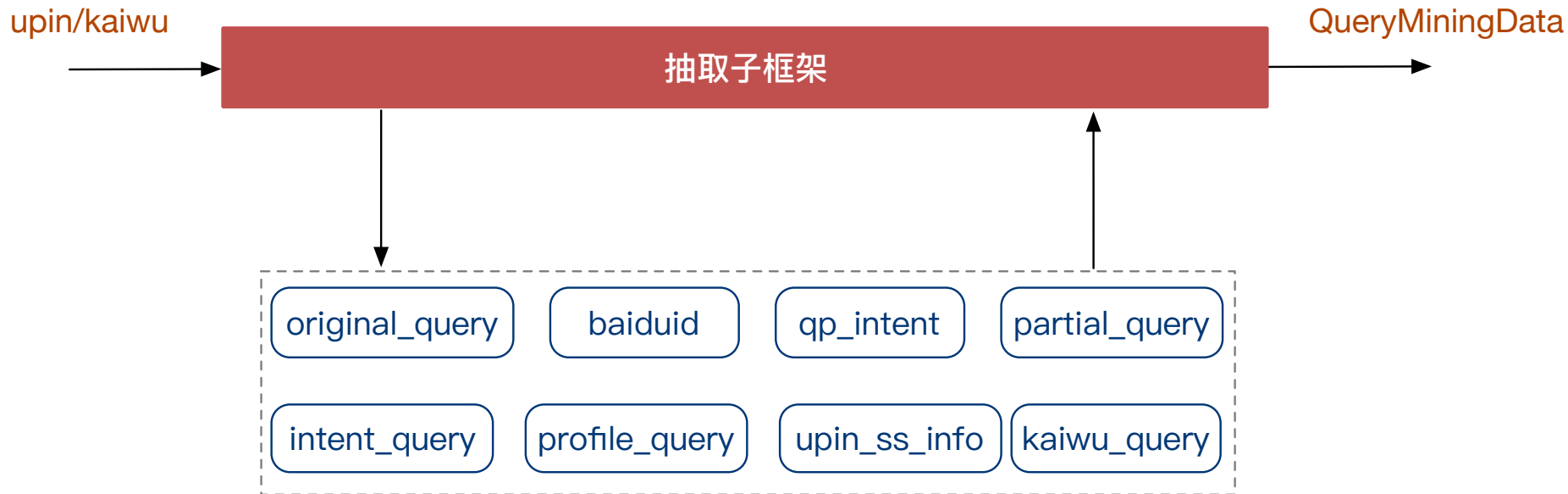
方案

- ✓ 统一意图分析，抽象通用query推荐平台
- ✓ 以信息流为切入点，一期实现feed的query推荐服务

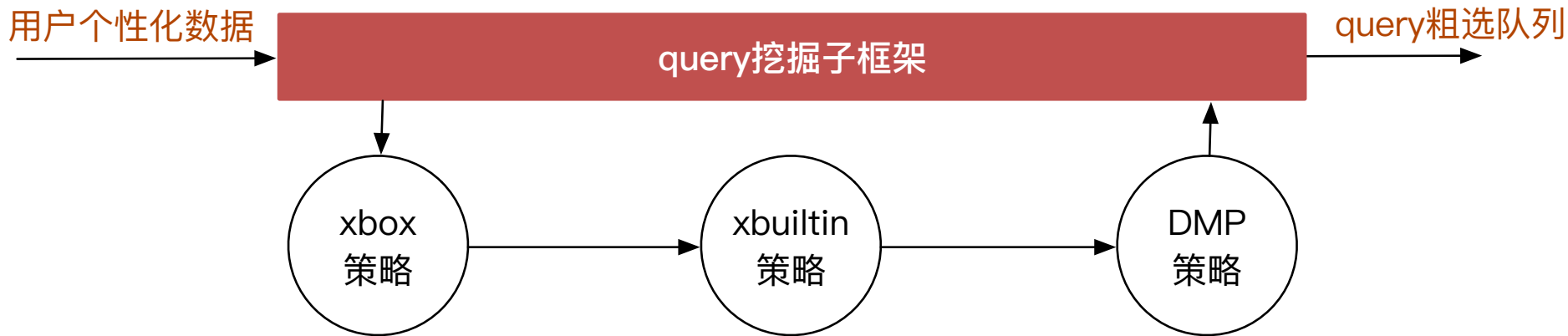
架构设计 | 整体架构



架构设计 | 抽取子框架——配置驱动，按需抽取



架构设计 | 挖掘子框架——挖掘策略配置化驱动



xbox配置化

✓ 添加新策略词表：名称 + 格式 + key

DMP配置化

✓ 添加新策略分支：指定字段名称

架构设计 | 挖掘子框架——xbox配置化

特点

- xbox交互格式固定
- key可枚举
- 请求处理，结果解析同质

抽象

- key数据统一放入PB
- 利用PB反射指定key
- 抽象请求填充、结果解析
- 框架统一调度，xbox请求并行化

使用

```
[@branch]
branch_name: intent_model
plugin_name: intent_xbox_query
[.branch_params]
branch: 6
rank: 5
max_num: 150
[.plugin_params]
[..@intent_xbox_params]
table_name: user2query
table_type: 2
key_type: baiduid
```

架构设计 | 挖掘子框架——DMP配置化

特点

- 数据解析自upin / kaiwu 第三方DMP
- 解析出的字段直接作为 query候选

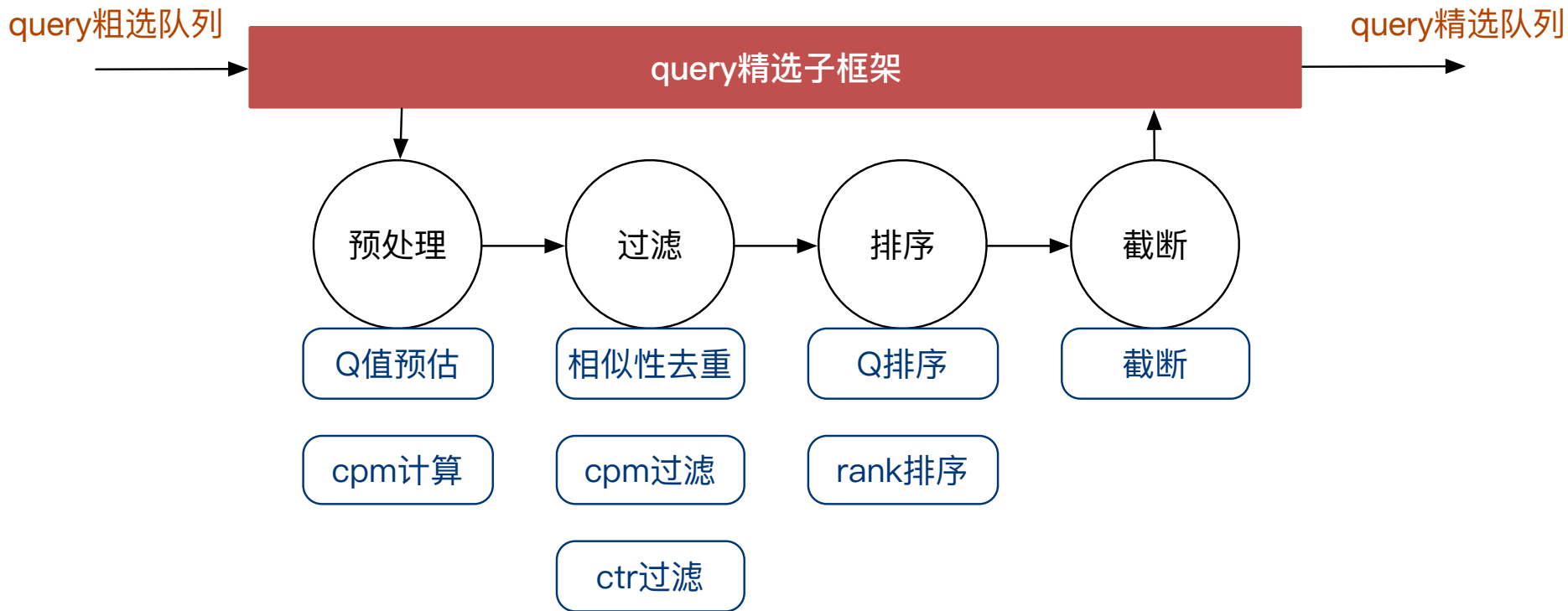
抽象

- 将第三方数据统一放入 PB
- 利用PB反射机制，配置指定使用字段

使用

```
[@branch]
branch_name: intent_query
plugin_name: dmp_query
[.branch_params]
branch: 2
rank: 6
weight: 0
max_num: 10
[.plugin_params]
@key_type:
upin_intent_query
```


架构设计 | 精选子框架——统一策略处理流程



Feed落地 | Query挖掘

策略名称	说明
hot_query	本地词表获取热词
tieba_query	本地词表获取贴吧词
session_query	upin拿到的session词
intent_query	upin拿的意图词
sequece_word	请求xbox拿时序词
cf_password	请求xbox拿协同过滤词

Feed落地 | Query筛选

策略名称	说明
predictor	intentq预估
排序	根据rank或者score排序
过滤	cpm、intentq阈值过滤
去重	相似query去重
频控	shown_info频次控制
截断	query返回个数截断

Feed落地 | 与feedcs的差别

- 业务角度

- 金门只做用户意图query的挖掘和扩展
- feedcs承载具体商业产品，这里为词包

- 技术角度

- query推荐和扩展本质上仍旧是对query的处理，与feedcs技术栈复用
- 金门统一输入输出，对query推荐场景进一步抽象，统一调度

Feed落地 | 收益

- 平台收益
 - 数据共享，策略复用
 - 支持策略快速上线，减少重复开发
- feedas收益
 - SOA / Microservices的架构拆分
 - query化策略的重组

END

Thanks