



隐私计算距离大规模商用 还差了什么？





辗转了政务、医疗等多个行业后，多数隐私计算公司最终还是选择将金融业作为优先落地的行业，理由很简单：作为国内数据化程度最高的行业之一，金融业有着最为成熟的技术基础、最强烈的数据场景需求以及雄厚的购买实力。

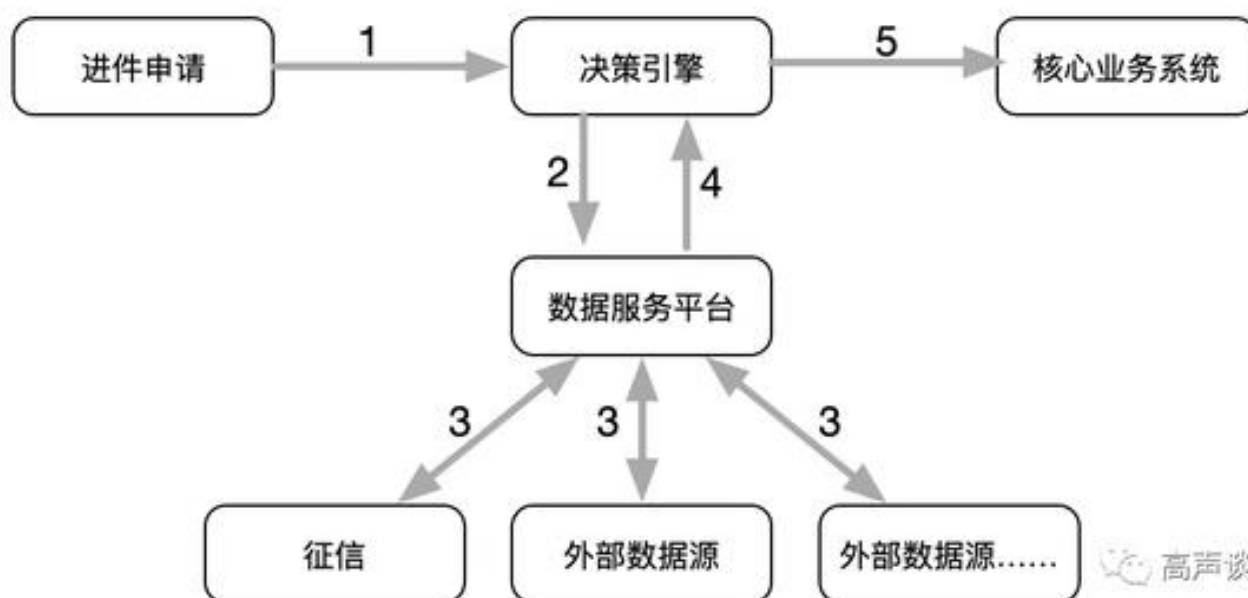
金融业也不负众望，去年下半年以来先后落地了数十个隐私计算项目，反向印证了上述观点。但是繁华背后存隐患：我们看到，上述落地案例中实验、探索项目居多，至今还没有真正上生产环境的案例出现。

其中原因有很多，从供给侧看，隐私行业的打法、产品性能与质量等方面存在这不小问题，笔者在《隐私计算公司开拓金融业的意见与建议》、《隐私计算的破局之道》等文章中有过深入分析，不再赘述。此次我们重点从需求侧分析银行等金融机构的疑虑和担心——对既有风控技术框架的

调整与改在成本。

银行传统风控技术框架

通常来讲，金融机构风控技术架构的设计思路见下图：“数据服务平台”是外部数据调用平台，承载与外部数据的统一报送与采集功能。为了方便风控建模，银行一般会单独设立针对性的风控集市用于模型训练和预测，不同目的的模型其风控集市差异较大，比如小微企业贷款与个人消费贷款的集市维度连主键信息都不一样，甚至更细致的，个人消费贷款的贷前信用评分卡模型和贷后风险表现模型也会分别建立集市进行支持。



我们知道，风控模型分为模型训练和模型预测（模型运行）两个阶段。

模型训练阶段是事前、非实时的，一般是先从数据集市中的存量数据按照一定规则提炼样本，按照主键信息从数据源接口批量查得结果后，结合数据集中银行已有的数据特征值一同建模，模型训练成型后部署在决

策引擎之中，准备上生产。

在模型预测（模型运行）时，从进件前端或是核心系统传输来的进件申请将客户主键信息推动至决策引擎，引擎实时掉起数据集市和数据服务平台中入模数据的 API 接口查得数据，一同进入模型中生成模型结果。然后，决策引擎根据核心系统需要展示的决策结果，将各类模型结果打包形成决策包后，统一推送至核心系统及进件前端进行展示。这便是最简单的模型训练、预测的全流程。



隐私计算重构了风控技术框架

我们首先分析隐私计算承载着何种职能。

由于部分设计个人隐私的数据的价值必须通过隐私产品交互获得，且这些数据无法移出隐私产品，必须在隐私产品内进行模型训练和预测，因此隐私计算产品替代了决策引擎的计算模块和部分数据服务平台职能。

模型训练阶段还是事前、非实时的，与之前不同的是，银行需将从数据集中提炼的样本（包含主键信息和银行端存量特征值）一同灌入隐私计算产品之中，在隐私产品中与不同数据源的节点互动迭代生成模型。需要强调的是，此时的成型模型不能移出，只能存放在隐私产品之中。

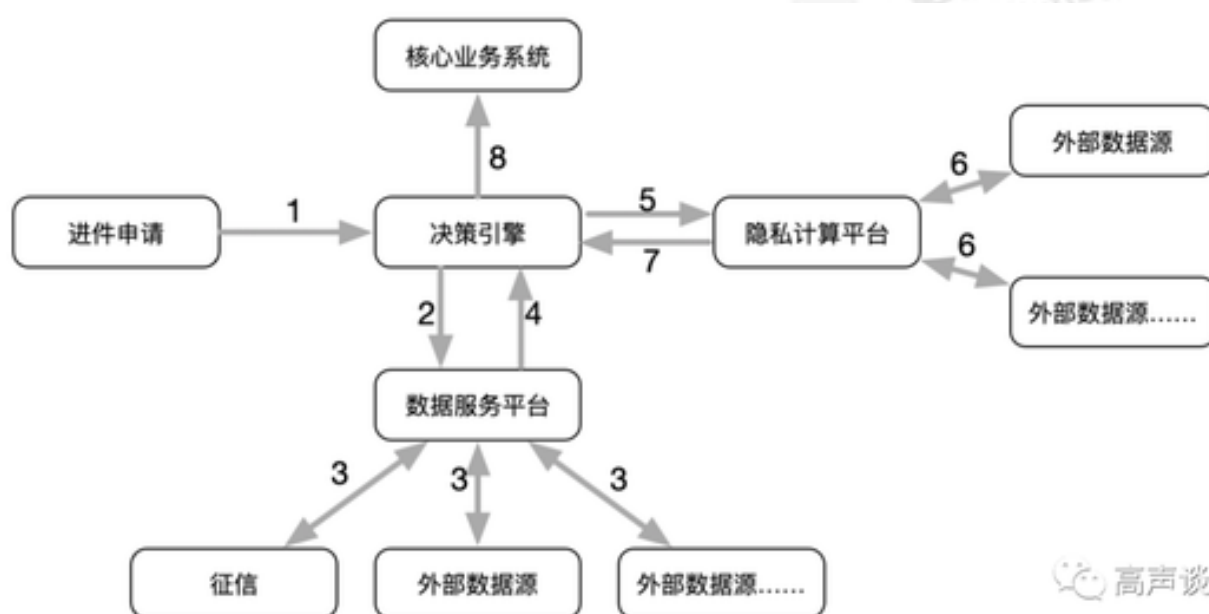
隐私行业所说的建模速度损耗就是指的该阶段，受算力和网络传输的影响，当前的隐私计算建模效率至少是明文状态的百倍。这还不包括从数据集市到隐私计算产品之间的传输环节。

模型预测（运行）时，由于银行接入了央行征信、百行征信等很多其它合规数据源（而这些数据源是明文差得的），因此还是不能绕开决策引擎和原有的数据服务平台。合意的做法是：还是将进件申请获得的客户主键信息推动至决策引擎，先由决策引擎通过数据服务平台，从央行征信等系统查得明文数据，从风控集中查得银行存量数据标签后，与主键信息一起灌入隐私计算产品之中，在隐私产品中实现与三方隐私数据节点的交互，并一同导入成型模型中计算出结果。

由于隐私产品很难结合核心系统和进件前段系统进行个性化改造，因

此将多个模型结果打包生成决策包的工作还需决策引擎承担。隐私计算平台要将计算得出的模型预测（模型运行）结果推给决策引擎，形成指令集后由决策引擎推送至核心系统及进件前端进行展示。

因此，调整后的风控技术架构见下图。与之前相比，合意的设计框架多了隐私计算平台这一支路。



隐私计算适合从简单、独立场景做切入

很多银行等金融机构采用的是统一风控平台架构，多年沉淀形成的风

预览已结束，完整报告链接和二维码如下：

https://www.yunbaogao.cn/report/index/report?reportId=1_42045

