CAICT 中国信通院　蚂蚁集团 ANT GROUP

数据价值释放与隐私保护计算应用研究报告

（2021 年）

中国信息通信研究院安全研究所
蚂蚁科技集团股份有限公司

2021 年 11 月

版权声明

编制说明

本蓝皮报告由中国信息通信研究院与蚂蚁科技集团股份有限公司牵头撰写，限于撰写组时间、知识局限等因素，内容恐有疏漏，烦请各位读者不吝指正。

本报告在撰写过程中得到了多家单位的大力支持，在此特别感谢参编单位北京瑞莱智慧科技有限公司、杭州锘崴信息科技有限公司、蓝象智联（杭州）科技有限公司、上海富数科技有限公司、深圳市洞见智慧科技有限公司（排名不分先后）的各位专家。

前　言

2020 年 10 月，《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》（以下简称“十四五”规划纲要）提出“加快迎接数字时代，激活数据要素潜能，推进网络强国建设，加快建设数字经济、数字社会、数字政府，以数字化转型整体驱动生产方式、生活方式和治理方式变革”，强调以数字化转型驱动**生产方式、生活方式和治理方式的变革**，以此来实现加快数字化发展、建设数字中国的远景目标，充分释放数字红利谱写数字中国新篇章。

政策红利释放，激活数据流通市场需求。数据作为数字化转型核心驱动力，其流通共享对**打造数字经济新优势、加快数字社会建设步伐、提高数字政府建设水平与营造良好数字生态**具有重要意义。着眼数据的高效共享与协同应用，我国密集出台《中共中央 国务院关于新时代加快完善社会主义市场经济体制的意见》《中共中央 国务院关于构建更加完善的要素市场化配置体制机制的意见》《全国一体化大数据中心协同创新体系算力枢纽实施方案》《网络安全产业高质量发展三年行动计划（2021-2023）（征求意见稿）》等多项政策，充分体现了国家对数据生产价值以及市场贡献的高度肯定。持续释放的政策红利有效激活了数据开放共享与利用的市场需求。

合规监管助力，拓展数据流通市场空间。立足于数字化转型发展实践中数据安全以及个人信息保护的迫切需求，我国陆续发布了《中华人民共和国民法典》《中华人民共和国数据安全法》（以下简称《数据安全法》）《中华人民共和国个人信息保护法》（以下

简称《个人信息保护法》）等法律，充分体现了数字时代国家对于维护数据安全及个人信息保护的信心、雄心和决心。各领域法律法规相互衔接补充纵深发展、逐步细化，构建了以《中华人民共和国网络安全法》《数据安全法》《个人信息保护法》为核心的数字安全法律体系，着重强调了在兼顾数据安全和个人信息保护，保护个人、组织的合法权益，维护国家主权、安全和发展利益的同时，促进数据的开发利用。日趋完善的数据安全合规监管框架进一步拓展了数据安全流通的市场空间。

多重部署加码，隐私保护计算前景可期。在日趋严格的合规监管、日渐强化的政策引导以及日益旺盛的市场需求等多重背景下，为有效打破数据流通壁垒、促进数据价值释放，隐私保护计算因从技术角度实现了数据价值的共享流通和协同应用，有效促进了数据这一新型生产要素经济价值最大程度地发挥而备受关注。当前，隐私保护计算技术在金融、医疗、政务等领域已具有初步的应用探索。但受制于隐私保护计算前沿技术了解不详、数据安全和个人信息保护政策法规理解不足、产业落地缺乏参考等诸多因素，使得隐私保护计算技术尚未实现规模化的应用。

在前期《隐私保护计算技术研究报告》《隐私保护计算与合规应用研究报告》研究基础上，本报告聚焦隐私保护计算技术产业落地缺乏参考的问题，对数据、数据价值、隐私保护计算如何助力数据价值释放以及在金融、医疗、政务领域场景的应用价值进行探讨与探索，为隐私保护计算技术的应用落地及数据价值释放提供参考。

目 录

图 目 录

表 目 录

一、数据概念内涵及价值

数据并非新生事物，但是数据的价值释放需以明确数据、数据价值、数据特征等相关概念为前提。本章尝试对数据、数据特征和数据价值进行定义，抛砖引玉，以供社会各界参考及讨论。

（一）数据定义

根据《数据安全法》定义，“数据，是指任何以电子或者其他方式对信息的记录。”该定义在法律层面明确了数据的记录方式，并将“数据”和“信息”进行区分。国际数据管理协会（DAMA）认为，“数据是以文本、数字、图形、图像、声音和视频等格式对事实进行的表现”，对“数据”存在的不同形态进行了列举，且指出“数据”是对事实的表现[1]。标准 ISO/IEC 11179-1:2015[2]将“数据”定义为“以适合于交流、解释或处理的形式化方式对信息进行可重新解释的表示”，该定义强调了“数据”的电子性质，其认为“数据”是对它代表的对象（信息）的解释；且该解释方式必须是权威、标准、通用的，只有这样才可以达到通信、解释和处理的目的。统计学将“数据”定义为“用于表示和解释而收集、分析和总结后的客观事实和数字符号”，并将“数据”分为定性数据和定量数据。根据我国权威科学技术名词审定机构全国科学技术名词审定委员会审定，在计算机科学技术中，“数据”是客观事物的符号表示，指所有可输入到计算机中并可被计

[1]《DAMA 数据管理知识体系指南》

[2]ISO/IEC 11179-1:2015(en) Information technology — Metadata registries (MDR) — Part 1: Framework reinterpretable representation of information in a formalized manner suitable for communication, interpretation or processing

算机程序处理的符号的总称；在管理科学技术中，“数据”是描述事件或事物的属性、过程及其关系的符号序列，比如自然语言符号、科学符号、数字以及图形图像等。

“数据”的定义虽未实现完全的统一，但结合上述定义分析，我们认为“数据”的定义包含了两个核心内涵：**一是描述客观事实**。数据是对感知到的客观事实进行描述或记录的结果，是对现实世界中的时间、地点、事件、其他对象或概念的描述。**二是须符号化表达**。数据须被符号化表达，方能被有效识别。如数字、文字、字母、声音、图片、视频等。综上，我们认为**数据是对感知到的客观事实进行描述或记录的符号或符号集合，如数字、文字、字母、声音、图片和视频等，是未经处理的原始素材**。

（二）数据的特征

事实相关性：数据是对客观事实的描述，是与客观事实相关的、无序的、未经加工处理的原始材料。

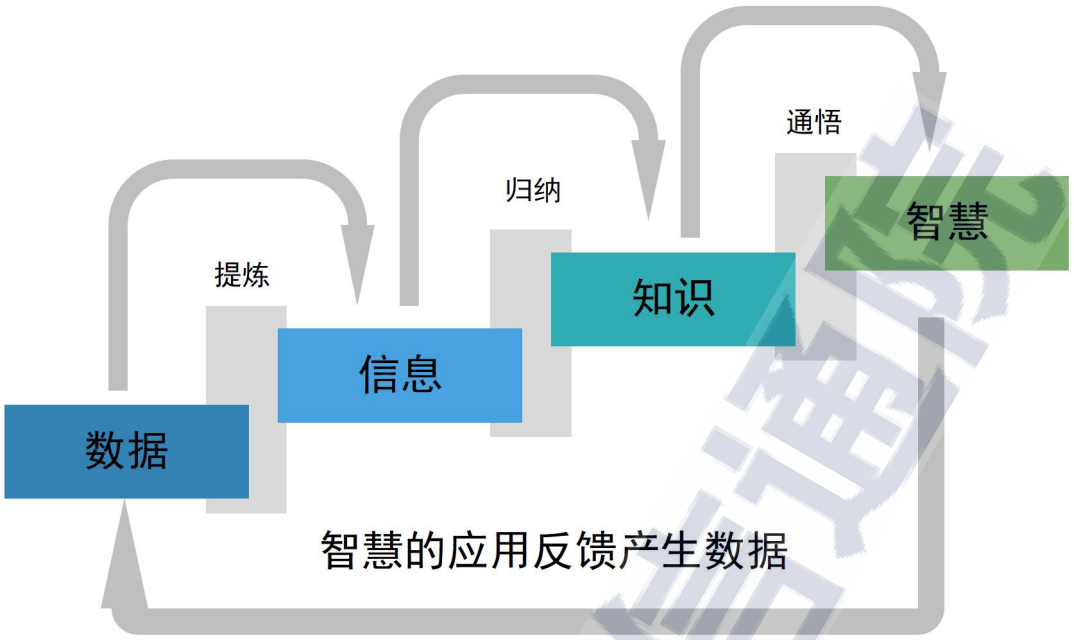
须符号化表达：数据本身是对事实的记录和描述，且必须以某种符号或符号集的形式进行表达。

可比特化记录：无论表达数据的符号是数字、文字、声音、图片或视频等，都可用二进制的比特符号统一记录。任何数据都可以被编码为一系列 0 和 1 组成的二进制序列。

蕴含价值性：数据本身并没有任何意义，其所蕴含的意义与价值是从数据本身当中“挖掘”“创造”而来的。因此，数据必须是可计算、可推理演绎、可解释、可分析、可挖掘的。

（三）数据的价值

数据通常是一个不言自明的概念，但数据的价值究竟应该如何体现，目前却少有研究。多数场景下，数据的价值往往被从数据资产的角度进行解释，但数据资产化仅仅表达了数据的经济价值，对数据助力社会服务、国家治理等公共价值的表达极为有限。知识管理体系中的“数据-信息-知识-智慧”（Data-Information-Knowledge-Wisdom，DIKW）模型对于数据价值的描述和理解为我们提供了参考（如图 1 所示）。基于该模型，我们认为“数据的价值”可以被直接理解为是由“数据”提炼的“信息”、由“信息”归纳出的“知识”、由“知识”通悟的“智慧”，并可用以指导我们的决策。通过决策来驱动生产方式、生活方式和治理方式的变革，进而间接实现数据的经济价值、社会价值、国家治理和安全价值的公共价值。虽然目前关于 DIKW 模型的起源尚未得到严谨的说明，但对于 DIKW 模型的理解（如表 1 所示），目前相关领域已达到初步共识，且被广泛应用在信息管理、信息系统和知识管理等领域。



来源：中国信息通信研究院

图 1 DIKW 模型

数据：数据是对感知到的客观事实进行描述或记录的符号或符号集合，如数字、字母、声音、图片和视频等，是未经处理的原始素材，解决**“知有无”**的问题。

信息：信息是对无序数据进行加工、提炼获得的**有意义的、有逻辑的、有关联性的**数据。信息比数据更“紧凑”，更“有用”，通常描述**何人、何时、何处、何事等**，解决**“知是何”**（Who，When，Where，What）的问题。

知识：知识是从积累的相关信息中通过过滤、总结等方式得到的，被用来解释和指导行动的信息，是经验所得的判断。知识可用来了解**“为什么”**以及**“怎样做”**，解决**“知为何”**（Why）和**“知何为”**（How to）的问题。

智慧：智慧是在知识的基础上，通过经验、阅历累积，试图理解过去未曾理解或未尝试过的事物，形成对事物的深刻洞察以及对事物的未来发展具有启示性、前瞻性的看法，体现为一种卓越的判断力，解决**“知最优”**（What is best）的问题。而智慧的应用又可以指导产生新的数据。

表 1 DIKW 模型解释

	数据	**信息**	**知识**	**智慧**
核心内涵	陈述 Representation	描述 Descriptions	解释和指导 Explanations & Instructions	预测和判断 Prediction
特征	事实依赖/ 无意义	有意义/ 逻辑性	本质性/原则性 /经验性/指导性	启示性 /前瞻性
解决问题	知有无	知是何 what，when， who，where	知为何、知何为 why，how to	知最优 what is the best
时间维度	过去和现在			未来
意义	**指导“正确”做事**			**规划做“正确”的事**

来源：中国信息通信研究院

数据、信息、知识和智慧同时兼具经济价值、社会价值、国家治理和安全价值等多重价值，其价值可以体现在由数据处理的信息、由信息升华的知识、以及由知识理解的智慧上。数据、信息、知识、智慧皆能够指导做出科学的决策，以此带来新的价值增值。值得注意的是，数据、信息、知识、智慧依赖于语境等背景知识，彼此之间并非割裂。在进行数据、信息、知识和智慧的研究与应用时，要结合相应的具体背景知识，不能简单、片面、割裂的理解。

（四）隐私保护计算助力数据价值释放

立足国家“十四五”规划纲要的“数字中国”远景目标，隐私保

护计算技术助力加快驱动生产、生活、治理方式的变革，以及营造良好数字生态。

1.打造数字经济竞争优势，驱动生产方式变革

在数字经济建设方面，数据凭借其可复制、可共享、可无限供给的特点，助力产业实现精细管理、精益生产、精准营销、精确规划等提升，以此降低经济运行成本、提高经济运行效率、赋能传统产业转型升级，催生大量新产业、新模式、新业态。依托数字经济中的海量数据规模和丰富应用场景优势，隐私保护计算技术助力打破“数据壁垒”，推动数据赋能全产业链协同转型，助力形成高质量供给创造新需求、需求牵引供给的动态平衡，促进国民经济良性循环。

2.加快数字社会建设步伐，驱动生活方式变革

在数字社会建设方面，隐私保护计算在助力推进学校、医院、养老院等公共服务机构实现资源数字化的基础之上，有助于进一步加大开放共享和应用力度，推动线上线下公共服务协同发展、深度融合，提高公共服务能力的便捷性，以及数字服务应用的普惠性。此外，对于基层、边远和欠发达地区，通过隐私保护计算技术，可在保护数据安全和隐私的情况下，助力扩大公共服务资源辐射覆盖范围，有效缓解“数字鸿沟”，提高“数字弱势群体”及人民群众对公共服务的获得感和满足感，促进社会运行方式变革，构建全民畅享数字红利的数字生活。

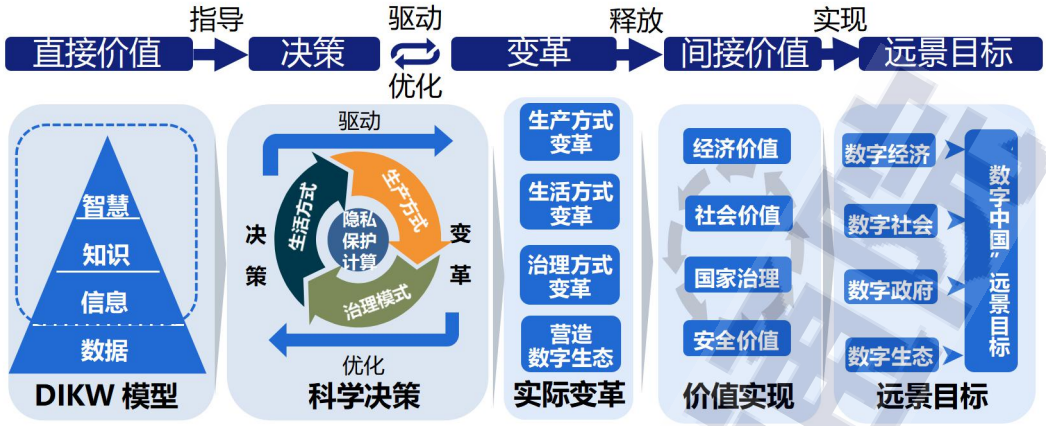
3.提高数字政府建设水平，驱动治理方式变革

在数字政府建设方面，隐私保护计算在助力确保公共安全的前提下，有助于提高数字化政府工作效能，实现数据跨部门、跨层级、跨地区汇聚融合、深度利用和高质量协作。具体体现在三方面，**一是**有效推动公共数据资源开放，以增强公共数据资源开放的透明度、增加政府公信力；**二是**有序推进政务数据资源共享，以提升协同治理能力；**三是**全面深化公共数据资源利用，以提高政府决策制定的科学性、专业性和时效性，助力政府精准施策。

4.激发安全技术创新活力，营造良好数字生态

隐私保护计算在提供数据安全和个人信息保护能力的基础上，**一是**从技术角度实现“数据”向“数据价值”流通的升维，破除既有数据壁垒；**二是**凭借其坚实的理论基础和安全性证明，加强数据应用透明度，增进数据价值利用下的安全保护信任，有效缓解数字经济发展中的数据安全和个人信息保护的信任危机，弥合信任鸿沟；**三是**助力实现数据合规应用中的数据最小化、数据分类分级和数据匿名化，促进数据应用的合规化发展。如隐私保护计算技术通过联邦学习的控制用法用量、安全多方计算实现目的受限，机密计算的授权代码运行等实现数据最小化。

综上，“十四五”规划纲要强调以数字化转型驱动生产方式、生活方式和治理方式的变革，而隐私保护计算技术的出现，为充分发挥海量数据和丰富应用场景优势，有力促进数字技术与经济社会发展各领域融合发展，加快实现数字化发展、建设数字中国的远景目标提供了重要的基础（如图 3 所示）。



来源：中国信息通信研究院

图 3 数据价值释放路径模型

二、隐私保护计算技术概述

（一）隐私保护计算及其关键技术

隐私保护计算（Privacy-Preserving Computation）是一套包含人工智能、密码学、数据科学等众多领域交叉融合的跨学科技术体系[3]。它能够在不泄露原始数据的前提下，对数据进行加工、分析处理、分析验证，其重点提供了数据计算过程和数据计算结果的隐私安全保护能力。随着数字技术的发展，隐私保护计算的内涵及主流技术不断演进。主流的技术研究焦点从早期的数据扰动和数据匿名化等演进至今，已经能够实现数据计算过程和数据计算结果的保护，形成一套包含众多领域的跨学科安全技术体系。**隐私保护计算具体涵盖了安全多方计算、联邦学习、同态加密、差分隐私和机密计算等技术。**

安全多方计算（Secure Multi-Party Computation，SMPC），由中国科学院院士姚期智于 1982 年通过“百万富翁问题”提出，旨在解

[3]中国信通院 《隐私保护计算技术研究报告》

决“一组相互独立且互不信任的参与方各自持有秘密数据，协同计算一个既定函数”的问题。安全多方计算保证了各参与方在获得正确计算结果的同时，无法获得计算结果之外的任何信息。

联邦学习（Federated Learning，FL），可被理解为是由两个或两个以上数据方共同参与，在保证数据方各自原始数据不出其定义的安全控制范围的前提下，协作构建并使用机器学习模型的技术架构。通常情况下，联邦学习需与其它隐私保护计算技术联合使用，才可在计算过程中实现数据保护。

同态加密（Homomorphic Encryption，HE），是一种允许在加密之后的密文上直接进行计算，且计算结果解密后与基于明文的计算结果一致的加密算法，可在不解密以实现数据机密性保护的同时完成计算。根据支持密文运算的程度，同态加密方案可以分为部分同态加密方案和全同态加密方案两类。部分同态加密方案能够支持有限的密文计算深度，常作为其他方案的组成部分之一进行使用。而全同态加密理论虽支持无限次任意给定函数的运算，但由于计算开销较大，目前尚未形成规模化的商用。

差分隐私（Differential Privacy，DP），是 Dwork 在 2006 年针对数据库的隐私问题提出的一种严格的、可量化的隐私定义和技术。差分隐私在保留统计学特征的前提下，去除个体特征以保护用户隐私。差分隐私具有两个重要的优点：**一是**提出与背景知识无关的隐私保护模型，实现攻击者背景知识最大化的假设；**二是**为隐私保护水平提供严格的定义和量化评估方法。

机密计算（Confidential Computing，CC），机密计算是指通过在基于硬件的可信执行环境中执行计算来保护数据应用中的隐私安全的技术之一。其中可信执行环境定义为可在数据机密性、数据完整性和代码完整性三方面提供一定保护水平的环境[4]。其基本原理是将需要保护的数据和代码存储在可信执行环境中，对这些数据和代码的任何访问都必须经过基于硬件的访问控制，防止他们在使用中未经授权被访问或修改，从而提高机构管理敏感数据的安全水平[5]。

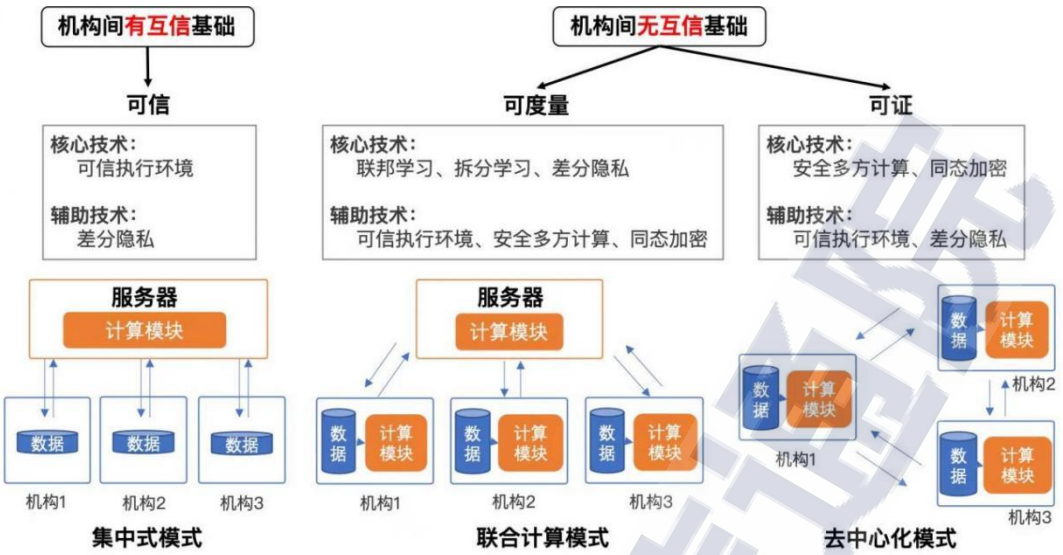
除上述技术外，隐私保护计算技术还包含了秘密共享、不经意传输、混淆电路、零知识证明等诸多技术方向，在此暂不一一赘述。

（二）基于隐私保护计算技术的数据流通模式

在实际应用中，根据数据流通方式、数据集中程度、模型复杂度等差异化的业务场景，从技术角度来说，基于隐私保护计算技术的数据流通方式可分为**可信环境模式**、**可证模式**和**可度量模式**三类（如图 2 所示）：

[4]Confidential Computing Deep Dive v1.0

[5]中国信通院 《隐私保护计算与合规应用研究报告》



来源：中国信息通信研究院

图 2 基于隐私保护计算技术的数据流通模式

1.可信环境模式

可信环境模式，是以机密计算技术为核心，在基于硬件的可信执行环境中执行计算，保护数据应用中的隐私安全的集中式计算模式。该模式本质上是一种集中式的数据计算模式，需以各参与方的强信任关系为前提，将各参与方的数据进行集中式汇总，并利用集中汇总的数据进行模型训练。

因该模式将数据进行了集中汇总，故可进行非常复杂的计算，具有效率高、网络延迟低等优势，但难点在于如何构建各参与方的强信任关系。该模式通过基于硬件的可信执行环境构建参与方的信任关系，其信任的基础是对可信执行环境的信任。目前市场上技术成熟的厂商主要有 Intel SGX，ARM TrustZone 等，较容易产生供应商锁定等供应链安全问题。**该模式的核心技术包括机密计算的可信执行环境等，辅助技术包括差分隐私等。**

2.可证模式

可证模式，是以安全多方计算和同态加密等密码技术为核心，支持在无可信第三方的情况下，各参与方协同计算一个既定函数的分布式计算模式。在该计算模式下，中间数据均以密态呈现。所谓“可证”是指数据的运算态或结果态的安全性可由其使用的密码算法的理论安全性来证明提供。

该模式的优势是其采用基于密码学的安全多方计算和同态加密等技术，凭借其坚实的理论基础和可证明的安全性，获得了较强的安全性保障。但是由于该模式包含复杂的密码学操作，实现相关技术需要付出较大的性能代价，对性能提出了严峻的挑战。对于一些计算复杂度较低的场景，该模式已取得良好的应用效果。**该模式的核心技术包括安全多方计算、同态加密等，辅助技术包括可信执行环境、差分隐私等。**

3.可度量模式

可度量模式，是以差分隐私技术为核心，可对数据计算过程中的隐私泄露风险进行量化评估的数据流通模式，该技术通常与联邦学习等其他技术结合使用。例如，在联邦学习中，中心节点需对各方模型更新的中间结果进行聚合，但此过程中存在数据重构时的攻击风险。差分隐私可在各方数据出域前，通过施加随机噪声的方式保护中间结果，并度量这些噪声带来的隐私保护效果。

该技术的优势是能够实现隐私风险的量化评估，但是会对数据的精度形成不可忽略的影响，因此对精度要求较高的场景需酌情使用。

该模式的核心技术包括差分隐私、联邦学习等，辅助技术包括可信执行环境、安全多方计算、同态加密等。

（三）基于隐私保护计算技术的数据流通场景

基于当前隐私保护计算技术的应用场景，其数据的流通场景主要包含单数据方的主动开放、无数据方的申请使用以及多数据方间的联合计算（如表 2 所示）。

表 2 基于数据流通的场景分类

序号	特点介绍	参与方式	需要保护的数据	隐私保护相关技术	典型应用场景
1	单数据方主动开放	数据拥有方主动发起（单方）。	发布数据的个人信息和敏感内容。	差分隐私等	公共管理和服务机构开放符合开放条件的公共数据等。
2	无数据方申请使用	无数据方提供查询条件、并申请查询；数据方根据查询条件进行查询并反馈查询结果。	无数据方查询条件中的隐私信息；数据拥有方的数据	PSI、PIR 等	征信查询、订单查询、敏感疾病查询等。
3	多数据方联合计算	双方或多方	各数据方的数据	安全多方计算、联邦学习、可信执行环境	联合风控、联合营销等。

来源：中国信息通信研究院

一是单数据拥有方主动开放数据。通常为公共管理和服务机构对符合开放条件的公共数据进行开放。为保障数据安全及个人隐私，在对数据进行脱敏处理或使用差分隐私等技术时往往会给数据加入噪声。如美国人口普查局会在发布人口数据时使用差分隐私技术进行保护处理，在保证数据的统计信息的基础上，避免泄露详细的个人信息，

保障了数据和个人隐私的安全。

二是无数据方申请使用数据拥有方的数据。在此场景下，无数据方需向数据拥有方提供查询条件，数据拥有方根据查询条件进行查询并反馈相关结果。借助隐私保护计算技术能够实现数据库数据及查询条件的“双盲”，以此保护数据和个人隐私的安全。相关的支撑技术包括隐私集合求交 PSI 和隐私信息检索 PIR 等。

三是多数据拥有方联合计算。两个或多个机构之间基于某种业务需求，将各方数据进行联合计算和分析。该类跨机构进行数据联合计算的场景是当前业界研究和应用最多的场景。

三、隐私保护计算技术落地应用案例

（一）金融行业应用案例

1.基于隐私保护计算纵向联邦的银行交易反欺诈案例

（1）业务背景

以云计算、区块链、大数据等为代表的新一代信息通信技术，正在加速金融业与信息科技的创新融合。金融科技（FinTech）在使支付、借贷、投资、保险等金融服务变得高效便捷的同时，也为银行业带来了申请欺诈、交易欺诈和营销欺诈等欺诈风险。申请欺诈是指在信贷申请阶段存在的恶意逾期、中介代办、内外勾结、团伙欺诈等行为；交易欺诈指第三方在客户不知情的情况下，非法利用他人账户进行的账户盗用、伪卡盗刷等行为，以及内部员工在支付和交易过程中的违规操作、骗取客户或行内资金等行为；营销欺诈是指黑产利用金

融机构发放新用户红利时的推广活动漏洞，进行非正常参与、非法获取营销红利，致使金融机构遭受损失的行为。

欺诈行为攻击对象不确定、犯罪主体难追踪、外部欺诈风险涵盖范围广、防控难度大等原因使得金融欺诈成为导致银行业受损最严重的风险之一。据国外研究机构统计，欺诈风险每年导致的银行业受损金额高达近千亿美元，国内银行每年因欺诈风险损失的金额也高达上百亿元。

（2）传统方案

近年来，基于机器学习和大数据的反欺诈风控技术迅猛发展，银行业在反欺诈风控领域取得一定的进步，大部分银行均构建了实时交易反欺诈系统，基于银行已有的业务数据，结合专家规则与机器学习模型来甄别交易欺诈行为。然而，在巨大的经济利益驱使下，金融交易欺诈黑色产业链愈发成熟，其技术和手段不断升级迭代，传统的基于机器学习和业务数据的反欺诈风控技术捉襟见肘，为金融行业的交易反欺诈工作带来了巨大的挑战。

（3）业务痛点

特征维度不足：对于绝大多数银行机构，反欺诈的最大难题是反欺诈模型建立过程中数据来源单一，单纯依靠自身业务数据构建出的反欺诈模型识别准确度极低。随着黑色产业链智能化与集团化发展，各类欺诈手段的特征越发隐蔽、难以察觉，且跨行业欺诈逐渐成为常态，单次欺诈行为贯穿社交媒体、银行 APP 等多个工具，各机构的单方数据无法应对。例如，在利用社交网络进行金融欺诈的场景中，

社交网络服务提供商掌握黑客针对用户的广撒网、常以中老年人群为目标等行为的特征；银行则掌握受害者在被欺诈后，向黑客转账以及后续资金转移时间、流向等特征；双方数据的特征割裂，均不足以独立实现对欺诈行为的有效识别。

数据安全与共享利用的矛盾：数据作为企业的核心竞争力之一，各企业不断加强对数据处理和利用的重视程度，但同时，日益频发的个人信息泄露和数据安全事件引发大众广泛关注，数据使用与隐私保护之间的矛盾日益突出。随着近几年国内外一系列数据安全与隐私保护相关政策法规的出台，以往的粗放式数据收集、使用与交易模式将被严格规范和限制，如何在数据安全框架内促进数据的共享利用成为下一阶段的重要议题。

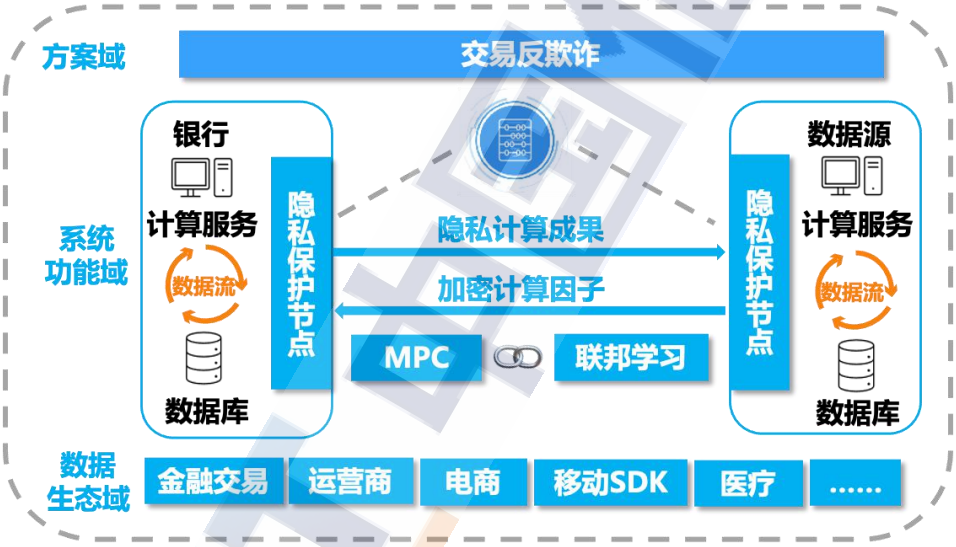
解决方案碎片化：为优化反欺诈效果，银行采取了诸多措施，例如，采购外部反欺诈评分类产品、构建全域反欺诈关联网络等。但各类措施大都聚焦于业务流程的单个环节（如身份识别认证、欺诈行为识别、信用等级识别等），无法覆盖业务全流程的欺诈风险。业务流程各环节的反欺诈解决方案部署割裂，整体协同配合困难，致使银行机构反欺诈能力发挥受限，业务反欺诈效果不佳。

综上，如何在有效保护数据安全的前提下，合法合规地利用内外部数据，丰富样本数据特征维度，构建更加精准的反欺诈风控模型，提升反欺诈能力，是当前各类银行的当务之急。

（4）实践案例

针对银行当前反欺诈管理中遇到的特征维度不足、安全与利用矛盾、解决方案碎片化等困境，基于联邦学习等隐私保护计算技术的“数据+平台+模型”一体化解决方案，通过实现银行与外部机构在反欺诈场景下的跨行业数据链接，联合了金融交易特征、社交行为特征和相关人员特征等多维度特征信息构建反欺诈模型，实现更精准有效的交易欺诈甄别，提升银行机构交易反欺诈能力。

在整体方案实施中，需在银行机构部署隐私保护计算节点，通过隐私保护计算节点与数据源生态完成对接（如图 4 所示）。



来源：隐私保护计算服务提供商

图 4 银行与外部数据源对接示意图

首先通过隐私保护计算平台的 PSI 功能[6]，以纵向联邦的方式，将银行机构准备的反欺诈样本数据与外部数据源进行隐私求交，获取多方的交集客户信息，在银行客户三要素信息（姓名、身份证号、手机号）不出库的前提下，完成银行与外部数据源之间的数据样本对齐。

[6]PSI 功能：允许持有各自集合的两方共同计算两个集合的交集。在协议交互的最后，一方或是两方应该得到正确的交集，而且不会得到交集以外另一方集合中的任何信息。（崔泓睿，刘天怡，郁昱等：《多方安全计算热点:隐私保护集合求交技术(PSI)分析研究报告》，2019.）

然后，运用隐私保护计算平台的特征工程与模型训练模块，完成反欺诈模型训练与调优工作。

从效果上看（如表 3 所示），该隐私保护计算反欺诈解决方案帮助银行安全引入客户的支付行为、设备信息、社交习惯等数据，提升了反欺诈模型的准确性和效率，通过隐私保护计算技术构建的反欺诈模型的模型评估指标AUC[7]可以达到0.82，模型风险区分能力指标KS[8]达到 0.51，模型效果有较大提升。

表 3 传统计算方案与隐私保护计算反欺诈方案对比

方案	反欺诈模型准确性	反欺诈效率	是否引入同行业标签	是否引入其他行业数据源	实时性
传统方案	低	低	否	否	低
隐私保护计算反欺诈方案	高	高	是	是	高

来源：隐私保护计算服务提供商

（5）实践价值

隐私安全：此实践中隐私保护计算方案使用的秘密分享、全同态、半同态加密等技术，确保了每个计算节点在整个计算过程中都无法看到其他参与方的任何隐私信息，最终结果输出只有发起方有权限查看，其他计算节点无法获取，从而确保了应用过程的隐私性。同时，在产品逻辑上，半诚实模型与恶意模型的实现，也保证了各方隐私信息的安全不泄露，不会在通信层面或者非数据方节点有任何隐私数据留存。

[7]AUC：机器学习领域中的一种模型评估指标：其值越接近 1 则代表模型效果越好。

[8]KS：机器学习领域中的一种模型风险区分能力评估指标：其值越大则模型的风险区分能力越强。

自动化编译引擎：首先，隐私保护 AI 编译器以底层数据流图[9]的视角揭示了机器学习算法和对应的分布式联邦学习算法的联系，可通过数据流图变换完成两者间的自动转换。**其次，**数据流图变换具有通用性，可以适配上层多种机器学习算法，如逻辑回归、贝叶斯分类、神经网络等。从数据流图的视角，分布式联邦学习变换可以理解为将整体数据流图切分为若干子图分布到各隐私保护计算参与方，并保证子图交互的部分（通信部分）以隐私保护的方式进行。简而言之，该引擎将隐私保护计算各参与方使用的不同算法“格式化”为统一算法，避免了各方分别转换算法所需的定制化改造带来的巨大工作量和时间成本。

安全可验证: 受限于密码学证明方式与联邦学习领域的结合不够深入，传统模式无法做到安全性的自动化验证。隐私保护计算方案可全方位实现事前、事中、事后的安全评估验证。同时，以底层数据流图的视角揭示算法和对应的分布式联邦学习算法的联系，使得运算流程透明可见、可审计。

高度适配产业需求: 相比人工编译模式，隐私保护算法的构造速度指数级提升，系统整体运行速度是典型架构模式的 20-40 倍，能够在实施难度、系统效率、安全可视等方面满足工程、业务、运维、安全等各方面综合需求，具备成熟的商用推广模式。

2.中小银行间横向反欺诈建模和黑名单共享案例

[9]数据流图（Data Flow Diagram）：简称 DFD，它从数据传递和加工角度，以图形方式来表达系统的逻辑功能、数据在系统内部的逻辑流向和逻辑变换过程，是结构化系统分析方法的主要表达工具及用于表示软件模型的一种图示方法。

（1）业务背景

风控能力一直被视为银行机构的核心能力，但目前行业内各梯队风控能力悬殊。大型银行在风控技术和经验上的优势极为明显，尤其在互联网平台的流量加持下，吸引了更多客户，从而积累了更丰富的数据，使得依托于客户数据规模的风控优势进一步扩大。与大型银行优势地位形成鲜明对比的是，中小银行的风控处境比外界所见的更加困难。**一方面**，大型国有商业银行或股份制商业银行的地方分行依托相对较低的利率和资金成本，更易吸引信用记录良好、资产结构良好、资产负债率较低等资质较优的客户，留给地域经营的中小银行的客户群体相对更容易存在征信不良、固定资产少、多头借贷等问题。**另一方面**，中小银行的服务客群更集中在不发达地区，这类客群很容易受到电信诈骗、钓鱼网站、木马病毒、黑客勒索等黑灰产影响，给中小银行的风控工作带来更复杂严峻的挑战。这些问题迫使中小银行必须持续完善自身风控体系，当前较为迫切的需求体现在反欺诈和识别不良客户两类风控业务上。

（2）传统方案

在反欺诈业务的传统解决方案中，两家或多家中小银行各有一批欺诈样本，分别构建反欺诈模型，拟合出客户特征和反欺诈样本之间的关系，但在样本和观察数据规模有限的前提下，本地建模学习的模型效果并不理想。若各行的客户分类、客户偏好、地理位置等分布差异较大，银行间共享样本并基于共享的样本合集构建模型，可显著的

提升模型效果，但出于对数据安全及个人信息保护的考虑，银行之间无法直接共享样本数据。

不良客户识别业务的风控处境更为严峻。传统条件下，数据分享只能通过明文方式，且被查询方能够获取查询方的记录，因此，不良客户识别业务在机构间的安全合作甚至无法在传统条件下实现。直至隐私保护计算技术出现后，银行机构间的数据融合才有了安全的实现方式。

（3）业务痛点

不良客户识别方面，各银行在开展业务过程都会积累业务黑名单，以此在前置风控环节识别并剔除不良客户。对于中小银行来说，因为其业务开展的时长及覆盖的客户有限，积累的黑名单无论在客户体量、客户地域分布上都较为局限，无法帮助银行精准高效地识别不良客户。

反欺诈方面，交易实时反欺诈系统对交易欺诈行为的甄别，很大程度上需要依靠专家规则与机器学习模型。对中小银行而言，一个棘手的问题是，积累的欺诈样本数量较少，不足以构建高准确度的交易反欺诈模型，导致反欺诈工作开展效果不理想。

政策合规方面，针对上述两个业务风控样本不足的共通性问题，中小银行迫切希望获得其他金融机构，尤其是同类银行的风控样本数据，作为自身风控样本数据的补充，以提升自身的风控能力。近年来，国内外出台的一系列数据安全与隐私保护相关政策法规，在数据收集、使用与交易模式等方面进行了规范，银行间的数据共享与流通将面临严格的合规限制。

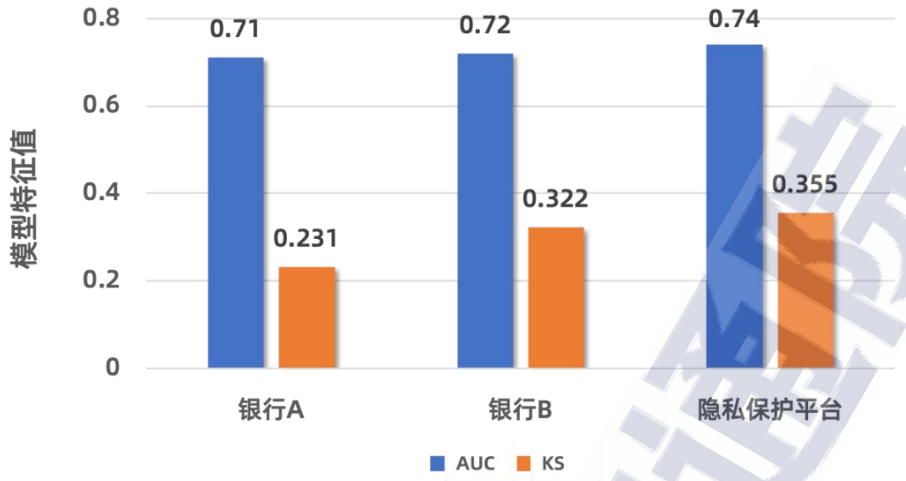
（4）实践案例

针对黑名单局限、建模所需数据样本不足、数据安全保护趋严等业务痛点，隐私保护计算技术为中小银行提供了解决方案。

在隐私保护计算平台实践案例中，通过在各银行部署隐私保护计算节点，实现了银行间的数据互联对接，在满足数据不出库、客户隐私不泄露的安全合规要求的同时，横向联邦功能为银行提供了反欺诈模型共建能力，匿踪查询功能实现了银行间的黑名单共享，帮助中小银行化解了上述的业务痛点。以下分别展开说明:

1）反欺诈模型共建

以银行 A 和银行 B 的反欺诈模型共建试点项目为例，在该反欺诈模型共建案例中，银行 A 和银行 B 分别提供欺诈样本数据，包括欺诈用户标签以及该用户的特征（如信用记录、消费习惯、常用手机设备等），并上传至隐私保护计算节点。隐私保护计算平台通过横向联邦的方式，对两银行的欺诈样本数据进行数据特征对齐及建模，在数据可用不可见的情况下，构建一个双方共用且效果更优的反欺诈模型。



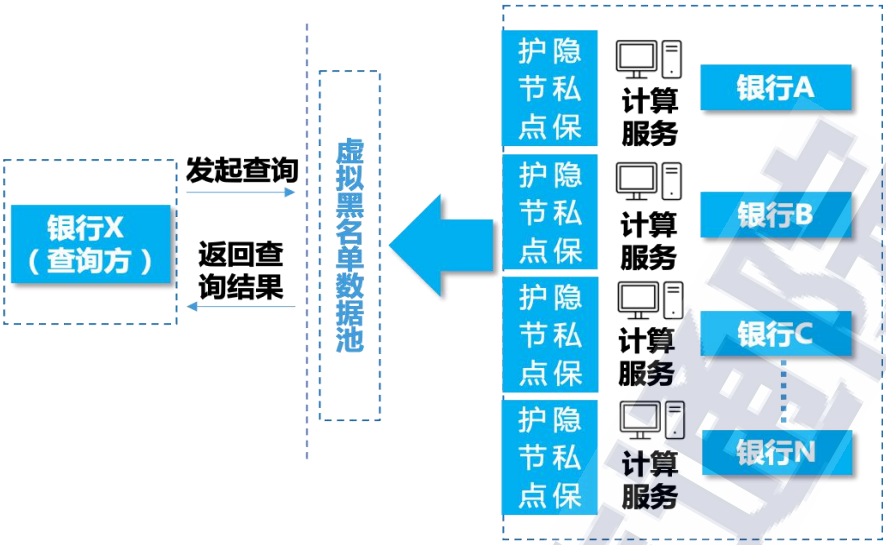
来源：隐私保护计算服务提供商

图 5 横向联邦反欺诈模型指标对比

在模型效果方面（如图 5 所示），传统方案中银行 A 和银行 B 分别在本地构建反欺诈模型，模型评估指标 AUC 值分别为 0.71 和 0.72，风险区分能力评估指标 KS 值分别为 0.231 和 0.322；隐私保护计算平台构建横向联邦学习得到的模型，AUC 值和 KS 值分别是 0.74 和 0.355。由此可见，基于隐私保护计算技术的横向联邦应用为金融反欺诈业务带来了一定程度的指标提升。

2）黑名单共享

除了反欺诈模型共建，隐私保护计算平台也可为金融机构提供金融黑名单共享的能力，打破金融机构间的“数据孤岛”。如某大型股份制商业银行牵头，与多家银行共同搭建了隐私保护计算平台。银行间基于匿踪查询技术为彼此提供隐私黑名单查询服务，各行在保护隐私安全的前提下实现了黑名单共享（如图 6 所示）。

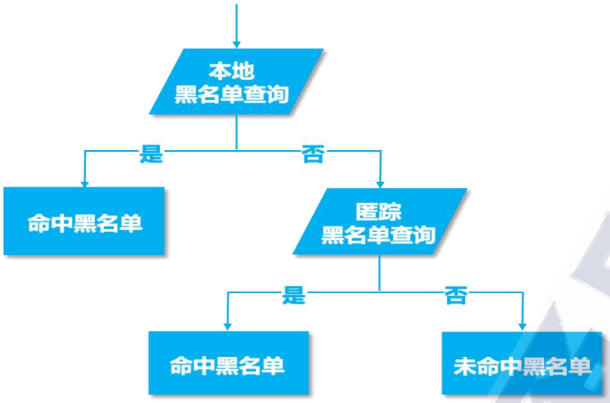


来源：隐私保护计算服务提供商

图 6 基于匿踪查询技术提供银行间隐私黑名单查询服务

匿踪查询技术能够保证查询发起方在不暴露被查询客户 ID 的前提下，获得该客户 ID 是否在其他机构的信息。在黑名单共享场景中，客户 ID 为客户身份三要素信息（姓名、身份证号、手机号），待查询的信息为是否在黑名单中，通过同态加密技术对客户 ID 和各维度金融信息进行加密，并对外提供黑名单服务。查询方即可知晓该用户是否在虚拟黑名单数据池中，并且仅能解密所请求客户 ID 的黑名单标签值。

在实际匿踪查询业务中（如图 7 所示），银行 X 要判断一个客户是否是黑名单客户，首先会在本地查询该客户 ID 是否存在于本行黑名单内，若命中黑名单，则业务流程结束；若未命中，则通过匿踪查询技术，向银行 A、银行 B、银行 C 分别查询客户 ID 是否在其黑名单内，并得到返回结果。



来源：隐私保护计算服务提供商

图 7 匿踪查询业务流程

（5）实践价值

在本案例中，银行通过部署隐私保护计算平台完善了自身的风控体系，一方面通过横向联邦实现了欺诈样本的安全共享与模型共建，另一方面通过匿踪查询实现了银行间的黑名单安全共享。

隐私保护计算方案在性能上也有较优的表现。在反欺诈模型横向联邦建模场景中，银行 A 和银行 B 分别提供百万级别的训练样本，平台每进行一次联邦建模的总耗时仅在分钟级别内，接近模型本地训练的性能；黑名单共享匿踪查询场景中，在银行 A、银行 B、银行 C 各自拥有数万量级黑名单的情况下，匿踪黑名单查询服务对于单个客户单次查询的平均耗时为 720ms，单次平均通信数据量为 420M，可以满足业务场景时效性的要求。

3.基于隐私保护计算技术的同业风控联盟案例

（1）业务背景

移动互联网及大数据技术的蓬勃发展加速了数字经济时代的到来，传统的金融业务模式正在随之不断发生变革，越来越多的金融机

构通过人工智能、云计算等技术拥抱数字化转型。互联网金融、数字化金融为金融信贷带来快捷便利的同时，网络犯罪的强隐蔽性和金融欺诈的低成本也使信贷行业面临着更严峻的欺诈风险，多头借贷恶意行为的发生几率随之升高。

多头借贷一般分为两类：**一类**是长时多头借贷，主要是用户消费、投资等需求超出现有收入水平和授信后寻求更多授信的行为，此类人群容易因资金链断裂导致逾期，抗风险能力较弱。**另一类**是短时多头借贷，主要是一些不法分子通过团伙欺诈、电信欺诈、杀猪盘等手段，利用金融机构间的信息差，在短时间内大量申请授信。统计数据表明，多头借贷客户的逾期风险可达普通客户的 3 倍以上，对金融机构的正常运转和金融稳定性带来极大隐患。

（2）传统方案

传统方案中，金融机构主要通过人行征信报告或外部三方数据的方式查询多头借贷人员名单。

人行征信报告主要通过汇总和加工各家金融机构主动上报的信息对外提供服务。**一方面**，在助贷模式中，平台端为提升用户体验，人为向多家机构推送同一用户的同一借贷申请，导致用户被动多头以致其征信被污染；**另一方面**，当前各家金融机构内部对借款的宽限期及征信上报的时间标准（包括时间点、上报频率）存在差异，最小颗粒度为天的征信报告无法帮助金融机构及时识别短时多头借贷的用户，存在利用征信信息收集和更新的时间差获得非法授信进行多头借贷的风险。

查询外部三方数据往往需要金融机构提供用户三要素信息，存在原始数据出库直接暴露个人信息和第三方数据来源合规风险，成为业务稳健发展的达摩克利斯之剑。

（3）业务痛点

安全隐私风险：传统的外部三方数据查询一般需要查询方将经 MD5 或者 SHA256 处理后的用户唯一标识发送至数据提供方，并通过 API 的方式获取查询结果。数据提供方基于已有用户信息规模优势，将自己所有用户的 ID 用 MD5 加密之后进行对比，相等则即可识别出对应用户。这种查询方法很容易解析到数据查询方的原始用户信息，并间接掌握被查询用户的身份信息，对于数据查询方而言存在用户信息泄漏风险。

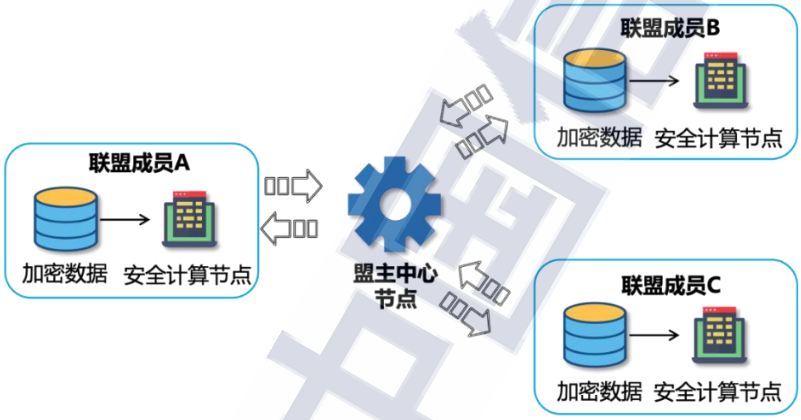
风险识别低效：征信机构的信息来源于各家金融机构的主动报送，参与信息报送的金融机构数量及其用户量影响着征信报告的覆盖度和准确度。**一方面**，一些非持牌的互联网金融机构用户体量较大但并未接入央行征信，接入机构有限导致征信机构覆盖用户不全面。**另一方面**，参与征信报送的金融机构因内部风险规则不同使得对宽限期的定义标准不一，且各家金融机构的报送时间点、频率存在差异，导致征信报告存在信息迟滞的问题。征信报告覆盖度和准确度因此受到影响，一定程度上降低了金融机构的贷前用户风险识别效率。

主观意愿不强烈，存在客户竞争：不同体量的金融机构在信息共享时存在数据共享的公平性问题和客户资源暴露问题，传统的银行联

盟模式下，客户名单共享易造成客户流失，以致金融机构彼此提防，数据共享意愿低。

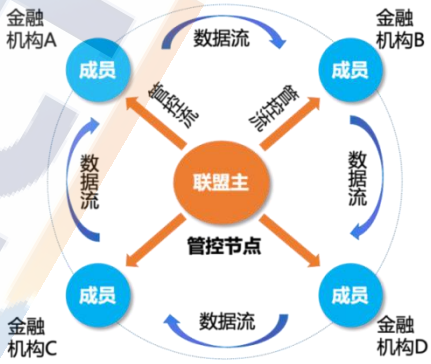
（4）实践案例

为有效应对上述问题，以将管控流和数据流分离为设计原则，基于隐私保护计算技术的隐匿查询双盲方案应运而生（如图 8 所示）。银联作为联盟主协调方部署盟主中心节点，实现中心化管控，参与联盟的银行为联盟成员，部署计算节点，联盟成员的数据均留存在本地。



来源：隐私保护计算服务提供商

图 8 隐匿查询双盲方案



来源：隐私保护计算服务提供商

图 9 数据流与管控流分离

联盟中某一成员发起查询请求到盟主中心节点（如图 9 所示），中心节点收到请求后转发到联盟内其他成员隐私保护计算节点，做到数据查询方与数据提供方身份的互盲。隐私保护计算节点通过隐匿查询技术保障数据提供方不能获知数据查询内容，保护数据查询方的输入数据，避免客户信息泄露，防止同业恶性竞争。

相比较传统数据共享方式，基于隐私保护计算技术的隐匿查询双盲方案在数据安全、数据实时性、数据质量、数据开放生态 4 个方面都有明显改善（如表 4 所示）。

表 4 传统计算方案与隐匿查询双盲方案对比

对比项	传统方案	隐匿查询双盲方案
数据安全	查询数据可被解析，存在数据泄露风险	数据不出本地的前提下多方联合计算，有效降低数据泄露风险
数据实时性	多头名单数据更新不及时，金融机构无法第一时间识别风险	可实时更新联盟多头数据
数据质量	助贷查询多资方模式导致查询被污染	参与方彼此直连，保证数据质量
数据开放生态	无法避免客户流失带来的同业竞争，参与方数据共享意愿低	身份双盲设计，参与方可放心进行数据共享，提升生态开放性

来源：隐私保护计算服务提供商

目前本案例联盟有 100 余家金融机构参与使用，包含多家头部金融机构，日均计算量达到 60 多万，满足实时多头数据预警应用场景的业务需求，后续可增设场景解决授权额度共享等痛点问题。

（5）实践价值

本案例将原始数据保留在本地，通过隐匿查询技术使得各参与方在查询过程中身份双盲，各参与方对数据“可用不可见”，只通过数

据使用的共享来实现数据价值共创。既满足各方业务需求，又保障了数据安全。

借助隐私保护计算技术能力，金融机构能够在贷前客户识别业务中及时更新客户的多头信息，实现了征信查询场景的实时响应，兼顾了业务准确性、实时性和数据安全性要求。同时，身份双盲设计消减了参与方的同业竞争顾虑，促进了金融机构间的数据安全、充分共享，进一步弱化了数据共享的壁垒。

4.基于区块链和隐私保护计算技术的小微企业智能风控产品案例

（1）业务背景

目前我国 95%以上的企业属于小微型企业，作为国民经济的重点之一，我国对小微企业扶植力度逐年加大，无论是政策上的支持、制度上的支撑、还是资金上的补助，都反映出国家大力发展小微企业的决心。为切实助力支持小微企业的发展，各商业银行将信贷业务的服务重心，从大中型企业逐渐转向小微型企业。

然而，随着大数据时代的快速发展，商业银行在对小微企业进行信贷风险控制时，为了得到更加精准的信用风险预测结果，往往会将小微企业的基本信息、资金流水等自有数据与工商、司法、税务、公安等外部数据相结合，以此来丰富数据维度、扩大数据规模，提升联合建模模型的准确度。但数据的合作过程涉及到多方数据的共享与利用，因此就存在隐私数据泄露、数据真实性等安全风险，导致商业银

行无法得到准确、全面的数据去判断小微企业的风险能力，面向小微企业的信贷业务面临严峻的考验。

（2）传统方案

商业银行结合外部数据对小微企业进行风险评估的业务场景下，传统解决方案是将经脱敏技术处理后的多方数据聚集在一起进行计算和建模。随着技术手段不断演进，基于脱敏技术的数据安全保护传统方案不再固若金汤：数据脱敏通过失真等变换在降低了数据敏感度的同时，又保留了一定程度的数据统计特征和可用性，但是攻击者仍可以通过如彩虹表[10]等特定技术手段对脱敏数据进行逆推处理，从而获取部分乃至全部原始数据，故仍存在原始数据泄露风险。

（3）业务痛点

商业银行方面的业务痛点，一是技术手段缺失，商业银行内部各部门、商业银行与工商、司法、税务、公安等外部大数据之间数据融合、风险信息共享程度低，机构间数据孤岛现象长期存在，各商业银行和小微企业有意愿打破这种桎梏，却缺乏有效技术手段，导致商业银行对小微企业的跨机构、多维度全景客户画像构建不全面、不准确，从而对小微企业客户风险和潜在价值的评估准确度降低，影响商业银行对小微企业信贷业务的风控评估与决策。**二是数据合作成本高**，由于商业银行与工商、司法、税务、公安等外部数据源在数据采集、统计标准等方面要求不一，在数据融合过程中就需要投入大量人力、物力重新整合数据，建立底层统一的数据资源框架，导致合作成本高昂。

[10]业界常用于加密散列函数逆运算的预先计算好的表,为破解密码散列值的工具。

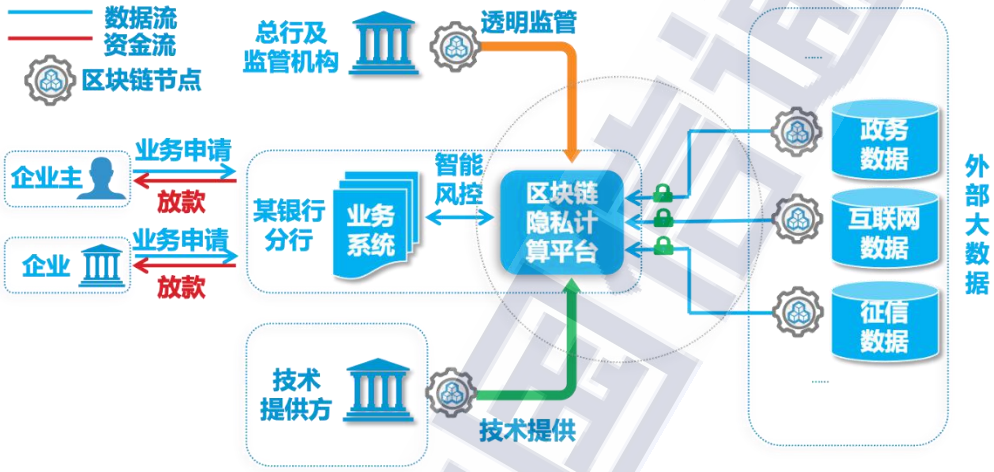
此外，在数据管理时，为解决内、外部数据安全问题，需要在组织架构、管理流程方面执行严格的管控，再次增加了管理成本。

小微企业方面的业务痛点，一是融资难，长期以来，商业银行对小微企业的印象是违规成本低、经营不规范、抗风险能力差，在社会各方尤其是与工商、司法、税务、公安等外部大数据中无法实现企业信息及时融合互通的背景下，商业银行无法**准确、客观**评估小微企业的风险等级，导致众多具备发展潜质的小微企业无法顺利借贷，形成了小微企业**融资难、融资贵、融资门槛高的现状。二是合作难，**受法律法规的约束，各企业无法进行各实体机构之间的数据共享，造成了相关企业之间因担保、营销以及发展规划差异形成的数据壁垒，企业间合作困难。

（4）实践案例

功能特点方面，基于区块链和隐私保护计算技术的小微企业智能风控产品，通过区块链上的隐私保护计算合约助力保护个人隐私和数据安全，实现大数据在各数据合作方之间的价值流通，具备相对安全的数据查询服务、风控数据分析、联合建模、多方数据规则和模型的部署与管理功能，能够联合外部大数据帮助商业银行信贷风控部门进行小微企业信贷客户的风险评估和决策，提升商业银行的风险识别能力和智能化水平。隐私保护计算技术在解决数据隐私保护和共享利用的平衡的同时，也面临着诸如数据真实性难确认、参与方身份难信任、可信数据共享协作网络难构建等挑战。本案例的风控产品通过区块链技术实现上链前数据具体来源、生成机制、存储过程的真实性交叉验

证，以及上链后数据使用可记录、源头可追溯、过程可审计、不可篡改等功能。通过区块链上的存证合约完成关键业务流程的上链记录，使数据应用、模型结果可信存储，同时支持对外开放接口提供给总行以及监管机构进行安全审计，解决了多方数据在融合过程中的安全性及真实性问题。



来源：隐私保护计算服务提供商

图 10 基于区块链隐私保护计算的大数据智能风控产品技术架构

技术架构方面，本案例通过安全多方计算和可信联邦学习技术将行内信贷客户的申请信息、存款、理财、行为偏好等数据和外部大数据进行安全融合（如图 10 所示），丰富了信贷用户风控数据特征维度，扩大了数据开放程度，实现了在保证商业银行与工商、司法、税务、公安等外部大数据源的原始数据不出各自私域的情况下，联合构建风控客户画像、风险规则和信用评分模型，帮助银行更加安全、全面、智能地评估信贷客户的风险状况。此外，通过区块链数字身份的建立，基于匿踪私密查询合约保护数据查询过程中行内信贷客户身份信息，采用切片决策引擎技术实现基于多方大数据的风控规则和模型

的安全部署和管理，并提供可视化监控分析展示系统，帮助银行建立贯穿信贷客户全生命周期的安全智能风控平台，提升多方大数据在行内的风控应用价值和效率（如表 5 所示）。

表 5 传统解决方案与隐私保护计算解决方案对比

对比项	传统解决方案	隐私保护计算解决方案
参与方	数据聚合一方后进行计算分析	多方联合分布式计算
特征维度	受制于数据安全，特征维度缺失	多方联合，几乎涵盖全部特征维度
计算效率	单一节点计算效率低	数据并行计算，多方联合，显著提高效率
计算精度	有损失	无损失
数据安全性	多方参与数据出域，存在极大安全隐患	数据不出域，实现数据的“可用不可见”

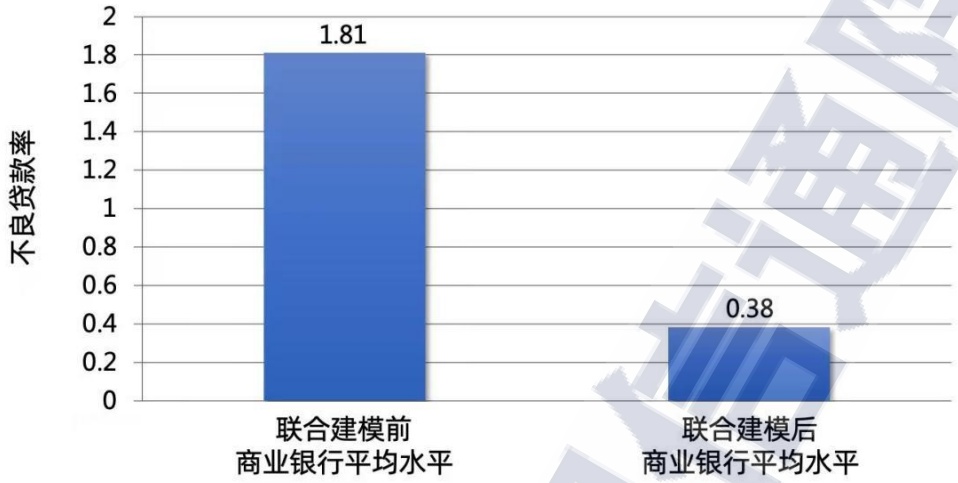
来源：隐私保护计算服务提供商

（5）实践价值

数据方面，隐私保护计算平台为银行提供了外部大数据安全融合能力，缓解了数据的泄露风险，提升了银行的大数据风控应用能力。本案例帮助银行联合了包括工商、税务、水电、司法、电信、征信机构等十余家跨行业数据源提供的上千个数据维度的外部大数据进行小微企业风控。在基于这些数据进行联合建模后，商业银行的不良贷款率从原有的 1.81 大幅下降至 0.38（如图 11 所示），经测算，基于区块链和隐私保护计算技术训练得到的信用评分模型，无损于传统方式得到的模型，其 AUC 提升 11%，F1 Score[11]提升 42%，精度（Precision）

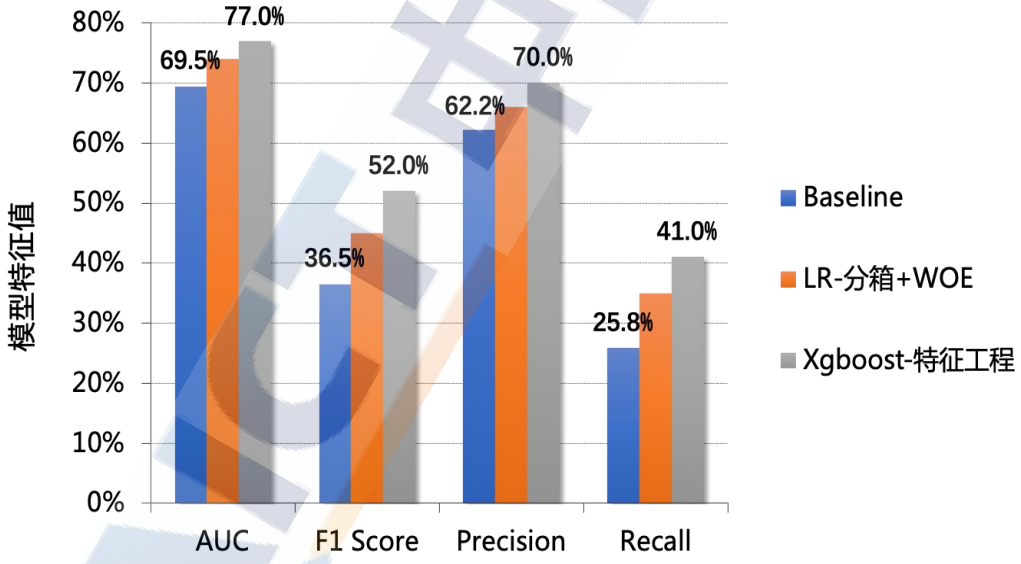
[11]F1 分数（F1 Score）：统计学中用来衡量二分类模型精确度的一种指标。

从 62.2%提升到 70.0%，提升幅度 13%，召回率（Recall）提升 59%（如图 12 所示），大大提升了银行的大数据风控应用能力。



来源：隐私保护计算服务提供商

图 11 联合建模前后不良贷款率对比



来源：隐私保护计算服务提供商

图 12 融合外部数据的建模效果

效率方面，本案例有效优化了银行建模分析决策路径和信贷风控流程。根据案例试点实践数据，隐私保护计算平台运行期间，小微企

业信贷业务平均审批效率较之前提升了 30%，不仅大幅提升了银行的风险管理水平，而且也极大优化了信贷客户的申请体验。

业务方面，商业银行结合外部数据的引入，深度挖掘自身数据，得以更加有效、低成本地触达小微企业客户，准确识别信用风险。区块链和隐私保护计算技术保护了数据提供者和数据使用者双方的数据隐私安全，使多方数据相对安全地应用于业务决策。

（二）医疗行业应用案例

1.全基因组安全联邦学习分析案例

（1）业务背景

随着数据要素价值释放的需求越来越强烈，现代医学研究、药物开发、公共卫生防疫以及临床医疗应用等生物医学科学的进步，也愈发倚重电子病历数据、基因数据、影像数据等生物医学数据的开放共享与利用。例如，近几年常被提及的“精准医疗”以及全基因组关联研究（Genome-Wide Association Studies，GWAS）等相关概念，都是数据应用价值在生物医学领域的直观体现，其本质是通过分析大样本的个体生物医学信息，鉴别特定疾病类型的生物标记物，辅助疾病的预防、诊断和治疗，提高疾病诊治与预防的效率及成本。

生物医学数据，尤其是基因数据，包含了大量涉及国家安全、个人隐私的敏感信息，数据泄露将对国家安全、公共利益造成难以估量的损失，这使得生物医疗数据的开放共享受到一定阻碍，基于数据规模化积累的价值释放方法遭遇瓶颈。

（2）传统方案

GWAS 是指在人类全基因组范围内筛选出与疾病相关的变异序列，即单核苷酸多态性（SNPs）。传统的 GWAS 解决方案需要以足够大的病例和对照样本数量为基础，对其所有感兴趣的 SNPs 进行基因分型，然后分析每个 SNP 与疾病的关联，计算其关联强度和 OR 值[12]。

在传统的解决方案中，GWAS 非常依赖大量基因数据的积累，样本量不足是各项 GWAS 研究中的常见问题和困难。即使多方以丰富病例和对照样本数量为目标展开数据合作，也很难保证合作过程的数据安全。传统方案在数据合作过程中需要各参与方将数据进行物理转移，汇总到一方后进行基因分析，面临着第三方不可靠带来的潜在数据隐私泄露、数据滥用、数据转卖等风险，以及数据分享意愿不强等问题。

（3）业务痛点

传统方案通过限制数据的流通，一定程度上保证了数据的安全性，但方案的落地实施仍存在其局限性。

一是样本割裂缺乏交互和共享，规模化积累不足。基于生物医疗数据的各项科学研究通常需要大量样本，单一数据源很难满足海量的数据需求。且受限于不同数据源所在国家和地区其数据安全和隐私保护法律法规的要求存在差异性，不同数据源可能在部分地区允许外传，而在其他部分地区禁止外传，各数据源也无法有效地直接和第三方分享自身数据，加剧了医疗数据孤岛困境，影响生物医学研究的合作。

[12]OR 值：优势比，流行病学研究中病例对照研究的一个常用指标。

二是超大数据量，高通量计算对技术与计算效率的要求较高。通过汇集多方医疗数据，数据量和数据维度的增加满足了样本规模需求，这虽然有利于提升模型精准度，但对计算效率也提出了更高的要求，传统方案中的单一计算节点力所不及，无法满足计算效率与精度之间的平衡。

三是统一大数据平台存在安全性不足等短板，严重限制了其发展。生物医学研究数据包含了大量敏感的个人信息，研究发现，基于几十个基因位点（SNPs）的数据就可以基本确定一个个体的身份。面对如此敏感的医疗数据，当前 GWAS 依赖统一大数据平台的实现方式其安全性更显不足。如何在保护医疗敏感信息、规避隐私泄露风险的前提下，广泛推行生物医学数据分享和联合分析、多元医疗数据融合，成为制约 GWAS 研究的关键挑战之一。

（4）实践案例

强直性脊柱炎（Ankylosing Spondylitis, AS）是最常见的自身免疫病之一，发病一般较早，且主要累及青壮年男性，如不能及时接受科学治疗，有较高致残率。在我国，至少有 1000 万强直性脊柱炎患者，人群庞大。研究发现，该病与 HLA-B27 等基因具有高达 90%的相关性，因此开展强直性脊柱炎的 GWAS 分析有很高的社会价值。

由某三甲医院牵头，在隐私保护计算技术的支持下，实现了在不分享明文数据(个体基因数据)的基础上，支持强直性脊柱炎的 GWAS 分析，为解决生物医学数据开放共享问题提供了思路。

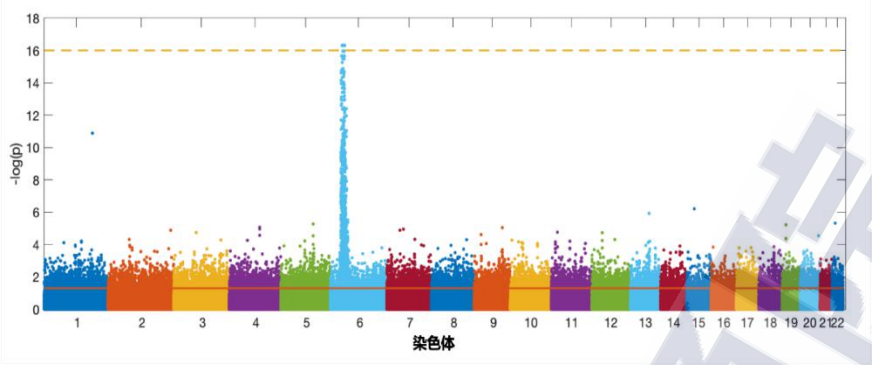
该方案基于隐私保护计算技术设计并开发了新框架，使用了具有隐私保护功能的安全联邦学习(Privacy-preserving Security Federated Learning)方法，整个数据共享的过程从始至终对患者信息进行保护，解决数据共享中存在的隐私安全问题。该框架以强直性脊柱炎作为切入点进行全基因组分析，以识别人类基因组中具有的潜在风险，即识别可能导致强直性脊柱炎的基因型。

表 6 P 值最高的单核苷酸多态性列表

单核苷酸多态性	染色体	位置	P 值
exm-rs886390	6	30334994	5.00E-17
exm-rs2844745	6	30343703	6.00E-17
exm-rs970270	6	30347306	5.00E-17
rs970270	6	30347306	6.50E-17
rs2516685	6	30361608	7.20E-16
rs12210947	6	30735105	9.10E-16
exm-rs4327730	6	30780936	5.00E-15
rs12192704	6	30792270	7.30E-15
exm-rs2254847	6	30933848	8.20E-15
exm-rs1634731	6	30955681	2.10E-14
exm-rs1619376	6	30983326	5.00E-14
exm529505	6	30993440	4.20E-13
rs2894179	6	31066671	6.30E-13
exm-rs3734854	6	31078836	4.33E-12
exm529653	6	31079264	5.20E-11

来源：隐私保护计算服务提供商

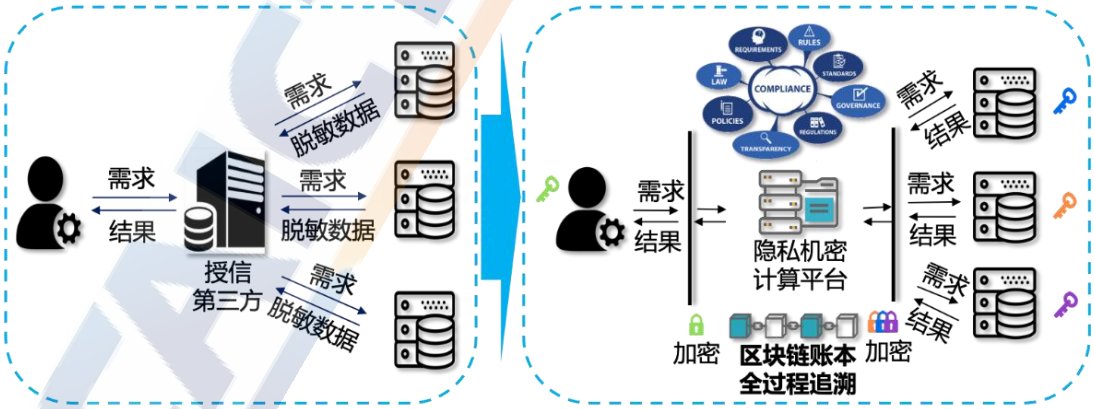
案例中（如表 6 所示）通过多中心全基因组关联分析，可得到部分 P 值最高的单核苷酸多态性数据，进一步以其为基础生成的曼哈顿图（如图 13 所示）可使得疾病相关的基因变异情况一目了然，为临床和实验提供了强有力的理论依据。



来源：隐私保护计算服务提供商

图 13 全基因组关联分析结果的曼哈顿图

相较于传统方案中，需要将数据拷贝移动到授信第三方，且还需面对由于不同机构间的不同隐私保护政策所带来的合规性挑战，隐私保护计算平台支持下的 GWAS 研究作为兼具隐私保护和跨机构数据共享的技术解决方案，连接多个数据源，实现了数据共享和有效利用（如图 14 所示）；在研究过程中只交换加密的经过处理中间计算结果，助力数据使用合规性，既保护了各方用户隐私、商业机密，又打破了数据孤岛，还使责任可追溯，让绝大部分计算在本地完成，有效减少数据冗余（如表 7 所示）。



来源：隐私保护计算服务提供商

图 14 传统方案和隐私保护计算平台技术方案架构对比

表 7 传统方案和隐私保护计算平台技术方案性能对比

对比项	传统方案	隐私保护计算平台
参与方	单一参与方或 数据汇到一方后分析	多方联盟式合作
数据样本量	受限于参与方或单一数据源样本量	多方合作显著提高样本规模及维度的丰富程度
计算效率	受限于单一节点计算效率	多方联邦模式计算，数据并行分析，显著提高效率
计算精度	基础参考标准	与传统技术方案比较均方误差在10^{-22}~10^{-28}之间
数据安全性	多方参与时需要数据物理转移，汇总分析，存在原始数据直接暴露的安全风险	数据不出域的情况下完成联合数据分析，数据“可用不可见”
结果安全性	计算结果不支持定向发放使用	计算结果定向发放使用

来源：隐私保护计算服务提供商

（5）实践价值

基于隐私保护计算平台打造的全基因组关联分析引擎，能满足 GWAS 研究所需的超大数据量（GB ~ TB 级数据）、多中心（10+中心）联合计算的技术要求。通过使用联邦学习框架，可以在不拷贝和移动原始数据的情况下实现多中心的联合 GWAS 研究，避免了传统数据共享过程中数据管理职责模糊的问题，使数据管理的职责清晰化。此外，部分计算在本地完成，有效减少了数据冗余问题，进一步提高了 GWAS 研究的计算效率，也有效解决了大数据平台安全性不足以及各参与机构分享意愿不强烈的痛点。

微观来看，隐私保护计算平台在此实践案例中是一个具有创新性的大数据流通共享和利用平台，不同于其他传统的大数据系统，该平

台在计算过程中不会泄露敏感的原始数据，充分保护生物医学隐私数据和医疗机构商业机密。在符合法律法规及相关管理部门监管要求的基础上，打破数据孤岛，建立了跨行业、跨部门、跨主体的安全、可控的大数据联合分析。

宏观来看，基于隐私保护计算技术的大数据分析管理平台，不仅在生物医学研究，在金融保险、商业营销等行业也具有广泛的应用前景。不仅可以应用在政府监管部门，也适合行业联盟、集团企业，在保护隐私安全、商业机密安全以及信息安全基础上促进数字产业发展，为建设数字中国提供动能。

2.基于“安全多方计算+联邦学习”的 DRG 付费

（1）业务背景

近几年，随着《关于进一步深化基本医疗保险支付方式改革的指导意见》《关于推进医疗保障基金监管制度体系改革的指导意见》等系列政策的出台，深化了医保支付方式的改革，成为促进我国医疗保障制度健康持续发展的重要内容。在系列政策的助力下，作为全球公认较为先进和科学的医保支付方式之一的“医疗诊断相关分组”（Diagnosis Related Groups，DRG）持续受到重视。

DRG 本质上是一种病例组合分类方案，即根据年龄、疾病诊断、合并症、并发症、治疗方式、病症严重程度以及转归和资源消耗等因素，将患者分入若干诊断组进行管理的体系[13]。传统医保费用支付方式是医保部门按照患者在院的实际费用（即按服务项目）支付给医疗

[13]《国家医疗保障 DRG 分组与付费技术规范》

机构，但在疾病诊断相关组-预付费（DRG-PPS）模式下，医保部门将根据患者所在诊断相关组的付费标准将费用预给医疗机构，以实现相关组内患者临床过程的相似，以及资源消耗的相近。

（2）传统方案

在传统方案中，各医疗机构需将患者病例信息按规范汇总至医保部门，由医保部门统一进行医疗诊断相关分组，并反馈至医疗机构。但由于医保部门给予医疗机构的反馈，往往是在医疗机构对患者完成诊疗之后，导致医疗机构在诊疗过程中对于患者的分组方法并不明晰，只能根据自身病例数据积累先进行预判，再根据预判进行分组诊疗。

（3）业务痛点

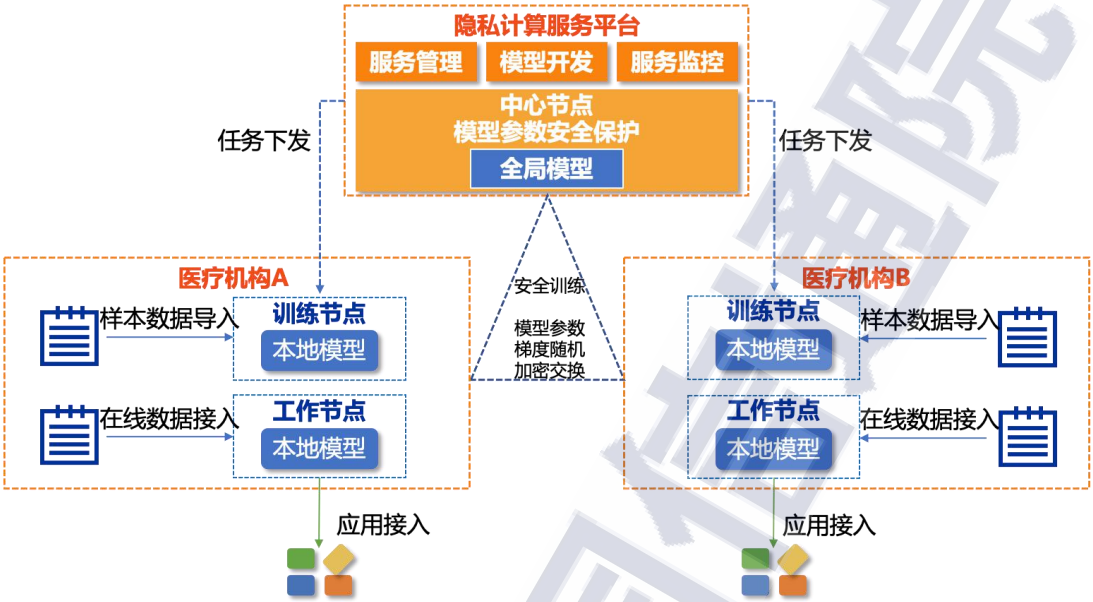
单个医疗机构建立模型，准确性不足：由于单个医院拥有的样本数量较少，以有限的数据难以进行模型训练，继而对分组的准确性产生影响。

多个医疗机构建立模型，安全性不足：各医疗机构及行业主管部门都将医疗数据安全作为监管重点，多方联合建模必要面对数据汇集可能导致的隐私泄露风险。由此，如何在确保各方医疗数据安全的前提下，充分挖掘数据价值，为医疗机构提供科学的参考、为人民群众就医提供便利和保障，成为医保 DRG 控费亟待解决的关键问题。

（4）实践案例

在本实践案例中，通过隐私保护计算服务平台将多家医疗机构的患者数据进行汇总训练，在保护患者隐私的前提下，增加患者样本数

量，扩大数据规模，最终获得了更准确的 DRG 分类模型，帮助医疗机构进行 DRG 预测。



来源：隐私保护计算服务提供商

图 15 基于隐私保护计算服务平台的联合 DRG 建模的流程

DRG 模型训练（以下简称平台）（如图 15 所示）基于隐私保护计算服务平台的联邦建模主要包含初始化、数据准备、隐私求交、模型训练、模型发布、服务集成、服务监控七个环节。

初始化：医疗机构 A 和医疗机构 B 在本地进行隐私保护计算节点部署，并进行网络授权和调试，待初始化完成后即开始具体联合建模项目的运营。

数据准备：医疗机构 A 和医疗机构 B 将本地样本数据加载到各自本地隐私保护计算节点上，在平台上进行对应样本的数据表结构注册并授权进入联合项目。

隐私求交[6]：平台上选择两方注册授权的数据集合进行隐私求交指令操作，实现两方样本数据对齐，形成虚拟宽表（数据存储表，列为属性，行为 ID）。

模型训练：平台上针对虚拟宽表进行模型训练，其中包括数据预处理、特征工程、特征筛选、算法调优以及模型评估，待模型训练完毕后产出模型评估报告并由联合项目机构进行线下模型评审，最终完成后即可进入模型服务部署阶段。

模型发布：机构针对提交的联合模型各自开发模型需要的对应机构的特征服务，一般以 API 形式对接本地隐私保护计算节点。完成特征服务后在平台进行特征定义（即注册），并将模型与特征绑定后进行发布。

服务集成：服务集成在平台进行操作，主要针对已发布的模型进行出入参配置，以及调用服务流程编排，并进行服务链路验证保证。待上述步骤完成后即可进行服务部署，一般以 API 形式由服务需求方（比如金融机构的决策系统）进行调用。

服务监控：服务正常运行时，平台提供全链路服务监控，用以监控联合模型服务的调用情况以及运行时模型稳定性情况。

两个医疗机构基于隐私保护计算服务平台，通过联邦学习技术实现了数据不出本地，使得数据隐私保护能力有所保障；同时，扩大了模型训练数据规模，提升了本地 DRG 模型准确度。

（5）实践价值

DRG 支付模式的优势**一是**能够减少对药品、耗材、大型建设设备的不合理使用，减少过度医疗，有效降低患者医疗成本、减轻患者经济负担；**二是**提高医疗机构医疗资源利用率，有利于促进医疗服务公开透明，有效规范医疗机构的医疗服务行为，有效提高医疗服务质量；**三是**医保基金不超支，助力医保控费。DRG 支付模式有助于实现医、保、患三方各自利益达到最大化，建立以患者为中心、使医保管理部门和医疗机构实现医保购买谈判、财务收支平衡，调动广大医务人员的积极性，优化临床路径、规范诊疗行为、提高服务效率，促进医疗卫生事业可持续发展。

基于隐私保护计算技术实现的联合 DRG 建模方式在患者方面，加强了数据授权和流转的立法保护和实际落地，保护了患者医疗数据和个人信息的安全性；医疗机构方面，本案例提供了更安全的数据不出本地的共享方式，保障机构数据利益的同时充分释放了数据价值。

（三）政务行业应用案例

1.基于电力联邦学习的城市电动汽车负荷分析与预测

（1）业务背景

面对全球气候变暖问题，我国明确在 2030 年前和 2060 年前分别实现碳达峰和碳中和。据统计，仅交通行业的碳排放量约占全国总碳排放量的 10%左右，其中道路交通在交通全行业碳排放中则高达 80%，推动新能源汽车产业发展已成为节能减排的关键抓手。

根据工信部统计，2020 年我国新能源汽车单年度销量约为 130 万台，与现存充电桩总数相当，新能源汽车与充电桩的保有量比例为

3.15:1，远高于《电动汽车充电基础设施发展指南》规定的 1:1。“公桩难找、私桩难设”是现阶段阻碍新能源汽车产业发展的痛点问题。要解决上述问题，亟需研发城市电动汽车负荷分析与预测技术，为开展充配电网协同布局规划提供技术支撑，实现增量基础设施优化配置，提升城市充电网的覆盖率和利用率。

（2）传统方案

在传统电力场景中，负荷、电量等用电数据来源单一，可在电力企业的数据中台进行汇聚、脱敏、分析和建模。而在电动汽车领域，电动汽车的充电可在公桩、私桩等不同渠道完成，在电动汽车负荷分析与预测场景中，电动汽车的充用电数据由电力公司、私桩个人等多方持有，出于商业利益的考虑和用户隐私保护等监管约束，往往形成各种数据壁垒。究其原因，一方面，用户充用电数据是持有机构的高价值资产，出于商业利益的考量，不会轻易对外开放；另一方面，这些数据关乎国家安全、涉及个人隐私，持有机构“不敢、不能”直接对外开放。

（3）业务痛点

安全事件危害范围广，程度重：由于电力网络存在结构复杂、业务特殊、系统繁多等特性，电力数据面临严峻的安全威胁与挑战，如若发生盗用、泄露、篡改、删除等安全事件，不仅会对电力企业自身的业务、信誉和经济利益造成严重损害，甚至可能影响能源供应，导致社会恐慌，威胁国家安全。

个人隐私数据高敏感，法律严：随着电动汽车大范围、高密度的推广和使用，电动车充用电数据几乎能够完整刻画出用户的行动轨迹和生活习惯，因此电动车充用电数据也是关乎用户隐私安全的高敏感个人数据，受到相关法律法规的严格保护。

数据孤岛使数据失真，存隐患：在传统的电动汽车负荷分析与预测过程中，电动汽车充用电相关数据往往分散在多个主体手中，例如新能源汽车数据、公共充电设施的用电数据、居民充电设施的用电数据等均由不同的机构收集与持有。传统方式使用单一数据源或高强度脱敏的数据，致使任意一方开展用户画像和数据建模时的模型效果不尽如人意，同时也存在巨大的数据安全隐患。

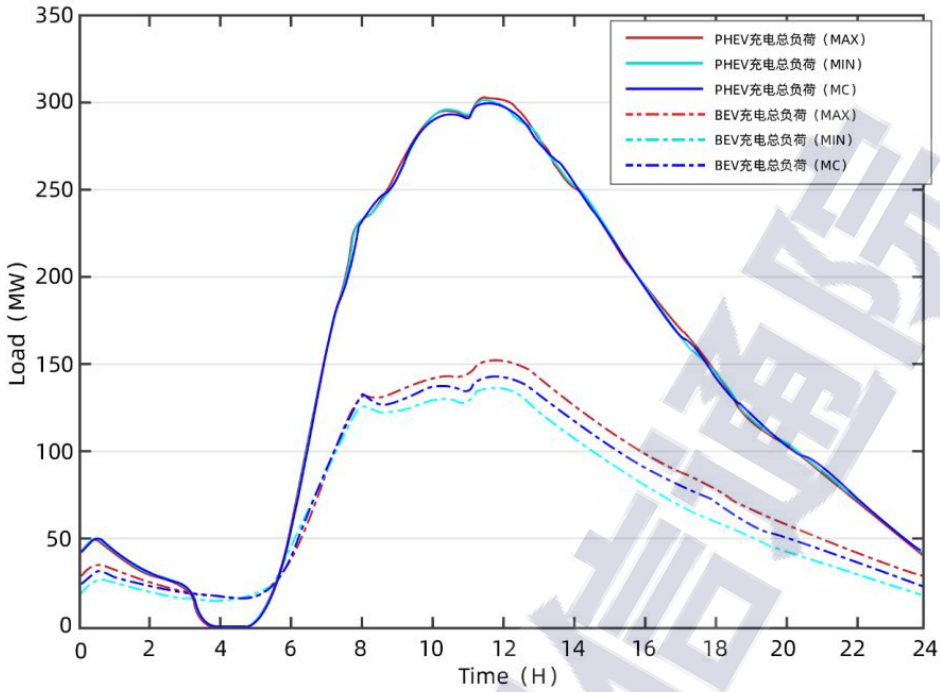
综上所述，如何通过技术手段来解决多方数据的共享问题，是城市电动汽车负荷的跨域分析与精准预测的一个关键难点。

（4）实践案例

针对以上问题，某科技公司联合电力公司，利用自主研发的安全计算平台，集成安全多方计算、联邦学习等隐私保护计算技术，提供面向配电网协同发展的电力场景安全计算解决方案。通过综合运用秘密分享、不经意传输、同态加密等密码学手段，提供满足实际电力业务场景需求的辅助联邦建模组件，包括安全数据对齐、安全多方统计与分析、联邦特征工程、联邦探索性分析、匿踪查询等功能；在此基础上，结合规模化的电动汽车场景，以上海电动汽车充用电数据为基础，构建充用电画像；最后，通过电力联邦学习算法建立城市电动汽车负荷分析与预测模型：

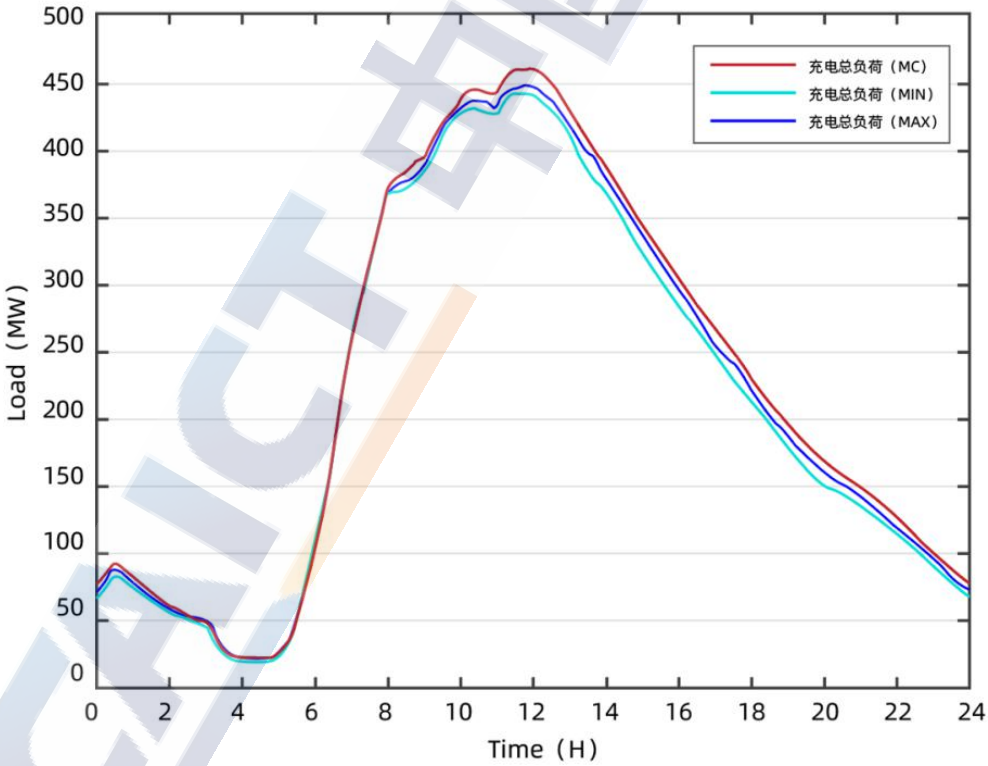
1）**城市电动汽车负荷分析：**基于横向电力联邦学习安全协同建模能力，融合公私充电桩运营数据，对全市充电桩、充电站、充电网进行负荷聚类分析，构建反映充电网的用电画像。首先，根据实际计算、存储、网络等资源条件，在数据持有者之间搭建支持多方安全协同建模的联邦学习平台。然后，通过多方协同的特征工程，建立分布式电动汽车负荷画像，并通过安全相关性分析方法量化评估各画像特征与充电桩、充电设施、充电网的负荷变化的相关性程度，从而对特征进行关联排序。最后，通过横向电力联邦学习算法建立融合多方同构电动汽车充用电数据的负荷分析模型（如图 16、图 17 所示），以充电桩、充电设施、充电网等多级粒度开展电动汽车负荷安全聚类分析，形成充电桩、充电设施、充电网负荷的聚类结果。

2）**城市电动汽车负荷预测：**基于纵向电力联邦学习安全协同建模能力，融合包含充电桩运营数据在内的电力、经济、社会、交通、规划等多源异构数据，以用电画像特征库为基础构建充电网负荷预测模型。首先，建立融合多源异构数据的负荷预测模型，针对充电桩、充电设施、充电网等不同粒度开展电动汽车负荷预测，不同粒度不同分类，使用多种长、短期负荷预测方法构建多方协同数据模型。然后，分析比较各种预测模型的性能，得到适用于各种情况的解决方案。



来源：隐私保护计算服务提供商

图 16 PHEV 与 BEV 充电负荷曲线



来源：隐私保护计算服务提供商

图 17 电动汽车充电总负荷曲线

该方案基于联邦学习的用电预测建模方法，挖掘用电时序数据的局部、全局变化特征，构建兼具线性和非线性拟合能力的用电预测模型，建立电力数据价值和数据安全之间的平衡。其中，短期用电预测模型可用于月末供售电量预测、配电网元件重过载预警、台区可开放容量计算，为相关专职开展日常工作提供决策支持；中长期用电预测模型可用于预测无重大突发事件影响下的规上工业和一般工商业用电量，助力政府量化评估行业景气状况和复工复产状况。通过实行该方案，成功地支撑了充配电网的协同布局规划，助力新能源汽车产业发展和营商环境的持续优化（如表 8 所示）。

表 8 传统技术方案与安全计算平台创新方案对比

对比项	传统技术方案	安全计算平台创新方案
参与方	单一电力机构 或汇聚于同一中心电力机构	分布式、跨域的多个电力机构， 无中心机构
样本量	单一样本或脱敏的聚合数据集	间接聚合的多源数据样本， 数据样本更丰富
安全性	采用脚本或人工脱敏的情况，数据关系被破坏，易遭受单点攻击	数据不转移不汇集，采用高困难性和复杂度的新型加密和安全计算技术
效率	受限于单一机构效率	多机构并行计算，此外可扩展硬件加速
准确性	单一数据源导致模型效果不佳	海量、高维电力大数据的聚合样本， 显著提升建模效果
审计监管	由于机构间系统差异和数据孤岛，主要依赖人工审计	数据安全开放共享，自动跨域授权，此外分布式的架构可协同区块链进行存证溯源

来源：隐私保护计算服务提供商

（5）实践价值

安全计算平台创新方案以“电力数据跨源协同”为核心，提出了集成联邦学习、安全多方计算等新型隐私保护计算技术的电力场景解决方案，建立开放环境下多主体安全协同建模框架，能够在原始数据不出域、不直接交换的前提下，以不可破解的加密方式实现电力数据的开放共享与多元协同应用，使各数据持有机构之间安全高效地协同使用各方数据，合法合规地进行多源数据协同建模与分析，确保了各方在模型训练、更新、应用等环节实现“数据不出门，算法满地跑”，解决了电力场景中数据隐私安全、跨域数据协同应用和数据价值挖掘困难的难题。

在此基础上，该方案通过构建充电桩、充电站、充电网的用电画像，建立城市电动汽车的负荷分析与预测模型，进一步实现了样本规模的扩大、特征显著性的提升、模型预测精度的提高以及对电力数据安全和个人隐私的保障。

四、隐私保护计算技术应用困境及建议

现阶段隐私保护计算技术在金融、医疗、电子政务等领域已有一些落地尝试。但总体来说，隐私保护计算技术仍处于大规模商业应用的早期，由于技术和解决方案还不够完全成熟，隐私保护计算在走向市场化、产业化的过程中，仍面临诸多挑战，需多方精诚协作。

“徒善不足以为政，徒法不能以自行。”法规制度的生命力在于执行。在我国日臻完善的数据安全治理体系下，《数据安全法》《个人信息保护法》强调了在兼顾安全的基础上，鼓励依法合理有效利用数据和个人信息。《金融科技发展规划（2019-2021）》《中国一体

化大数据中心协同创新体系算力枢纽实施方案》《网络安全产业高质量发展三年行动计划（2021-2023 年）（征求意见稿）》等相关政策文件中，也提及强化安全多方计算、联邦学习、机密计算等技术的研究攻关和部署应用，促进数据要素安全有序流动。那么，如何合规地使用隐私保护计算技术，建立安全合规与正当商业利用相平衡的制度框架，亟需对除法律之外的行政法规、部门规章、准则、指南等“软性”措施进行细化、优化，强化与法律之间的衔接，为依法合理有效释放数据要素价值奠定基础。

“欲知平直，则必准绳。”当前围绕隐私保护计算已开展一系列的标准化工作，但由于技术路线丰富、场景强相关、轻量化与定制化无法兼得、安全性不统一等问题，尚未形成兼顾权威性、适用性、科学性的标准。隐私保护计算的标准化工作，不仅包含标准制定，更需注重标准的实施效果。建议立足市场需求和社会发展实际，坚持问题导向、结果导向，围绕隐私保护计算技术、应用和测评，加强研制可操作性强、量化指标明确、与实际发展水平衔接紧密、更加科学的标准。尤其在安全方面，当前隐私保护计算产品存在质量水平参差不齐的情况，随着《数据安全法》《个人信息保护法》的陆续生效，隐私保护计算技术标准化建设工作需求愈加迫切，以引导安全标准高质量发展，促进创新成果转化，加速市场化和产业化步伐，切实发挥安全标准“固根基、兜底线”的功能，实现安全能力的“优胜劣汰”，切实提升我国数据安全治理的水平。

“志不求易者成，事不避难者进。”隐私保护计算技术的性能和安全性难以兼得成为制约隐私保护计算技术应用落地发展的重要因素之一。建议**一是效率方面追“高线”**，通过硬件加速、算法优化、代码加速、通信优化等多种组合手段实现算力和效率的提升，实现满足高吞吐率、低延时、高实时性的隐私保护计算技术解决方案。**二是功能方面行“基线”**，隐私保护计算技术的细分技术应用在保护效果、计算性能、计算精度、硬件依赖、支持场景等方面都存在着较大的差异。因此在落地应用过程中，需积极开展面向产业需求的工程探索，充分发挥隐私保护计算技术各技术路线优势，科学融合各细分技术路线，实现满足场景覆盖广、产品成熟度高的隐私保护计算技术解决方案，筑牢隐私保护计算技术应用功能“基线”。**三是安全方面守“底线”**，建立健全隐私保护计算产品和应用安全性测试评估标准，通过权威性强、专业性强、认可度高的第三方评估机构对其安全性进行评估检测，守住隐私保护计算技术应用的安全“底线”。此外，随着量子计算的发展，基于计算复杂度的公钥密码及相关的隐私保护计算技术的安全性受到严峻挑战，密码学家亟需未雨绸缪，积极准备好更安全的抗量子密码算法。

“同舟共济扬帆起，乘风破浪万里航。”强化网络安全意识，提升网络安全素养。隐私保护计算技术应用过程中，存在技术门槛高、安全理解难、信任建立障碍、法律知识匮乏等困境。建议发展数据流通共享和相关术语的概念，提升公众对隐私保护计算技术、数据安全和个人信息保护的法律法规等相关内容的整体认知程度，增强对数据

要素市场培育及数据流通共享应用的信任水平，推进数字安全文化理念建设，在降低供需对接成本的同时，树立正确的安全价值观，营造良好的数据安全氛围。此外，加强专业人才培养，发挥人才优势，有效推动隐私保护计算技术成果落地，促进隐私保护计算技术成果交易。

中国信息通信研究院 安全研究所

地址：北京市海淀区花园北路 52 号

邮编：100191

电话：010-62300264

传真：010-62300264

网址：www.caict.ac.cn