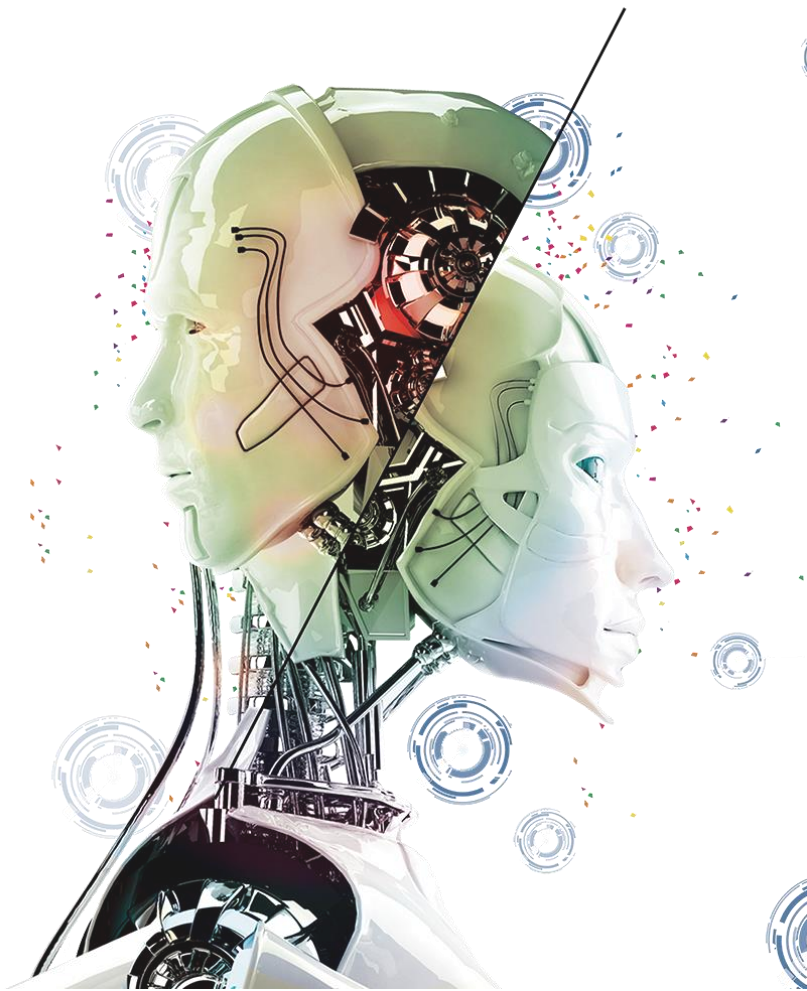


# 隐私计算三剑客在营销中的应用

陈治宇 百度 资深安全工程师



# 目录 CONTENT

## 01 数字营销与隐私

此部分内容作为文字排版占位显示  
(建议使用主题字体)

## 04 安全沙箱

此部分内容作为文字排版占位显示  
(建议使用主题字体)

## 02 联邦学习/多方安全计算

此部分内容作为文字排版占位显示  
(建议使用主题字体)

## 05 百度点石产品矩阵

此部分内容作为文字排版占位显示  
(建议使用主题字体)

## 03 可信执行环境

此部分内容作为文字排版占位显示  
(建议使用主题字体)

## 06 未来趋势分析

此部分内容作为文字排版占位显示  
(建议使用主题字体)

# 01

## 数字营销与隐私

替换文字内容，点击添加相关标题文字



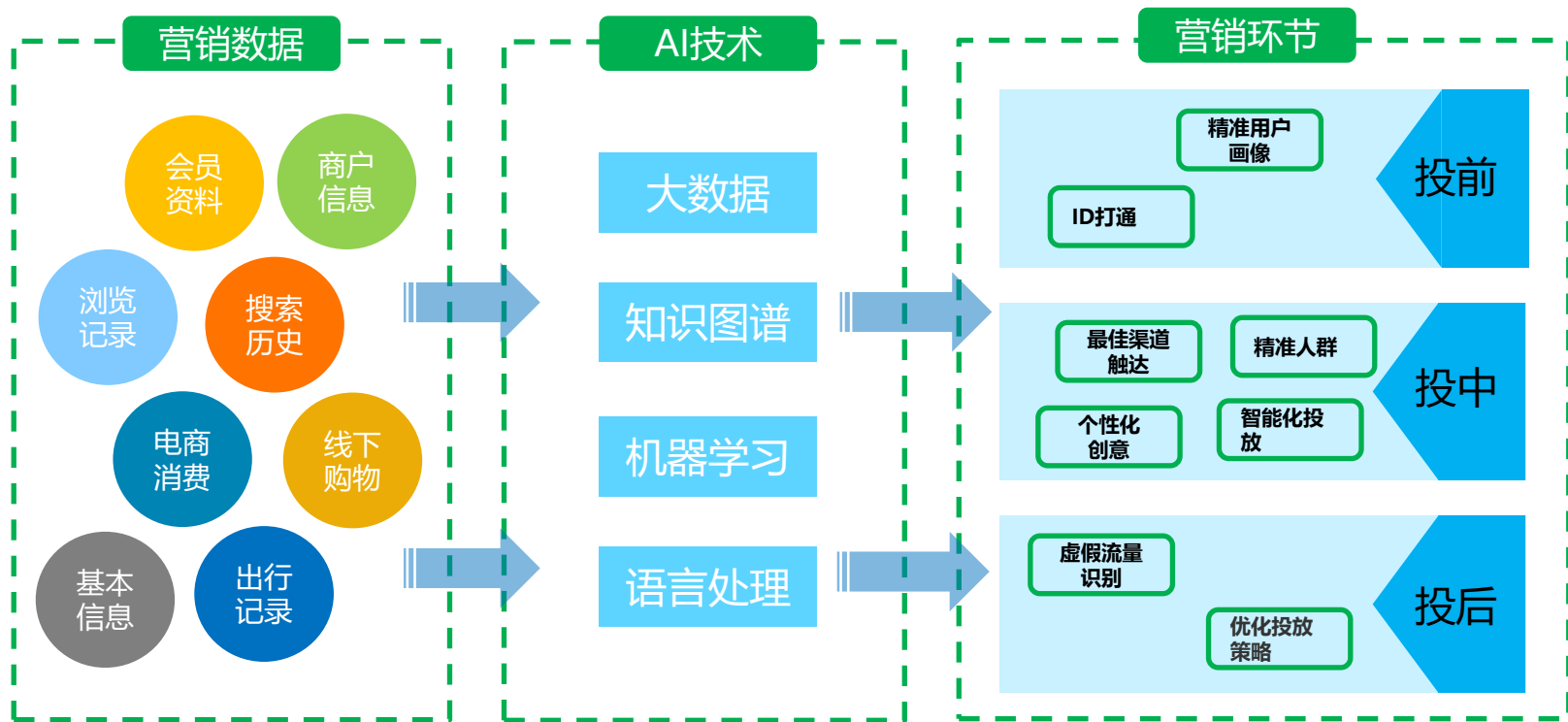
# 营销技术的发展

阶段	传统营销	互联网+营销	大数据+营销	AI+营销	AI+全域营销
核心	触达	交互	精准	效率	生态
特点	相对于传统线下营销模式，大众媒体的出现使得营销信息可以覆盖十分广大的消费者群体	在广泛覆盖性的基础上有了更多交互的可能性，不再是单向信息传播，而是与消费者互动与沟通	大数据技术成熟应用，精准营销概念和平台不断出现，关注营销对用户触达的精准度	人工智能技术在营销领域的逐渐渗透，营销各个场景和环节更加智能化，营销效率不断提高	实现用户全场景覆盖、全链路数据采集、全域用户洞察、全渠道精准触达的AI+全域营销
技术	<p>1920 广播广告</p> <p>1941 电视广告</p> <p>1995 门户广告</p> <p>2000 搜索广告</p> <p>2004 社交媒体广告</p> <p>2004 AdExchange</p> <p>2006 信息流广告</p> <p>2010 DSP/DMP</p> <p>2012 营销云</p> <p>2016 智能创意</p> <p>2018 营销机器人</p> <p>2019及以后 数据中台, 5G, AIOT, 数字人, 开放域对话, AR, VR</p>				
目标	获取增量用户	获取增量用户	获取增量用户	获取增量+存量运营	客户全时全场景价值挖掘
理念	以产品为中心		以客户为中心		客户全时全场景价值挖掘

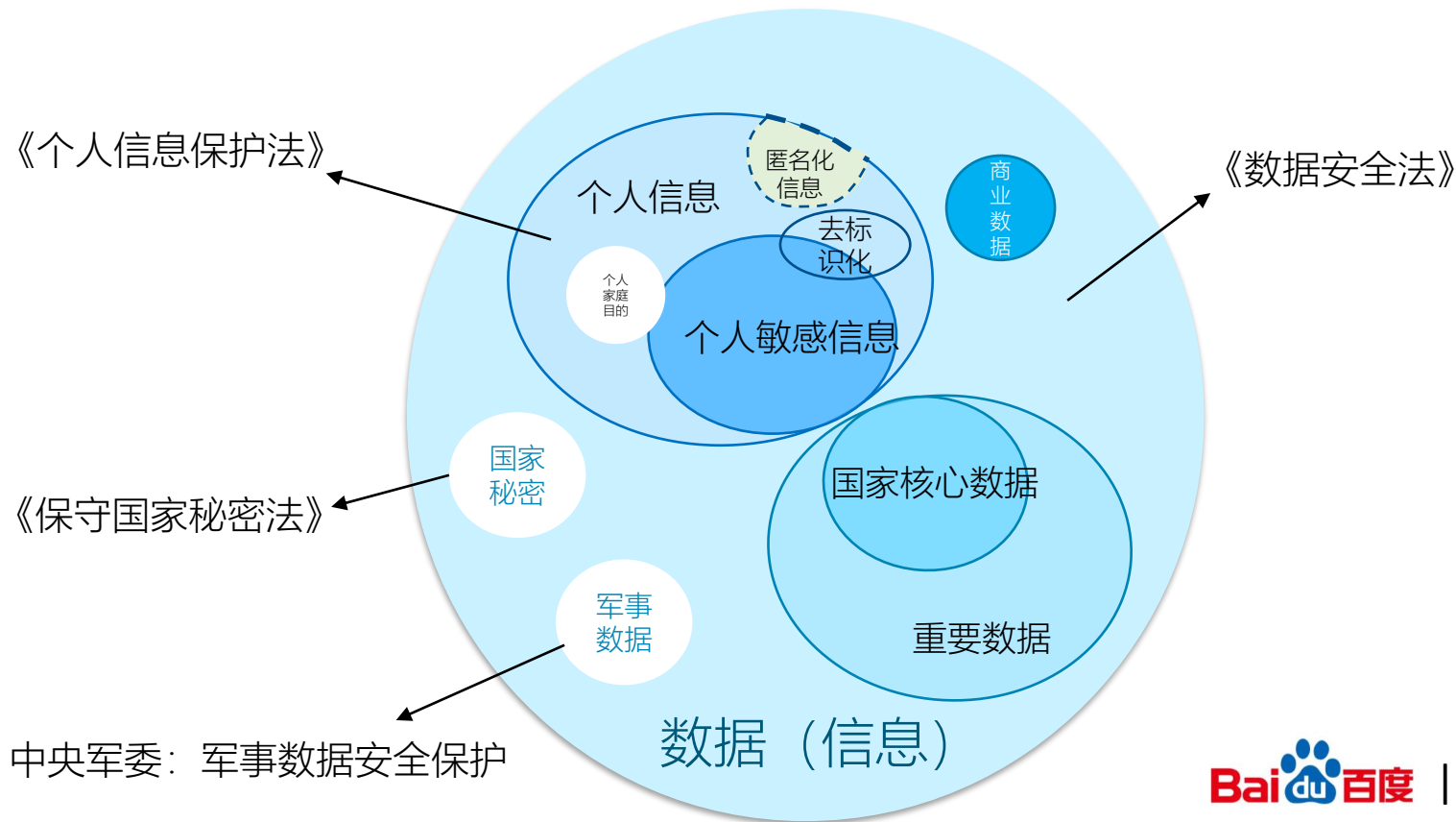
资料来源:知萌咨询公共资料整理。

图表引自: 中国广告协会《2022AI营销白皮书》

# AI+营销：涉及隐私数据



# 隐私数据

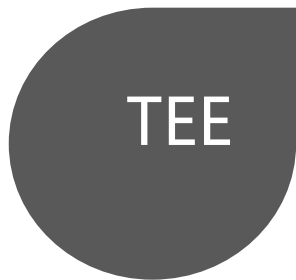


# 什么是：隐私计算？

- 解决的问题：“数据可用不可见”
- 什么效果：仅获取到“结果”，但得不到以外的信息
- 价值：“使用权”和“所有权”分离
- 怎么实现：软件，硬件，不软不硬



# 隐私计算：三剑客





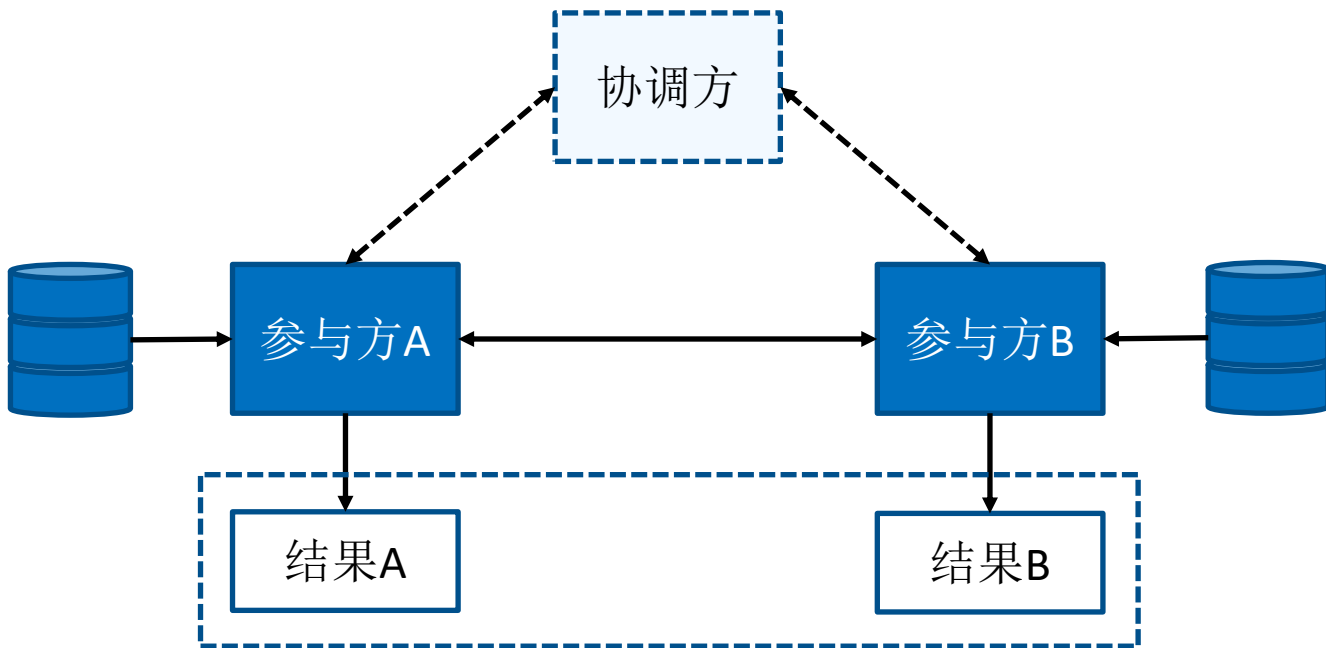
# 02

## 联邦学习/多方安全计算

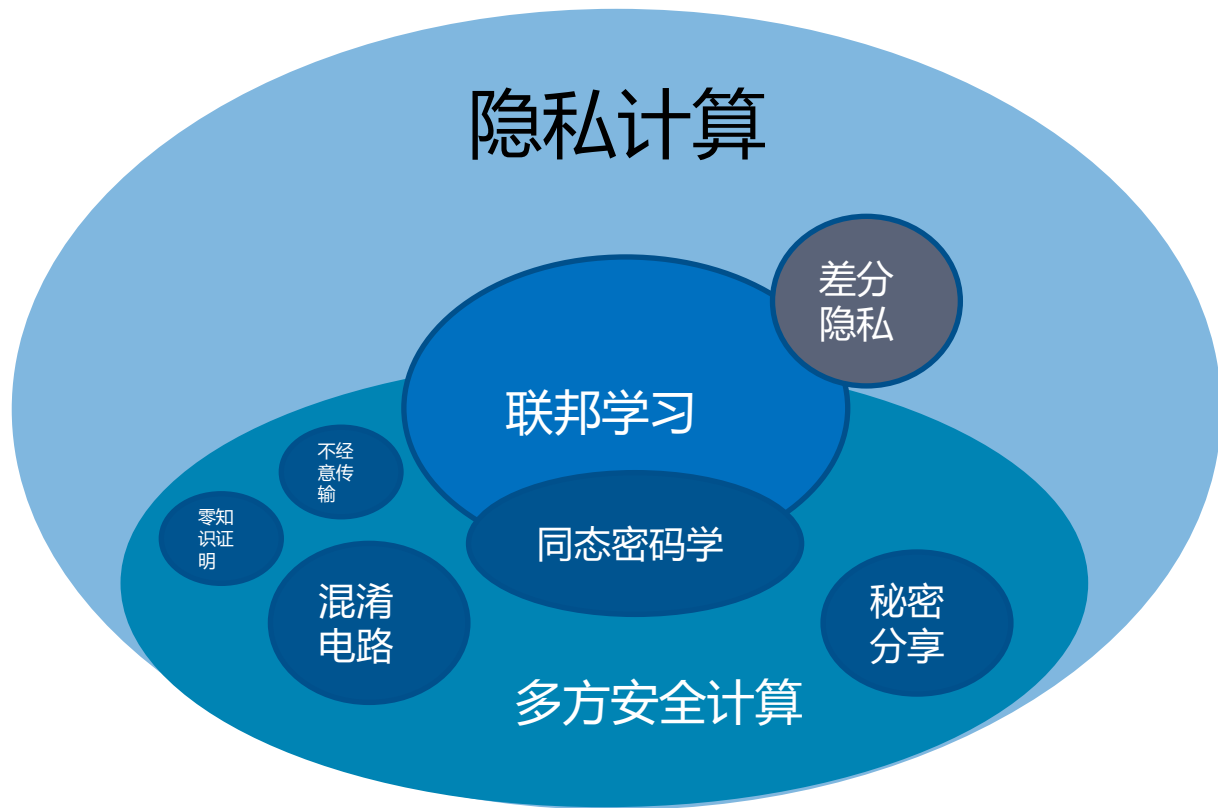
替换文字内容，点击添加相关标题文字



# 联邦学习 (FL)



# 联邦学习和安全多方计算

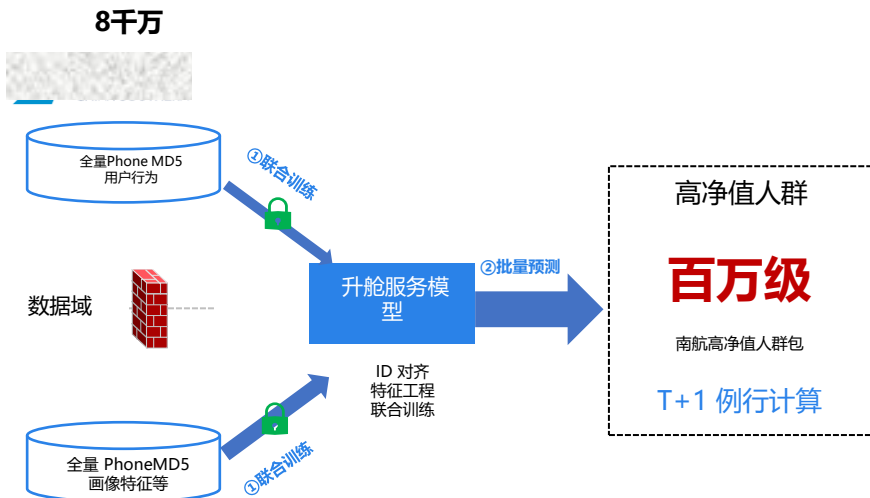


# 优势

- 不依赖硬件
- 原始数据不出区域
- 点对点链接
- 知名度高

# 航空领域营销

## 联合建模流程

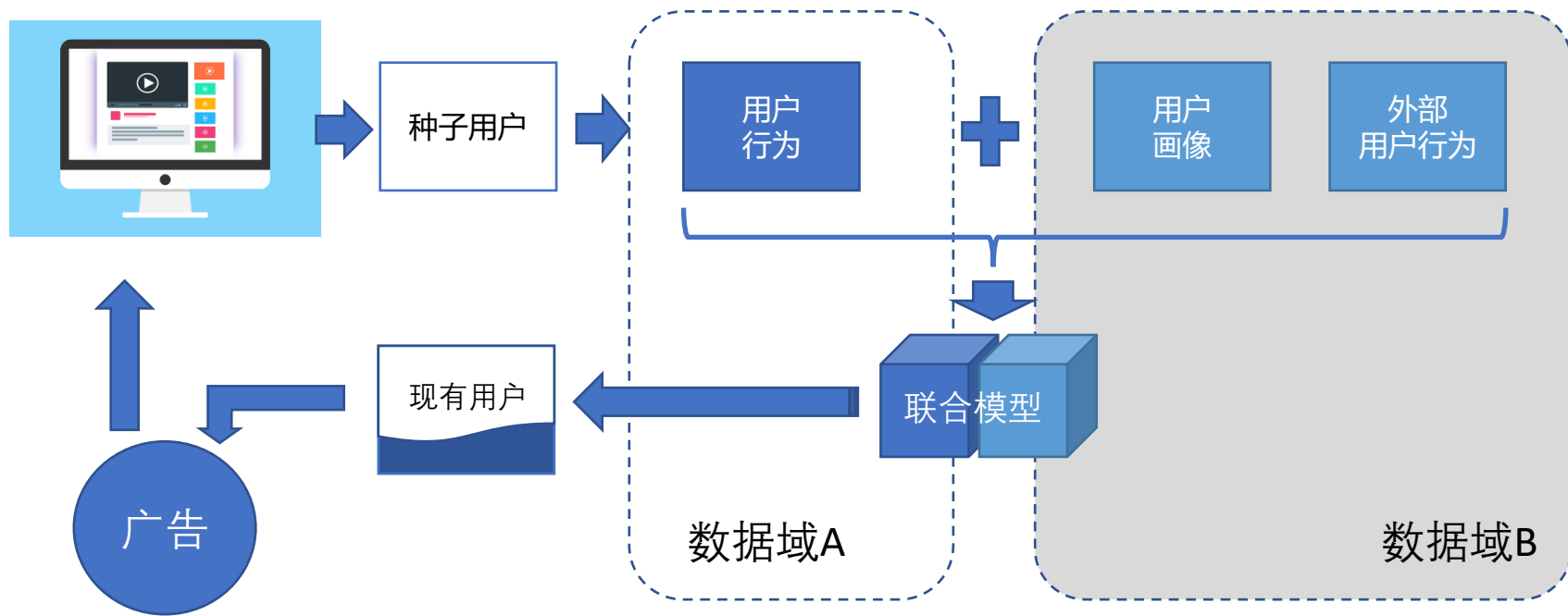


## 精准投放流程

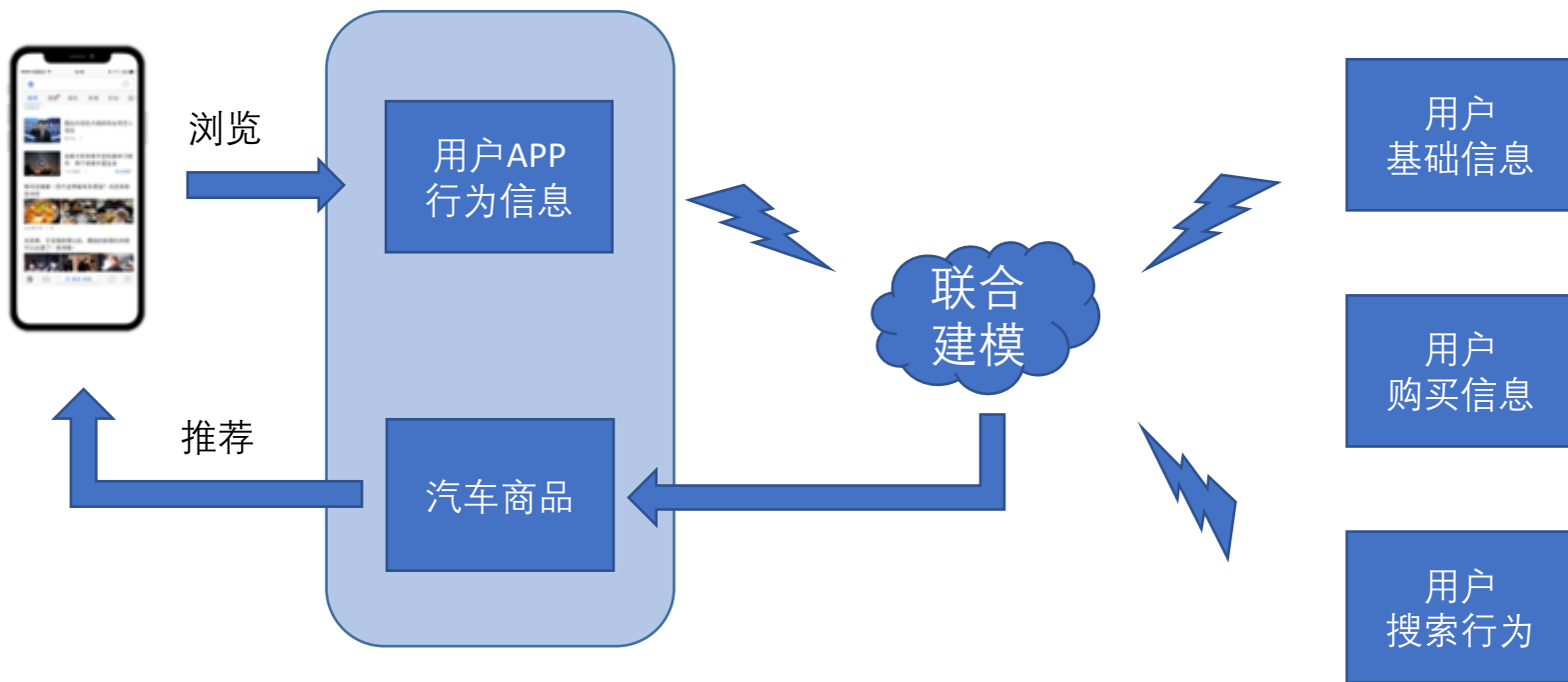


\*针对高净值人群精准投放，提高了转化率，降低了成本，同时降低对低消费人群打扰，提升用户体验。

# 教育领域营销



# 汽车营销领域（推荐）



# 实现多样性

## 同态类型:

部分同态 (Partially Homomorphic)  
有点同态 (Somewhat Homomorphic)  
全同态 (Fully Homomorphic)

## 安全模型:

无攻击模型  
半诚实模型 (Semi-Honest Adversary)  
恶意模型 (Malicious Adversary)  
隐蔽敌手模型 (Covert Adversary)

## 调度方式:

算子层面: Operator Layer  
算法层面: Algorithm Layer  
组件层面: Component Layer  
引擎层面: Engine Layer

## 计算平台:

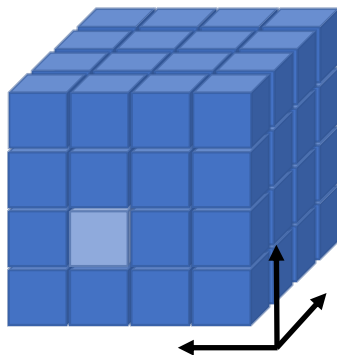
Local  
Docker/Kubernetes  
Spark  
MPI

## 通信方式:

Socket通信  
RPC通信  
MQ通信  
P2P通信

## 存储方式:

File  
DataFrame(vTable)  
DataBase





# 劣势

- 速度慢 (计算复杂度, 网络波动)
- 原理复杂, 安全性难以度量(多样性)
- 可用算法有限

# 期待，有没有：

1. 速度快
2. 简单易懂
3. 有充足的算法？

是否可以通过“硬件”解决？



# 03

## 可信执行环境

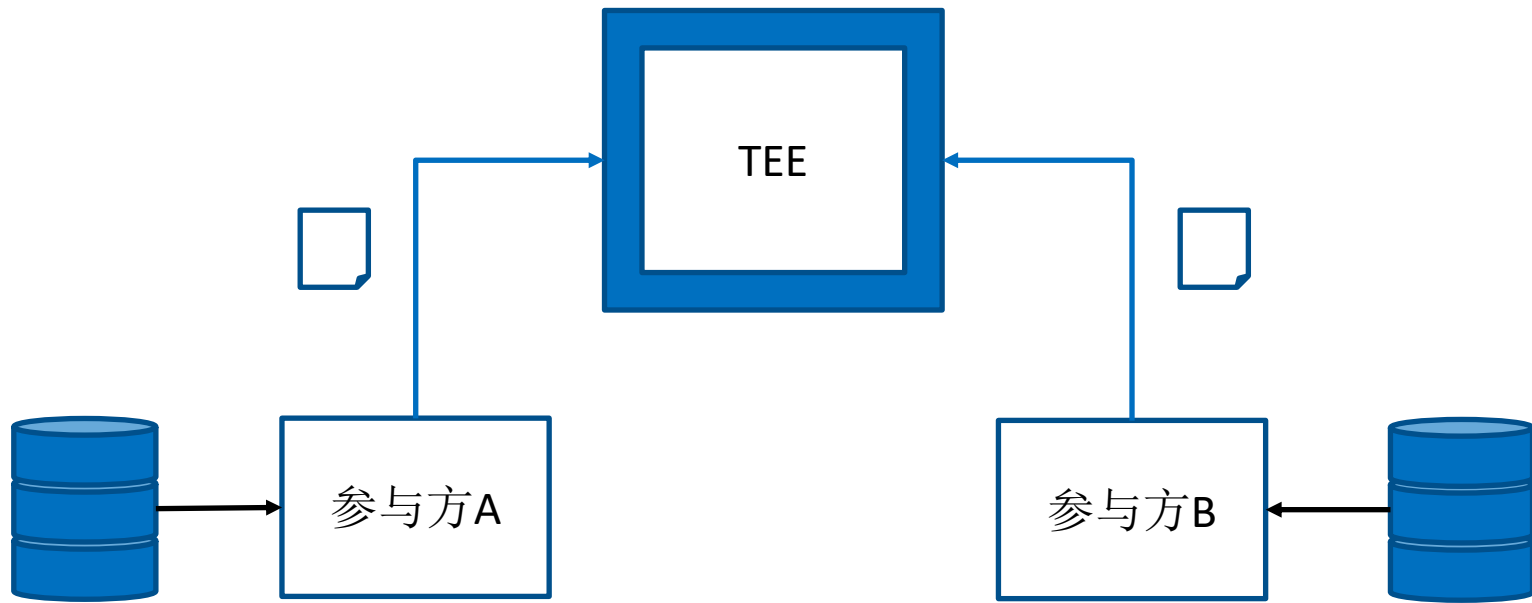
替换文字内容，点击添加相关标题文字



# 优势

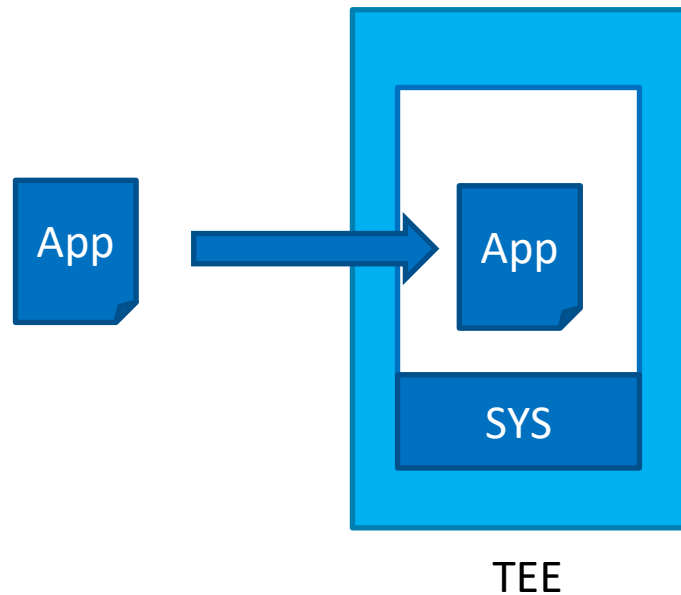
- 计算速度快：大数据，低延迟
- 执行逻辑较清晰，容易度量其安全
- 移植算法成本较低，可以用的算法很多

# 可信执行环境 (TEE)

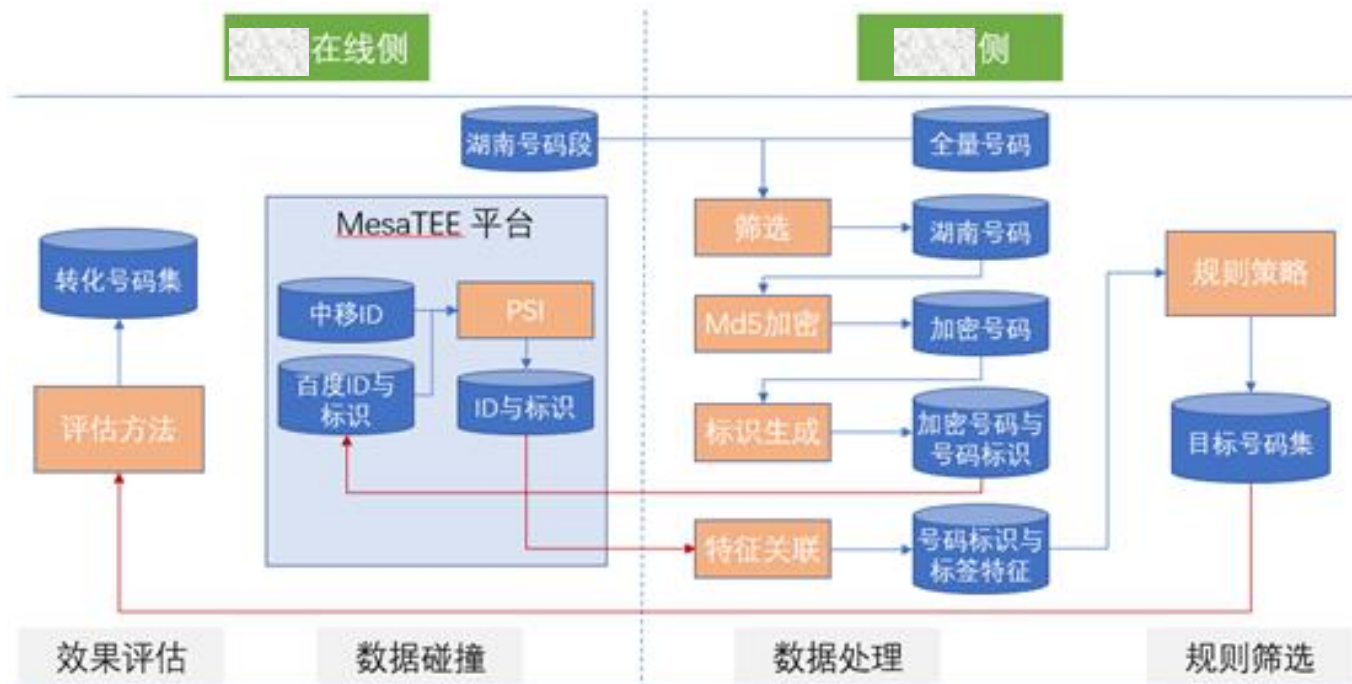


# 更加简单化的TEE

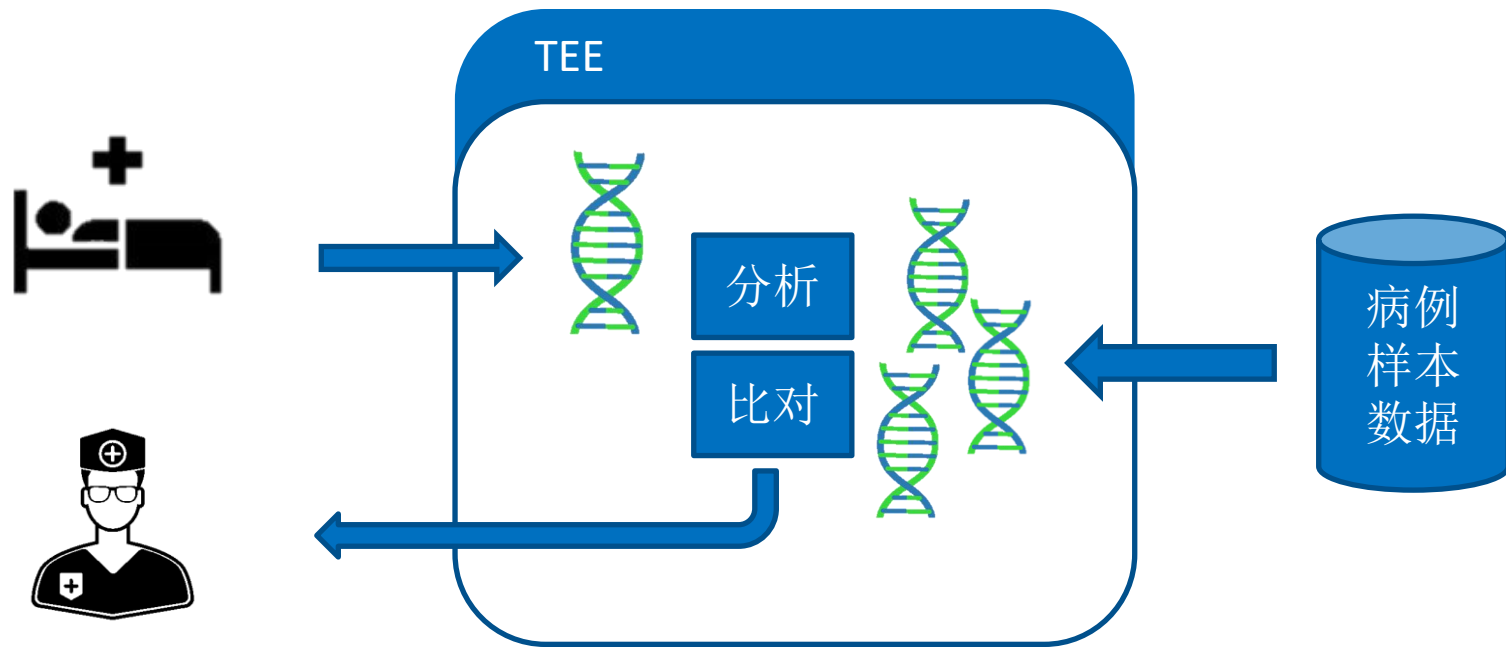
- 基于LibOS的可信执行环境
  - Occlum
  - Gramine(GrapheneSGX)
- 基于可信虚拟化技术
  - AMD: SEV,SEV-ES
  - INTEL: TDX
  - ARM: CCA



# 通信领域营销



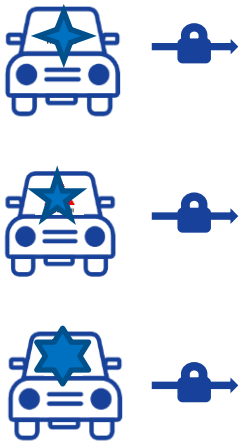
# 生物基因领域



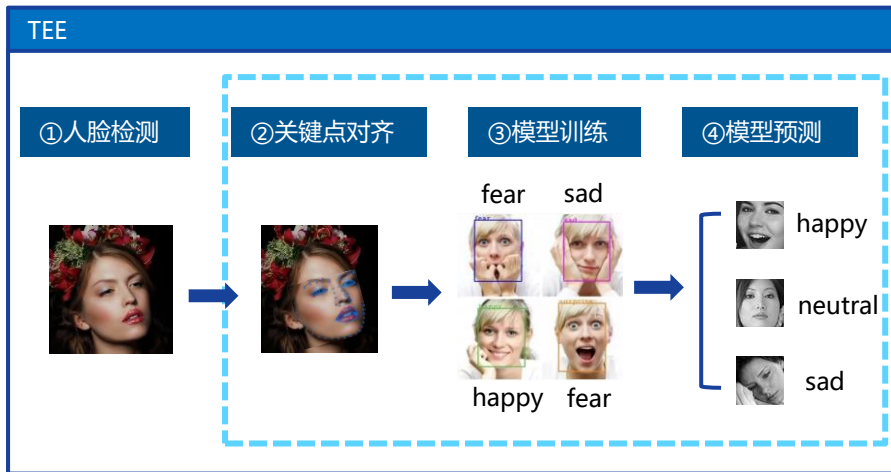


# 汽车音乐推荐

## 数据源



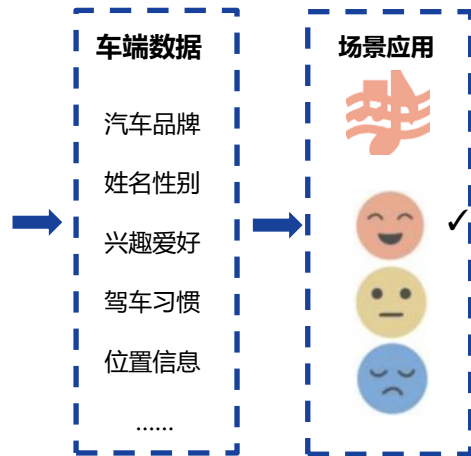
## 数存储、处理与建模



- 人脸检测：有效识别图片中的人脸，精确定位人脸位置，返回高精度人脸；
- 关键点对齐：对图片中的人脸进行关键点定位，并返回人脸关键点坐标位置，包括人脸轮廓、眼睛、眉毛、嘴唇以及鼻子轮廓等；

- 模型训练：通过读取训练集海量图片，进行情绪识别模型多轮迭代，输出模型参数；
- 模型预测：输入测试集人脸图片，调用情绪识别模型，输出情绪识别结果（happy/sad/neutral）

## 数据融合与应用



- 数据融合与应用：结合用户情绪识别结果与车端其他数据，可以用于车载音频切换，随着客户的情绪，切换相关主题音乐

# 劣势

- 数据需要加密后传到TEE
- 依赖硬件
  - INTEL: SGX1/2, TDX
  - AMD: SME, SEV, SEV-ES
  - ARM: TRUSTZONE, CCA
- 算法需要移植适配
- 需要权衡安全性和高可用: 数据重放

# 期待，有没有：

1. 速度更快
2. 简单易懂
3. 所有算法和工具可用
4. 不依赖硬件

是否可以通过“隔离”解决？



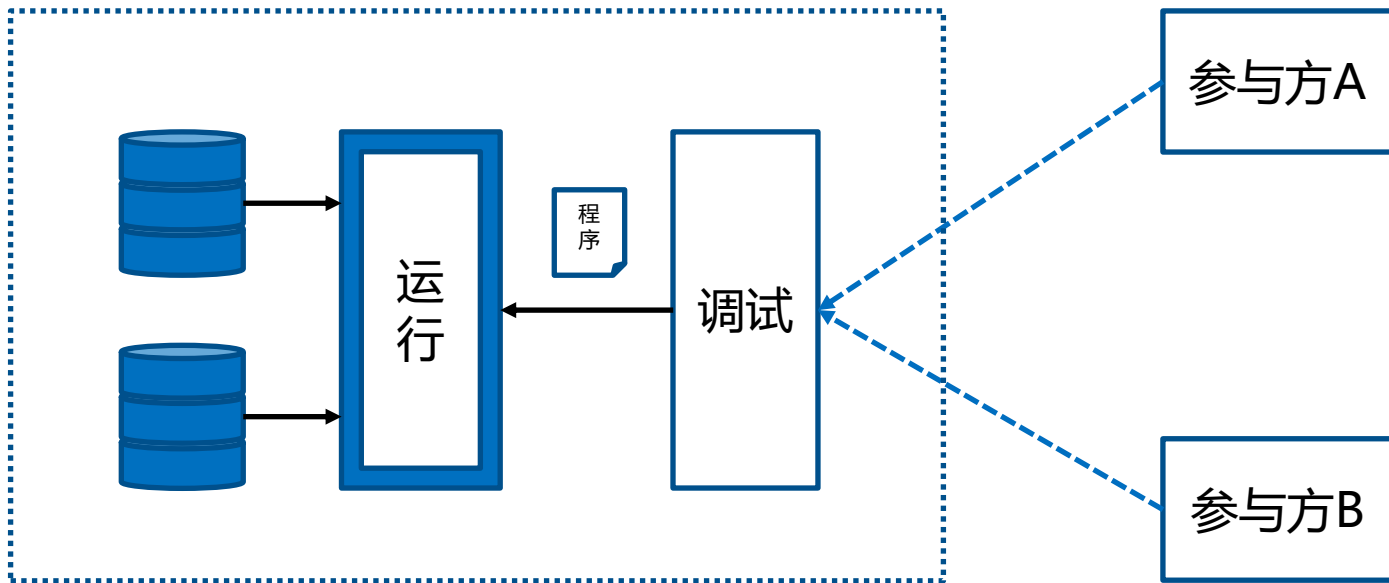
# 04

## 安全数据沙箱

替换文字内容，点击添加相关标题文字



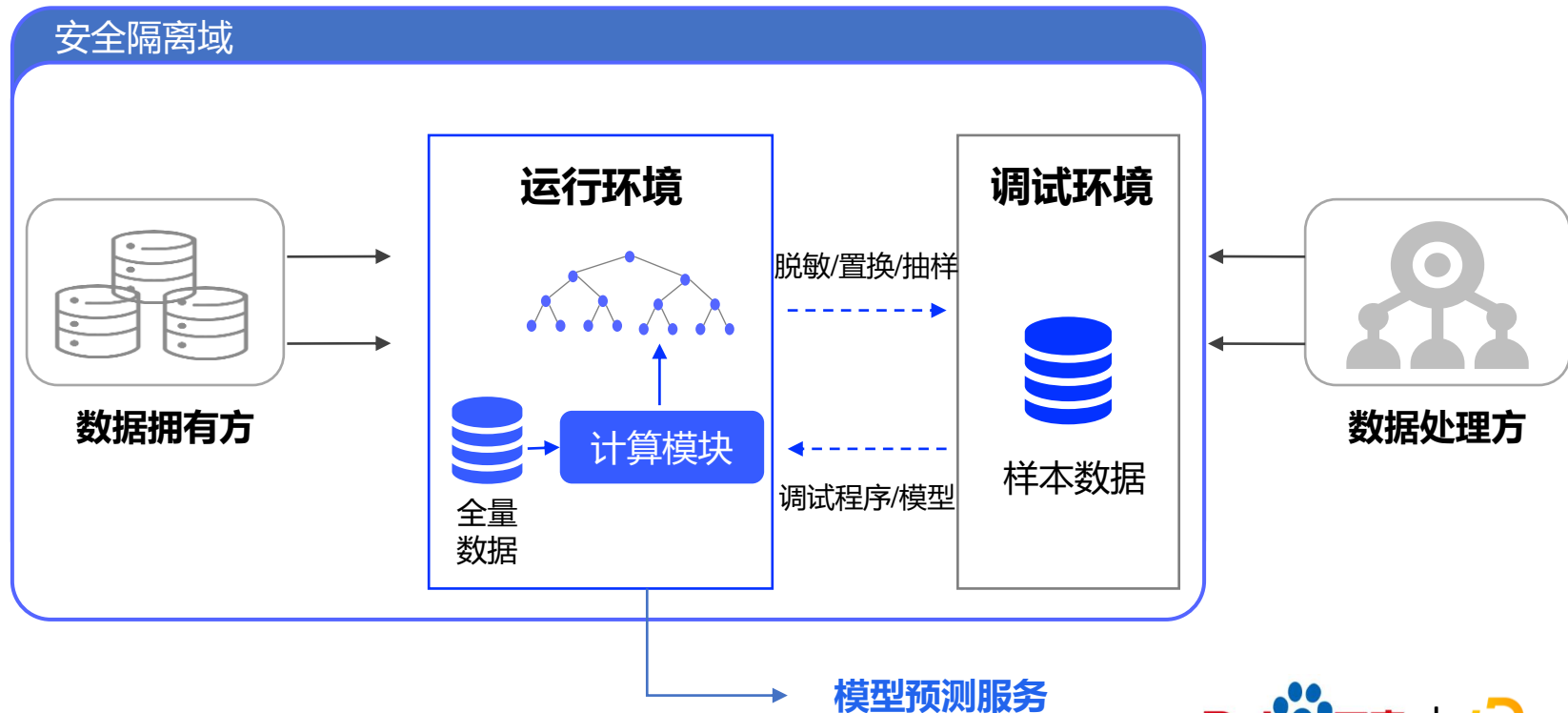
# 安全数据沙箱 (SANDBOX)



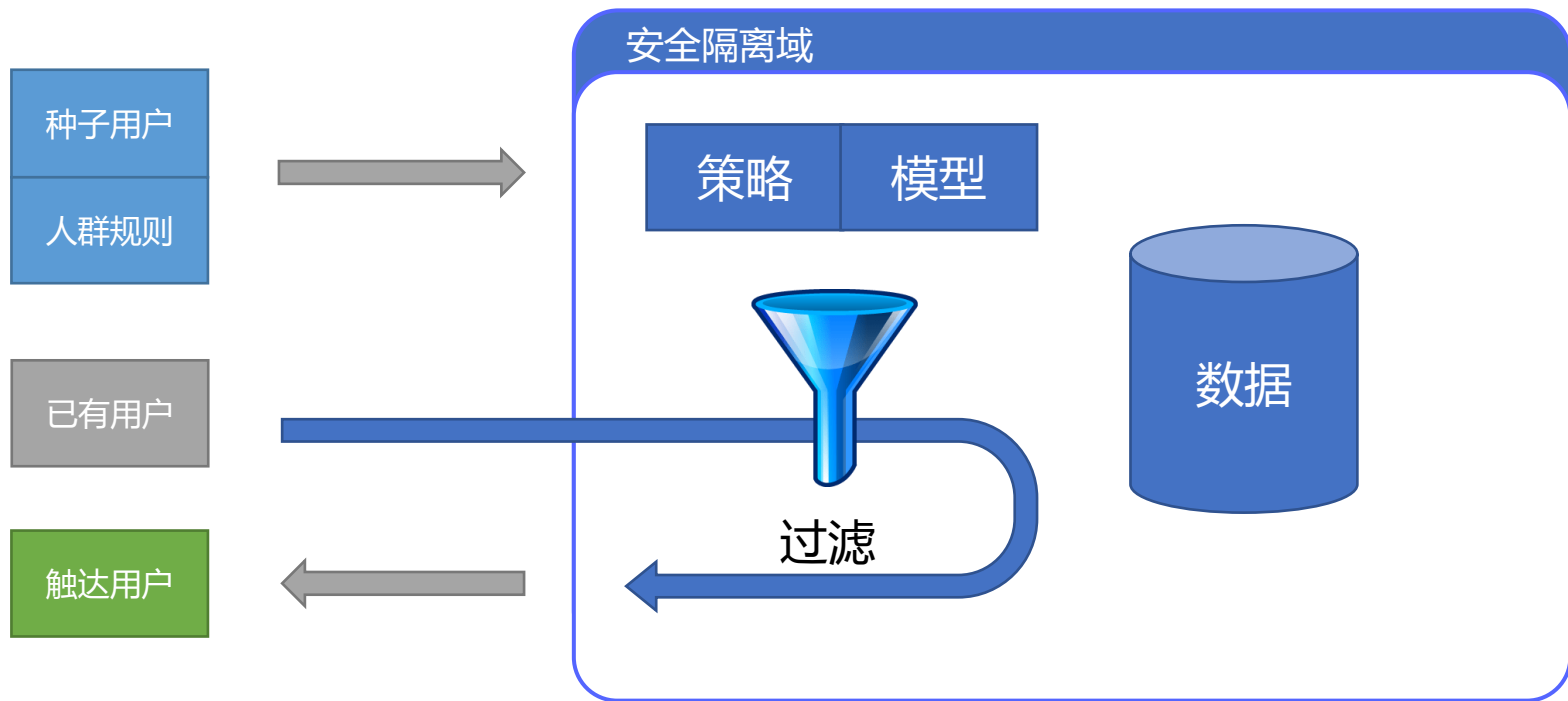
# 优势

- 简单，速度快，几乎无损耗
- 可大规模应用，适用特大的数据场景
- 不依赖物理硬件
- 所有算法都可以使用

# 广告传媒领域

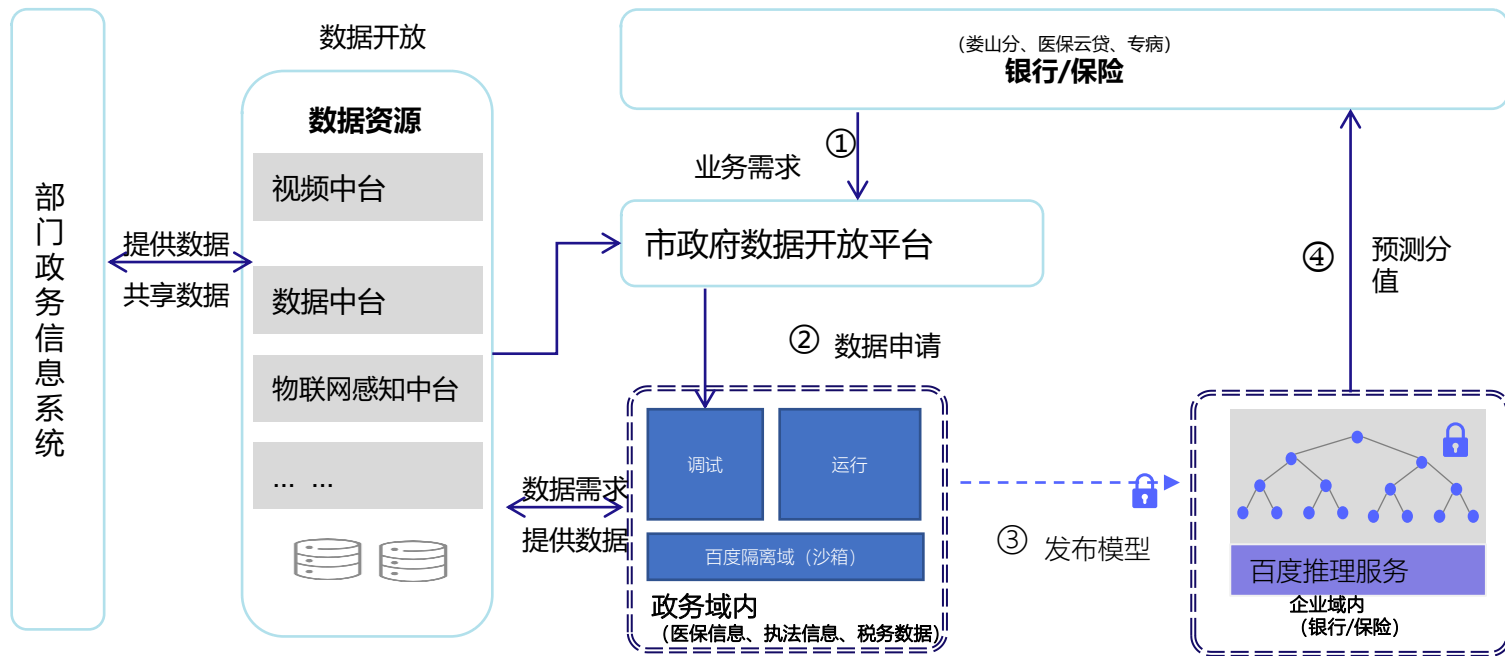


# 金融理财营销





# 政务数据在营销里的应用



# 劣势

- 仅保护平台方的数据安全

# 期待，有没有：

1. 不依赖硬件
2. 速度快
3. 简单易懂
4. 所有算法和工具可用
5. 两方的数据可进出

是否有“全覆盖”的产品？



# 05

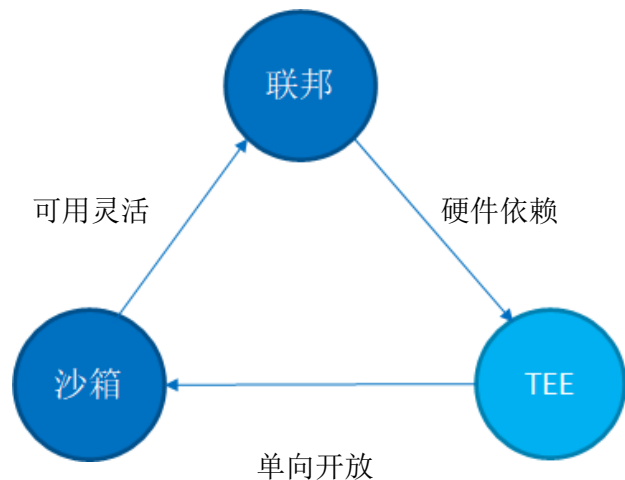
## 百度产品矩阵

替换文字内容，点击添加相关标题文字



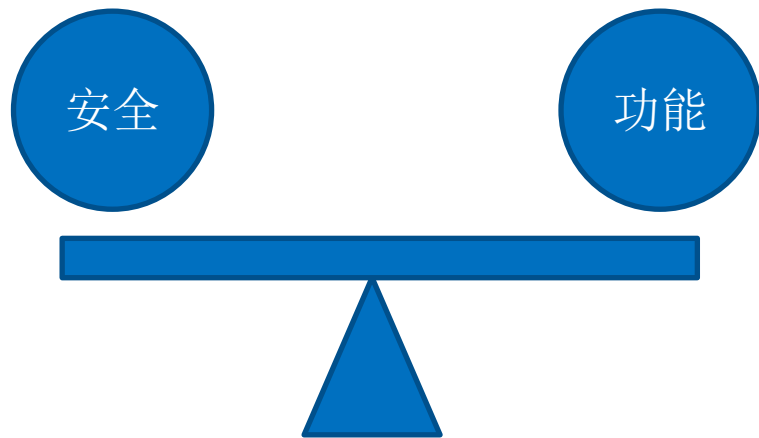
# 隐私计算中的舍得

- 联邦学习FL
  - 舍：计算效率，可用的算法
  - 得：无硬件要求，原始数据不出域
- 可信执行环境TEE
  - 舍：硬件的通用性，加密原始出域
  - 得：更快的计算效率，一定算法的可扩展
- 安全数据沙箱SBX
  - 舍：仅保护数据方的安全
  - 得：超大规模的数据计算，复用所有的算法和工具



# 恰当的应用隐私技术

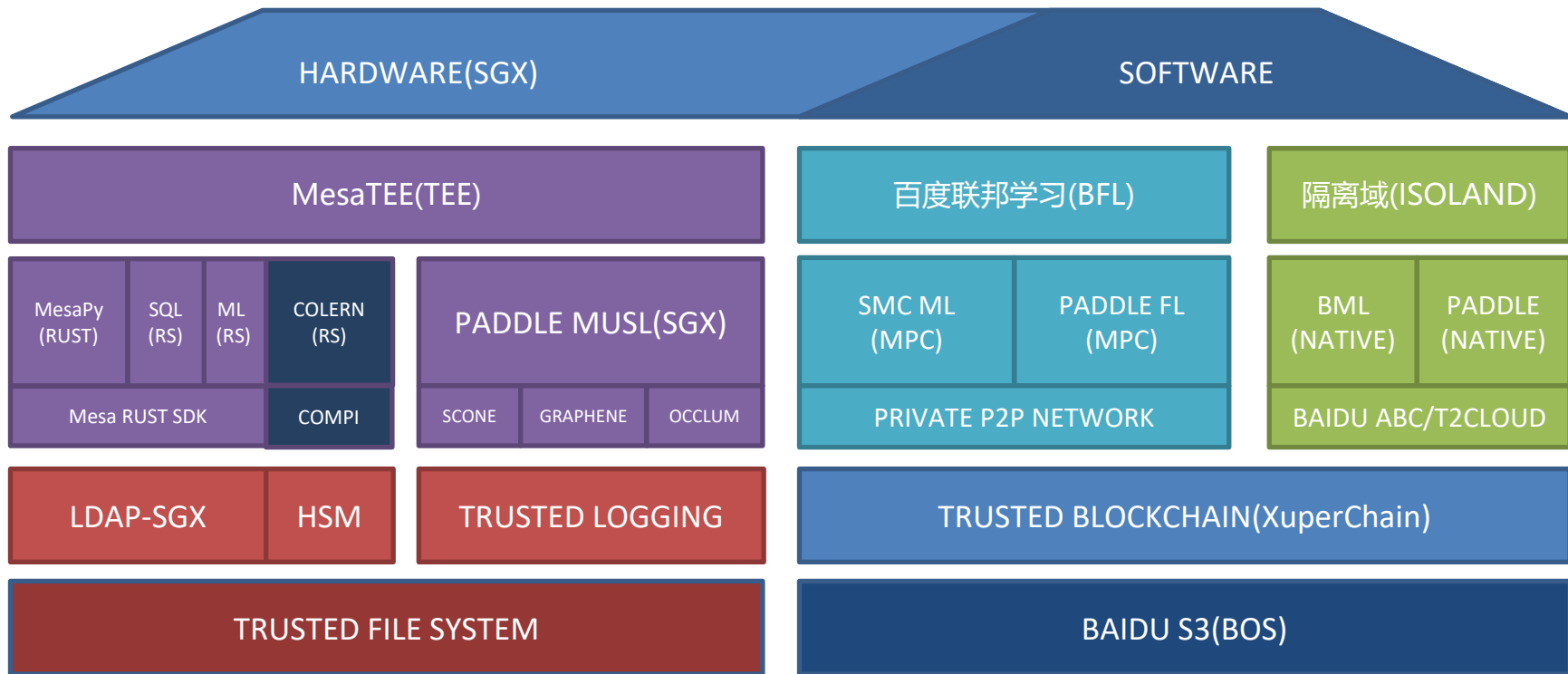
- 明确保护/防御对象
- 明确数据的规模
- 明确操作的范围
- 明确实施的运行环境



# 隐私计算一指禅

- 数据小，两方不出，常规算法：**联邦学习**
- 数据中，两方汇聚，专用算法：**可信执行环境**
- 数据大，一方不出，任意算法：**安全沙箱**

# 百度隐私计算：技术框架图





# 06 未来趋势分析

替换文字内容，点击添加相关标题文字



# 趋势1：技术融合

- 联邦FL/MPC + TEE -> 降复杂，提效率
- SANDBOX + TEE -> 使用方数据安全问题
- SANDBOX + MPC -> 外部数据引入

# 趋势2：软硬结合

- FL/MPC + GPU/FPGA -> 深度学习提速

Paddle/PaddleFL, CryptGPU/CrypTen

MPC: CPU(200X) 降低到 GPU(35X), 提升10X以上

- SANDBOX + TrustedVM + Clear Container

启动加速, 不同数据, 不同等级

启动速度: 100s 降低到 10s以内

# 趋势3：互联互通

- 隐私计算平台的互通

- 中国信通院云大所《隐私计算 跨平台互联互通》系列标准
- IEEE SA《P3117 - Standard for Interworking Framework for Privacy-Preserving Computation》



实现系统应用层面的互联互通：主要包括节点管理、任务编排、任务执行、监控管理等内容。



实现各类协议层面的互联互通：主要包括算法协议、资源协议、节点交互协议三方面的内容。



实现通信层面的互联互通：主要包括加密传输机制、通信框架与接口、数据传输格式等内容。

# 非常感谢您的观看

---

Baidu 百度 | DataFun.

