隐私计算产品白皮书

V1.0

亚信科技隐私计算产品，帮助企业构建可信数据流通与交易，激活数据要素价值，释放数据要素红利。

声明

任何情况下，与本软件产品及其衍生产品、以及与之相关的全部文件（包括本文件及其任何附件中的全部信息)相关的全部知识产权(包括但不限于著作权、商标和专利）以及技术秘密皆属于亚信科技（中国）有限公司（“亚信科技”）。

本文件中的信息是保密的，且仅供用户指定的接收人内部使用。未经亚信科技事先书面同意本文件的任何用户不得对本软件产品和本文件中的信息向任何第三方（包括但不限于用户指定接收人以外的管理人员、员工和关联公司）进行开发、升级、编译、反向编译、集成、销售、披露、出借、许可、转让、出售分发、传播或进行与本软件产品和本文件相关的任何其他处置，也不得使该等第三方以任何形式使用本软件产品和本文件中的信息。

未经亚信科技事先书面允许，不得为任何目的、以任何形式或任何方式对本文件进行复制、修改或分发。本文件的任何用户不得更改、移除或损害本文件所使用的任何商标。

本文件按“原样”提供，就本文件的正确性、准确性、可靠性或其他方面，亚信科技并不保证本文件的使用或使用后果。本文件中的全部信息皆可能在没有任何通知的情形下被进一步修改，亚信科技对本文件中可能出现的任何错误或不准确之处不承担任何责任。

在任何情况下，亚信科技均不对任何因使用本软件产品和本文件中的信息而引起的任何直接损失、间接损失、附带损失、特别损失或惩罚性损害赔偿（包括但不限于获得替代商品或服务、丧失使用权、数据或利润、业务中断），责任或侵权（包括过失或其他侵权）承担任何责任，即使亚信科技事先获知上述损失可能发生。

亚信科技产品可能加载第三方软件。详情请见第三方软件文件中的版权声明。

亚信科技控股有限公司（股票代码：01675.HK）

亚信科技是中国领先的软件产品及服务提供商，拥有丰富的软件产品开发和大型软件工程实施经验。公司深耕市场 30 年，在 5G、云计算、大数据、人工智能、物联网、数智运营、业务及网络支撑系统等领域具有先进的技术能力和众多成功案例，客户遍及通信、广电、能源、政务、交通、金融、邮政等行业。

2022 年，亚信科技完成收购商业决策服务领域的领先企业艾瑞市场咨询股份有限公司（「艾瑞咨询」），并整合形成新的"艾瑞数智"品牌。通过此次收购，亚信科技的核心能力从产品研发、交付服务、数据运营、系统集成延伸至咨询规划、智能决策，成为领先的数智化全栈能力提供商。

亚信科技始终致力于将 5G、AI、大数据等数智技术赋能至百行千业，与客戶共创数智价值。公司以"产品与服务双领先"为目标，产品研发围绕数智、云网、IT 及中台产品体系持续聚焦，实现行业引领，其中云网产品保持国际引领，数智产品实现国内领先，部分国际先进，IT 领域产品处于国内第一阵营。

面向未来，亚信科技将努力成为最可信赖的数智价值创造者，并依托数智化全栈能力，创新客户价值，助推数字中国。

部分企业资质

能力成熟度模型集成 CMMI5 级认证

信息系统建设和服务能力评估(CS4 级)

云管理服务能力评估证书卓越级

数字化可信服务 - 研运数字化治理能力认证

1S09001 质量管理体系认证证书

150200001T 服务管理体系认证证书

1S027001 信息安全管理体系认证证书

企业信用等级（AAA 级）证书

信息系统安全集成服务资质（二级）

信息系统安全开发服务资质（二级）

部分企业荣誉

连续多年入选中国软件业务收入百强榜单

连续多年入选中国软件和信息服务竞争力百强企业

中国软件行业最具影响力企业

中国软件和信息服务业最有价值品牌

中国软件和信息服务业最具影响力的行业品牌

中国数字与软件服务最具创新精神企业奖

中国电子信息行业社会贡献 50 强

中国人工智能领航企业

新型智慧城市领军企业

IDC 未来运营领军者

目录

1 摘要

隐私计算是采用多方安全计算、密码学和分布式技术等手段，以实现数据隐私保护和共享计算为目标的技术。它具备联邦学习、多方安全计算和可信计算等功能，可以在满足数据隐私和法规要求的前提下，打破数据壁垒，深入挖掘数据，同时保护原始数据的隐私性。

亚信科技推出了基于自主研发、技术领先的隐私计算产品，目的是让数据在技术信任机制下，以“可用不可见”的安全方式释放融合价值。产品采用业界首创的“1+X”隐私计算平台架构，提供了联邦学习建模、数据安全求交、匿踪安全查询和多方安全计算等功能。

亚信科技创新开发了端到端的数据要素市场化解决方案，该方案完全基于芯片能力可信计算能力的资源底座，并结合自主研发的软件应用，为治安、新能源、金融、智慧城市、电信、医疗等行业提供完善高效的解决方案，能够有效平衡数据安全和数据共享的需求，保护用户隐私，促进数据的可持续利用和创新。

隐私计算产品的应用场景包括联合营销、联合风控、智慧医疗、电子政务等，可以帮助企业、机构等实现更安全、合规的数据处理和数据分析，提升业务效率和创新能力。同时，该产品还可以为数据安全和隐私保护提供更可靠的技术保障，促进数字经济的发展。

本白皮书将针对隐私计算产品从产品概述、功能架构、产品基础功能、产品特色功能、产品差异化优势、场景解决方案、客户成功故事几个方面进行介绍。

2 缩略语与术语解释

隐私计算产品常见术语如表 2-1 所示。

表 2-1术语解释

缩略语或术语	英文全称	解释
FL	Federated Learning	联邦学习是一种分布式机器学习技术，通过在多个拥有本地数据的数据源之间进行分布式模型训练；在不需要交换本地个体或样本数据的前提下，仅通过交换模型参数或中间结果的方式构建基于虚拟融合数据下的全局模型，从而实现数据隐私保护和数据共享计算的平衡，即“数据可用不可见”、“数据不动模型动”的应用新范式。
MPC	Secure Multi-Party Computation	多方安全计算指参与者在不泄露各自隐私数据情况下，利用隐私数据参与保密计算，共同完成某项计算任务。
PIR	Private Information Retrieval	匿踪查询是安全多方计算中非常实用的一门技术与应用，可以用来保护用户的查询隐私，进而也可以保护用户的查询结果。其目标是保证用户向数据源方提交查询请求时，在查询信息不被感知与泄露的前提下完成查询。
PSI	Private Set Intersection	安全求交技术，全称为“隐私保护集合交集”；是一种能够使持有数据的两方计算得到双方数据集合的交集部分，而不暴露交集以外任何数据集合信息的技术。

缩略语或术语	英文全称	解释
SQL	Structured Query Language	结构化查询语言，是一种计算机语言。用来存储、检索和修改关系型数据库中存储的数据。
TEE	Trusted Execution Environment	可信执行环境是一种能确保代码和数据在执行过程中安全、完整且不被篡改的环境。
区块链	Block chain	区块链是一种由多方共同维护，使用密码学保证传输和访问安全，能够实现数据一致存储、防篡改、防抵赖的技术体系。
互联互通网络	interconnection network	不同隐私计算技术平台部署后相互连接，通过交互与协同形成的提供跨平台联合隐私计算服务的网络。
隐私计算	privacy-preserving computation	在保证数据提供方不泄露原始数据的前提下，对数据进行分析计算的一系列信息技术，保障数据在流通与融合过程中的“可用不可见”。

3 产品概述

AISWare PEC 亚信科技隐私计算产品（以下简称“本产品”），应用于数据可信流通场景，实现数据可用不可见，用途可控可计量。产品依托多方安全计算、联邦学习、区块链等技术，支持面向通信、金融、政务、交通、能源等各行业的数据流通应用构建，激活数据要素价值。

3.1 趋势与挑战

目前，隐私计算技术正处于高速发展阶段，可解决企业和机构面临的数据合规难题，为数据安全制度落地提供有力的技术支撑。然而，隐私计算在安全、性能和数据的互联互通等方面仍存在挑战。隐私计算的发展，可以从技术、应用和法规三个层面来分析。

- 技术层面

虽然隐私计算技术在近年来有了很大的进步，但是还没有达到完全的工业化水平，很多技术还处于实验室或原型阶段，需要更多的研究和验证，以提高性能、可扩展性、兼容性和易用性。隐私计算技术还缺乏统一的标准和规范，不同的技术和平台之间难以互相操作和协同，导致数据流通分析的效率和质量受到影响。

- 应用层面

在实际应用中，获取足够多且具有高质量的数据并不总是容易的。某些领域可能面临数据稀缺的问题，而某些数据集可能存在缺失值或不完整的信息。这会对机器学习算法和数据驱动的方法的准确性和鲁棒性造成负面影响。

- 法规层面

在制定隐私计算相关法规时，需要平衡个人隐私保护和企业创新发展的关系。过于严格的法规可能会限制企业的创新和发展，而过于宽松的法规则可能难以有效保护个人隐私。因此，需要制定具有针对性和可操作性的法规，以促进隐私计算技术的发展和应用。

3.2 产品定义

亚信科技隐私计算产品（AISWare PEC），基于首创的“1+X”架构，能够实现异构算子之间开放互联，并在云原生架构上实现算子组件透明化统一管控；产品将复杂密码学技术统一封装，提供图形化的开发方式，降低使用门槛。企业可通过隐私计算产品接入数据资产，无缝对接数据生产流程，快速构建隐私计算跨行业应用。

本产品主要以软硬一体化形式（AISWare PEC-Appliance）进行交付，支持x86、ARM 架构服务器；含可插拔安全增强模块，包括硬件密码模块、TEE 硬件模块、可信计算模块；可按需选择 GPU/FPGA/ASIC 等加速模块。此外，也支持以软件平台形式（AISWare PEC-Worker）交付，通过容器化、虚拟机等方式进行部署。

3.3 产品定位

本产品支持两方以上的数据进行隐私计算。由数据「提供者」通过与隐私计算「服务商」的合作，将数据接入本产品；数据「服务商」向数据「使用者」提供隐私计算技术的调用服务。

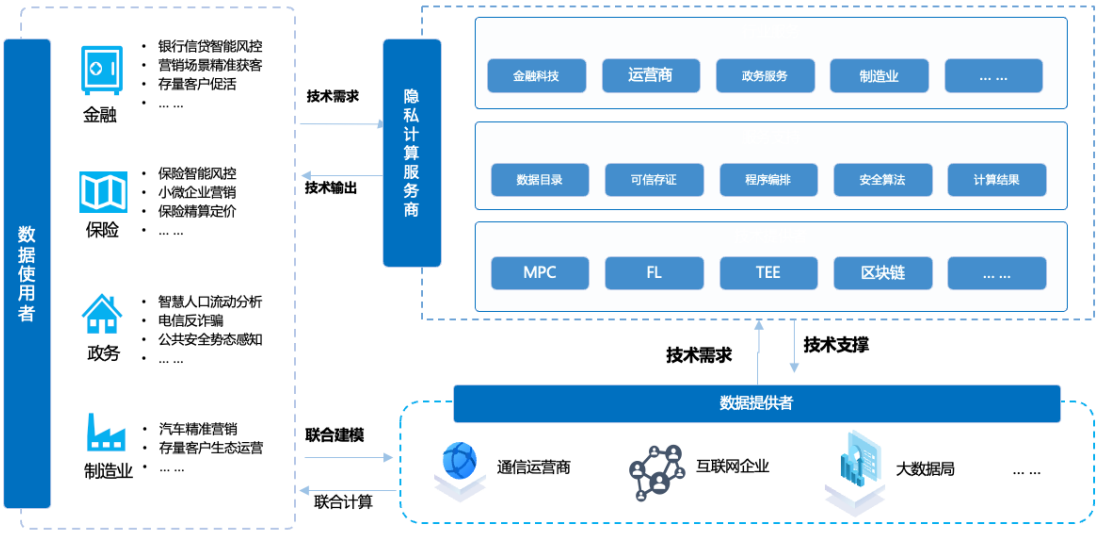


图 3-1隐私计算产品业务图

- 数据提供方：拥有大量数据的企业，如通信运营商、互联网企业、大数据局等，通过本产品提供数据。
- 隐私计算服务商：通过 MPC、FL、TEE、区块链等实现技术应用，服务于金融、运营商、政务、制造业等行业。

- 数据使用者：各个行业可根据自身的业务需求，与数据提供方进行端到端的隐私计算。

4 产品功能架构/产品体系

本产品基于数据资产管理、多方安全计算、联邦学习、区块链等数字化技术，实现数据的“可用不可见”；连接企业及行业中的数据孤岛，帮助企业构建可信数据流通与交易，激活数据要素价值，释放数据要素的巨大红利和能量。

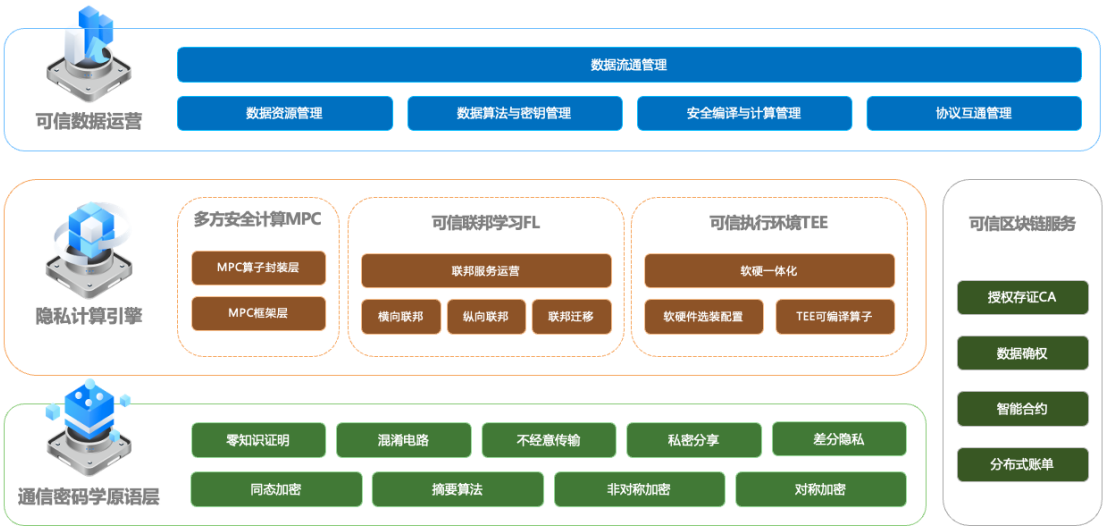


图 4-1隐私计算产品构架图

- 通信密码学原语层：基于同态加密、不经意传输、秘密分享等密码学协议，完成密码学原语的封装。
- 隐私计算引擎：将核心能力封装为多方安全计算、可信联邦学习、可信执行环境等隐私计算应用。
- 可信区块链服务：在隐私计算过程中，对计算前、计算后的数据进行上链，提供数据授权、数据可信追溯的作用。
- 可信数据运营：包括数据资源管理、安全管理、算法管理、协议互通管理等功能，以满足数据安全流通需求。

5 产品基础功能

AISWare PEC 的基础功能包括异构互联管控、高性能安全求交、联邦学习模型开发、运营服务门户、数据分级管控、智能运营驾驶舱等。

图 5-1隐私计算平台功能构架图

- 异构互联管控：基于互联互通协议管理、隐私计算业务构建功能，以实现隐私计算产品的业务应用，以及各平台间的数据流通。
- 高性能安全求交：通过优化计算资源（如：内存的优化），实现安全求交算子的高性能处理能力，提供亿级的数据计算功能。
- 联邦学习模型开发：基于同态加密实现各参与方“数据不出库”的联合模型开发，支持横向联邦、纵向联邦，支持逻辑回归、决策树、k 均值、神经网络等常用机器学习算法。
- 运营服务门户：根据数据集的信息进行参与方的合作入驻申请，同时支持入驻进度的查询等功能。
- 数据分级管控：数据分类分级、安全策略等数据安全管理体系建设，实现数据的安全治理和全生命周期的防护。

- 智能运营驾驶舱 对合作运营的整体情况、双方机构的运营情况进行分析。支持对各方合作运营的数据、项目、算子算法、任务监控等内容的统计。

作为融合了软件平台和硬件技术于一体的专用设备，本产品还提供硬件密码模块、TEE 模块、以及可信计算模块，支持可插拔的联邦学习、多方安全计算、可信执行环境、区块链能力。通过与运营商大数据能力的结合，提供场景化服务模板，提供开箱即用的隐私计算业务服务，赋能生产金融、政务、制造、医疗、运营商等行业。

图 5-2隐私计算一体机功能架构图

本产品支持 Mini、Standard、Jumbo 三种规格 ,不同规格的产品定位如下。

图 5-3隐私计算一体机产品规格

- Mini :主要面向小批量隐私数据处理场景，内置多方安全计算或联邦学习一种，可以为客户快速提供隐私计算基础环境。
- Standard :面向金融、政务、零售等企业级隐私计算场景，支持多方安全计算、联邦学习、区块链能力，为企业提供开箱即用的隐私计算服务。
- Jumbo :支持所有企业级隐私计算场景，并提供专业安全加固及硬件协同加速，为企业提供复杂算法的大规模数据量隐私计算应用需求。

6 产品特色功能

本产品的特色功能主要包括，首创的隐私计算“1+X”架构、互联互通递进式纳管、基于场景化服务打造隐私服务产品和一站式安全 AI 协作等，具体如下。

6.1 首创隐私计算“1+X”架构

亚信科技首创的 1+X 隐私计算平台集成架构，可实现异构算法的互联互通、快速集成，通过可视化编排界面快速组装跨行业的隐私计算应用。该架构包含技术底座、核心功能、对外开放能力三部分，以下进行详细介绍。

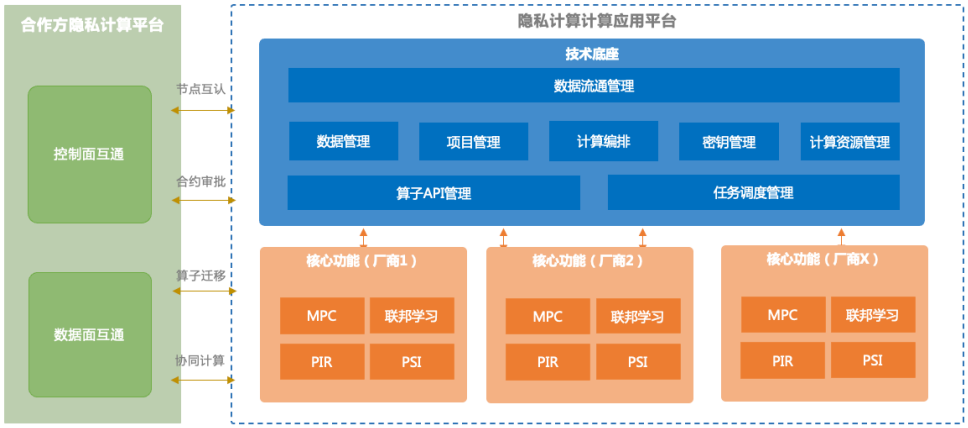


图 6-1 1+X 隐私计算平台集成架构

- 技术底座负责数据流通、数据管理、合约管理、密钥管理、算子 API 管理、资源管理与计算任务的调度管理。
- 核心功能负责 MPC、PSI、PIR、联邦学习的核心算法能力。

- MPC、PIR、PSI、联邦学习等主要场景通过开放基础类算子能力并以 RESTAPI/RPC 形式对外开放；技术底座通过对算法任务的编排将基础类算子能力进行封装形成对外场景。

在应用方面，该架构集成了行业内多数主流隐私计算企业的核心算法功能，能够在保持各平台异构自治的同时达成全局性数据智能的效果。

6.2 互联互通递进式纳管

本产品从业务价值出发，推动隐私计算互联互通的标准化进程。在确保数据隐私安全的前提下，从技术层面实现互联互通。

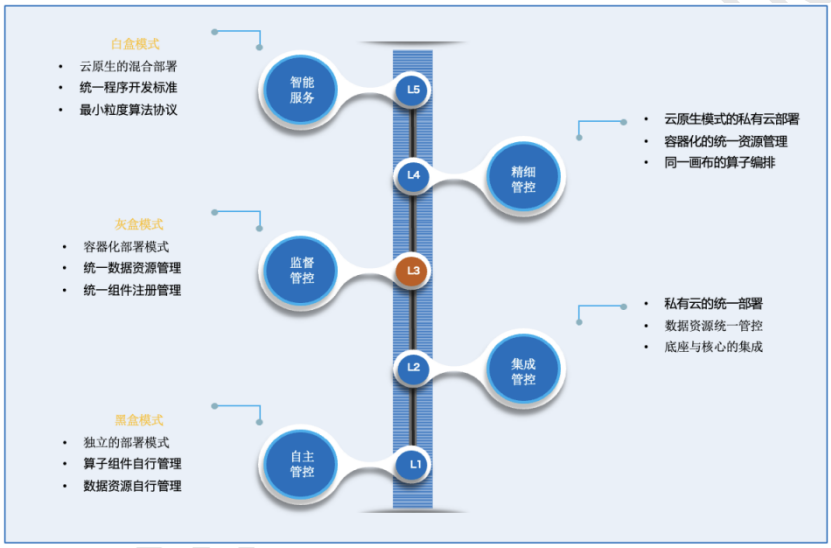


图 6-2互联互通递进式纳管

通过 L1-L5 的递进式纳管，将异构隐私计算产品的数据资源、算子组件、算法协议间的互联模式完成从黑盒到灰盒、从灰盒到白盒的转变；从而实现多边信任增强，提升各行业隐私计算互通效率。

6.3 场景化隐私数据服务

本产品通过采用多方安全计算、联邦学习等技术，可以在保证数据隐私的前提下进行数据共享和计算，从而满足金融行业的不同需求；可应用于多个场景，如联合风控、联合营销、反欺诈等。

图 6-3场景化服务

6.4 场景化一站式安全 AI 协作

本产品提供从数据准备、联邦建立、联合训练到模型部署、联合推理的全流程可信联邦学习拉通能力，通过低门槛、开放普惠的联合模型的开发、应用、服务，助力联合建模能力在垂直行业低门槛落地。

- 可视化的联邦模型编排：提供拖拽式的联邦模型开发功能，开发者无需再编辑复杂的配置文件，通过简单的界面操作和配置完成联邦训练任务开发。
- 开放普惠：企业可平滑接入生产环境自有数据，大幅降低数据对接成本。多种存储类型、多种规模数据合规接入。支持异构计算引擎互联互通，支持与异构平台算法组件对接和扩展，算法组件配置可热插拔。

6.5 软硬融合、协同加速

本产品采用软硬件结合的协同计算，模型训练速度、密态计算速度性能大幅提升，算法性能提升了 5～10 倍。支持异构算力协同加速，将复杂运算转移至硬件设备执行，支持高并发、低延迟，大幅提升算法并行处理效率。

- 实现即插即用的异构硬件加速：支持 FPGA、DCU、ASIC 等加速卡协同加速，实现即插即用的异构硬件加速。
- 显著的端到端性能提升：通过外挂硬件加速卡，整体的端到端性能提升了 5~10 倍，单个算子性能提升 10 倍以上。

- 灵活易用的应用层开发：支持多种同态计算算子，构建 Numpy 的异构加速算子接口，支持算子模块化 API 调用。
- 软硬协同加速多种场景：支持多种隐私计算计算任务包括隐私查询、安全求交、特征工程、联合建模、联合预测等。

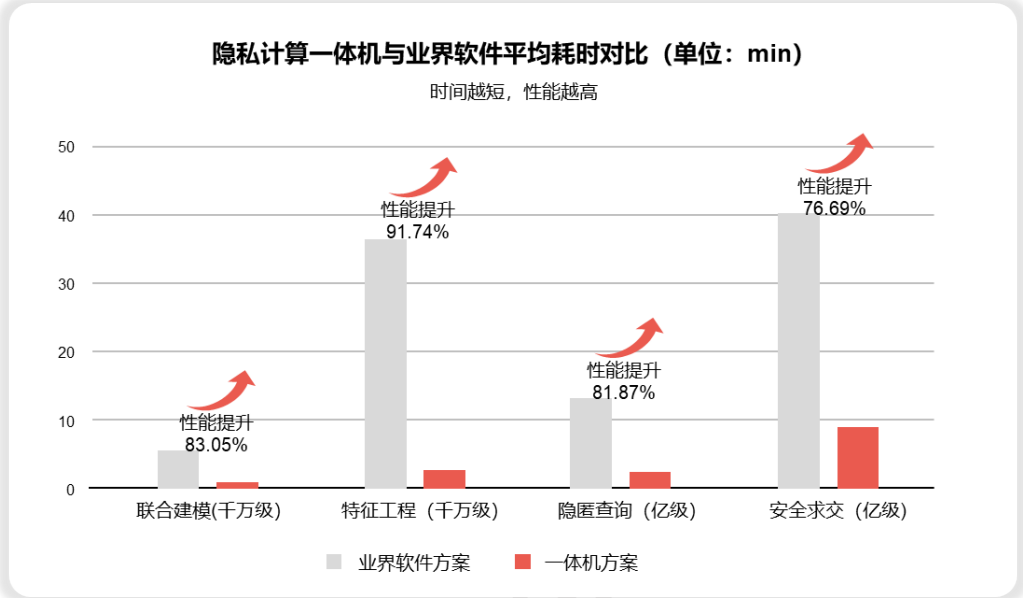


图 6-4隐私计算一体机性能对比

6.6 芯片适配、自研可信

本产品全栈、全流程覆盖支持硬件芯片、操作系统、中间件。并自主研发隐私计算算法软件，基于密码学原语之上实现多方安全计算、安全求交、匿踪查询、联邦学习等隐私计算解决方案；基于自主研发的隐私计算平台软件（AISWare PEC-Worker），可实现高性能、高可靠性和高安全性。

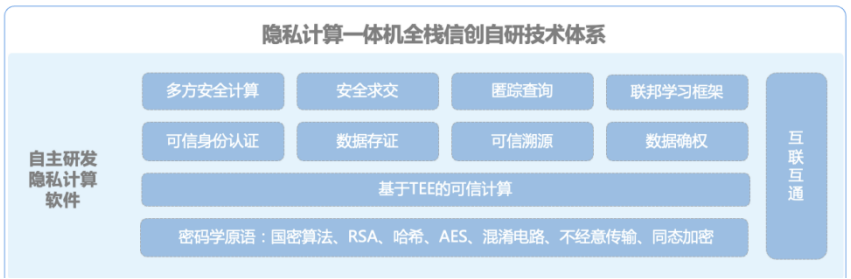


图 6-5隐私计算一体机全栈自研技术体系

7 产品差异化优势

本产品在隐私计算领域的专业优势可以从开放互联、引领标准、开箱即用、高性能加密等方面体现。

7.1 开放互联

开放式 1+X 隐私计算框架是本产品提供的一个创新解决方案，它支持业界多种算法的可插拔式一键集成，为隐私计算领域提供了一个灵活、高效、安全的平台。通过该框架，用户可以方便地集成不同厂家的算子，实现跨平台、跨设备的隐私计算。

- 灵活性：1+X 架构允许在保证核心功能的前提下，根据不同的需求和场景，灵活地选择和组合不同的算法和组件，这可以极大地提高平台的适应性和灵活性。
- 高效性：1+X 架构通过优化算法和组件之间的交互，最大限度地提高了计算和通信的效率，从而实现了更高效的数据处理和分析。
- 开放性：1+X 架构支持广泛的开放性和可扩展性，可以方便地与其他平台、系统和应用进行集成和交互，从而可以充分利用现有的技术和资源，降低开发和集成的成本。

7.2 引领标准

本产品首创的 1+X 隐私计算平台集成架构，为隐私计算领域的互联互通提供了统一的标准和指导，有助于实现不同平台之间的数据共享和计算，推动跨行业的数据合作和创新发展。

该标准集成了行业内多数主流隐私计算企业的核心算法功能，可以在保持各平台异构自治的同时达成全局性数据智能的效果。这种分级纳管模式支持管理系统、算法协议及计算原语的松耦合设计，进而实现递进式互联互通的安全可视化，提高了安全层面的可解释性，让用户掌握更强的系统运营能力。

此外，该规范进一步细化接口标准体系，形成“1+X”跨平台互联互通 L1 级~L5 级五种分级纳管模式。每个级别都有相应的要求和规定，以满足不同场景下的互联互通需求。

7.3 开箱即用

本产品预装了多领域行业样板间服务，包括针对不同行业、不同数据类型和应用需求的隐私计算方案，如企业风控、精准营销、客户流失等；可实现一键部署、开箱即用的使用体验。

7.4 高性能加密

产品采用了多种高效加密技术，以确保数据的安全性和隐私性。这些加密技术包括同态加密、零知识证明、多方安全计算等，可以实现对数据的加密、解密、计算和验证等操作，同时保证计算结果的准确性和安全性。

- 同态加密是一种能够实现加法和乘法运算的加密技术，可以在不暴露明文数据的情况下进行计算，并保证计算结果的准确性。
- 零知识证明是一种能够验证某些数据或信息的真实性和有效性的加密技术，可以在不泄露数据或信息的情况下进行验证。
- 多方安全计算是一种能够实现多个参与方在不泄露各自数据的情况下进行联合计算的加密技术，可以保证计算结果的真实性和安全性。

高效加密技术是平台的重要特点之一，可以实现对数据的全面保护和安全计算，为数据隐私和安全保驾护航。

7.5 应用场景快速复制

本产品提供跨行业协作的快速复制能力，在作为生产级工具的基础上，基于对通信领域的业务积累及数据理解，通过开箱即用的场景化模板和灵活轻量的部署能力，实现跨域协作应用的横向拓展。

图 7-1可快速复制的行业样板间

8 场景解决方案

亚信科技隐私计算产品解决方案，覆盖运营商、金融、政务、能源、车企等业务领域，提供丰富的应用场景，主要包含隐匿查询、安全求交、联合统计、联合建模几大类应用场景，助力产业实现精细管理、精益生产、精准营销、精确规划，赋能数智转型。

8.1 匿踪查询

将从应用场景、业务需求、方案建设这三方面介绍匿踪查询的场景解决方案。

8.1.1 匿踪查询应用场景

匿踪查询作为隐私计算领域中安全多方计算下的子分支，使数据持有方无法获知具体查询对象，从而很好地保护查询方的隐私信息，打消安全顾虑，促进数据安全有序流通。在应用场景上，匿踪查询主要适用于标签查询、评分查询、名单查询、信息核验等场景。

8.1.2 匿踪查询业务需求

数据隐匿查询时，引入安全计算的方式，查询方只能从数据服务方这里得到一个查询结果，而且数据服务方也无从知晓查询本身的踪迹，由此可以有效保障双方的隐私安全。

- 保持匿名性：客户或用户在查询过程中不需要提供真实的身份信息，而是通过匿名的方式进行查询，以便保护客户的隐私和安全。
- 提高查询效率：在保护隐私和保证安全的前提下，提高查询效率是另一个重要的业务需求。
- 保证数据安全：在查询过程中，需要保证数据的安全性和完整性，防止数据被篡改或泄露。
- 精准查询结果：尽管是匿名查询，但查询结果需要精准地反映出客户的实际状况和需求，以便更好地为客户提供服务。
- 支持实时查询：客户或用户需要能够实时查询到最新的数据信息，以便及时做出决策和响应。

8.1.3 匿踪查询方案

匿踪查询是一种保护数据隐私的查询方法，其核心思想是在不透露查询需求和身份信息的前提下，实现数据的查询和利用。

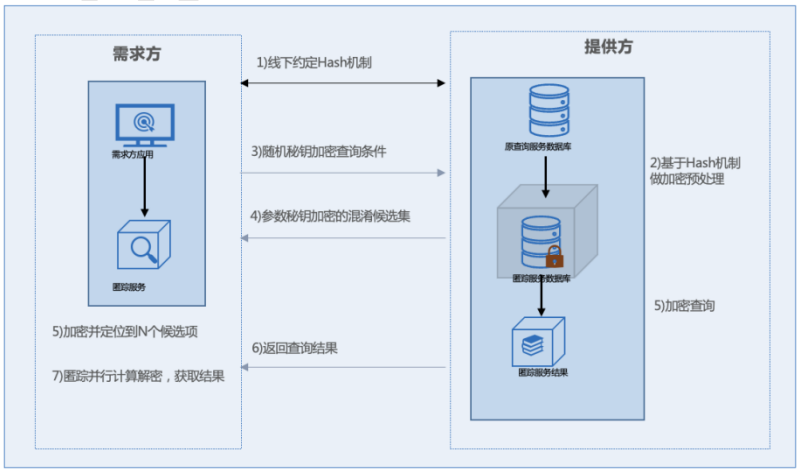


图 8-1匿踪查询方案

- 查询方生成基础密钥：查询方根据数据持有方的公钥、自身的查询方公钥和待查询数据的数据序号，生成基础密钥。
- 向数据持有方发送基础密钥：查询方向数据持有方发送基础密钥，使数据持有方基于预设衍生算法的逆运算，确定不同可查询数据的数据序号的还原密钥。
- 数据持有方确定还原密钥：数据持有方根据预设衍生算法的逆运算和基础密钥，确定不同可查询数据的数据序号的还原密钥，进而确定不同可查询数据的数据加密密钥。
- 数据持有方加密数据：数据持有方采用各数据加密密钥分别对相应可查询数据加密，得到各可查询数据密文。
- 获取数据持有方反馈的各可查询数据密文：查询方获取数据持有方反馈的各可查询数据密文。
- 查询方解析密文：查询方根据持有方公钥和还原密钥，解析出查询的具体内容。

8.2 安全求交

将从应用场景、业务需求、方案建设这三方面介绍安全求交的场景解决方案。

8.2.1 安全求交应用场景

各参与方对所拥有数据进行求交处理。求交计算过程只保护交集外的信息私密性，即求交双方均会获得对方数据集中哪些与己方数据集相同，无法获得不相同部分。应用场景主要包括了目标用户对齐、安全交集运算、安全联合运算、标签数据扩充等。

8.2.2 安全求交业务需求

隐私计算的安全求交业务需求包括数据隐私保护、数据控制权保持、交集运算准确性、计算效率、安全性和合规性等方面。这些需求需要在使用安全求交技术时综合考虑并满足。

- 数据隐私保护：在安全求交过程中，需要保护原始数据的隐私信息，避免数据泄露和滥用。
- 数据控制权保持：隐私计算的安全求交应保持各方的数据控制权，即各方只能访问自己的数据，不能泄露给其他方。
- 交集运算准确性：安全求交的结果应该是准确的，即计算出的交集结果应该是正确的，不出现错误或遗漏的情况。
- 计算效率：安全求交需要保证计算效率，以满足业务需求。

8.2.3 安全求交方案

安全求交属于隐私计算领域的特定应用，它的实现是基于隐私保护集合求交技术，允许持有各自集合的两方或多方来共同计算各自集合的交集运算；在协议交互的最后，一方或是多方根据预先约定得到正确的交集，而且不会得到交集以外其他方集合中的任何信息。

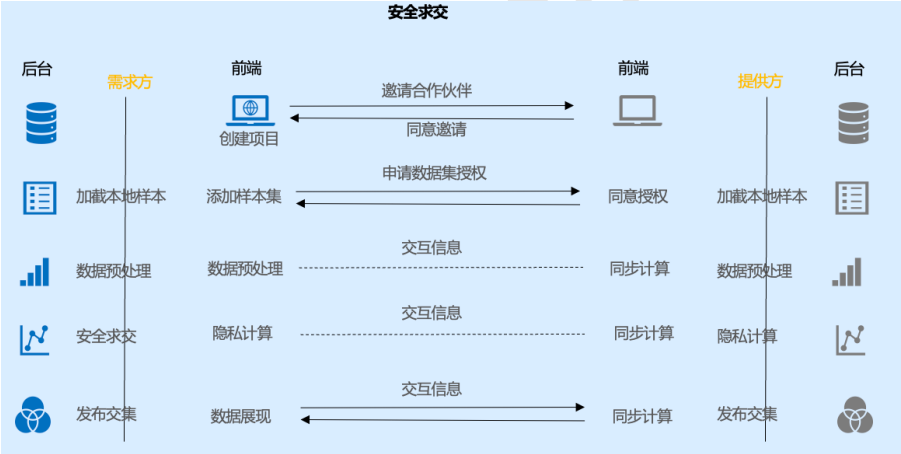


图 8-2安全求交方案

- 构建基于不经意传输扩展的伪随机函数用来完成数据比对，并使用布谷鸟 hash 算法减少传输数据量。
- 与基础的不经意传输算法实现的 PSI 相比，性能大幅提升，网络开销大幅下降。

8.3 联合统计

将从应用场景、业务需求、方案建设这三方面介绍联合统计的场景解决方案。

8.3.1 联合统计应用场景

联合统计可以帮助参与方在不泄露各自数据的前提下进行统计计算，从而得到更准确的结果。例如，在金融总资产认证场景中，银行和证券公司可以联合进行资产统计，以便更好地评估客户的资产状况和信用风险；在医疗统计分析场景中，医院和科研机构可以联合进行疾病发病率和流行趋势的统计研究，以更好地制定预防和治疗方案；在政务常住人口迁徙场景中，政府可以联合进行人口统计和分析，以便更好地规划城市发展和公共资源分配。

8.3.2 联合统计业务需求

在不可信第三方情况下，通过多方共同参与，安全地完成某种协同计算。即在一个分布式的网络中，多个数据方进行统计，使用方只获取统计结果（标签），无其他额外信息。保护各方原始数据不泄漏。

- 金融行业：在金融行业中，联合统计可以用于风险评估、投资分析和市场预测等方面。例如，银行和证券公司可以利用联合统计计算客户的总资产和投资回报率，以便更好地评估其信用风险和投资价值。
- 医疗行业：在医疗行业中，联合统计可以用于疾病发病率和流行趋势的统计研究，以及医疗资源的利用和分配等方面。例如，医院和科研机构可以利用联合统计研究某地区的流感发病率和传播途径，以更好地制定预防和治疗方案。
- 政务领域：在政务领域中，联合统计可以用于城市规划、公共资源配置和人口迁徙等方面。例如，政府可以利用联合统计计算城市各区域的常住人口数量和分布情况，以便更好地规划公共设施和资源配置。
- 市场营销领域：在市场营销领域中，联合统计可以用于消费者行为分析、市场调研和广告效果评估等方面。例如，电商平台可以利用联合统计计算用户的购买行为和偏好，以便更好地推荐商品和优化广告投放。
- 社交媒体领域：在社交媒体领域中，联合统计可以用于用户行为分析、话题趋势研究和社交关系分析等方面。例如，社交平台可以利用联合统计计算用户的话题关注度和活跃度，以便更好地推荐内容和优化用户体验。

8.3.3 联合统计方案

根据需求，从各个数据提供方收集数据，并进行预处理，以准备进行联合统计。为了保护数据的隐私，需要对数据进行加噪和加密，可以使用一些隐私保护技术如差分隐私、同态加密等。通过安全通道和加密传输等技术，实现数据的安全传输和共享，保证数据的安全性和隐私保护。根据需求，采用适当的统计方法和算法实现联合统计，如联合分布、联合均值等。

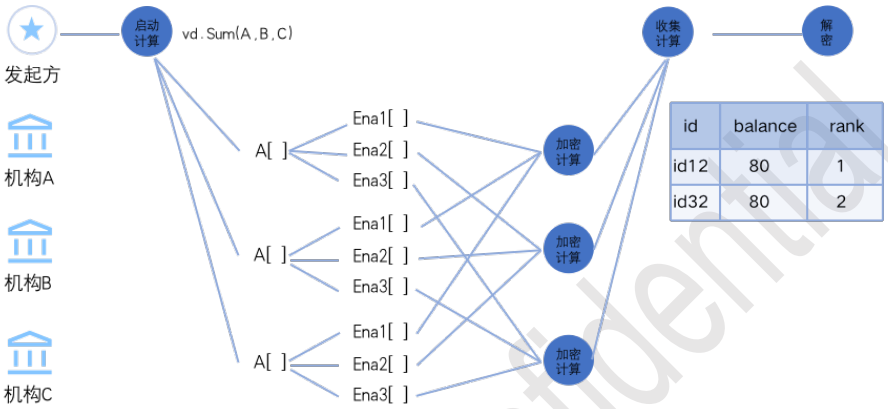


图 8-3联合统计方案

实现流程：线下约定统计分析算法；预置数协商后，分别处理；预置数下发；模型分组计算，随机拆分结果；秘密共享份额；结果份额；协商秘密份额，恢复。

8.4 运营商+汽车：联合营销

车企和运营商分别接入数据集后，通过安全数据对齐和纵向联邦学习完成建模，可通过关键指标（如 IV 值）评估运营商特征贡献度；发布模型后，增换购营销系统通过认证后的 API 进行服务调用，通过联合建模提升营销精准度。

图 8-4智慧车企增换购营销解决方案

8.5 运营商+银行：贷款营销

满足数据隐私安全的前提下，通过对潜在贷款用户模型的纵向联合建模，运营商和银行双方在不共享数据的基础上实现了联合建模 ,从技术上打破数据孤岛，实现 AI 协作。

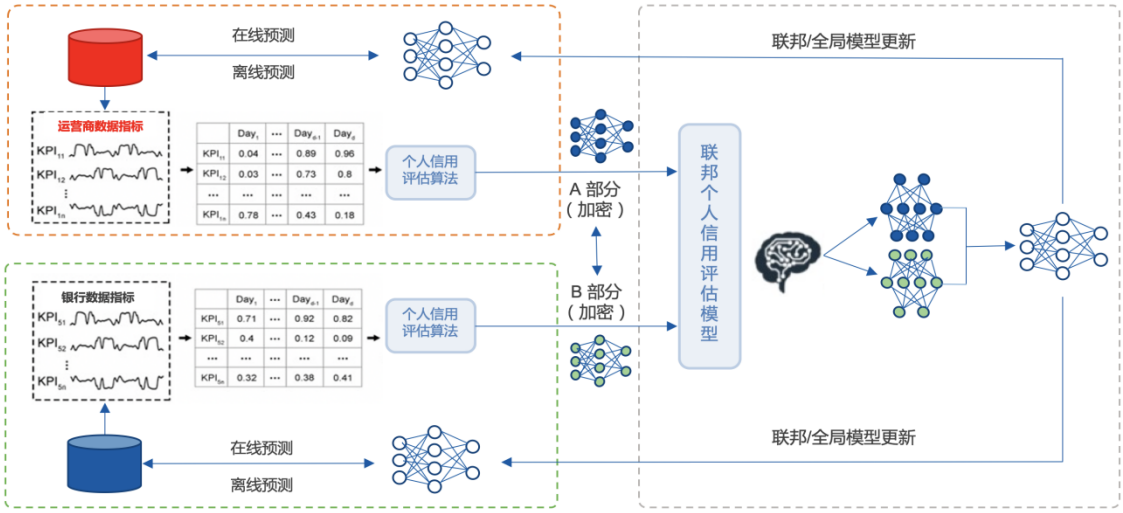


图 8-5 “运营商+银行”信用评分场景方案示意图

联邦学习建模远远超越单域建模，接近建模的理论最优解性能，在查准率上，联邦学习建模相比银行单域提升 10%，相比运营商单域提升 30%，在查全率上，联邦学习建模相比银行单域提升 5%，相比运营商单域提升 10%。

8.6 运营商+运营商：智能反诈

电信诈骗检测难主要原因，一方面是诈骗数据相对稀少，更重要的一方面是运营商间数据不互通，形成孤岛。通过联邦学习技术，在保证不泄露数据隐私（128bit 安全性）的前提下，利用三方运营商的用户数据，提取欺诈特征，基于横向联邦学习算法联合训练分类模型。

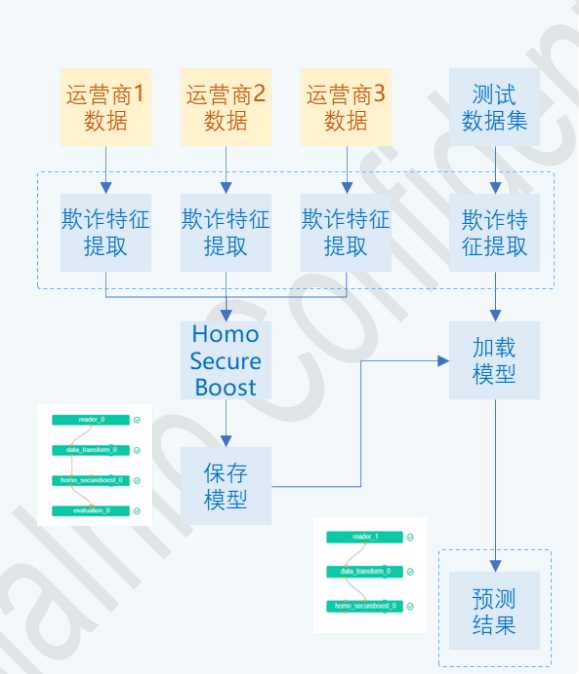


图 8-6智能反诈方案

8.7 运营商+保险：保险代理人挖掘

针对保险公司增员质量较低导致的新人留存率较低、流动大的问题，通过运营商与保险公司双方数据挖掘出最适合保险销售的准增员群体，提升保险公司业务收入。

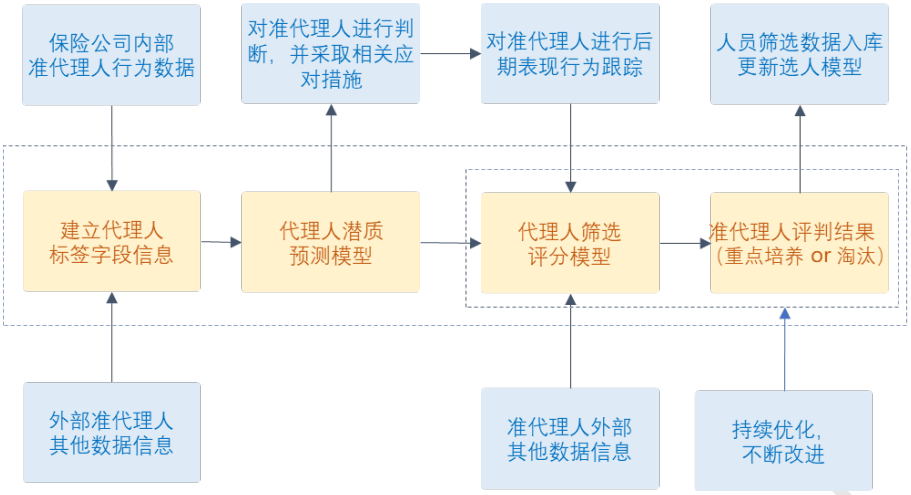


图 8-7保险代理人挖掘方案

8.8 运营商+医疗：智能推荐

基于联邦学习模型架构，打造运营商和医疗机构在保证数据隐私安全前提下的跨域建模能力，赋能医疗智能推荐场景，实现对不同用户推荐专家咨询、极速问诊、体检等。

图 8-8 医疗智能推荐场景解决方案

9 产品客户成功故事

隐私计算产品可以应用于多个行业，如运营商、金融、政务、医疗、制造业等。

9.1 某集团数联网之数据交付平台

某运营商在助力经济社会转型的同时，也积淀了非常有价值的数据资产，呈现出超全覆盖、超多维度、连续不间断、超大数据资产等特征。在电信运营商层面，为形成对外的整体数据服务能力提供隐私计算平台互联共通的基础。

9.1.1 客户需求

战略与业务双驱动，推动某运营商隐私计算平台建设。一是国家层面鼓励数据流动、数据要素市场化，且数据安全监管日趋严格；二是运营商具有天然的数据优势，垂直行业亟需运营商数据进行行业数据融合深度分析，支撑决策。

图 9-1项目背景

数据融合需求强烈，现阶段不同行业及企业之间数据属于割裂，数据无法发挥最大价值，数据要素市场因数据隐私、数据安全性、性能、安全合规等因素导致数据流通存在困难。

图 9-2客户痛点与需求

9.1.2 建设方案与成效

基于隐私计算技术，融合“敏捷理念”，打造运营商“1+X”跨行业数据融合服务，助力集团与各行行业携手联合创新，共筑大数据生态，共谋大数据发展。

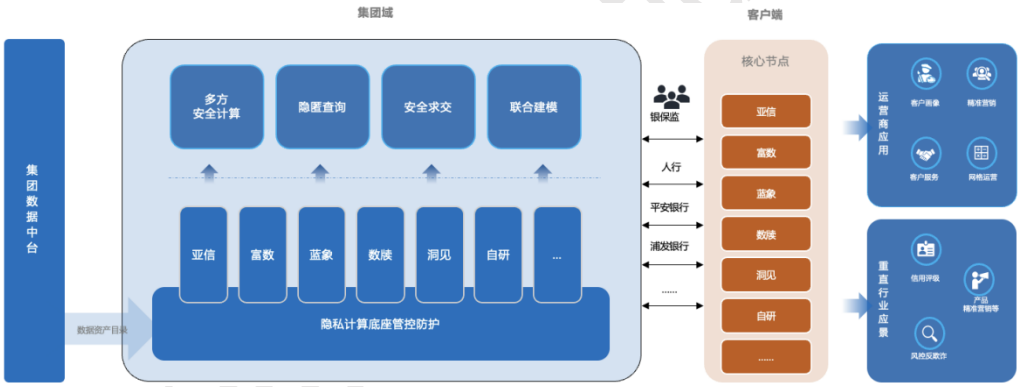


图 9-3建设方案

平台于 2022 年初步建设完成，至今已服务了 8 个行业，支撑了 11 类跨行业应用场景，逐步培育并扩大中国移动数据要素对外开放合作新生态，创造了巨大的经济效益，并形成了良好的社会效益。

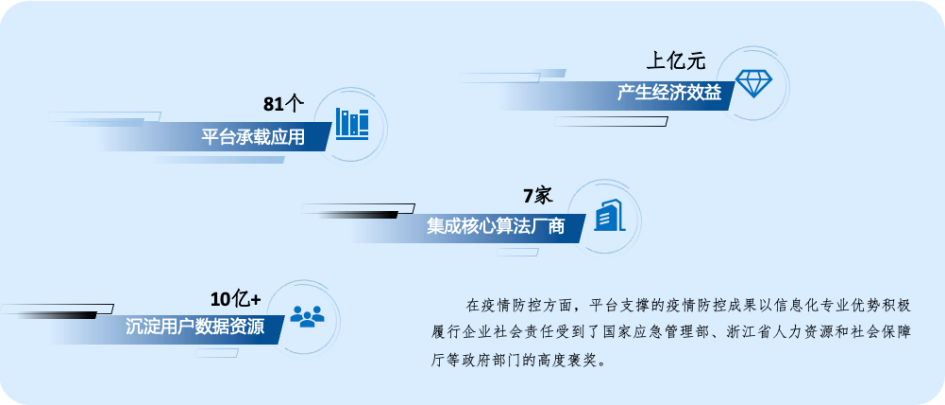


图 9-4项目成效

9.2 某金融机构隐私计算平台

某金融机构以营销、运营、风控等为核心组成的业务全流程，可基于隐私计算技术，将机构自身数据与他方机构进行融合，可极大程度上丰富金融场景、扩展金融业态。

9.2.1 客户需求

某金融机构自身积累了大量高价值的数据比如交易数据、银行数据、市场数据、风险数据、客户数据等。但针对营销、运营、风控等场景，自身数据比较单一、有限。为了扩展自身的业务场景，因此对外部数据的需求较强烈，同时也需要满足国家对行业数据安全的监管要求。

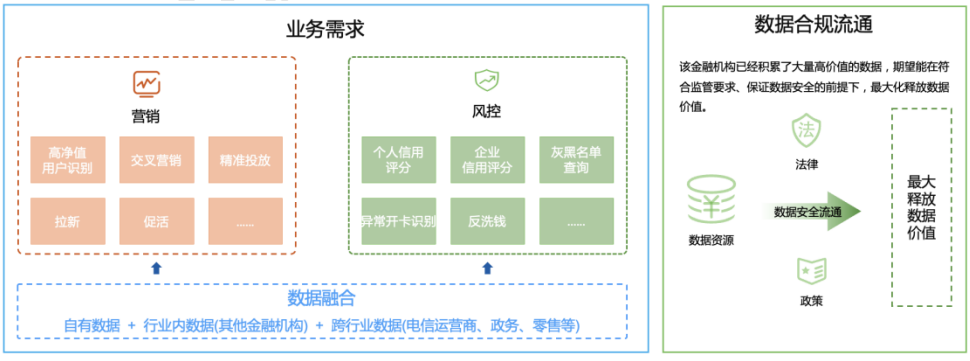


图 9-5客户需求

- 在营销方面 ,金融机构需要了解客户需求和偏好 ,提供个性化的营销服务。通过隐私计算平台 ,金融机构可以在保证数据隐私的前提下 ,联合外部数

据进行客户画像分析和风险评估，从而更好地了解客户需求，提供更精准的营销服务。同时，通过隐私计算平台，金融机构可以与外部合作伙伴进行联合营销，共享数据和资源，提高营销效果和效率。

- 在风控方面，金融机构需要对风险进行评估和管理，防止风险损失的发生。通过隐私计算平台，金融机构可以在保证数据隐私的前提下，进行风险评估和反欺诈分析，及时发现和处理风险事件。同时，通过隐私计算平台，金融机构可以与外部合作伙伴进行风险共担和风险分散，提高风险管理能力和效率。

9.2.2 建设方案与成效

该金融机构依托亚信科技隐私计算平台，可以引入横向(跨行业)和纵向(行业内部)的机构作为合作方，建立自己主导的数据开放生态。平台采用 1+X 架构，通过集成其它数据核心厂商的算法功能实现平台开放互通。

图 9-6建设方案

- 资源层：是基于隐私计算平台底座的机构、节点、算法、项目、程序、任务等进行统一管控，从而实现算力、数据存储等应用。
- 算子算法层：提供亚信自研的安全求交、匿踪查询、联合统计、联邦学习的算法，同时基于“1+X”架构，对外部厂商算法的集成。
- 应用层：通过算法算子对底座的数据进行场景化的应用，如电信反欺诈、联合营销、三要素验证、智能风控、联合统计等场景。

金融机构作为数据需求方，可丰富用户画像，灰黑名单等，提升营销风控效果；作为数据提供方，向生态中的其他成员开放数据，并获取相应的收益，有助于企业最大化释放自身的数据价值；为平台运营方，可快速抢占数据开放市场先机，提升企业的行业影响力，为企业发展开辟新的道路。

- 满足数据安全融合需求：隐私计算产品保证在数据不流出金融机构端的情况下，由多方提供数据，助力金融机构的信贷风控与精准营销的效率提升。
- 实现精准营销服务：帮助金融机构更好的识别、分析客户需求，实现精准营销，打造良好客户体验，提升综合竞争力。
- 增强风控管理能力：融合他方数据，进行全量汇聚分析，识别风险欺诈人群，强化对风险的预判和防控能力，在使用更少的风控人员的条件下，带来更高效可靠的风控管理。

9.3 某车企增换购业精准营销

本节主要介绍“运营商+汽车”在联合营销场景的应用案例。

9.3.1 客户需求

精准营销是汽车产业市场竞争中非常重要的一环，以往传统的汽车营销推荐基于车企自有数据进行建设，存在客户信息实时性和准确性难以保证、数据维度不够全面、数据样本体量不足等局限，导致模型精度不足、跟进营销效率低下，造成营销人力浪费并错失商机。在车企数字化转型加速、数据应用安全要求加强的多重因素推动下，传统车企迫切需要安全可信的创新营销路径。

9.3.2 建设方案与成效

基于以上背景，某车企通过寻找外部数据进行跨域合作赋能，在保障各方数据隐私安全的前提下，帮助该企业识别增换购高意向需求客户，联动其营销业务板块满足增换购业务需求。

本案例整合运营商与车企双方的行业数据优势，持续实时对购车用户、购车意向进行模型推理及预测。非平衡条件下联合分析，运营商侧涉及 13 亿数据及 1000 余个模型标签的分析及建模。基于真实样本的推理结果，模型表现出较好的预测能力，支撑某车企进行应用触达等。

图 9-7建模过程示例

通过该案例的实施，某车企保客增换购营销的到店率、领券率、有效线索占比均得到明显提升，实现了营销活动的闭环评估。应用效果体现在：

- 增换购客户意向率提升 60%，AUC 提升 20%；
- 客户领券率提升 30%，到店率提升 56%；
- 有效推荐线索占比提升 68%，营销成本节省超 20%。

图 9-8案例应用效果

9.4 某医疗机构智能推荐

本节主要介绍“运营商+医疗”在智能推荐场景的应用案例。

9.4.1 客户需求

某运营商省公司作为全集团的标杆企业、省内规模最大的电信运营商，在数据价值挖掘及模型提升方面，面临以下痛点：

- 数据变现收到制约，有大量数据资产，但受隐私法规的制约无法变现。
- 模型准确性提升遇到瓶颈，因为数据维度、数据量不够需要提升自身模型性能。
- 无法实现海量训练、海量连接，物联网、5G 等数据量特别大，并且不易于汇聚起来训练。

该客户需要足某移动与合作医疗机构在保证数据隐私与安全合规的前提下，进行多方数据联合建模、数据协同利用，赋能智能推荐场景。主要需求包括：

- 能够进行在保证数据隐私安全前提下的跨域建模；
- 要求提供可扩展的多方协作建模、联合学习的能力；
- 需具备系统管理功能，包括但不限于项目管理、组织管理、用户管理、角色管理等相关功能。

9.4.2 建设方案与成效

通过新型分布式的机器学习范式，打造基于纵向联邦学习的“运营商+医疗”智能推荐模型；基于联邦学习框架打造医疗行业智能推荐模型架构，实现某移动与合作医疗机构的本地化部署联调。基于某移动与合作医疗机构的数据特点构建专家咨询/极速问诊联邦推荐模型，并应用于该机构预约挂号 APP 的推荐场景。

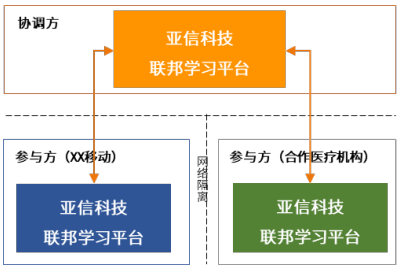


图 9-9 部署方案示意图

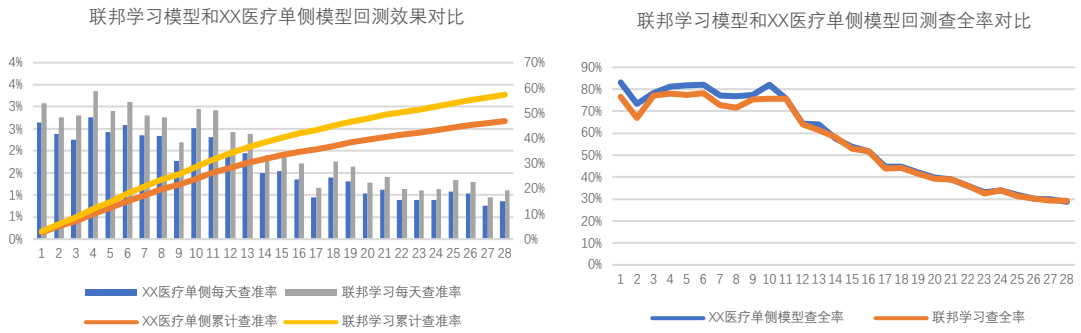


图 9-10 联邦学习智能推荐模型效果

通过基于纵向联邦学习的智能推荐模型，达到了以下应用效果：

- 在数据不出库的前提下，建立联邦学习模型架构，实现了某移动与合作医疗机构的数据虚拟打通；
- 打造专家咨询/极速问诊联邦推荐模型，实现精准营销，解决所有用户推荐同样内容的问题，点击率提升 10%，转化率提升 50%，累计查准率提升 10%；
- 后续逐渐扩充体检推荐、挂号医院推荐、医生推荐等，最终联邦移动数据实现不同用户登陆 APP “千人千面”的目标。

10 资质与荣誉

产品在汽车行业的应用案例于 2023 年入选 Forrester《亚太地区隐私保护技术现状》报告。

图 10-1入选 Forrester 亚太地区隐私保护技术典型用例

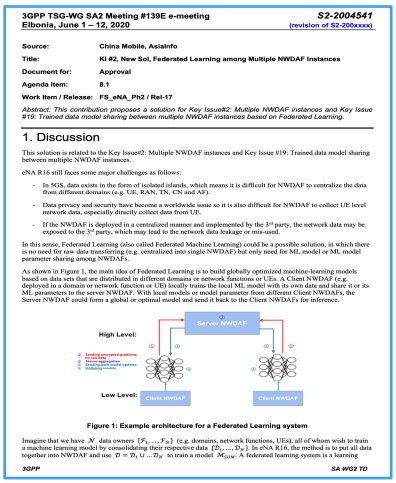
亚信科技首次将隐私计算引入 TMF 体系框架；入选 TMF Catalyst 优秀项目。

图 10-2TMF 体系标准贡献

亚信科技在以下标准中作为牵头单位：

- IEEE P3117™ -隐私保护计算互通框架标准草案
- IEEE P3127 区块链联邦学习
- IEEE P2986 隐私保护与反击；

亚信科技建议的 5G 核心网元 NWDAF 的新型联邦学习技术标准，已被全球权威标准化组织 3GPP 所采纳。



3GPP TSG-WG SA2 Meeting #139E e-meeting
Elbonia, June 1 – 12, 2020

S2-2004541
(revision of S2-200xxxx)

Source: China Mobile, AsiaInfo
Title: KI #2, New Sol, Federated Learning among Multiple NWDAF Instances
Document for: Approval
Agenda Item: 8.1
Work Item / Release: FS_eNA_Ph2 / Rel-17

Abstract: This contribution proposes a solution for Key Issue#2: Multiple NWDAF instances and Key Issue #19: Trained data model sharing between multiple NWDAF instances based on Federated Learning.

1. Discussion

This solution is related to the Key Issue#2: Multiple NWDAF instances and Key Issue #19: Trained data model sharing between multiple NWDAF instances.

eNA R16 still faces some major challenges as follows:

- In 5GS, data exists in the form of isolated islands, which means it is difficult for NWDAF to centralize the data from different domains (e.g. UE, RAN, TN, CN and AF).
- Data privacy and security have become a worldwide issue so it is also difficult for NWDAF to collect UE level network data, especially directly collect data from UE.
- If the NWDAF is deployed in a centralized manner and implemented by the 3rd party, the network data may be exposed to the 3rd party, which may lead to the network data leakage or mis-used.

In this sense, Federated Learning (also called Federated Machine Learning) could be a possible solution, in which there is no need for raw data transferring (e.g. centralized into single NWDAF) but only need for ML model or ML model parameter sharing among NWDAFs.

As shown in Figure 1, the main idea of Federated Learning is to build globally optimized machine-learning models based on data sets that are distributed in different domains or network functions or UEs. A Client NWDAF (e.g. deployed in a domain or network function or UE) locally trains the local ML model with its own data and share it or its ML parameters to the server NWDAF. With local models or model parameter from different Client NWDAFs, the Server NWDAF could form a global or optimal model and send it back to the Client NWDAFs for inference.

Figure 1: Example architecture for a Federated Learning system

Imagine that we have N data owners $\{\mathcal{F}_1, ..., \mathcal{F}_N\}$ (e.g. domains, network functions, UEs), all of whom wish to train a machine learning model by consolidating their respective data $\{\mathcal{D}_1, ..., \mathcal{D}_N\}$. In eNA R16, the method is to put all data together into NWDAF and use $\mathcal{D} = \mathcal{D}_1 \cup ... \mathcal{D}_N$ to train a model $\mathcal{M}_{SUM}$. A federated learning system is a learning

3GPP SA WG2 TD

图 10-3 3GPP TSG-WG SA2 / Rel-17

11 联系我们

亚信科技（中国）有限公司

地址：北京市海淀区中关村软件园二期西北旺东路 10 号院东区亚信大厦

邮编：100193

传真：010-82166699

电话：010-82166688

Email：5G@asiainfo.com

网址：www.asiainfo.com

Thank you

依托数智化全栈能力，创新客户价值，助推数字中国。