隐私计算联盟
Privacy Preserving Computing Alliance

云大所
CAICT

隐私计算白皮书

（2021 年）

隐私计算联盟

中国信息通信研究院云计算与大数据研究所

2021 年 7 月

版权声明

编写委员会

❖ **主要编写单位**（排名不分先后）：

中国信息通信研究院云计算与大数据研究所、上海富数科技有限公司、蚂蚁科技集团股份有限公司、华控清交信息科技(北京)有限公司、北京数牍科技有限公司、零知识科技（北京）有限公司、深圳市腾讯计算机系统有限公司、联易融数字科技集团有限公司

❖ **参与编写单位**（排名不分先后）：

洞见智慧科技有限公司、北京冲量在线科技有限公司、天翼电子商务有限公司（电信翼支付）、天翼云科技有限公司、交通银行股份有限公司、上海阵方科技有限公司、京东科技控股股份有限公司、北京瑞莱智慧科技有限公司、华为云计算技术有限公司、蓝象智联（杭州）科技有限公司、北京融数联智科技有限公司、神谱科技（上海）有限公司、翼健（上海）信息科技有限公司、杭州锘崴信息科技有限公司、杭州趣链科技有限公司、优刻得科技股份有限公司、医渡云（北京）技术有限公司、深圳致星科技有限公司（星云Clustar）、上海浦东发展银行股份有限公司、中国光大银行股份有限公司、同盾科技有限公司、零幺宇宙（上海）科技有限公司、牛津（海南）区块链研究院有限公司、西安纸贵互联网科技有限公司

本报告的法律合规部分还得到了世辉律师事务所、腾讯研究院、广东君信律师事务所法律专家的审阅。

❖ **编写组主要成员**（排名不分先后）：

袁　博	闫　树	吕艾临	王思源
仵姣姣	魏　凯	姜春宇	刘嘉夕
靳　震	叶锦梅	李雪妮	卞　阳
黄翠婷	孙小超	王　磊	殷　山
昌文婷	任维赫	李　艺	王国赛
黄丽成	金银玉	单进勇	蔡超超
苏冠通	徐茂桐	李　南	程　勇
刘　江	刘淑珍	许　焱	陈　曦
姚　明	李　博	王湾湾	陈浩栋
周岳骞	章　庆	徐　潜	王益斌
刘天琪	王光中	谢　谨	夏伏彪
龚自洪	宋红花	孙中伟	雷明禹
徐世真	张　煜	林佳萍	毛仁歆
王　超	薛瑞东	黄小刚	祝文伟
黄程韦	张莺耀	范家璇	王　爽
李　帜	徐　静	汪小益	马　强
何永德	包仁义	黄　尧	杨　柳
马文婷	黄小芮	黄登玺	刘　巍
李晓林	彭宇翔	张　威	王江凌
于　昇	杨文韬	杨　珍	

前言

2020 年 4 月，中共中央、国务院发布《关于构建更加完善的要素市场化配置体制机制的意见》，将数据同土地、劳动力、资本、技术等传统生产要素并列，作为一种新型生产要素参与分配。作为释放要素价值的关键环节，数据资源的开放共享、交换流通成为重要趋势，其需求日益增加。

然而，近年来数据安全事件频发，数据安全威胁日益严峻。特别是《中华人民共和国数据安全法》的颁布和实施，对企业合规安全地发挥数据价值提出了更高的要求。既要应用数据，又要保护安全，如何兼顾发展和安全，平衡效率和风险，在保障安全的前提下发挥数据价值，是当前面临的重要课题。以多方安全计算、联邦学习、可信执行环境等为代表的隐私计算技术为流通过程中数据的“可用不可见”提供了解决方案，已在一些领域开始推广应用。可以说，隐私计算是在实现保护数据拥有者的权益安全及个人隐私的前提下，实现数据的流通及数据价值深度挖掘的一类重要方法。

近两年来，在政策驱动和市场需求同时作用下，隐私计算技术、产业、应用迅速发展，成为商业和资本竞争的热门赛道。2020 年底，中国信通院在工业和信息化部相关司局的指导下，联合业界六十余家技术企业和应用单位成立隐私计算联盟，成为隐私计算领域的重要行业组织。2021 年，中国信通院云大所联合隐私计算联盟的三十余家企业共同完成了这本《隐私计算白皮书（2021 年）》。本白皮书试图回答以下这些问题：

- 隐私计算是什么：为什么会有隐私计算技术？它能发挥什么价值？面临什么样的政策环境？
- 隐私计算技术发展情况：隐私计算的技术体系是怎样的？各类隐私计算技术的方案架构和特点有哪些？每种隐私计算技术擅长解决的问题是什么？其成熟度和缺陷有哪些？技术融合与扩充的情况如何？
- 隐私计算应用场景：隐私计算常用的应用场景有哪些？在每个场景里，隐私计算解决了什么痛点、如何应用？
- 隐私计算产业发展情况：国内外隐私计算主要有哪些企业？隐私计算行业的商业模式、论文情况、技术开源情况、标准建设情况如何？
- 隐私计算合规性情况：从法律视角看，隐私计算解决了哪些数据流通的合规性问题？应用隐私计算过程中，面临哪些合规性风险？如何解决这些风险？
- 隐私计算面临的问题与挑战：隐私计算的发展面临哪些问题？这些问题该如何改善？

道阻且长，行则将至；行而不辍，未来可期。面对这个日新月异、快速发展的行业，我们期待与业界共同守正创新，推动隐私计算行业健康发展，让隐私计算在数据要素市场建设和数据流通过程中发挥更大的价值！

目录

第一章

隐私计算概述

隐私计算（Privacy-preserving computation）是指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”。

站在数据成为比肩石油的基础性关键战略资源的当下，隐私计算为需求强烈但瓶颈重重的数据流通提供了破局思路。Gartner 发布的 2021 年前沿科技战略趋势中，将隐私计算（其称为隐私增强计算）列为未来几年科技发展的九大趋势之一。随着各方关注度的提升，隐私计算已成为发展火热的新兴技术，跻身商业和资本竞争的热门赛道。

（一）数据流通需求推动隐私计算势头火热

数字经济时代的特点之一便是将数据视作关键的生产要素，并通过跨领域、跨行业、跨地域的机构间数据流通释放要素价值。但是，目前我国数据要素市场化配置尚处于起步阶段，规模小、成长慢、制约多，机构间的数据流通仍存在诸多阻碍。

一是数据权属的界定仍不明确，在相关立法和制度尚未健全的当下，实践中并未能形成具有共识性的权属分割规则，产权争议、难以监管的风险令供需双方望而却步。**二是数据流通的安全风险高**，数据安全事件频发，出于对国家安全、个人信息和商业秘密的保护，企业参与数据流通的主动性、积极性因此降低。**三是如何确保流通过程的**

安全合法仍然较难把握。现有监管要求并未给出数据对外提供和处理的明确合法依据与参考指引，企业依然困惑于数据可流通的对象、范围、方式等一系列问题。除此之外，数据流通在数据质量、数据定价等方面也都面临着诸多挑战。

为解决上述障碍，政府部门和大数据行业从业者进行了艰辛的探索，寻求通过技术手段解决个人信息保护、权益分配、数据安全保障、追溯审计等难题。针对较为核心的个人信息保护，业界通过数据标识加密技术、数据标识关联技术和有效授权技术等为确保敏感信息不可实别和确保数据仅在授权范围内使用提供了一定的思路。

但是，以上技术仍不能抵御数据流通后被反推和滥用的风险，而“可用不可见”的隐私计算正是解决这一问题的技术突破口。从技术原理讲，隐私计算并不能简单归属于某一个学科领域，而是一套融合了密码学、安全硬件、数据科学、人工智能、计算机工程等众多领域的跨学科技术体系，包含了多方安全计算、联邦学习和可信执行环境等不同的代表性技术方案。从应用目的讲，一方面隐私计算可以增强数据流通过程中对个人标识、用户隐私和数据安全的保护；另一方面隐私计算也为数据的融合应用和价值释放提供了新思路。

（二）政策环境为隐私计算发展提供新机遇

近年来我国数据立法进程不断加快，尤其强调数据应用过程中的数据安全。《中华人民共和国网络安全法》《中华人民共和国数据安全法》和《中华人民共和国个人信息保护法（草案）》逐步完善了国家数据相关立法的顶层设计，着重强调了流通过程中的数据安全和个人

信息保护。

隐私计算是平衡数据利用与安全的重要路径。自 2016 年，工业和信息化部、中国人民银行、国家发改委、中央网信办、国家能源局等各部委先后在相关政策文件中提出加强隐私计算相关技术的攻关和应用。人民银行于 2021 年 5 月组织金融机构开展包括应用隐私计算进行数据共享在内的金融数据综合应用试点。在地方政府层面，广东省于 7 月发布的《数据要素市场化配置改革行动方案》中提出构建包含隐私计算在内的数据新型基础设施。政策的提前布局对于我国抢占隐私计算技术和应用关键领域奠定了良好基础。

表 1：相关法律及政策文件梳理

	时间	文件名	发布单位	简述
法律	2016 年 11 月	《中华人民共和国网络安全法》	十二届全国人大常委会第二十四次会议	强调对收集的用户信息严格保密，维护网络数据的完整性、保密性和可用性，实行网络安全等级保护制度
	2021 年 4 月	《中华人民共和国个人信息保护法(草案)》	十三届全国人大常委会第二十八次会议	强调个人信息在数据流通过程中的安全合规
	2021 年 6 月	《中华人民共和国数据安全法》	十三届全国人大常委会第二十九次会议	强调数据安全与开发利用并重，确立数据分类分级管理制度，多种手段保证数据交易合法合规
政策文件	2016 年 12 月	《大数据产业发展规划（2016-2020 年）》	工业和信息化部	支持企业加强多方安全计算等数据流通关键技术攻关和测试验证
	2019 年 9 月	《金融科技(FinTech)发展规划(2019-2021 年)》	中国人民银行	提出利用多方安全计算技术提升金融服务安全性
	2019 年 9 月	《工业大数据发展指导意见（征求意见稿）》	工业和信息化部	提出在工业领域积极推广多方安全计算技术，促进工业数据安全流通
	2021 年 5 月	《全国一体化大数据中心协同创新体系算力枢纽实施方案》	国家发改委、中央网信办、工业	提出“试验多方安全计算、区块链、隐私计算、数据沙箱等技术模式，构建数据可

			和信息化部、国家能源局	信流通环境，提高数据流通效率”
	2021 年 7 月	《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）》	工业和信息化部	提出推动隐私计算等数据安全技术的研究攻关和部署应用，促进数据要素安全有序流动
	2021 年 7 月	《广东省数据要素市场化配置改革行动方案》	广东省人民政府	提出构建包含隐私计算在内的数据新型基础设施

技术价值的凸显，再加上政策环境的助力，隐私计算在数据相关产业内悄然兴起，相关的学术会议和论文在近几年呈现大幅增长，相关研究从技术原理逐步转向应用实践。在算法协议不断优化、硬件性能逐步增强之下，隐私计算的可用性大大提升，越来越多的企业入局隐私计算的研发和产品化，金融风控、互联网营销、医疗诊治、智慧城市等越来越多的场景落地应用。目前，隐私计算已成为数据流通领域内最受关注的技术热点，市场一片火热。

第二章

隐私计算技术发展态势

从 20 世纪 70 年代一直到近年，隐私计算交叉融合了密码学、人工智能、计算机硬件等众多学科，逐渐形成以多方安全计算、联邦学习、可信执行环境为代表，混淆电路、秘密分享、不经意传输等作为底层密码学技术，同态加密、零知识证明、差分隐私等作为辅助技术的相对成熟的技术体系，为数据安全合规流通提供了技术保障。

（一） 隐私计算技术体系基本建立

从技术角度出发，隐私计算是涵盖众多学科的交叉融合技术，目前主流的隐私计算技术主要分为三大方向：第一类是以多方安全计算为代表的基于密码学的隐私计算技术；第二类是以联邦学习为代表的人工智能与隐私保护技术融合衍生的技术；第三类是以可信执行环境为代表的基于可信硬件的隐私计算技术。不同技术往往组合使用，在保证原始数据安全和隐私性的同时，完成对数据的计算和分析任务。

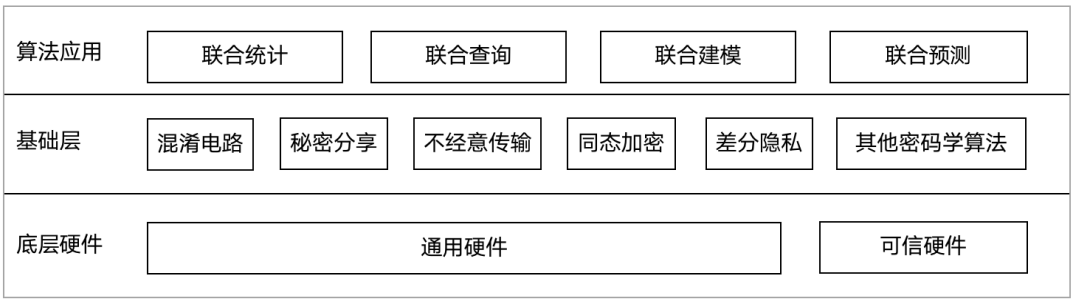

图 1 隐私计算技术体系

隐私计算技术为数据的隐私保护与计算提供丰富的解决方案，可从底层硬件、基础层和算法应用等不同角度加以区分。

如图 1 所示，**从底层硬件来说，**多方安全计算与联邦学习通常从软件层面设计安全框架，以通用硬件作为底层基础架构；可信执行环境则是以可信硬件为底层技术实现的隐私计算方案。

从算法构造来说，多方安全计算技术基于各类基础密码学工具设计不同的安全协议；联邦学习除可将多方安全计算协议作为其隐私保护的技术支撑外，基于噪声扰动的差分隐私技术也广泛应用于联邦学习框架中；可信执行环境通常与一些密码学算法、安全协议相结合为多方数据提供保护隐私的安全计算。

从算法应用来说，以不同技术为基础，隐私计算逐渐演化出丰富的算法应用场景。这些应用往往为了实现特定计算目的而组合应用了多种隐私计算技术，可更直接用于实际生产。联邦学习技术方案主要应用于联合建模和预测场景中；多方安全计算和可信执行环境则可作为更加通用的技术方案，可设计用于联合统计、联合查询、联合建模及联合预测等诸多场景。

还需要指出的是，隐私计算技术体系还在快速发展中。以上划分只是一种业界常用的分类方法。目前各类技术也在互相融合，有望在更广泛的场景中发挥作用。

（二）多方安全计算基于密码学原理实现通用计算能力

1. 基本方案架构

多方安全计算（Secure Multi-party Computation，MPC）由图灵奖

获得者姚期智院士于 1982 年通过提出和解答百万富翁问题而创立，是指在无可信第三方的情况下，多个参与方共同计算一个目标函数，并且保证每一方仅获取自己的计算结果，无法通过计算过程中的交互数据推测出其他任意一方的输入数据（除非函数本身可以由自己的输入和获得的输出推测出其他参与方的输入）。

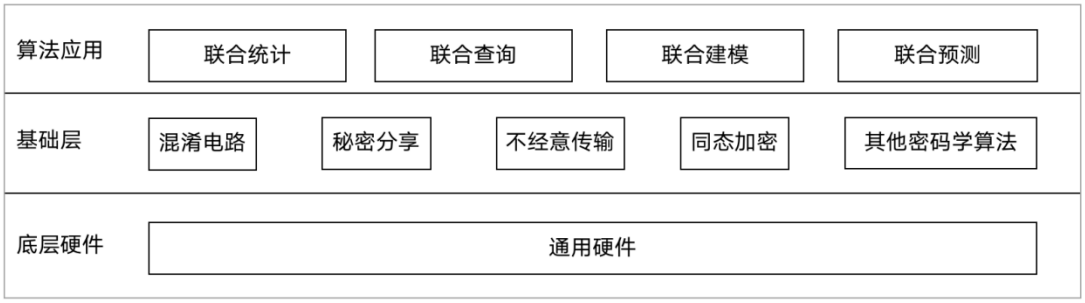

图 2 多方安全计算技术框架

如图 2 所示，**从底层硬件来说**，不同于可信执行环境基于可信硬件来保证数据的隐私计算，多方安全计算以通用硬件作为底层架构设计基于密码学的算法协议来实现隐私计算。

从算法构造来说，多方安全计算是多种密码学基础工具的综合应用，除混淆电路、秘密分享、不经意传输等密码学原理构造的经典多方安全计算协议外，其它所有用于实现多方安全计算的密码学算法都可以构成多方安全计算协议，因此在实现多方安全计算时也应用了同态加密、零知识证明等密码学算法（鉴于同态加密、零知识证明在隐私计算中的特殊地位，后面我们将单独叙述），有时也与可信执行环境等基于可信硬件的隐私计算技术结合提供安全加强的方案。

从算法应用来说，多方安全计算根据其可在各方不泄露输入数据的前提下完成多方协同分析、处理和结果发布这一技术特点，广泛应

用于联合统计、联合查询、联合建模、联合预测等场景，也可以支持用户自定义计算逻辑的通用计算需求。

2. 相关分析

从协议实现角度分析，在基于多方安全计算的隐私计算领域，被广泛应用的有混淆电路、秘密分享和不经意传输等基础密码学技术。

混淆电路（Garbled Circuit，GC）协议的思想起源于姚期智院士针对百万富翁问题提出的混淆电路解决方案，因此也被称为“姚氏电路”。混淆电路使用布尔电路构造安全函数计算，保证一方输入不会泄漏给其他方，计算出结果，并能指定结果由哪方获得或者是两方以分片形式共有。该技术可实现各种计算，常用于通用计算场景，通信量大但通信轮数固定，适用于高带宽高延迟场景。

秘密分享（Secret Sharing，SS）协议最早由 Shamir 和 Blakley 在 1979 年提出，是指将秘密信息拆分成若干分片，由若干参与者分别保存，并且通过参与者的合作，对分布式存储的各分片进行安全计算，全部分片或达到门限数的分片根据多个份额可重新恢复秘密信息。秘密分享计算量小、通信量较低，构造多方加法、乘法以及其他更复杂的运算有特别的优势，能实现联合统计、建模、预测等多种功能。

不经意传输（Oblivious Transfer，OT）协议由 Rabin 于 1981 年首次提出，指数据发送方有 n 个数据，数据接收方接收其选定的一个数据，且不能获取其他数据，同时数据发送方无法知道接收方的选择。不经意传输常用构造多方安全计算协议，是 GMW 协议、混淆电路设计、乘法三元组的基础构件，还可用于实现隐私集合求交（Private Set

Intersection, PSI）、隐私信息检索（Private Information Retrieval，PIR）等多种多方安全计算功能。

3. 技术特点

多方安全计算能够在不泄漏任何隐私数据的情况下让多方数据共同参与计算，然后获得准确的结果，可以使多个非互信主体在数据相互保密的前提下进行高效数据融合计算，达到“数据可用不可见”。最终实现数据的所有权和数据使用权相互分离，并控制数据的用途和用量，即某种程度上的“用途可控可计量”。多方安全计算具有很高的安全性，要求敏感的中间计算结果也不可以泄漏，并且在近 40 年的发展中其各种核心技术和构造方案不断接受学术界和工业界的检验，具有很高的可信性，其性能在各种研究中不断提升，现在在很多场景下已经达到了产业能实际应用接受的程度。

然而，多方安全计算也面临一些问题，例如：密码学复杂的运算过程造成的计算性能问题，不同技术间的加密数据不能互通造成的新的数据孤岛问题以及一些传统的安全问题等。这些问题都是制约多方安全计算发展的瓶颈。

（三）联邦学习变革机器学习范式广泛应用于联合建模

1. 基本方案架构

联邦学习（Federated Learning，FL），又名联邦机器学习、联合学习、联盟学习等。联邦学习是实现在本地原始数据不出库的情况下，通过对中间加密数据的流通与处理来完成多方联合的机器学习训练。联邦学习参与方一般包括数据方、算法方、协调方、计算方、结果方、

任务发起方等角色，根据参与计算的数据在数据方之间分布的情况不同，可以分为横向联邦学习、纵向联邦学习和联邦迁移学习。

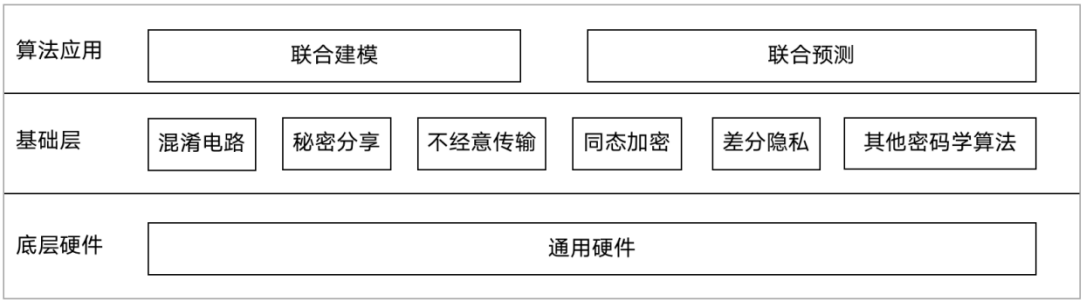

图 3 联邦学习技术框架

如图 3 所示，**从底层硬件来说**，区别于部署可信硬件的可信执行环境等技术，联邦学习一般以通用硬件作为底层基础设施。

从算法构造来说，常应用同态加密、差分隐私技术以及包括基于秘密分享、不经意传输、混淆电路等密码学原理的各类多方安全计算协议和其它用于保证隐私计算的密码学技术来提升安全性。

从算法应用角度来说，随着产业应用的需求，联邦学习框架也可与可信执行环境配合使用，提供安全性、应用性更强的综合解决方案。由于联邦学习是基于人工智能的技术工具，为提升用户隐私和数据安全前提下的联合 AI 模型训练效果而提出，因此广泛应用于联合建模、联合预测等场景中。

2. 相关分析

如何有效协调数据参与方协同构建模型是联邦学习的一项主要工作。因此，根据协调方式的不同，联邦学习从拓扑架构的角度分析，可分类为集中式拓扑架构和对等网络拓扑架构[1]**。**

[1] 引用文献：杨强等.联邦学习实战[M].北京:电子工业出社,2021.5:12.

对于集中式的拓扑结构，一般存在一个聚合各方本地模型参数信息的中心计算节点，该节点经过联邦平均等相应算法更新后，将结果返回各方。其中，中心计算节点既可能是独立于各参与方的第三方服务器，也可能是某一特定的参与方。它的优势在于易于设计与实现，往往被认为效率更高，但在一定程度上牺牲了安全性。

对于对等式网络拓扑结构，不存在中心计算节点，各参与方在联邦学习框架中的地位平等。相比在集中式的拓扑结构中需要考虑中心计算方存在泄露隐私或者遭受恶意攻击等的安全问题，分布式的网络拓扑结构安全性更高。但分布式拓扑需平等对待联邦学习中的每个参与方且能够使所有参与方有效更新模型并提升性能，设计难度较大。

3. 技术特点

联邦学习针对传统的由建模方（计算方）收集明文数据并进行人工智能模型训练存在的泄露训练数据隐私的问题而提出，通过对各参与方间的模型信息交换过程增加安全设计，使得构建的全局模型既能确保用户隐私和数据安全，又能充分利用多方数据，是解决数据孤岛和数据安全问题的重要框架，其强调的核心理念是“数据不动模型动，数据可用不可见”。

然而，联邦学习作为一门跨密码学、机器学习等领域的人工智能学科，其在应用过程中不可避免的会出现许多新的问题和挑战，例如：联邦学习过程中出现的数据和模型的隐私泄露和安全攻击如何防护；如何对非独立同分布、参差不齐的质量的数据建模；如何降低通信复杂度以及计算复杂度；如何评估各参与方的贡献，即联邦奖励机制问

题；如何实现不同联邦学习平台间的互联互通以及联邦学习的可解释性等等问题。这些仍需要进一步解决和完善。

（四）可信执行环境依托于可信硬件提供高效计算方案

1. 基本方案架构

可信执行环境（Trusted Execution Environment，TEE）通过软硬件方法在中央处理器中构建一个安全的区域，保证其内部加载的程序和数据在机密性和完整性上得到保护。TEE 是一个隔离的执行环境，为在设备上运行的受信任应用程序提供了比普通操作系统（Rich Operating System，RichOS）更高级别的安全性以及比安全元件（Secure Element，SE）更多的功能。

目前主要的通用计算芯片厂商发布的 TEE 技术方案包括 X86 指令集架构的 Intel SGX（Intel Software Guard Extensions）技术、AMD SEV（Secure Encrypted Virtualization）技术以及高级 RISC 机器（Advanced RISC Machine，ARM）指令集架构的 TrustZone 技术。而国内计算芯片厂商推出的 TEE 功能则包括兆芯 ZX-TCT（Trusted Computing Technology）技术、海光 CSV（China Security Virtualization）技术，以及 ARM 架构的飞腾、鲲鹏也已推出自主实现的 TrustZone 功能。

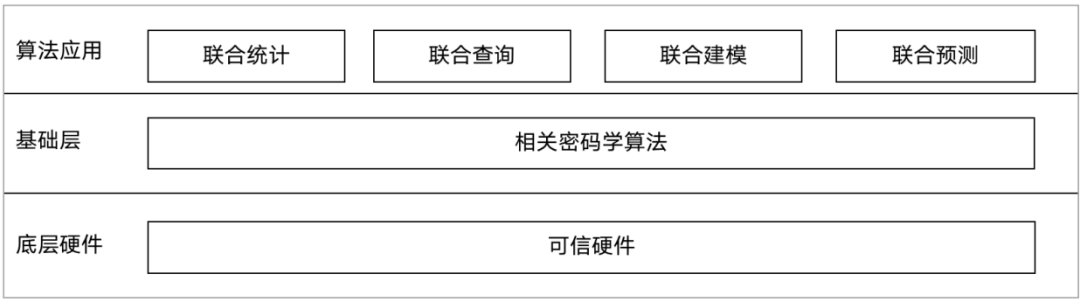

图 4 可信执行环境技术体系

如图 4 所示，**从底层硬件来说，**不同于多方安全计算和联邦学习，TEE 将多方数据集中到可信硬件构建的可信执行环境中一起进行安全计算。TEE 中可信硬件一般是指可信执行控制单元已被预置集成的商用 CPU 计算芯片[2]。**从基础算法来说，**为了保证传输至可信环境中的数据的安全性，TEE 常结合相关密码学算法来实现加密和验证方案[3]。**从应用角度来说，**作为通用的计算平台，TEE 可以在可信执行环境中对多方数据完成联合统计、联合查询、联合建模及预测等各种安全计算。

2. 相关分析硬件实现

目前主流的 TEE 技术以 X86 指令集架构的 Intel SGX 技术和 ARM 指令集架构的 TrustZone 技术为代表。

Intel SGX 技术是一组预置在 Intel 商用计算芯片内的用于增强应用程序代码和数据安全性的指令，主要面向 PC 端。开发者使用 SGX 指令把计算应用程序的安全计算过程封装在一个被称为飞地（Enclave）的容器内，保障用户关键代码和数据的机密性和完整性。

[2] 芯片设计厂商除提供通用指令集外，针对 TEE 单元会提供单独的 TEE 指令集用于驱动 TEE 设备。
[3] 虽然标准定义可以通过软件方式或硬件方式实现 TEE，但实际生产场景下，行业内更多通过软硬结合的方式进行安全性的保障与支持。

Intel SGX 将应用程序以外的软件栈（如 OS、BIOS 等）都排除在可信计算基（Trusted Computing Base，TCB）以外，一旦软件和数据位于 Encalve 中，即便是操作系统和虚拟机监视器（Virtual Machine Monitor，VMM）（也称 Hypervisor）也无法影响 Enclave 里面的代码和数据，从而在安全隔离的情况下保证软件功能的通用性。

ARM TrustZone 技术基于 ARM 芯片，主要面向移动设备，是用于 ARM 指令集体系结构的 TEE。ARM 通过对原有硬件架构进行修改，在处理器层次引入了两个不同权限的保护域——安全世界和普通世界，任何时刻处理器仅在其中的一个环境内运行。TrustZone 通过中断路由以及对内存总线和内存管理单元的限制来提供隔离保护。

3. 技术特点

TEE 通过隔离的执行环境，提供一个执行空间，该空间有更强的安全性，比安全芯片功能更丰富，提供代码和数据的机密性和完整性保护。另外，与纯软件的密码学隐私保护方案相比，TEE 不会对隐私区域内的算法逻辑语言有可计算性方面的限制，支持更多的算子及复杂算法，上层业务表达性更强。利用 TEE 提供的计算度量功能，还可实现运行在其内部的身份、数据、算法全流程的计算一致性证明[4]。

TEE 因支持多层次、高复杂度的算法逻辑实现，运算效率高以及可信度量保证运行逻辑可信等的特点，被广泛认可，但其技术本身依

[4] 基于可信度量方式，单个 TEE 实例内可以整合封装身份签名逻辑、数据 Hash 逻辑与计算逻辑，可提供身份、数据、算法三者关联的一致性证明。

赖硬件环境，CPU 相关实现属于 TCB，由芯片设备的设计生产厂商提供，必须确保芯片厂商可信。此外使用 MPC 等密码学技术与 TEE 技术相结合可以增强其安全性，强化 TEE 实例之间机密通信和组网的安全性，进一步防止隐私数据泄露。

（五）相关技术扩充隐私计算技术体系

除了上述关键技术，同态加密、零知识证明、差分隐私、区块链等技术也常应用或辅助于隐私计算。

同态加密（Homomorphic Encryption，HE），能实现在密文上进行计算后对输出进行解密，得到的结果和直接对明文计算的结果一致。该概念最早在 1978 年由 Ron Rivest、Leonard Adleman 和 Michael L. Dertouzo 提出，已发展出各种半同态加密和全同态加密算法。同态加密算法以通信量小、通信轮数少的特点，已在多方安全计算、联邦学习、区块链等存在数据隐私计算需求的场景落地应用。

零知识证明（Zero-Knowledge Proof，ZKP），由 S.Goldwasser、S.Micali 及 C.Rackoff 在 20 世纪 80 年代初首先提出，指的是证明者能够在不向验证者泄漏任何有用信息的情况下，使验证者相信某个论断是正确的。零知识证明是一种两方或多方的协议，两方或多方通过一系列交互完成生成证明和验证。在实际应用中，零知识证明能实现证明者向验证者证明并使其相信自己知道或拥有某一消息，而证明过程不会向验证者泄漏任何关于被证明消息的信息。

差分隐私（Differential Privacy，DP）技术是 Dwork 在 2006 年针对数据库的隐私泄露问题提出的一种新型密码学手段。该机制是

在源数据或计算结果上添加特定分布的噪音，确保各参与方无法通过得到的数据分析出数据集中是否包含某一特定实体。差分隐私包括本地差分隐私和计算结果差分隐私。本地差分隐私指在汇聚和计算前数据就加入噪声，用于数据收集方不可信的场景；计算结果差分隐私是指最终计算结果发布前对其加噪声。

隐私计算最核心的是计算，但整个过程还有完整的系统需要用的辅助技术很多，主要有区块链和证书授权中心（Certificate Authority，CA）等。一方面，区块链隐私计算框架能在数据共享过程中有效保护个人信息，并为数据真实性、数据确权等问题提供可行解决方案，实现全流程可记录、可验证、可追溯、可审计的安全、可信数据共享网络，并为进一步建设高效、高安全和高流动性的数据要素交易市场打下基础。另一方面，隐私计算过程中的每一方都需具有相同信任根的证书链，各参与方之间通信使用证书链建立 SSL 安全通道，认证授权隐私计算框架实现参与方间的双向认证，从而确保参与方身份真实准确，实现对隐私计算任务定向授权，验证后执行任务。

下表对隐私计算相关的技术进行了主要对比。

表 2：隐私计算相关技术主要对比

技术	性能	通用性	安全性	可信方	整体描述	技术成熟度[5]
多方安全计算（MPC）	低~中	高	高	不需要	通用性高、计算和通信开销大、安全性高，研究时间长，久经考验，性能不断提升	**已达到技术成熟的预期峰值**
可信执行环境（TEE）	高	高	中~高	需要	通用性高，性能强，开发和部署难度大，需要信任硬件厂商	**快速增长的技术创新阶段**
联邦学习（FL）	中	中	中	均可	综合运用 MPC、DP、HE 方法，主要用于 AI 模型训练和预测	**快速增长的技术创新阶段**
同态加密（HE）	低	中	高	不需要	计算开销大，通信开销小，安全性高，可用于联邦学习安全聚合、构造 MPC 协议	**快速增长的技术创新阶段**
零知识证明（ZKP）	低	低	高	不需要	广泛应用于各类安全协议设计，是各类认证协议的基础	**快速增长的技术创新阶段**
差分隐私（DP）	高	低	中	不需要	计算和通信性能与直接明文计算几乎无区别，安全性损失依赖于噪声大小	**快速增长的技术创新阶段**
区块链（BC）	低	中	中	不需要	基于带时间戳的块链式存储、智能合约、分布式共识等技术辅助隐私计算，保证原始数据、计算过程及结果可验证	**逐渐接近技术成熟的预期峰值**

[5] 来源：中国信通院调研、Gartner

第三章

隐私计算主要应用场景

根据中国信通院统计，目前典型的应用场景包括联合风控、联合营销、智能医疗、智能政务等热点应用，也包括智慧能源、智慧城市、工业互联网等探索性应用。

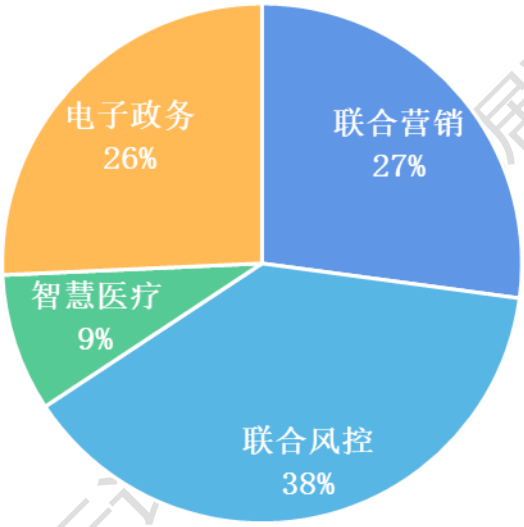

图 5 隐私计算应用行业

当前隐私计算应用主要集中在数据驱动的金融、互联网领域和拥有大量数据源和数据流通需求的医疗、政务领域，同时跨机构、跨行业应用需求强烈，目前最主要的应用集中在联合营销、联合风控、智慧医疗、电子政务等场景。

（一）联合营销：跨行业数据融合重构用户画像

当前营销业务进入到智能时代，应用于营销的数据维度不断丰富，应用场景也不断增加。然而，用户画像的数据往往是相互割裂的，只有通过整合多机构间多维度的数据才能构建更立体的用户画像，实现

资源的优势互补、开拓市场广度和挖掘市场深度的营销目的。利用隐私计算可以帮助机构在不输出原始数据的基础上共享各自的用户数据进行营销模型计算，根据建模结果制订营销策略，实现双赢的联合营销目的。

在构建营销模型中，可通过隐私计算技术，对交互的标签、特征、梯度等数据进行等密码学处理，保证密文接收方或外部第三方无法恢复明文，直接基于密文进行计算并获得正确的计算结果，从而达到各参与方无需共享数据资源即可实现联合构建营销模型，进一步丰富用户画像，从而进行精准营销。

在高价值用户识别中，可以利用隐私计算技术，通过联合统计、隐匿查询等方式将内部和外部数据进行安全融合，打通多方数据孤岛，利用外部数据更精准的对用户客群进行分类，识别高价值用户，制定更精准的营销策略。

银行机构利用隐私计算技术，可对运营商、政务、征信等数据实现应用场景所需的价值融合，从而为用户提供聚合金融服务。**保险公司**将用户基本信息、购买保险、出险赔付和电商、航旅等其他合作方的消费、出行、行为偏好等数据进行安全融合。通过匿踪查询技术可信地获取客户的黑名单、消费能力、画像标签等信息，用于识别消费者的潜在风险等应用。**电信运营商**通过融合金融机构数据在共有的用户群中找到对理财产品、保险产品有兴趣的用户群，筛选找到更精准的目标用户进行营销，提升交叉销售效果，获取更多的新客。**互联网公司**利用自身拥有的大量用户行为信息和基础画像数据，与广告数据

方拥有的深度转化链路数据（如付费信息）进行安全求交，并通过多方安全计算或联邦学习技术联合训练、建模、优化广告模型效果。在游戏、金融、教育、电商行业的广告应用案例中都能提升广告投放效果和用户体验。

（二）联合风控：引入外部数据优化金融风控模型

联合风控是隐私计算在金融领域的一个重要应用场景。一般而言，用户在本机构的金融业务数据难以满足金融风控的需求，但由于不同机构间数据分散、数据保护等原因，金融机构之间、金融机构与其他行业机构之间的数据融合壁垒较高，“数据孤岛”现象严重，提升了金融机构的风险识别难度，难以降低融资成本。

利用隐私计算技术，可以实现跨机构间数据价值的联合挖掘，更好地分析客户的综合情况，交叉验证交易真实性等业务背景，降低欺诈及合规风险，从而综合提升风控能力。

在构建风控模型时，一方面可通过融合多个金融机构数据，解决单个金融机构样本量有限的问题，形成在相关场景中的全局认知，提升模型精准度；另一方面，可以综合利用金融机构同其他行业数据，在各方原始特征不出域的前提下建立风控模型，形成对业务的多维度认识，提升风控质量。

在信息核验时，可通过隐私计算实现多方黑名单数据共享，对电诈、洗钱、骗贷等行为的黑名单用户进行匿踪识别，数据方不能获知查询的具体内容，提升客户背景调查的安全可信程度。

（三）智慧医疗：数据互通发挥医学数据价值

医学研究、基因分析等工作非常依赖大量数据的积累，然而，医疗相关机构的这些数据割裂，离散在不同机构及业务系统内，机构间的数据难以互通互联，严重制约了临床科研成果的产出。

在智慧医疗领域，利用隐私计算技术，可实现在数据隐私保护下医学数据安全统计分析和医学模拟仿真和预判，从而进行跨机构的精准防疫、基因分析、临床医学研究等应用。

在疫情防控中，通过隐私计算保障个人数据的安全性，对高危人群进行筛选疫情传播仿真分析，通过防控筛查模型精准筛查高风险易感人群，构建潜在传染的关系网结合病患信息，快速追溯传染路径方向和传播源。

在基因分析中，要依赖大量隐私数据，可利用隐私计算技术在原始基因数据不出库的基础上，实现基因数据的安全共享，进行全基因组的联合计算及关联分析等，在隐私保护的前提下挖掘多样的基因资源。

在临床医学研究中，在数据不出本地的情况下，可以实现分布式的统计分析算法，对数据进行联合建模、分析，从而获得临床科研的成果，例如临床研究可行性分析、大样本量队列研究、疾病预测模型、药物市场洞察等。应用隐私计算，将会大幅提升医疗研究效率，加速科研成果转化。

（四）电子政务：促进政务数据安全共享开放

隐私计算技术为政务数据的开放提供了有效解决方案。在企业自

有数据、第三方数据或政府共享数据都需要保护且不能离开本地节点的场景下，基于隐私计算进行数据安全利用。

在政务数据共享上，政务公共数据分布在各部门，通过隐私计算技术搭建政务公共数据密文开放共享交换平台，打通跨域数据的应用价值链，使得数据基于业务应用需要在各业务条线之间，安全地共享和流通，实现数据安全共享融合而不泄密。

在政务数据开放上，政府机构建设保护各方隐私安全的公共数据开放平台，使用隐私计算技术融合政府数据和社会、企业数据进行安全计算，联合统计，联合建模，实现数据融合价值，可以广泛应用在信用评估、服务选址、健康医疗、家政服务、旅游投资、营销设计等众多领域，让政府部门掌握的数据在安全保护前提下，最大限度造福社会。

除了上述主要集中应用场景之外，隐私计算技术应用也呈现出向更多行业扩散的态势，在智慧能源、智慧终端、智慧城市等更多场景均有探索性应用。

第四章

隐私计算产业发展态势

作为数据流通的关键技术，隐私计算发展火热，备受政策关注和市场瞩目。自 2018 年开始，国内隐私计算进入快速启动期，投入研究和发布相关产品企业数量激增，当前市场一片火热。在最早一批互联网大厂和专精型创业团队之后，许多大数据、AI、泛区块链、金融科技和传统数据安全企业开始转型，纷纷入局隐私计算。但总体来看，隐私计算发展的商业环境尚未成熟，市场整体还有很大的成长空间。

（一） 隐私计算市场发展迅速

全球范围内看，国内国外布局隐私计算的大型企业和创业团队都有很多，但市场仍处于蓬勃发展的早期阶段，竞争格局尚未确定，更多典型和杀手级应用还有待拓展，仍然有孵化创新独角兽的机会。

1. 国外：研究活跃，但商业化形态较为局限

从隐私计算本身的发展历程来看，国外企业布局隐私计算较早。早在 2008 年第一家专攻多方安全计算解决方案的技术厂商 Partisia 就已在丹麦成立，为商务合同、加密拍卖等场景提供安全方案。

科技巨头中，微软从 2011 年开始深入研究多方安全计算、谷歌在全球率先提出联邦学习的概念、Intel 打造 SGX 成为绝大部分可信执行环境实现方案的底座，均已成为各条技术路线主要的领路人。其他如 IBM 致力于将同态加密与云服务结合，帮助用户数据安全上云；

Facebook 则是专攻基于隐私计算的机器学习。

创业公司中，Sharemind、Privitar 致力于搭建自研的多方安全计算平台；Duality 基于密码学开发的 SecurePlus 平台在新冠疫情中支撑了医学机构进行病毒基因分析。此外，AI 公司 Zama、区块链公司 Enigma 等均在推进多方安全计算、同态加密等方向的技术研发。

但从总体的应用场景来看，目前国外隐私计算项目中的很大一部分都是面向区块链和加密虚拟货币的场景。如美国的 Unbound Tech 和丹麦的 Sepior 均集中于将多方安全计算应用于分布式密钥管理领域。

2. 国内：百花齐放，众多厂商加入竞争赛道

跟国外相比，我国企业开始布局隐私计算的时间要更晚，大致在 2016 年之后才开始出现独立的隐私计算商业项目，但国内产业化发展的速度较快。伴随着各行业企业对合规数据流通的需求日益强烈，越来越多的行业客户开始愿意进行尝试，整体行业从概念验证到全面实施趋势明显。根据调研，目前超过 81%的隐私计算产品进入了试点部署或实施阶段[6]。

[6] 数据来源：中国信通院根据调研整理

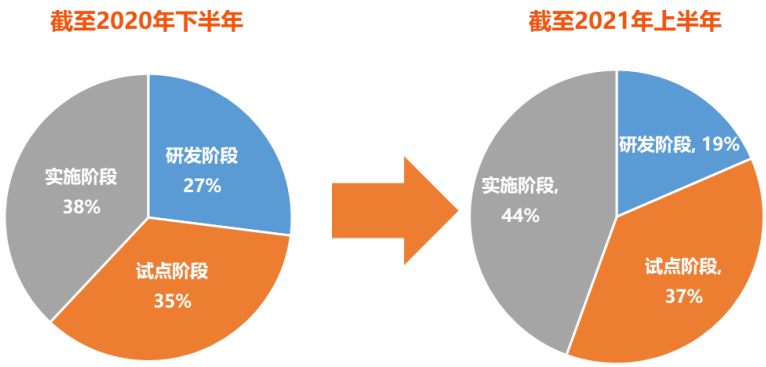

图 6 国内隐私计算平台应用情况

从技术路线上看，多方安全计算的复杂度高、开发难度大，龙头企业多致力于此，力图打造以多方安全计算为底座的数据流通基础设施，26%的企业布局了这类技术方案[7]；可信执行环境对于硬件的局限及国外芯片的强依赖，使得其在国内的产品选型相对较少，提供此类方案的企业占比约为 21%，较集中于互联网大厂和部分初创企业，但目前已出现一些技术企业与芯片企业在国产化硬件研发上的合作探索；对于联邦学习，由于机器学习类应用需求的突出，且有较成熟的开源社区为基础，开发难度相对轻松，因而，运营商、金融科技公司等自营业务需求方大多专注在基于联邦学习的隐私计算产品化中，提供联邦学习方案的企业数量占比约为 52%。此外，由于各类技术方案各有优势，面对用户的不同应用需求，21%的企业提供多种技术方案供用户选择。

[7] 中国信通院调研统计结果。

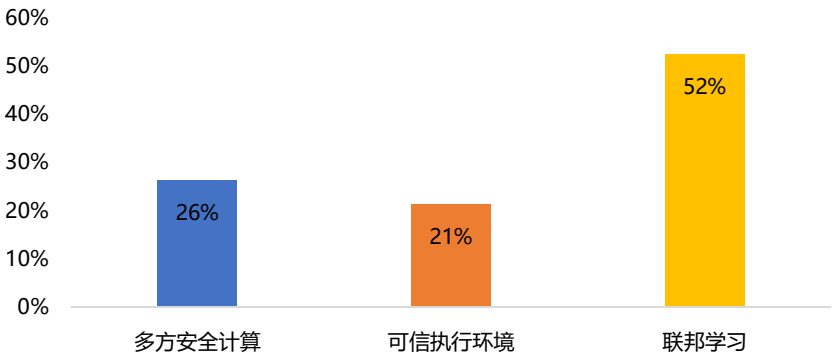

图 7 国内隐私计算平台技术路线

（二） 产业发展配套环境正在逐步完善

从产业发展的配套环境来看，相关企业和机构在学术研究、开源生态、标准体系等方面积极探索，推动着隐私计算向上蓬勃发展，但现阶段，技术大规模推广所需的成熟商业模式尚未形成，仍然需要各方力量共同努力。

1. 学术研究领域关注度持续提升

作为一门融合了多学科的新兴技术，发展与应用隐私计算对于技术理论的研究有较强依赖。2011 年以来的十年间，隐私计算领域共发表论文 5280 篇[8]，论文数量始终保持着不低于 10%的增速逐年上升。其中 2019 年数量上升幅度较大，目前隐私计算领域每年发文量稳步超越 1000 篇。可见理论研究层面对隐私计算的关注在持续增强。

[8] 数据来源：中国信通院基于 Web of Science 核心合集统计整理，检索关键词包括：隐私保护计算（Privacy-Preserving Computation 和 Privacy-preserving computing）、多方安全计算（Secure Multi-party Computation）、可信执行环境（Trusted Execution Environment）、联邦学习（Federated Learning）、机密计算（Confidential Computing）。

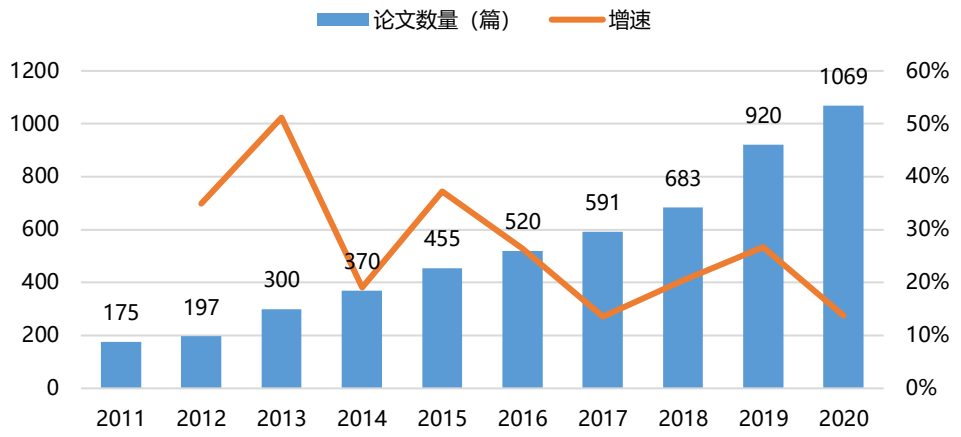

图 8 隐私计算领域论文发表数量

从发表论文的国家和地区来看，中国和美国的论文数量最多的，分别占了总数的 34.8%和 27.6%。此外，从研究机构来看，国内西安电子科技大学、中国科学院、电子科技大学、北京邮电大学、上海交通大学、武汉大学这些深耕在信息技术安全领域的高校已在隐私计算领域有了比较多的研究成果。

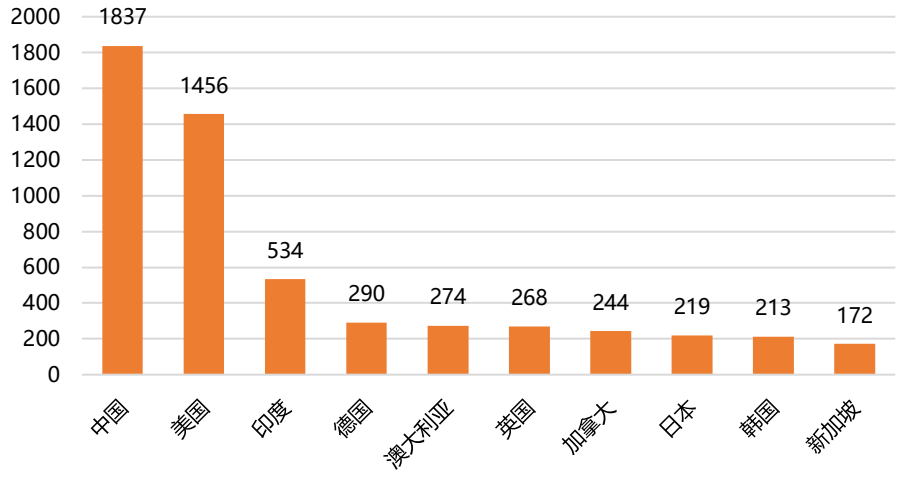

图 9 各国隐私计算领域论文发表数量

2. 技术企业积极拥抱开源生态

开源社区的知识共享和多方协同有利于加速技术升级和商业化

项目落地。近两年，国内外很多大厂和创业团队都在积极开源，表 3 归纳了目前国内外隐私计算领域的主要开源项目情况。

表 3：目前主要的隐私计算开源项目

序号	项目名	开源时间	机构	技术路径
1	PySyft	2017 年 7 月	OpenMined 开源社区	多方安全计算、联邦学习
2	TF-Encrypted	2018 年 3 月	DropoutLabs, Openmined, 阿里巴巴	多方安全计算
3	EzPC	2018 年 4 月	微软	多方安全计算
4	Asylo	2018 年 5 月	谷歌	可信执行环境
5	MesaTEE	2018 年 9 月	百度	可信执行环境
6	FATE	2019 年 2 月	微众银行	联邦学习
7	TF-Federated	2019 年 8 月	谷歌	联邦学习
8	Private Join & Compute	2019 年 8 月	谷歌	多方安全计算
9	PaddleFL	2019 年 9 月	百度	联邦学习
10	CrypTen	2019 年 10 月	Facebook	多方安全计算
11	Fedlearner	2020 年 1 月	字节跳动	联邦学习
12	Rosetta	2020 年 8 月	矩阵元	多方安全计算
13	KubeTEE	2020 年 9 月	蚂蚁集团	可信执行环境

从开源项目的活跃度和影响力来看，联邦学习的开源生态为工业化的落地应用贡献了强劲力量，特别是 FATE，2020 年及之后出现的很多联邦学习类产品都或多或少的吸收和借鉴了 FATE 供给的营养。在中国信通院调研统计中，55%的国内隐私计算产品是基于或参考开源项目开发的[9]，这其中开源项目就以 FATE 为主。

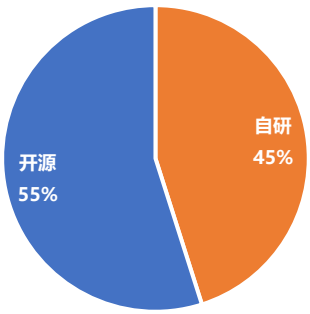

图 10 国内隐私计算平台自研情况

[9] 根据中国信通院云大所隐私计算测试调研

3. 配套技术标准体系日渐完善

技术的最终使命是业务应用，作为技术的载体，产品如何构建、如何落地将对用户的业务形态产生重要影响，这就需要统一规范的技术标准，为产品的可用、易用划出基线。目前，从 IEEE、ISO、ITU-T 等国际组织到中国通信标准化协会（CCSA）、全国金融标准化技术委员会（金标委）等均在组织力量制定或发布隐私计算相关的技术标准。

国际标准内容以框架和功能为主。除了早期 ISO 基于传统信息安全视角针对秘密分享和同态加密制定的国际标准外，各国际组织在隐私计算领域的标准工作从 2018 年起步，现有标准内容已经覆盖了多方安全计算、可信执行环境和联邦学习三条技术路线，但目前已完成并成功发布的仅有 IEEE P3652.1《联邦学习架构框架与应用指南》。

表 4：隐私计算相关国际标准情况

组织	标准名称	状态	发起单位
IEEE	P3652.1《Guide for Architectural Framework and Application of Federated Machine Learning》（联邦学习架构框架与应用指南）	2018 年立项 2021 年发布	微众银行
	P2842《Recommended Practice for Secure Multi-Party Computation》（多方安全计算参考框架）	2019 年立项	阿里巴巴
	P2830《Standard for Technical Framework and Requirements of Shared Machine Learning》（共享学习系统技术框架及要求）	2019 年立项	蚂蚁集团
	P2952《Standard for Secure Computing Based on Trusted Execution Environment》（基于可信执行环境的安全计算）	2020 年立项	蚂蚁集团
ISO/IEC JTC1 SC27	ISO/IEC 19592-1 《Information technology - Security techniques - Secret sharing》(信息技术-安全技术-秘密分享)	2016 年发布	——
	ISO/IEC 18033-6 《Information technology - Security techniques - Encryption algorithms - Part 6：Homomorphic encryption》（信息技术-安全技术-加密算法-同态加密）	2019 年发布	——
	ISO/IEC 4922-1《Information security - Secure multiparty computation - Part 1: General》（信息安全-多方安全计算-第 1 部分：通用）	2020 年立项	德国标准化学会

	ISO/IEC 4922-2《Information security - Secure multiparty computation - Part 2: Mechanisms based on secret sharing》（信息安全-多方安全计算-第 2 部分：基于秘密分享）	2020 年立项	德国标准化学会
ITU-T SG16	《Technical Framework for Shared Machine Learning System》（共享学习系统技术框架）	2019 年立项	蚂蚁集团
ITU-T SG17	《Technical Framework for Multi-Party Computation》（多方安全计算技术框架）	2019 年立项	阿里巴巴

国内标准开始向性能和安全扩展。相比国际标准，国内隐私计算相关标准的制定和迭代更加高效，且已经开始从基础的功能标准向产品性能、安全性等方向拓展，加速构建更加完善的隐私计算技术标准体系。中国通信标准化协会大数据技术标准推进委员会（CCSA TC601）自 2018 年开始制定隐私计算领域的相关团体标准，由中国信通院云大所牵头，目前已完成多方安全计算、可信执行环境、联邦学习和区块链辅助的隐私计算技术工具四项针对产品基础功能的标准，正在加速进行产品性能和安全性标准的制定。同时针对解决不同厂商提供的产品间的技术壁垒，实现跨平台间协作的互联互通系列标准也已提交行标立项，正在快速制定中。

表 5：中国通信标准化协会隐私计算相关标准情况

	标准名称	标准类别	进展
功能标准	基于多方安全计算的数据流通产品 技术要求与测试方法	团体标准	已完成
	基于联邦学习的数据流通产品 技术要求与测试方法	团体标准	已完成
	基于可信执行环境的数据计算平台 技术要求与测试方法	团体标准	已完成
	区块链辅助的隐私计算技术工具 技术要求与测试方法	团体标准	已完成
性能标准	隐私计算 多方安全计算产品性能要求和测试方法	行业标准	已完成
	隐私计算 联邦学习产品性能要求和测试方法	行业标准	已完成
	隐私计算 可信执行环境产品性能要求和测试方法	行业标准	制定中

安全性	隐私计算 多方安全计算产品安全要求和测试方法	行业标准	制定中
	隐私计算 联邦学习产品安全要求和测试方法	行业标准	制定中
	隐私计算 可信执行环境产品安全要求和测试方法	行业标准	制定中
互联互通	隐私计算 跨平台互联互通（系列标准）	行业标准	已提交立项

全国信息安全标准化技术委员会 TC260、中国人工智能产业发展联盟 AIIA、中国人工智能开源软件产业发展联盟 AIOSS 等组织也针对隐私计算编写或发布了相关标准。此外，2020 年 11 月，由金标委归口，中国人民银行科技司提出并负责起草的《多方安全计算金融应用技术规范》（JR/T 0196—2020）金融行业标准正式发布，标准主要规范了如何在金融场景部署和应用以多方安全计算。2021 年 7 月，中国支付清算协会发布《多方安全计算金融应用评估规范》（团体标准）。这些标准都是在技术本身的基础上，对隐私计算如何在金融行业应用进行了拓展。

4. 成熟商业模式还需继续孵化

从商业模式上看，隐私计算厂商的收入主要可以分为两个部分。一是提供技术应用的系统平台或解决方案，类似于传统的软件销售和服务收入，提供平台部署、调试和后续配套运维的服务，同时，在面对不同的企业或机构用户亦会有一定程度的个性化和定制化需求；二是基于隐私计算平台上的数据流通活动产生利润分成，此时技术厂商将更加深入到数据流通服务的角色中，与数据流通平台的运营者共享收益，运营商、传统大数据转型企业和金融科技公司等掌握大量数据资源的隐私计算技术厂商将在这种模式下发挥出优势。

但是，目前隐私计算行业整体处于规模商用的前期。2019 年主要集中于技术科普和市场教育，2020 年开始出现较多的产品 POC，直到近期才开始出现规模化的招标与应用。产业发展依赖成熟的商业环境尚未形成，因此仅有少量项目可以全部确认收入，且均来自于提供独立的平台产品和解决方案。但根据 KPMG《隐私计算行业研究报告》预测，接下来几年国内隐私计算市场规模将快速发展，三年后技术服务营收有望触达 100-200 亿人民币的空间，甚至将撬动千亿级的数据平台运营收入空间。

第五章

隐私计算合规探讨

在人工智能联合建模、金融行业联合风控、电商行业精准营销对数据融合的需求日益旺盛的同时，各国对数据隐私保护的法规规制和监管力度也日趋收紧。在这样的大背景下，隐私计算由于其降低了参与方的授权风险、实现数据“可用不可见”、增强参与方对数据流通的安全控制和遵循个人数据最小必要原则等特性，有助于在满足合规和监管要求的前提下实现数据的互联互通，促进数据产业的发展。

（一）隐私计算有助于提升数据流通的合规性

对处理个人数据而言，目前我国个人信息处理的合规路径主要是匿名化和授权同意。对匿名化而言，隐私计算由于可以控制原始数据不出本地，只输出切片、标签化、脱密后的梯度和参数等信息，从而成为满足匿名化“不可识别、不可复原”方案的重要一环，同时也因此与“最小必要原则”相一致。对授权同意而言，如果对输入模型的数据进行的脱敏、加密处理满足了匿名化要求，此后联合建模的其他参与方也可能减轻获得用户重复授权带来的负担。

对非个人数据而言，隐私计算原始数据不出域、只传递梯度等数据的特质也有助于满足《数据安全法》和《网络安全法》等要求的安全保护义务。同时也有助于控制数据泄露的风险，进而减轻数据主体的顾虑，充分挖掘数据的流通价值。

（二）隐私计算方案设计需要关注合规要求

与传统数据流通方案相比，隐私计算的数据保护效果更有利于满足合规要求。同时，我们也需要充分审视技术方案的合规风险，寻求技术和制度合规的方向。

1.原始数据的合规瑕疵仍需关注

就匿名化而言，当存在反向推演出原始数据等可能性时，隐私计算将可能无法满足“不可识别不可复原”的要求。另外，将未经授权的数据输入隐私计算模型也可能为其他参与方增加风险。

在实践中为了规避合规风险，参与方仍有动力在采用匿名化、去标识化之外获取用户授权同意。然而获取个人完全符合法律要求的授权同意难度极高，例如联合建模的数据使用目的很难被涵盖在授权协议中，二次获取用户授权则是实践中普遍的难题，超范围数据使用的风险也可能会在数据经过多手流转之后愈发扩大。

2. 模型泄露风险有待控制

为了构建联合模型，参与方仍然需要上传或共享模型参数或梯度信息。已有研究证明[10]，隐私计算存在因衍生信息、梯度等模型信息泄露而推知原始数据的可能性。当发生数据泄露或存在安全隐患时，隐私计算可能无法满足匿名化的要求，相关参与方也可能因未充分履行数据安全保护义务而承担法律责任。

3. 参与方的安全隐患需要排除

参与方可能会承担生成和发放公私钥、加解密结果等任务。但某

[10] Zhu L, Liu Z, Han S. Deep Leakage from Gradients[C] // WALLACH H M, LAROCHELLE H, BEYGELZIMER A, et al. NeurIPS 2019. 2019 : 14747 – 14756.

些参与方可能会违约获取额外信息，可能会暴露其他方的数据隐私；部分参与方可能恶意合谋获取其他参与方的数据等等。

4. 输出结果需要确认不再敏感

在输出最终计算结果时，各参与方也需要确保输出的结果不存在隐私风险。例如在金融机构和征信机构预测借款人信用的场景，如果在输出预测结果时泄露了借款人的 ID，则有可能泄露借款人本身的借款需求。

（三）隐私计算合规实践路径的探索

尽管隐私计算的合规红线仍不明确，我们仍建议相关企业厘清技术方案和管理制度的风险点，在现有域内外法律、标准和行业最佳实践的基础上，探索平衡合规、效率和可用性要求的合规实践路径。

1. 搭建合规基准框架和管理制度

企业可依照业务适用的法律法规、标准和行业最佳实践基础搭建合规框架。在建立制度后，需通过相应的技术工具把制度落实到业务流程中，从而减少合规风险及其带来的潜在损失。

2. 根据数据类型选择合规基础

企业可根据隐私计算的使用场景、输入模型的数据类型和技术手段，选择匿名化、基于同意或二者结合的合规基础。例如在获取个人充分、明确的同意成本过高、二次征得个人同意的难度过大、一手数据源存在合规瑕疵等情况下，建议企业考虑匿名化和授权同意相结合的方法，尽可能覆盖数据处理各流程的风险。

3. 控制参与方带来的风险

建议相关参与方在考虑其他参与方的可问责性、恶意参与风险和泄露风险等因素的基础上，通过合同尽可能明确各参与方在各环节、各场景下的合规义务和责任承担。例如在共享去标识化数据时，共享方可通过协议明确数据去标识处理的具体方法和义务承担主体，禁止信息接收方发起对数据集中个体的重标识攻击和将数据关联到外部数据集，禁止参与方未经许可共享数据集等等，并为每种可能遇到的情形约定相应的保证金和违约责任等。

4. 对计算的过程和结果的合规性进行证明和存证

以多方安全计算为例，系统中每个参与方除了各自的预期输出之外，无法获得其他额外的信息。因此，多方安全计算的技术方案至少需要能够确保和说明参与方无法通过其他参与方的输入来构建自己的计算输入，也无法从当前参与方的输入结果推导出原始数据。

在能对技术方案的安全性和合规性进行证明后，技术厂商和使用隐私计算产品的参与方也需要通过日志等形式对数据的输入输出、安全系统阻断非法请求的证据等进行记录，从而满足监管或外部认证的合规要求，并为可能的诉讼或其他法律行动提供证据。

第六章

隐私计算的挑战和难题

目前隐私计算技术正处于快速迭代和发展的阶段，可解决企业和机构当前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。但在隐私计算安全、性能和数据的互联互通等方面仍存在挑战，这些难题在一定程度上限制了隐私计算的推广和应用。

（一） 安全性挑战影响市场信任

隐私计算产品与其他的数据处理产品不同，其本身肩负着保护隐私数据安全的重要功能，技术服务厂商与产品使用者都应当谨慎对待隐私计算产品的安全性挑战，而算法协议安全、开发应用安全和安全共识正成为当前隐私计算推广应用亟需面临的挑战。

算法协议尚无法实现绝对安全。一方面，隐私计算产品的算法协议差异化较大，难以形成统一的算法安全基础。隐私计算产品所使用的算法协议多种多样，各自的协议安全根基也不相同：多方安全计算、同态加密等密码学算法基于数学与密码学基础；联邦学习等隐私机器学习安全基于机器学习理论、差分隐私和相关密码学协议；可信执行环境则更多依赖于硬件厂商的安全技术。另一方面，隐私计算产品安全协议依赖安全假设，仍存在安全风险。隐私计算产品的安全基础通常都会设定安全假设，以此为基础进行协议和算法的设计，比如假设硬件提供商的可信性、假设计算参与方会遵循协议流程、假设多个参

与方之间互不共谋等。但实际中这种假设并不一定完全成立，往往都需要通过博弈论、现实约束等方法进行加强。

开发应用安全同样存在挑战。在假定算法协议安全达成的情况下，一方面隐私计算产品面临生产化过程中产生的安全问题，例如密码学算法通常遇到的侧信道攻击、错误注入攻击，硬件通常遇到的侵入式攻击，或者类似其他信息系统遇到的恶意黑客攻击。由于隐私计算产品的安全要求较高，其要求整个计算全过程安全，那么木桶效应会导致最薄弱的环节成为整个产品的最易被攻击部分。另一方面，第三方机构的介入也会引来安全风险。在实际使用中，诸如证书管理中心，通信与授权的协调节点等任何第三方机构的介入，都会打破技术信任的完整性，引入不确定的风险因子。

安全性共识有待形成。隐私计算的核心逻辑是通过数学原理、密码学原理和硬件技术建立技术保障机制，让多个数据参与方在技术信任的共识下开展协同计算。但是，隐私计算涉及的隐私保护技术和算法非常多，算法复杂度、性能、优势场景等并不相同，隐私计算参与者很难通过直观的方法验证所用产品的安全性。业内有待建立涵盖主流隐私计算技术产品的系统性安全分级标准。同时，真实应用中的信任共识通常难以达成，使得隐私计算技术的部署和使用也进展缓慢。

（二）性能瓶颈阻碍隐私计算规模化应用

密文计算需要更大的计算和通信负载，导致遇到性能瓶颈。在保证了参与节点的可用性之后，隐私计算依然面临计算和网络性能的限制。为了保证计算过程的安全性，隐私计算从理论层面上而言一定要

比明文计算付出更大的计算和存储代价，比如同态计算的密文扩张规模可达 1 到 4 个数量级。而考虑到隐私计算是一种多方同步计算，性能的瓶颈会出现在最薄弱的环节，即计算或通信资源最受限的参与方会直接限制整个计算平台的性能。

同步性和可用性对隐私计算参与方的资源要求较高。隐私计算产品通常由多方共同运行，而与其他的数据处理方式不同，由于多主体的特性，其存在对同步性与可用性的要求。隐私计算一般是为解决多个数据源因为隐私问题而无法进行跨数据源明文数据处理的问题，这种情况下，需要多个数据源或计算节点同时在线、同步计算、实时通信，当出现一方因网络或计算资源不足无法继续参与的情况时，可能会引起整个计算被迫停止。因而，保障同步性和可用性是隐私计算面临的关键挑战之一，尤其是在大规模应用情况下。

（三）互联互通壁垒或使数据“孤岛”变“群岛”

对于数据提供机构和数据应用机构来讲，普遍存在与不同机构合作时需要部署不同的隐私计算平台的问题，导致系统建设重复和运营成本浪费。由于不同的隐私计算平台是基于各自特定的算法原理和系统设计实现的，且目前闭源的平台很多，平台之间很难完成信息的交互。因此隐私计算平台互联互通壁垒成为了隐私计算正在面对的新挑战，或使得“数据孤岛”变成了“数据群岛”。

算法原理的差异性为互联互通带来挑战。隐私计算常用技术方案涵盖多种多样的具体算法。这些算法设计之初从底层的数据加密、数据计算的逻辑、数据交互的流程上已经截然不同，在理论层面上的协

议连接或混用都存在较大的互联互通挑战，至于技术方案层面的连通问题就更是困难重重。

系统设计过程中的功能组件多样性增加互联互通成本。为了使得隐私计算协议可以应用于生产环境，技术服务厂商需要开发相应的通信模块与加密组件，以及数据、任务、模型、节点管理等诸多功能组件。这些组件均是不同的技术服务厂商结合自身技术积累和场景应用而实现的，存在很大的差异化。这导致不同厂商间的隐私计算平台在功能组件层面就难以实现互联互通，为用户的部署增加成本，并且存在重复建设的情况。

第七章

隐私计算发展展望

随着大数据发展和应用的不断深入，市场各方对跨源、跨领域、跨用户的数据流通共享需求日益增大，隐私计算技术在近几年得到广泛的关注和迅速的发展，在金融、运营商、医疗、政务等场景也开展了应用试验。然而，隐私计算技术要实现大规模的落地应用，仍需要在性能、技术融合、安全等方面进一步提升，以及在一些非技术因素上形成相关配套。

（一）算法优化和硬件加速将成为隐私计算可用性提升的重要方向

隐私计算普遍借助了密码学技术来实现多方协同计算，效率是影响其能否被广泛应用的一个重要因素。例如，隐私计算联合建模的耗时是传统集中式机器学习的数十倍甚至数百倍以上，联合统计的耗时也是传统集中式明文计算的数百倍以上。因此，隐私计算平台在实际落地应用中需要关注性能的优化，来提升可用性。

性能由算法协议、计算流程、系统架构、数据规模、软硬件环境、网络带宽等多种因素共同决定。**在算法优化层面**，一些常用方式包括：算法加速，尽可能的降低子模块耦合度，对算法流程重新进行深度编排；通信加速，最大程度的减少节点间通信次数及通信量；代码加速，使用更底层的语言（例如 C/C++）来构建基础算子，通过调整字符串

和循环体等方式来降低计算开销等。**在硬件加速层面**，通过新的密码学技术和算法协议，结合硬件加速技术（如 GPU、FPGA、ASIC 加速）和专有算法实现硬件来加速计算量较大的环节和步骤，也能够有效提高性能。

此外，在工程化层面也需要进行大量的优化工作，例如做好计算流程的调度，数据的读取、加密、传输、计算、解密、存储等各个阶段实现最优化，进而将整体性能提升到最优状态，以满足高吞吐、低时延，及某些特定场景的实时性要求。

（二）多元技术融合有望拓展隐私计算应用边界

一方面，隐私计算分支技术间的加速融合满足更多应用场景。以联邦学习为例，与多方安全计算融合能够满足对等网络无可信第三方的联合建模应用需求；与差分隐私融合能够增强对梯度参数的保护程度，进一步防止中间梯度信息泄露；与可信执行环境融合能够提升隐私数据或模型的安全等级等。这些不同隐私计算技术相互融合能够发挥技术的最大优势，更好满足业务场景的多样性需求。**另一方面，隐私计算与区块链等其他领域技术的融合拓展应用边界。**例如区块链可应用到隐私计算各个环节，实现全闭环的安全和隐私服务：隐私计算各流程的操作和处理记录上链保存，可实现记录的防篡改；基于区块链解决数据共享参与者身份及数据可信问题，能够在一定程度上避免主观作恶、数据造假等；此外，区块链还非常适合隐私计算多边信任关系建立，例如使用联盟链来建立隐私计算群体激励机制，通过多个标准化智能合约为参与方提供可信服务。安全审计智能合约的引入，

使得隐私计算在保护隐私数据的合规性方面更加容易验证，将合规监管变成一种服务。

（三）标准体系制定有望助力隐私计算应用落地

当前，国内外众多标准化组织已开始制定或发布以框架和功能为主的隐私计算相关技术标准。通常隐私计算产品只能以自证清白的方式来证明安全性，对隐私计算产品的安全问题难以全面系统有效地评估。此外，隐私计算的算法具有多样性和复杂性，普遍需要繁杂的交互和计算流程，使得某些隐私计算技术缺乏可解释性，降低了隐私计算产品需求方的接受度，增加了评估难度。

一方面，完善的隐私计算相关标准有助于产品规范。隐私计算的安全性规范化过程，可引入安全专家的建议，并通过安全性验证技术、审计、形式化证明等方式使得行业需求方在选择隐私计算平台时有据可依，降低对隐私计算的安全性顾虑。同时，不同应用场景中的安全需求具有多样性，而隐私计算的性能、准确性与安全强度往往也是强关联。因此隐私计算需要进行安全等级划分，在产品开发和实际落地应用中，形成安全与性能、准确性的平衡。**另一方面，成熟的检测和验证手段有助于产品落地应用，**隐私计算产品的安全性，除了需考虑算法协议安全性、通讯安全性、密码安全性、系统安全性等常规的安全性问题，还应确保应用算法逻辑的安全性，尤其是针对应用算法逻辑实时生成的场景（例如基于多方安全计算基础算子的组合实现多方联合统计分析场景），需要新的技术手段来实现自动化的安全检测、识别与预警。

（四）多方生态融合有望推进隐私计算行业发展

隐私计算发展需要法规体系、技术体系、应用体系等多方生态的融合。**一是**法规体系需加速完善，作为数据安全治理和建设的顶层指导，有助于更好地理解安全场景与需求，进而有利于将隐私计算技术实际落地与应用。全球各国已纷纷颁布相关法规，对数据安全与隐私保护相关问题进行严格的规范与引导，例如欧盟 GDPR、美国 CCPA、中国《网络安全法》、《数据安全法》等。**二是**应用体系需进一步加强，目在金融、运营商、医疗、政务数据等行业，存在成功的隐私计算应用案例，实现基于隐私计算的数据安全开放共享，但多数领域仍处于试点应用阶段，还未进入规模化推广阶段，需要产学研用各界加强隐私计算布局。**三是**开源协同加速隐私计算技术迭代，技术开源，已经全面渗透到信息技术的各个领域，微软、谷歌、Facebook、腾讯、阿里、百度等全球知名巨头都在积极拥抱开源。隐私计算作为保障跨机构数据安全合作的关键基础，也注定包含开源模式。开源的成本优势不仅体现在技术复用，降低开发门槛，还体现在问题发现和修复敏捷性上，具有快速迭代优势。隐私计算未来发展趋势必将是开源平台与自研平台并存，形成既开放又独特的多元生态。

在健全完善的法律法规、丰富多样的应用实践、成熟可用的开源技术等多方生态共同作用下，隐私计算行业将迎来蓬勃发展，隐私计算平台有望成为数据合规流通基础设施的关键一环，在保证安全的前提下有效持续释放数据要素价值，促进数字经济高质量发展。

附录

国内主要隐私计算平台

以下为截至 2021 年 7 月依据中国通信标准化协会隐私计算相关标准，通过中国信通院云大所隐私计算产品测试的技术产品，以通过测试时间为序。

序号	企业名称	产品名称	通过的测试	通过时间
1	蚂蚁区块链科技(上海)有限公司	蚂蚁链摩斯安全计算平台（MORSE）	多方安全计算	2019.12
			多方安全计算	2020.12
		蚂蚁链数据隐私服务	可信执行环境	2020.12
2	腾讯云计算（北京）有限责任公司	腾讯神盾沙箱	多方安全计算	2019.12
		腾讯云联邦学习应用平台软件	联邦学习	2020.12
		腾讯神盾 Angel PowerFL 联邦计算平台	多方安全计算	2020.12
			联邦学习	2020.12
			多方安全计算（性能）	2021.6
			联邦学习（性能）	2021.6
3	华控清交信息科技（北京）有限公司	华控清交多方安全计算平台	多方安全计算	2019.12
		清交 PrivPy 多方计算平台	联邦学习	2020.12
			多方安全计算（性能）	2021.6
			联邦学习（性能）	2021.6
4	北京百度网讯科技有限公司	百度点石	多方安全计算	2019.12
		联邦计算平台	多方安全计算	2020.6
		百度智能云度信金融安全计算平台	多方安全计算	2020.6
		点石安全计算平台（MesaTEE）	可信执行环境	2021.6
5	上海富数科技有限公司	富数安全计算平台	多方安全计算	2019.12
		阿凡达安全计算平台	多方安全计算	2020.12
			联邦学习	2020.12

序号	企业名称	产品名称	通过的测试	通过时间
			多方安全计算（性能）	2021.6
			联邦学习（性能）	2021.6
6	杭州趣链科技有限公司	趣链联邦计算软件	多方安全计算	2020.6
			区块链辅助隐私计算	2021.6
			多方安全计算（性能）	2021.6
7	北京数牍科技有限公司	Tusita 多方安全隐私计算平台	多方安全计算	2020.6
8	同盾科技有限公司	同盾智邦学习平台	多方安全计算	2020.6
		同盾智邦知识联邦平台	联邦学习	2020.12
9	厦门渊亭信息科技有限公司	DataExa-Insight 人工智能中台系统	多方安全计算	2020.6
			联邦学习	2020.12
10	深圳市洞见智慧科技有限公司	洞见安全多方数据智能平台	多方安全计算	2020.6
		洞见数智联邦平台（INSIGHTONE）	联邦学习	2020.12
			多方安全计算	2021.6
			区块链辅助隐私计算	2021.6
			多方安全计算（性能）	2021.6
			联邦学习（性能）	2021.6
11	蚂蚁智信（杭州）信息技术有限公司	共享智能平台	多方安全计算	2020.6
			可信执行环境	2020.12
12	天翼电子商务有限公司	密流安全计算平台	多方安全计算	2020.6
		CTFL 天翼联邦学习平台	联邦学习	2020.12
		PrivTorrent 密流安全计算平台	可信执行环境	2021.6
		大禹-天翼数据融通平台	区块链辅助隐私计算	2021.6
13	北京融数联智科技有限公司	UPAI 安全计算平台	多方安全计算	2020.6
14	蓝象智联（杭州）科技有限公司	GAIA·Edge	多方安全计算	2020.12
15		联邦学习云服务平台	多方安全计算	2020.12

序号	企业名称	产品名称	通过的测试	通过时间
	深圳前海微众银行股份有限公司		联邦学习	2020.12
		多方大数据隐私计算平台 WeDPR-PPC	区块链辅助隐私计算	2021.6
16	矩阵元技术（深圳）有限公司	矩阵元隐私计算服务系统	多方安全计算	2020.12
17	翼健（上海）信息科技有限公司	翼数坊 XDP 隐私安全计算平台	联邦学习	2020.12
			可信执行环境	2020.12
18	京东云计算有限公司	京东智联云联邦学习平台	联邦学习	2020.12
19	京东数科海益信息科技有限公司	联邦模盒	联邦学习	2020.12
		万象+隐私计算平台	区块链辅助隐私计算	2021.6
20	杭州锘崴信息科技有限公司	锘崴信联邦学习平台	联邦学习	2020.12
		锘崴信隐私计算平台	可信执行环境	2020.12
21	深圳前海新心数字科技有限公司	新心数述联邦学习平台	联邦学习	2020.12
22	中国电信股份有限公司云计算分公司	天翼云诸葛 AI-联邦学习平台	联邦学习	2020.12
23	光之树（北京）科技有限公司	云间联邦学习平台	联邦学习	2020.12
24	神谱科技（上海）有限公司	神谱科技 Seceum 联邦学习系统	联邦学习	2020.12
25	星环信息科技（上海）有限公司	星环联邦学习软件	联邦学习	2020.12
26	北京冲量在线科技有限公司	冲量数据互联平台	可信执行环境	2020.12
			区块链辅助隐私计算	2021.6
27	上海隔镜信息科技有限公司	天禄多方安全计算平台	可信执行环境	2020.12
28	华为云计算技术有限公司	可信智能计算服务 TICS	可信执行环境	2020.12
			联邦学习	2021.6
29	浙江天猫技术有限公司	DataTrust 阿里云隐私增强计算软件	多方安全计算	2021.6
			联邦学习	2021.6
			可信执行环境	2021.6
			联邦学习（性能）	2021.6
30	上海凯馨信息科技有限公司	凯馨多方安全计算平台	多方安全计算	2021.6
31	深圳市云计算科技有限公司	ELF 隐私计算服务平台	多方安全计算	2021.6

序号	企业名称	产品名称	通过的测试	通过时间
32	杭州金智塔科技有限公司	金智塔隐私计算平台	多方安全计算	2021.6
33	南京三眼精灵信息技术有限公司	智力共享平台·数链	多方安全计算	2021.6
34	北京瑞莱智慧科技有限公司	隐私保护机器学习平台 RealSecure	多方安全计算	2021.6
			联邦学习	2021.6
35	联易融数字科技集团有限公司	蜂密隐私计算平台	多方安全计算	2021.6
			区块链辅助隐私计算	2021.6
		蜂隐联邦学习平台	联邦学习	2021.6
36	医渡云（北京）技术有限公司	多方安全计算平台（YIDUMANDA）	多方安全计算	2021.6
			联邦学习	2021.6
37	苏州同济区块链研究院有限公司	梧桐隐私计算平台 WPC	多方安全计算	2021.6
38	北京火山引擎科技有限公司	火山引擎隐私计算平台	联邦学习	2021.6
39	深圳致星科技有限公司（星云 Clustar）	星云隐私计算平台	联邦学习	2021.6
40	云从科技集团股份有限公司	云从隐私计算平台	联邦学习	2021.6
41	北京九章云极科技有限公司	DataCanvas FL 联邦学习平台	联邦学习	2021.6
42	天冕信息技术（深圳）有限公司	天冕联邦学习平台	联邦学习	2021.6
43	度小满科技（北京）有限公司	貔貅隐私计算平台	联邦学习	2021.6
44	北京神州泰岳智能数据技术有限公司	数联盈	联邦学习	2021.6
45	中移系统集成有限公司(雄安产业研究院)	中移联邦计算服务平台	联邦学习	2021.6
46	阿里云计算有限公司	阿里云机器学习 PAI	联邦学习	2021.6
47	零幺宇宙（上海）科技有限公司	光笺可信执行环境	可信执行环境	2021.6
48	西安纸贵互联网科技有限公司	纸数魔方-基于区块链的可信执行环境数据计算平台	可信执行环境	2021.6
49	光之树（杭州）科技有限公司	天机可信计算平台	可信执行环境	2021.6
50	杭州安恒信息技术股份有限公司	安全岛数据共享访问控制系统 DAS-SMPC	区块链辅助隐私计算	2021.6

联系方式：

中国信息通信研究院 云计算与大数据研究所

地址：北京市海淀区花园北路 52 号

邮编：100191

邮箱：yuanbo@caict.ac.cn

网址：www.caict.ac.cn