隐私计算与区块链

融合应用白皮书

同盾科技有限公司

北京航空航天大学杭州创新研究院

北京航空航天大学

2022 年 3 月 1 日

目录

第一章 数据要素流通背景与政策

1.1 数据流通政策

近年来，数字经济蓬勃发展，已经成为带动中国经济增长的核心动力。数据的开放流通政策可追溯至2015年9月国务院发布的《促进大数据发展行动纲领》，正式提出了加快政府数据开放共享，推动资源整合，提升治理能力，稳步推动公共数据资源开放和统筹规划大数据基础设施建设。而在2019年10月，党的十九届四中全会首次明确数据可作为生产要素按贡献参与分配，提出“健全劳动、资本、土地、知识、技术、管理、数据等生产要素由市场评价贡献、按贡献决定报酬的机制”。2020 年 4 月，中共中央国务院发布了《关于构建更加完善的要素市场化配置体制机制的意见》，正式将数据与土地、劳动力、资本、技术等传统要素并列为生产要素，并将提升社会数据资源价值作为核心目标之一，强调要加快培育数据要素市场，推动政府数据开放共享，提升社会数据价值，加强数据资源整合和安全保护，以及制定数据隐私保护机制和安全审查制度。这表明数字经济时代，数据成为新的关键生产要素，已成为社会基础性战略资源，蕴藏着巨大潜力和能量。2020年10月党的十九届五中全会进一步强调，推进数据要素市场化改革，加快数字化发展，建立数据资源产权、交易流通、跨境传输和安全保护等基础制定和标准规范，推动数据资源开发利用。2021年3月，十三届全国人大四次会议表决通过的“十四五”规划中明确提出，要迎接数字时代，激活数据要素潜能，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、营造良好数字生态，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。而在2021年12月国务院发布的《“十四五”数字经济发展规划》，把“建立数据要素市场体系”放在了发展目标的第一位。数据成为了比肩石油的基础性关键战略资源，随着政策红利的相继释放，激活了数据流通、高效共享与协同应用的巨大市场需求。2022年1月国务院办公厅进一步印发《要素市场化配置综合改革试点总体方案》，推动要素市场化配置改革向纵深发展，围绕土地、劳动力、资本、技术、数据等要素市场化配置该给提出8个方面的任务，其中在“探索建立数据要素流通规则”方面中提出了四项任务：完善公共数据开放共享机制；建立健全数据流通交易规则；拓展规范化数据开发利用场景；加强数据安全保护。

各省市也积极参与数字产业的发展与建设，陆续发布相关的数据条例以及数据要素市场化配置改革方案。2020 年 9 月北京市发布《北京市促进数字经济创新发展行动纲要(2020-2022 年)》，提出要体系化构建数字经济发展体制机制，聚焦基础设施建设、数字产业化、产业数字化、数字化治理、数据价值化和数字贸易发展六大方向。上海市 2021 年 11 月通过《上海数据条例》并于 2022 年 1 月 1 日正式施行，成为国内首部省级人大制定的数据条例，把数据纳入法治的轨道，推动上海城市数字化转型和增强城市竞争力。同年，上海是还发布了《推进上海经济数字化转型赋能高质量发展行动方案（2021-2023）》和《推进上海生活数字化转型构建高品质数字生活行动方案（2021-2023）》两个行动方案，推动数据新要素专项行动，更好激活数据要素价值，探索建立数据要素市场化体系，深化数据资源市场化配置、资产化管理、场景化开放、便利化流通的新格局。广东省在 2021 年 7 月正式印发《广东省数据要素市场化配置改革行动方案》，作为全国首份数据要素市场化配置改革文件，明确五大方面 24 项任务，从释放公共数据资源价值、激发社会数据资源活力，加强数据资源汇聚融合与创新应用、促进数据交易流通，强化数据安全保护等方面着力。此外还有《浙江省数字经济发展“十四五”规划》、《天津市加快数字化发展三年行动方案（2021—2023 年）》、《海南省国民经济和社会发展第十四个五年规划和二〇三五年远景目标纲要》等一系列的相关数据要素政策指引条例在 2021 年相继颁发。与此同时，各地的大数据交易中心相继建立，贵阳大数据交易所 2015 年 4 月正式挂牌运营，率先探索数据流通交易价值和交易模式。同年 11 月，华东江苏大数据交易中心成立。上海数据交易所和北京国际大数据交易所分别于 2021 年成立。据不完全统计，目前我国的大数据交易场所已有 20 余家。根据国家工信安全中心测算，2020 年我国的数据要素市场规模已达到 545 亿元，未来几年预估年均复合增长率将超过 30%。

1.2 数据隐私安全保护政策

数据在为人们的生活带来了种种便利的同时，也使得大家对于个人的数据隐私和安全产生了担忧，这俨然已经成为世界性的问题。国内和国际均对数据方面

的监管日趋严厉，各种数据保护的法律法规相继制定和完善，严格定义数据的控制权和监管数据的流动。

在国际上，全球已有 100 多个国家和地区制定了数据保护方面的政策和法规，并纷纷建立了国家数据保护机构或监管机构。2018 年 5 月在欧盟生效的《通用数据保护条例》(GDPR)，强化数据主体权利，明确个人数据处理的合法性、合理性和透明性原则，并规定用户可以要求经营者删除其个人数据并且停止利用其数据进行建模，而违背该条例的企业将会面临巨额罚款，据统计，2021 年共有 412 家企业因为违反 GDPR 被罚款，包括亚马逊、WhatsAPP 等互联网头部机构，罚款总额高达 10 亿欧元。在 GDPR 正式实施一个月后，2018 年美国加利福尼亚州颁布了《加州消费者隐私法案》(CCPA)，规定了数据收集、使用、传输和出售应履行的一系列义务，强调个人对数据的控制权，加强保障消费者隐私权和数据安全保护。在 CCPA 实施后，美国加利福尼亚州在 2020 年 11 月又通过了《加州隐私权利法案》(CPRA)，CPRA 为加州居民确立了新的数据隐私权，对企业和服务提供商施加了新的义务和责任，并创建了一个独立的数据监管机构。2021 年 7 月，美国统一法律委员会宣布通过了《统一个人数据保护法》，首次为美国各州提供了一个统一的综合数据保护法。

图 1：国际数据安全与隐私保护相关法规

2016 年 11 月，我国通过了《中华人民共和国网络安全法》，旨在通过多项举措加强个人信息和数据保护。2019 年 5 月 28 日，我国国家互联网信息办公室发布了《数据安全管理办法(征求意见稿)》，对数据收集、数据处理使用等方面进行详细规定，并提出了收集重要数据的备案制以及向第三方提供重要数据的批准制的新要求。中国人民银行 2020 年 2 月正式发布了《个人金融信息保护技术规范》，从安全技术和安全管理两个方面，对个人金融信息保护提出了规范性要求。而随着 2020 年 3 月《信息安全技术个人信息安全规范》修订版正式获

批发布，2021 年 1 月施行的《民典法》在现行法律规定的基础上对隐私权和个人信息保护作出了专门规定、以及 2021 年 9 月正式施行的《数据安全法》和 2021 年 11 月施行的《个人信息保护法》等一系列数据安全、隐私保护要求的法令法规的制定与施行，逐步完善了国家数据相关立法的顶层设计，强调数据流通过程的数据安全和个人信息保护，数据安全和隐私保护将迎来新时代。

图 2：国内数据安全与隐私保护相关法规

这些法令法规的施行，成为国家大数据战略中至关重要的法制基础，也成为数据安全保障和数据经济发展领域的重要基石。在数据安全、用户隐私需要合法合规保护的同时，企业如何在兼顾数据安全的同时保障业务的高效运行，进一步发挥数据要素的价值，成为了当前企业和社会关注的重点。

1.3 行业发展需求

随着互联网、大数据、人工智能、云计算等技术的创新发展，数字科技成为世界科技革命和产业变革的重要驱动力，也是新一轮国际竞争的关键领域。随之而来，数据也成为了继土地、劳动力、资本、技术之后的第五大生产要素，数据成为推动数字经济深化发展的核心引擎，也为数字经济发展提供不可或缺的动力源泉。目前数据要素市场建设正在加速推进，并催生了新产业、新业态和新模式。

国家工信安全中心测算数据显示，2020 年我国数据要素市场规模达 545 亿元，预计到 2025 年，规模将突破 1749 亿元，整体将进入群体性突破的快速发展阶段。从宏观经济增长层面来看，数据要素对 2021 年 GDP 增长的贡献率和贡献度分别为 14.7%和 0.83 个百分点；从行业发展层面来看，数据要素对各个行业的产值影响具有较大差异，其中，信息传输、软件和信息技术服务业产出对数据要素最为敏感；从企业绩效层面来看，数据要素显著提升企业总资产净利润率，数字化转型对于制造业企业的影响最大。

数据要素对增加各个领域的效益有明显的成果。在制造业中，通过数据要素而开展的分析与运筹优化，使得工业企业业务增长，提高企业的生产效率，缩短产品的研发周期，并同时提高了能源利用率，真正为企业做到了开源节流。在金融领域，数据要素的应用提升资源配置效率。通过多数据来源的分析挖掘，提升金融产品风险定价的有效性和精准度，并通过在各业务领域的实际场景中，捕捉个人和企业潜在需求，拓展金融服务的领域。在医疗领域，大数据医疗的市场蓬勃发展离不开数据要素的支持，在数字化临床诊疗、医院管理、医疗科技创新、医疗保险、新药研发等多个业务领域，医院和患者等都可以从中获益良多。

从需求侧来看，数据要素市场需求旺盛。企业通过数据库采集、互联网采集和系统日志采集产生数据要素。通过数据要素市场化配置提升数据交换共享、数据交易的发展速度，对促进经济的稳定增长有着重要的作用。从目前来看，培育统一开放、竞争有序的数据要素市场，促进数据要素的安全高效流通，关乎数字经济持续实现高质量发展，也是数据作为社会生产要素在市场中产生更大价值的重要前提。

在目前的数据要素流通过程中，仍存在诸多的风险和挑战，对数字经济发展和培育数据要素市场造成困扰和障碍，其中最突出的问题是数据安全、隐私保护的问题。传统的数据流通模式一般是以数据包形式或明文数据 API 接口的方式进行数据共享，直接交换了原始数据的所有权，容易导致隐私信息的泄露和数据的滥用等问题。在日趋严厉的数据安全、隐私保护的合规监管要求，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，传统的数据交换模式不再可行，迫切需要通过技术的手段来化解数据要素流通过程的数据安全和隐私保护难题，以平衡数字经济发展多样化场景的需求和数据要素合规、合理使用的要求。

隐私计算技术被认为是在确保隐私保护的基础上能够有效打破数据流通壁垒、促进数据价值释放，从技术角度实现数据价值的共享流通和系统应用的重要手段。隐私计算通过多方安全计算、联邦学习、可信执行环境等一些列的技术进行创新应用，实现数据的“可用不可见”，实现数据要素安全、高效流通，并消除数据所有方因为数据所有权交换而导致的数据权利的转移的顾虑，保护数据提供方的数据权益不受伤害。从应用层面看，隐私计算使得数据要素流通过程中的

不同主体在满足数据合规要求的前提下，充分调动数据资源方、数据使用方、数据运营方、数据监管方等各方的积极性，实现不同数据资源的汇集、交易和流通，提升数据要素价值，进一步促进数据要素市场化配置。

此外数据要素流通过程中还存在确权难、定价难、交易难、监管难等挑战，以及数据要素流通的可信性也是亟待解决的核心问题，需要通过区块链等技术手段配合《网络安全法》、《数据安全法》、《个人信息保护法》等制度，实现数据要素安全可信流通。区块链作为一种创造信任的机制，可解决多方之间数据以及业务协同中的可信难题，可在监管前提下，通过融合智能合约、共识机制、密码学机制等，在多方业务协同之间构建信任机制。区块链技术去中心化、不可篡改以及对数据资产可进行智能合约编程等特性，决定了区块链技术在数据确权、交易、定价、交割、监管等环节均可以发挥其作用，具有广阔的应用空间。区块链技术应用到数据要素流通全流程，可以使得数据要素流通更加透明、安全与高效，能够重塑数据要素交易规则，激活数据要素流通的活力与挖掘更大的数据要素价值。

1.4 数据要素流通面临挑战

我国数据要素流通的尚处于起步阶段，仍面临诸多挑战，需要政策、法律、技术与监管协同发力，共同促进数据要素以高效、安全、可信、可控的方式进行有效流通与价值释放。目前数据要素流通面临的挑战，主要包括流通机制、技术成熟度、数据安全、流通制度、国际合作等方面。

(1) 数据要素的流通机制不健全：我国现有的法律制度多是从保护和监管的角度出发，强调对数据的规范利用和安全隐私保护，并未就具体的流通实践形式、流通市场准入、市场监管等方面给出清晰的法律界定。所以在数据流通立法体系尚不完善、数据流通行为缺乏统一监管机制下，各类参与主体难以把握监管和处罚尺度，对责任的判断没有具体的判断，在数据要素流通的具体过程中没有明确的合规依据指引，进而因顾虑而放缓了发展的脚步。

(2) 数据要素流通的技术和应用不成熟：众多领域普遍存在业务驱动力跟不上数字化转型需要，数字应用意愿差、流通水平低，产生的数据垃圾、数据烟囱

越来越多。多数行业主体对数据应用还仍处于初始阶段，投入成本太高、技术太复杂等，提升了数据应用的门槛，加大了其参与数据应用的难度。

(3) 数据安全与隐私保护有待提升：数据安全是数据要素的保障与基石。当前技术手段由于无法保证数据绝对安全，尤其是涉及秘密和隐私数据的安全，阻碍了参与方的主观参与意志。数据泄露、越权滥用等事件，使得人们对数据安全和隐私保护产生了强大的不信任感与抵触感。

(4) 数据流通的制度和规则不明确：当前，数据资产化定价流通的理论基础不扎实，数据要素权属界定、分类分级、估值定价、收益分配等方面缺乏系统框架，数据要素流通难以制定明确的规则。数据流通所涉及来源者、收集者、持有者、使用者等多方主体的权限和利益没有具体说明，如何保护参与主体的利益分配公平公正，仍需建立相关的规则指引。

(5) 国际合作尚不明显：我国数据要素交易规则尚未与国际通行规则相匹配。我国在数据交易、跨境支付、海关监管等方面的法律规范与国际其他主体，还存在较大差异，导致跨境数字贸易难以实现自由畅通，这不仅提高了企业和消费者的成本，并造成人才跨境、服务跨境、数据跨境、技术跨境等方向发展的流通壁垒。对此，我们应尽快与国际接轨，在保证安全与合法权益的前提下，加强国际间合作，共同促进国际数据要素流通与应用。

第二章 隐私计算与区块链技术发展现状

2.1 隐私计算技术发展现状

最近几年，学术界和工业界都已经开始在数据安全和隐私保护方向的探索。其中隐私计算能在不暴露原始数据的情况下进行数据计算，而且数据结果可以被验证，可以实现在确保数据安全的基础上，让数据价值跨主体地使用，放大数据的价值，被广泛认为是平衡数据安全、隐私保护和跨机构间数据协同的重要技术手段。隐私计算是密码学、数据科学、人工智能与分布式计算等学科交叉融合的成果，是数据要素互联互通不可或缺的关键技术。

从技术角度出发，目前的隐私计算是涵盖众多学科的交叉融合技术，主流分为三大技术路线：第一类是以多方安全计算为代表的基于密码学的相关技术；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。隐私计算相关技术的发展历程如下图所示：

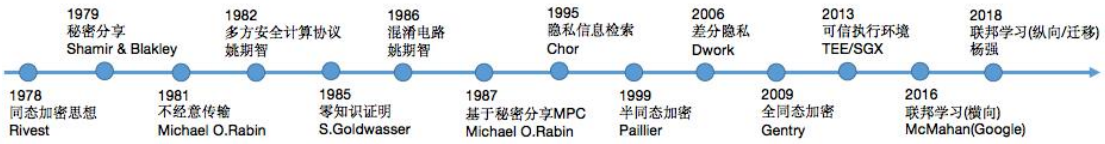

图 3：隐私计算相关技术发展历程

这些不同的隐私计算技术，其起源、定义和研究初衷各有差异，但目前逐渐呈现融合统一的趋势，现在基本都聚焦在研究如何在保证数据安全的前提下打破数据孤岛，实现数据可用。在工业界，随着国内外政策法规的不断完善，再加上对于某些数据高敏感行业强监管的需求，具体的解决方案基本上沿着两个方向在演化：

中心化向分布式或去中心化过渡。现有的大数据平台基本上都是中心化的，对数据进行集中的存储、管理、分发等操作。中心化方式的缺点是数据存储在第三方平台，脱离数据提供方的控制，违背了数据隐私保护的规定。同时，随着数据规模的不断变大，直接在中心服务器上计算或 学习的压力也会不断增加。为

了减轻这种压力，计算或学习过程需要分 散到数据提供方或终端设备上进行，这种分布式计算或学习的过程则是人工智能领域更关心的问题。而如果没有中心节点的存在，这种智能化的过程则变为去中心化的形式。这时的数据是分而治之，各自为数据所有者控制，每个节点上的数据相对只是小数据，但是由于可以触达更多的数据，其性能甚至会超越有限数据的中心化聚集方式。

数据向知识化升级。为了保护节点数据安全和隐私，直接共享使用显然是不可行的，要做到数据对外不可见才是关键，这就需要密码学。通过加密方法(如:哈希编码、同态加密等)对数据脱敏和去标识化，让数据转化成为安全的信息或者知识，再对分散的信息计算或知识聚合，来保证数据不直接共享但是可用的。进而解决参与方互不信任的问题，最终对在小数据上生成的信息或知识进行联合以实现大智能。

在技术实现上，主要包括了联邦学习、多方安全计算、可信执行环境、同态加密、差分隐私、零知识证明等，其技术路线相互交错也互相支撑。

2.1.1 联邦学习

联邦学习(Federated machine learning/Federated Learning)，又名联邦机器学习，联合学习，联盟学习，被认为是解决数据安全与人工智能两难问题的一个重要技术。联邦机器学习是一个机器学习框架，能有效帮助多个机构在满足用户隐私保护、数据安全和政府法规的要求下，进行数据使用和机器学习建模[1]。联邦学习作为分布式的机器学习范式,可以有效解决数据孤岛问题,让参与方在不交换原始数据的基础上联合建模,能从技术上打破数据孤岛,实现 AI 协作。

联邦学习在如下两种情况可以很好解决企业的人工智能应用难题：一是涉及到保护数据隐私和核心价值的场景，因为联邦学习的整个学习训练过程，没有传输任何原始数据；二是多方数据补充的场景，这可能存在单方样品数量不够充分或单方数据维度不够丰富的情况。例如在金融风控场景中，银行希望引入外部数据源做特征补充来建立联合模型。基于用户授权，联邦学习技术可以在保证数据安全不出库的同时，整合不同机构间对用户行为特征不同维度的捕捉，以用户为

基础，形成对个人的较为全面的描述。对比传统模型方式，模型可以学到更多用户信息，从而提升模型效果，促进业务发展，实现降本增效。

从整个数据产业看，联邦学习可以增加可用数据的总量，能很好的解决现存数据孤岛的问题；对企业自身而言，使用联邦学习能简单、合法且低成本的获取外部有效的数据信息，快速解决某些因数据量或数据维度不足而导致的困扰，而且也不会造成合作企业间数据或商业机密的泄露。规范数据使用可以在汇聚更多数据的基础上迎来数据价值挖掘的下一个爆发点，带动人工智能的数据基础设施进步，联邦学习未来会逐步成为人工智能的基础设施。

2.1.2 多方安全计算

安全多方计算（Secure Multiparty Computation）是解决多源数据隐私保护问题的另一个重要方法。多方安全计算[2]是指在无可信第三方情况下，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，每个参与者都各自持有秘密输入，希望共同完成对某个函数的计算，但要求每个参与者除计算结果外均不能得到其他参与实体的任何输入信息。也就是参与者各自完成运算的一部份，最后的计算结果由部分参与者掌握或公开共享。也就是说多方安全计算技术可以获取数据使用价值，却不泄露原始数据内容，保护隐私。因此在大数据发展并立法保护的今天，多方安全计算技术从多年的学术研究，变得具有巨大的商业价值。

多方安全计算是由一系列密码学安全计算协议组成，常采用的技术有秘密共享(Secret Sharing)[3]、不经意传输(Oblivious Transfer)[4]、混淆电路(Garbled Circuit)[5]、同态加密(Homomorphic Encryption)[6]等。安全多方计算技术在需要隐私保护的场景中具有重要意义，其主要适用的场景包括联合数据分析、数据安全查询、数据可信交换等。

多方安全计算框架可以让多方安全地计算任何函数或某类函数的结果。自1986年姚期智提出第一个通用的多方安全计算框架以来，30多年间已经有BMR、GMW、BGW、SPDZ等多个多方安全计算框架陆续提出。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能泄露，在近几十年的发展中，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也不断提升，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。

2.1.3 可信执行环境

可信计算(Trusted Computing)是指系统提供的计算服务是可信赖的，是一种运算与防护并存的信息安全技术，保证了计算的行为与预期一致，同时保证全程是可检测可监控。可信执行环境（Trusted Execution Environment）是计算平台上由软硬件方法构建的一个安全区域，可保证在安全区域内部加载的代码和数据在机密性和完整性方面得到保护。其目标是确保一个任务按照预期执行，保证初始状态和运行时状态的机密性、完整性。相较于其他基于密码学的隐私计算技术，TEE的优势在于性能强大，能够解决复杂场景的业务诉求，同时开发成本低，能够迅速做出满足业务诉求的产品。

目前各家厂商的可信执行环境有不同的实现方式，功能也有所不同。Intel 的 TEE 主要是基于 SGX(Software Guard Extension)指令集扩展[7]，旨在以硬件安全为强制性保障，不依赖于固件和软件的安全状态，提供用户空间的可信执行环境，通过一组新的指令集扩展与访问控制机制，实现不同程序间的隔离运行。应用程序和数据放在置信区域 Enclave 内执行，从而保证应用程序和数据的隐私性和安全性。ARM 则主要通过 TrustZone[8]来实现安全扩展，TrustZone 是用于高级 RISC 机器指令集体系结构的 TEE，可在安全模式和不安全模式之间切换，但没有内置的硬件密钥可以支持数据的密封和证明提供信任根。AMD 则通过安全加密虚拟化(SEV)提供内存流量的自动内联加密和解密，为虚拟机(VM)使用的数据保证机密性，但 SEV 本身不提供内容完整性和时效性保证。这些基于不同技术实现的可信执行环境，凭借其芯片级的保护方案，具有高安全性、高性能、高通用性等优势，在数据联合应用、交易隐私保护、区块链、人工智能、医疗等领域有着广阔的应用前景。

2.1.4 同态加密

同态加密(Homomorphic Encryption)是一类具有特殊属性的加密方法，是 Ron Rivest, Leonard Adleman，以及 Michael L. Dertouzo 在 1978 年提出的概念。与一般加密算法相比，同态加密除了能实现基本的加密操作之外，还能实现密文间的多种计算功能，即先计算后解密可等价于先解密后计算。这个特性属性对于保护信息的安全具有重要意义，利用同态加密技术可以先对多个密文进行计算之后再解密，不必对每一个密文解密而花费高昂的计算代价；利用同态加密技术可以实现无密钥方对密文的计算，密文计算无须经过密钥方，既可以减少通信代价，又可以转移计算任务，由此可平衡各方的计算代价；利用同态加密技术可以实现让解密方只能获知最后的结果，而无法获得每一个密文的消息，可以提高信息的安全性。

同态加密技术在分布式计算环境下的密文数据计算方面具有比较广泛的应用领域，比如安全云计算与委托计算、多方保密计算、匿名投票、文件存储与密文检索等。例如在云计算方面，虽然目前云计算应用中，从安全角度来说，用户还不敢将密钥信息直接放到第三方云上进行处理，通过实用的同态加密技术，则大家可以放心使用各种云服务，同时各种数据分析过程中也不会泄露用户隐私。加密后的数据在第三方服务处理后得到加密后的结果，这个结果只有用户自身可以进行解密，整个过程第三方平台无法获知任何有效的数据信息。另外一个应用，在区块链上，使用同态加密技术，智能合约也可以处理密文，而无法获知真实数据，能极大的提高隐私安全性。

2.1.5 差分隐私

差分隐私(Differential Privacy - DP)[9]是 Dwork 在 2006 年提出的一种隐私计算技术，目的是提供一种技术保证在数据库查询和统计时，能够最大化数据查询的准确性，同时最大限度来减少识别单个记录的机会，也就是让查询结果对于数据集中单个记录的变化不敏感，从而攻击者就无法通过加入或减少一个记录，观察查询结果的变化来推测个体的具体信息。

差分隐私主要通过在输入输出中增加适量的随机噪音以确保修改数据集中一条个体记录不会对统计结果造成显著影响，差分隐私的主要噪声机制包括：拉普拉斯噪音(Laplace Noise)，高斯噪音(Gaussian Noise)，指数机制等。基于拉普拉斯机制和高斯机制的差分隐私是通过向查询结果中加入服从拉普拉斯分布和高斯分布的噪声来实现差分隐私保护，适用于数字型结果的差分隐私。基于指数机制的差分隐私，则采用指数机制来引入一个打分函数来实现差分隐私保护，也就是当接收到一个查询后，不是确定性的输出一个结果，而是以一定概率值返回结果，这个概率值是由打分函数确定，指数机制适用于离散型结果的差分隐私。

差分隐私常用于多方参与的统计分析中，也用于联邦机器学习算法，在中间参数交互过程通过采用差分隐私，避免恶意参与方通过中间结果反推出训练集中的个体特征信息，增强联邦算法的安全性。

2.1.6 零知识证明

零知识证明(Zero-Knowledge Proof)[10]是由 S.Goldwasser、S.Micali 及 C.Rackoff 在 20 世纪 80 年代初提出的一种基于密码学的验证技术，旨在证明者可以不向验证者泄露任何关于被验证信息的信息的情况下，使得验证者相信某个结论是正确的。

零知识证明通常包括两个参与方，宣称某一结论是真的证明者(Prover)和确认结论是真的验证者(Verifier)。零知识证明的实现机制分为交互式和非交互式两种类型。交互式零知识证明要求证明者和验证者之间通过交互来完成对结论的验证，这种交互式输入通常以一个或多个挑战的形式进行，使得证明者的回答能够使验证者确信结论是真的。非交互式零知识证明，证明者创建一份证明，任何使用这份证明的参与方都可以进行验证，非交互式零知识证明通过大幅减少通信步骤而提高效率。零知识证明的典型算法有 ZK-snark、ZK-stark、Plonk、Bulletproof、Aztec 等。

零知识证明主要应用于验证场景，借助零知识证明“不泄露信息”的特性，可以在不泄露交易信息的情况下证明区块链上资产转移的有效性，可以用于验证数据的真实性、算法分析结果的可靠性等。

2.2 区块链技术发展现状

2.2.1 区块链主要特性

区块链[11]最早在 2008 年由中本聪提出，作为比特币的底层技术。区块链是利用块链式数据结构来验证与存储数据、利用分布式节点共识算法来生成和更新数据、利用密码学的方式保护数据传输和访问的安全、利用由自动化脚本代码组成的智能合约来编程和操作数据的一种全新的分布式基础架构与计算范式。由于这种技术体系实现了防伪造、防篡改、多方参与、可溯源的数据记录方式，能够解决传统数据记录与存储方面的部分问题，得到社会的广泛关注。随着区块链在其他多个领域应用潜力的凸显，其逐渐成为不可忽视的新一代信息技术之一。近年来，区块链在金融、保险、医疗健康、政务、供应链等领域落地应用不断增加，区块链技术的价值得到全社会认可。

区块链的主要特性包括：去中心化、不可篡改性、可溯源性和开放性等。

(1) 去中心化：在区块链网络中分布着众多节点，节点与节点之间可以自由连接进行数据、资产、信息等的交换，而无需通过第三方中心机构。

(2) 不可篡改性：区块链使用共识机制和密码学技术来保证链上信息不被篡改，主要用到的是密码学中的哈希函数以及非对称加密。

(3) 可溯源性：指区块链保存了从第一个区块开始的历史数据，连接形式是后一个区块拥有前一个区块的哈希值，区块链上任一一条记录都可以通过链式结构追溯本源。

(4) 开放性：区块链技术基础是开源的，除了交易各方的私有信息被加密外，区块链的数据对所有人开放，任何人都可以通过公开的接口查询区块链数据和开发相关应用，因此整个系统信息高度透明。

2.2.1 区块链核心技术

共识机制和智能合约是区块链技术的两个核心概念。在区块链体系中，共识机制和智能合约，保证了数据的真实性和合约执行力，实现“去中心化”。

共识机制是区块链系统中实现不同节点之间建立信任、获取权益的数学算法。是一种区块链治理体系，是通过结合经济学、博弈论等多学科设计出来的一套保证区块链中各节点都能积极维护区块链系统的方法。区块链共识机制首先由中本聪在比特币白皮书中提出，逐渐发展成为一种维护分布式账本多中心化的重要机制，是保持区块链安全稳定运行的核心。借助于共识机制，在尽可能短的时间内完成分布式数据的记录，做到整个流程安全、明确及不可逆，由此形成一种难以攻破的、不可篡改的、诚实可信的分布式多节点数据记录系统，最终实现对人的去信任，以及对技术的信任，实现节点上的价值转移，即基于区块链的交易。共识机制涉及通过多个甚至全部区块链节点就区块信息达成全网一致共识，保证最新区块被准确添加至区块链，每个节点存储的区块链信息一致、不分叉，甚至可以抵御恶意攻击。

区块链的另一个核心概念，是智能合约。智能合约是一套以数字形式定义的承诺，包括合约参与方可以在上面执行这些承诺的协议，一套承诺指的是合约参与方同意的（经常是相互的）权利和义务，这些承诺定义了合约的本质和目的。一个合约就是存在区块链里的程序。合约的参与双方将达成的协议提前安装到区块链系统中。在双方的约定完成后，开始执行合约，不能修改。至于合约执行所需要的“燃料”，也就是手续费，也需要提前支付。智能合约可以解决日常生活中常见的违约问题，如果应用到各行业中，可以避免违约的信用问题。

在区块链出现之前，商业领域的信任关系通常要依赖于正直、诚信的个人、中介机构或其他组织才能建立起来。在区块链这个新兴的领域中，信任关系的建立是基于网络，甚至是网络上的某个对象。由区块链驱动的智能合约将会要求双方遵守他们的承诺。

2.2.3 区块链发展历程

从区块链诞生至今的十几年发展历程来看，区块链技术整体概况可以分为三个阶段：区块链 1.0 时代、区块链 2.0 时代和区块链 3.0 时代。

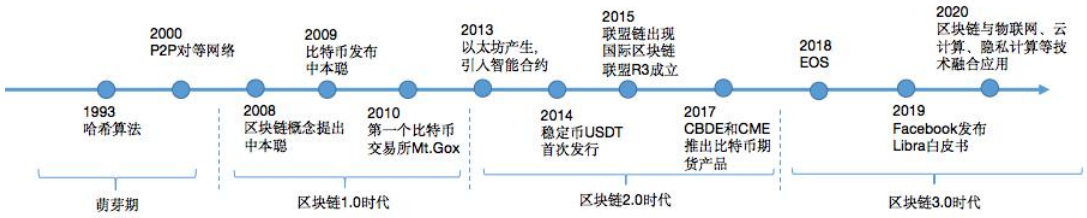

图 4：区块链技术发展历程

(1) 区块链 1.0 时代：比特币的诞生标志着区块链 1.0 时代到来的标志，这时期区块链技术主要应用在数字货币的兑换、转移和支付等方面。区块链技术的早期发展与数字货币密切相关，区块链技术创建了一套去中心化、公开透明的交易记录总账，由所有的网络节点共享数据库，并防止对数据内容的篡改。在这个时期，随着区块链技术的完善和应用的扩展，区块链在金融领域掀起了一股巨浪。在转账汇款和数字化支付相关领域,区块链技术备受关注在这些领域,传统方式要通过银行等中心机构进行开户行、对手行、清算组织、境外银行(代理行或境外分支机构)等烦琐的处理过程,时间长,成本高。应用区块链技术后,支付可以实现端对端的交易,去掉了繁冗的中间机构处理环节,不仅快捷,而且交易成本非常低廉。尤其在境支付方面,基于区块链的支付系统可为用户提供全球范围的跨境、任意币种的实时支付清算服务,跨境支付将以低成本方式瞬间完成。

(2) 区块链 2.0 时代：以太坊的出现标准着区块链 2.0 时代的到来。以太坊（Ethereum）[12]是一个开源的有智能合约功能的公共区块链平台，通过其专用加密货币以太币提供去中心化的以太虚拟机来处理点对点合约。而区块链 2.0 解决的问题则是市场的去中心化，这个时期的关键词是“合约”。“智能合约”是以太坊的显著特点之一，是可编程货币和可编程金融的基础。智能合约是以数字形式定义的一系列承诺，一旦合约被设立,在区块链系统上无须第三方的参与便可以自动执行智能合约，充分体现了程序员一直信奉的“代码即法律”。在以太坊出现,智能合约被广泛应用,以坊为智能合约提供了一个友好的、可编程的基础系统。而智能合约顺利执行的前提条件是已定的合约内容不能被篡改,且执行过程要公开透明，所以值得信任。区块链技术出现以后,非中心化、防篡改、集体维护、可追溯等特性成为智能合约天然的共生环境。以太坊为代表的新一代区块链应用与智能合约紧密结合后,区块链技术得以再次提升。在以太坊上,一份智能合约被创建之后便依靠程序自动执行,并且没有人能够阻止其运行,以太坊上的智能

合约能够控制系统中的各种数字资产,进行复杂的算法和操作。也因此区块链 2.0 的应用延伸至期货、债券、对冲基金、私募股权、股票、年金、众筹、期权等金融衍生品。公证文件、知识产权文件、资产所有权文件等电子化的进程与区块链的结合,也让有形或无形的资产在区块链上都找到了可能的运行环境。

(3) 区块链 3.0 时代：区块链 3.0 时代则是区块链全面应用的时代，区块链技术和实体经济、实体产业结合，将链式记账、智能合约和实体领域结合起来，由此构建一个大规模协作社会，实现去中心化的自治，更大范围的发挥区块链的价值。区块链技术在这一时代的应用将超越金融领域，可以广泛应用于政务、物流、医疗等各个领域，显著改进这些领域的服务流程，甚至颠覆这些领域内的传统商业模式。区块链 3.0 可提供包括身份认证、公证、仲裁、审计、域名、物流、医疗、邮件、签证、投票等应用辅助能力，应用范围扩大到了整个社会，区块链技术有望成为“万物互联”的一种最底层的协议。此外，区块链 3.0 的出现也面向当前行业有关隐私、可扩展性和互操作性等问题提出解决方案，进一步提升区块链技术的普适性。

第三章 隐私计算与区块链融合价值

隐私计算实现数据可用不可见，区块链确保计算过程和数据可信性，两种技术相互融合，相辅相成，实现更广泛的数据可信、安全协同。隐私计算与区块链的融合应用，由区块链技术解决信任问题，隐私计算技术提供数据安全、隐私保护的保障，有利于未来形成各大规模的数据要素流通网络，进一步推动数据资产化的发展，同时可以重新确定数据责权以及利益分配机制。

3.1 隐私计算增强区块链应用数据隐私性

区块链通过采用智能合约、共识机制、密码学等技术融合应用，实现不可篡改、可追溯、去中心化、公开透明等特性，这些特性有效解决信任问题。但同时也给区块链带来开放性。区块链本质上是一个分布式账本系统，需要不同节点对交易和状态进行验证达成共识，而在共识层面没有涉及隐私保护，参与方可以访问公开的区块链上的数据，这种开放性的同时也带来了隐私保护的问题。在面向需要保护的隐私数据（例如资产信息、交易金额等）情况下，区块链技术与隐私计算技术的融合应用，可以通过隐私计算技术的数据安全保障来增强区块链应用的数据隐私性。

区块链通过与多方安全计算等隐私计算技术融合应用能够保证链上数据的安全性，例如将多方安全计算应用在授权、管理密钥、共识机制中，增强输入数据的隐私性，可以提升区块链的安全性，加强跨机构的协作信任。使用可信执行环境可对链上数据进行保密计算，并对智能合约进行隐私保护，相较于其他的传统加密技术，计算损耗更少，可以同时提高区块链的效率和安全性。

此外，区块链受限于链上计算资源，数据处理能力有限，往往导致难以承载大数据量的计算要求，尤其是在高吞吐量和高时效性要求的情况下，通过与隐私计算技术结合，可以实现链上链下协同，实现在保证数据隐私安全前提下优化链上计算效能。复杂的跨机构、跨节点计算任务可以通过链下的隐私计算技术运行，区块链则主要侧重于提供存证、确权、利益分配等可信性增强能力。

3.2 区块链技术增强隐私计算应用可信性

隐私计算是以联邦学习、多方安全计算、可信执行环境为主要技术路线的跨学科融合技术，通过广泛采用现代密码学技术和信息安全技术，实现在保证原始数据隐私性的同事，完成跨机构之间的数据联合建模、联合统计、联合分析等协同应用。通过隐私计算技术，保障了跨机构间数据协作的数据安全性，但在数据要素流通共享的全流程中，还存在数据确权、数据定价、存证监管、可验证性等问题，这些问题正好通过区块链技术来进行有效互补，来提升隐私计算技术应用过程的可信性。

区块链技术可以通过共享账本、智能合约、共识机制、加密机制、分布式数字身份等区块链特性，进行区块链与隐私计算技术融合，可以实现原始数据的链上存证核验、计算过程关键数据和环节的上链存证回溯，确保计算过程的可验证性，从而赋能可信隐私计算具有不可篡改、可追溯、可验证性、可控性、开放性，保证在实现数据隐私保护的同时增强隐私计算过程的可信性和可验证性。

隐私计算的任务信息、数据申请信息、授权信息、参与计算的数据信息、流程信息、过程数据、计算结果等可通过区块链进行记录与存储，存储在链上的信息可通过其他参与方对数据进行签名确认的方式，进一步提高数据可信性，同时可通过对哈希值的验证匹配，实现信息的不可篡改性，此外通过智能合约实现按唯一标识对链上数据进行关联，通过区块链的这些特性来增强隐私计算的可追溯性和可审计性。

隐私计算需要第三方可信协调方的计算任务可通过区块链完成，例如密钥生成与分发、模型的聚合、模型贡献度计算等，通过区块链的智能合约逻辑公开透明增强隐私计算任务的计算信任。在实际落地应用中，往往很难找到可充当可信第三方的机构，而区块链的去中心化共识机制可以很多的承担这角色，有效解决部分隐私计算技术需要可信第三方的难题。

在多方参与的隐私计算任务中，可通过区块链实现分布式数字身份和可验证声明功能，形成多方参与的共识任务联盟，实现对参与方的数字身份管理，通过区块链实现数据合作参与方身份的真实可信。

数据确权是隐私计算技术需要解决的一个问题，通过区块链技术，对隐私计算任务使用的数据和数据提供方的数据进行关联，实现数据提供方对数据资产权益的公开确权，数据提供方通过签名双方声明数据的所有权，通过加密算法规定数据的使用权，并通过区块链智能合约将数据使用限制到具体的隐私计算算法和使用次数，实现可信隐私计算的可控可计量。

此外，通过区块链技术进行可信隐私计算的分布式数据资源共享，参与机构将需要共享的数据资源摘要，通过区块链的共享账本技术进行参与方之间的可信共享，并基于不同应用场景的数据交换共享需求，建立面向隐私计算的可信安全数据共享交换机制，增加可信隐私计算的数据资源的可控性。

第四章 隐私计算与区块链融合典型方案

4.1 区块链和隐私计算链上链下双层融合架构

隐私计算与区块链融合应用有多种实现方式，其中目前最常见的一种方式是链上链下双层异构结构。其中链上主要是区块链网络层，链下主要是隐私计算层，通过这种方式，既保持了区块链技术架构和隐私计算技术架构的独立性，有实现了那种技术能力的融合互补。

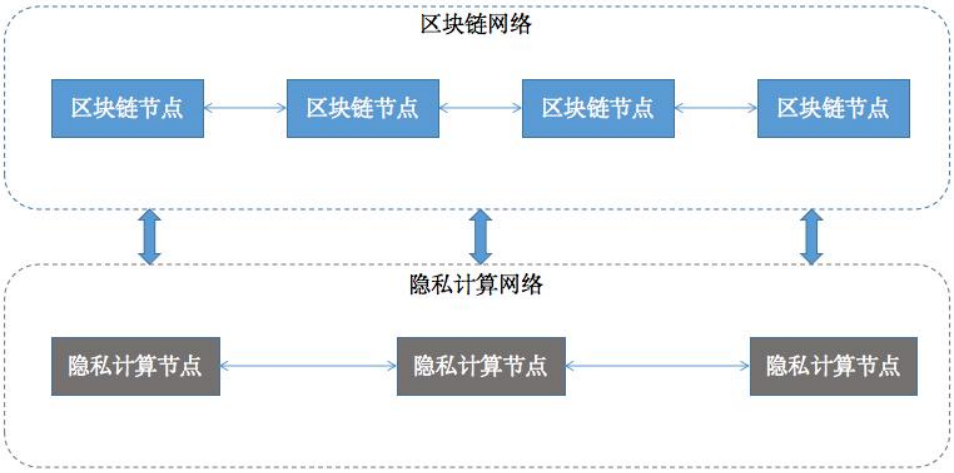

图 5：区块链和隐私计算的链上链下双层融合架构

“链上+链下”分层机制，将数据信息流通和数据使用记录存证等放在链上实现，链下实施多方数据的安全联合计算、安全联合建模、安全联合预测、安全联合查询等，链上存放安全计算的可验证信息。这种链上数据可信存证协同、链下数据安全加密计算双层异构组网的分层机制将数据安全共享流转全流程合理地分配到链上和链下，最大化发挥区块链和隐私保护计算各自独特的性能优势，从而提升网络数据共享流通协作效率，并降低应用成本。

4.2 区块链与隐私计算技术的协同机制

隐私计算可以通过密码学来保证计算过程中原始数据的隐私安全性，但隐私计算的安全保护知识实现数据安全合规共享流通中的一环，仍无法解决各参与方的信任问题，还存在数据确权、数据定价、数据流通存证监管、数据交易恶意节

点探查等问题。通过隐私计算与区块链的融合，让计算前身份真实、授权有效，计算中节点账本同步更新、实时监控，计算后关键数据要素和操作可审计可追溯、账户结算公正合理，从而保证了数据要素链上链下流转全过程的安全可信。

区块链与隐私保护计算通过设计“链上可信存证协同、链下安全加密计算”双层异构组网协同机制，通过区块链提供分布式信任机制，隐私计算提供数据安全、隐私保护的协同计算任务的安全保障。通过整合区块链和隐私计算的双方优势，合理分配链上链下计算能力和安全能力，从而高效可信地实现业务流程协同和计算过程协同，降低了多节点间大规模复杂计算的部署和使用成本。

从整体融合应用来看，可分为业务流程协同和计算过程协同两方面：

a.业务流程协同：数据持有者将共享数据目录、数据使用申请、数据使用审批、数据使用审计等功能在区块链上完成，利用智能合约技术来实现业务自动化处理逻辑，提升数据共享流通、业务有序运转的协作效率。

b.计算过程协同：对于零知识证明这类证明过程性能要求低但算力消耗大的隐私计算技术，可采用链下证明链上验证的方法，由发起方先在链下计算出证明内容，利用硬件加速、可信执行环境、算法优化等手段来提升证明效率，再将证明传递到区块链上的接收方节点，在链上验证该证明内容的正确性。

隐私计算技术和区块链技术的融合应用，包括链下数据隐私保护协同计算，链上通过数字身份、加密算法、数字账本、激励机制等实现数据要素安全流转。

4.2.1 隐私保护链下数据协同计算

链上链下双层结构下的数据隐私保护和协同应用，可以采用“链下计算、哈希上链”的轻量级实现方案。链下采用多方安全计算、联邦学习、可信执行环境等隐私计算技术将区块链上的隐私数据可见信息与使用价值剥离，在数据采集、传输、存储、使用展示、开放共享、销毁的数据全生命周期实现闭环保护。而数据信息或计算任务对应的摘要信息、过程信息等则保存在链上，并采用哈希算法等保证只有摘要信息且无法反推出原始数据或任何中间结果。

数据协同应用的过程，由链上智能合约发起，再通过链下隐私计算系统完成跨机构间的数据价值共享的计算任务，并通过链上摘要信息完成计算过程的存证和审计，完成链上、链下资源相结合的数据流转。通过这种方式，可实现将原始数据和计算过程实现链上、链下的隔离，补齐区块链的隐私保护能力，而链下的隐私计算平台只需要负责数据价值的共享和高开销的计算任务。

隐私计算技术以“数据可用不可见”的方式来完成跨机构间的数据协同计算，尤其是多方安全计算、联邦学习等技术，保证了原始数据不出本地，增强了数据的可控性。这些技术本质上是侧重于计算过程中的保密性，而无法确保数据的正确性，而在隐私保护协同计算基础上，通过区块链可以帮助隐私计算任务实现全流程的可追溯性，协助跨机构间完成可信、安全、可控的协同数据应用。

4.2.2 分布式数字身份数据资产确权与控权

通过分布式数字身份技术对人、机构、设备、数据资产等进行标识，将现实事物实体数字化，身份信息传输和使用秉持最小化原则。系统先使用数据标签、数据指纹等方式为每个实体生成统一标准的唯一身份标识符，身份信息可在不同联盟机构之间进行传递。然后通过技术手段和共识机制对参与共享计算的数据或结构主体进行交叉真实性验证，在链上将数据身份与数据持有者数字身份进行关联，实现数据持有者对数据资产权益的公开透明，一定程度避免了主观作恶、合谋推导、数据造假等。数字身份支持对实体身份属性数据进行增删改查维护，也可支持对身份凭证进行发行、验证、撤销、模板维护以匹配监管、审查机构要求。通过将分布式数字身份与数据资产捆绑，利用交叉校验来确保参与方身份和数据正确性与真实性，唯一身份标识利于各参与方建立良性秩序，并对区块链上数据资产进行有效确权和控权。

4.2.3 数字账本实现可信存证与审计

利用区块链可追溯、不可篡改特性，在整个隐私保护计算过程中将关键操作记录和数据要素加密后在区块链的各节点本地数字账本中记账存证，便于后续外

部监管机构或内部审计部门监管审核，也便于发生矛盾时责任方追责和系统故障恢复。

(1) 操作记录账本：将关键环节如数据生成及采集合法性验证、数据处理存证和共识、数据使用授权、数据可信流转、数据安全加工和协作以及数据跟踪和审计等详细记录在区块上，并在所有节点的本地账本上做实时同步更新。

(2) 数据要素账本：将涉及的中间数据摘要、判断结果、计算结果等关键数据要素在区块链上存档，支持被授权的节点进行查询和校验。

4.2.4 分布式数据目录共享

利用区块链的不可篡改性、可追溯性、多方共识的特性，可以通过各方参与的数据资源目录的建设、存储和维护，来有效解决隐私计算中数据资源发布与共享问题，并增强数据目录的可控性。

参与机构将需要共享的数据资源信息，如数据名称、所属机构、访问方式、发布时间、数据元信息等，发布到区块链的共享账本上，并提供数据目录信息在链上的查询、管理、维护，增强目录的可控性和提高数据治理能力和信息标准化服务水平，形成一个具有不可篡改性的去中心化可信数据资源共享目录。

区块链的共享账本可避免数据提供方作出随意变更数据的行为，同时区块链保存了数据资源目录的历史发布和变更记录，实现可追溯性，增强了数据资源共享的可控性。

4.2.5 可信群体激励机制

隐私计算主要是用于多方的数据协同应用，解决的是多方协同计算过程的数据安全、隐私保护的问题。其自有的两方或多方协作的特点，必定要求参与方按照约定好的协议或规则执行，协同任务的成果以及后续的收益也需要通过互利共赢的方式进行分配，只有在此基础上才可能推动技术的持续应用。而隐私计算目前主要依赖半诚实的安全模型，参与方之间需要有够高的信任度，对于这种复杂的多边关系，目前的信任成本较高，合作摩擦系数较大，容易增加合作的难度，

或硬性合作范围的扩大，难以形成规模效应。区块链技术去中心化合作机制，有利于建立多边信任机制，适合作为隐私计算技术的补充来建立可信的群体激励机制。

在隐私计算的数据提供、结果查询、算力支持等环节设计激励机制，与区块链账户关联，按各参与方的贡献量或消费量对与之捆绑的区块链节点账户进行数字积分的奖励或支付。公正合理的激励机制能够激励各节点贡献数据、模型或算力，正向诱导潜在恶意节点，例如：贡献量按数据提供的数量和维度、参与方本地数据对最终模型效果的影响占比、每天计算参与次数等综合进行计算;消费量按每天数据查询次数、任务发起条件等进行核算。利用区块链技术实现可信的奖励机制，可以鼓励更多高质量的参与方参与到模型训练、联合计算任务中贡献算力、算法、数据，提高模型准确度或计算结果可用性。

隐私计算与区块链技术的融合互补，有效拓展数据要素协作的应用边界，解决数据要素流通全流程中的可信性、安全性、隐私性、可监管性等问题，是技术发展的必然趋势。

4.3 区块链与隐私计算融合应用案例

4.3.1 隐私计算与区块链融合平台助力小微企业风控

小微企业联合风控是隐私计算与区块链链融合在金融机构的一个重要应用场景。近年来，大数据和人工智能技术广泛应用于银行借贷的多方风控场景。银行仅拥有客户的行为数据、征信数据、外部三方数据等，难以满足业务风控的需求。此外，由于不同机构间拥有的数据相互独立不互通，银行机构之间、银行机构与运营商、银联、政务数据平台等机构之间的数据无法直接打通。而此类数据包括个人通话情况、位置信息、企业用水用气等数据无法直接落地于银行场景，出现一系列信息不对称、风险识别不精准、融资成本高等问题。

通过隐私计算与区块链融合技术，可以有效打破数据孤岛，实现跨机构间数据价值的联合挖掘，管理和控制数据权限，全方位评估客户的综合情况，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。由各

数据提供方自定义信用计算模型，在隐私计算技术支持下安全聚合多方数据，计算出准确可信的信用评分。同时，基于区块链技术对链上计算存证、确权、交易记录上链，保证数据不可篡改，数据使用和流转可追溯、可审计。

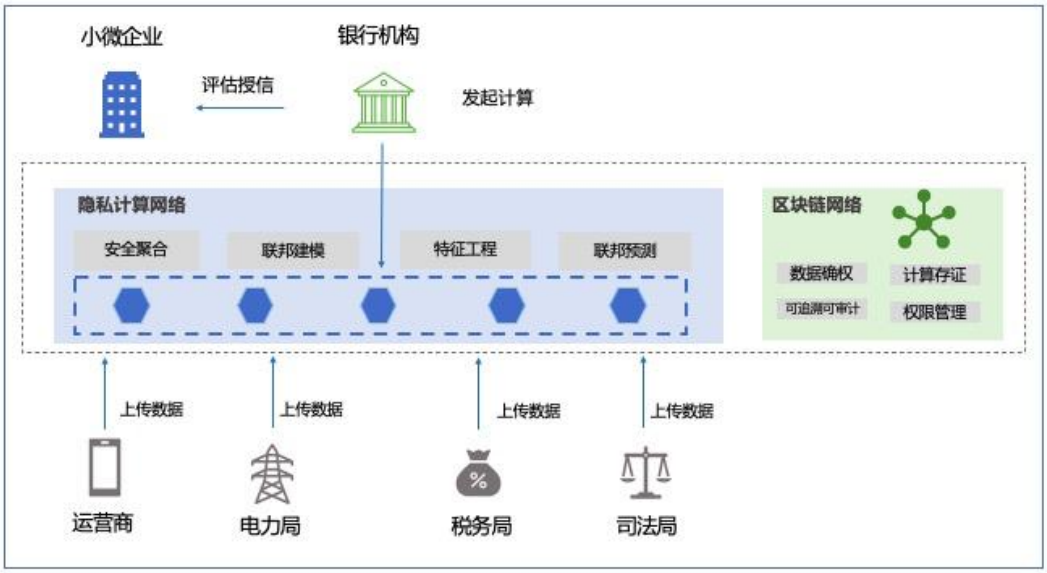

图 6：隐私计算与区块链融合平台助力小微企业风控

4.3.2 隐私计算与区块链融合政务数据开放应用

自 2020 年数据被定义为生产要素，探索公共信息资源开放，更好促进资源的社会化利用，成为智能政务的一个重要方向。隐私计算与区块链融合应用可助力于政务数据的开放共享。区块链作为一项新型信息技术，具有多中心、防篡改、可追溯、非对称加密等特性，正好契合数字政务对于数据共享以及数据流通的安全性和可信性的需求。通过共识机制构建一个多方参与的信任网络,进一步实现互联网与政务的深度融合,优化政府业务流程,使得政务公开真正走向阳光、透明、可信。通过进一步融合隐私计算技术，能够以安全合规的方式接入政府数据、企业自有数据和第三方数据。对于数据需求方，通过隐私计算技术实现“数据可用不可见”的所有权和使用权的分离，并通过一整套产品技术及权限管控保证政务数据共享、数据流通、融合计算全流程的数据使用安全、高效、可控。政务数据的开放，可以使用联邦学习及多方安全计算，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，进行数据使用和机器学习建模。

基于隐私计算和区块链的智能政务数据开放平台，可以赋能小微金融场景，例如以安全的方式帮助金融机构评估企业信用，为信用良好企业提供信用贷款，破解企业融资难、融资贵痛点；可以应用于商业选址应用场景，例如向第三方商业服务选址专业机构开放人口、交通、规划、兴趣点、房价、教育等空间数据、融合第三方机构自有数据（如手机信令）、选址模型算法（综合评估客流、年龄、消费、关注门店分布等因素），为连锁企业提供科学选址服务。利用政务开放的数据，类似的做法还可以广泛应用在健康医疗、家政服务、AI 智能客服、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，引入外部技术人才，与企业数据融合计算，最大限度造福社会。

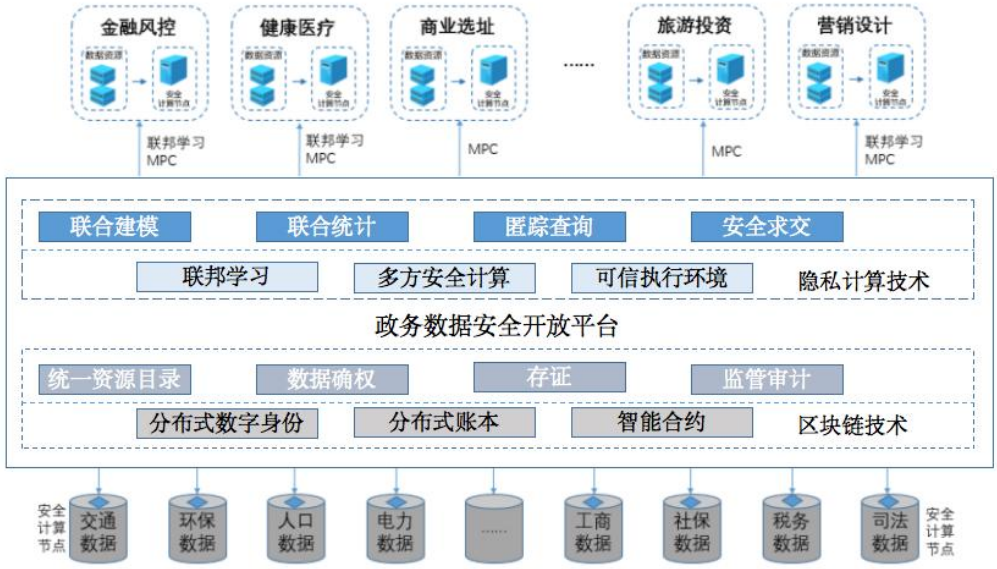

图 7：隐私计算与区块链融合政务数据开放应用

第五章 展望与建议

区块链和隐私计算技术在近几年快速发展，但目前在各自领域分别面临着隐私性、可信性等挑战，通过融合区块链和隐私计算技术，实现技术优势互补，可以有效解决数据孤岛、数据安全等多方合作中的数据安全问题，同时解决多方协作的信任问题，形成“区块链+隐私计算”的融合模式，解决实际落地应用中的可信性、安全性、可监管性挑战。

区块链技术具有“去中心化”、“分布式数据存储”、“可追溯性”、“防篡改性”、“公开透明”等优势特点，能够有效解决数据领域的数据真实性、安全性与开放性问题，通过建立可信任的数据管理环境，防范和避免各类数据造假、篡改、遗失等数据管理问题，促进数据的高效共享与应用。与此同时，可行的隐私保护改进方案也是目前区块链 3.0 需要解决的一个重要发展方向，但无论从市场商业竞争角度还是个人信息安全角度来看，都没有人希望自己的数据完全公开、透明。而隐私保护合规也是数据管理领域的一条重要安全保障，一方面保护着数据所有者的隐私安全，另一方面也影响着数据流通共享的效率与发展。通过这两种技术的有效结合，利用区块链技术建立了数据时代的信任机制，结合隐私计算在数据共享中确保安全的隐私保护防线。

通过多方安全计算、联邦学习、可信执行环境等隐私计算技术与区块链技术联合应用，能有有效联接“数据孤岛”，以技术手段来解决数据隐私保护，以数据“可用不可见”的方式来实现数据价值流通和释放，守住安全合规底线，这也是数据要素流通的前提与基础。这些技术可以实现机构间数据能力的安全开放、汇聚融合、协同洞察，以匿踪查询、安全计算、联邦学习等方式提升服务智能和效能的上限。同时利用区块链技术对数据从产生到处理，从交易到计算进行全生命周期的记录与存证，保证过程的可验证性和可信性，从技术层面实现数据的“可控可计量”。隐私计算与区块链的融合互补，拓宽数据协同应用的边界，有效解决数据流通过程中的隐私保护难题，明确数据所有权、使用权、收益权等权属边界，为数据要素流通提供可信、安全的全链路保障，必将成为实现多方可信协作、数据价值互联互通的基础技术设施。

基于以上技术搭建实现隐私保护计算的区块链数据协同基础平台，突破区块链网络高性能、高安全性和高可扩展性瓶颈，在金融、医疗、政务、能源、通信、教育、工业互联网等众多领域应用，能够在满足用户及企业信息隐私保护、安全存储的同时，大幅度推进相关技术的应用落地，以数据安全赋能支持国家发展战略。

目前区块链和隐私计算的融合应用处于初期阶段，暂未形成大规模广泛化的数据流通网络。一方面是由于目前的实际应用需求尚未形成规模性的爆发，另一方面区块链和隐私计算的技术成熟度仍有进一步提升空间。由于两种技术均面向数据安全性和数据可控性问题，国产自主可控是一个重要的研究方向，与信创领域的协同发展必不可少。此外，性能提升是促进广泛化应用的重要基础，区块链和隐私计算均具有分布式特性和采用了大量密码学技术，多节点间的交互带来的通信效率问题以及算力问题，尤其是未来广泛化应用后带来的数据量爆发式增长，和参与节点的急剧增多，隐私计算和区块链技术应用的性能问题将面临巨大的挑战，当前迫切需要从算法层面、通信层面、软硬结合层面以及多技术融合层面进行全维度的性能提升探索，以推动隐私计算和区块链技术的融合应用进入新一轮的技术与应用创新周期。

参考文献

[1] Yang Q , Liu Y , Chen T , et al. Federated Machine Learning: Concept and Applications[J]. ACM Transactions on Intelligent Systems and Technology, 2019, 10(2):1-19.

[2] Evans D , Kolesnikov V , Rosulek M . A Pragmatic Introduction to Secure Multi-Party Computation[J]. Foundations & Trends® in Privacy & Security, 2018, 2(2-3):70-246.

[3] R, L, Rivest, et al. A method for obtaining digital signatures and public-key cryptosystems[J]. Communications of the Acm, 1978.

[4] Siim A S , Pullonen S P , The Simplest Protocol for Oblivious Transfer. International Conference on Cryptology and InformationSecurity in Latin America. Springer International Publishing, 2015.

[5] Kolesnikov V , Schneider T . Improved Garbled Circuit: Free XOR Gates and Applications[J]. Springer-Verlag, 2008.

[6] Gamal T E . A public key cryptosystem and a signature scheme based on discrete logarithms[J]. IEEE Transactions on Information Theory, 1984, 31:469-472.

[7] Chakrabarti S , Knauth T , Kuvaiskii D , et al. Trusted execution environment with Intel SGX[J]. 2020.

[8] Pirker M , Slamanig D . A Framework for Privacy-Preserving Mobile Payment on Security Enhanced ARM TrustZone Platforms[C]// IEEE International Conference on Trust. IEEE, 2012.

[9] Dwork C . Differential Privacy: A Survey of Results[C]// Springer, Berlin, Heidelberg. Springer, Berlin, Heidelberg, 2008.

[10] Goldwasser S , Micali S , Rackoff C . The Knowledge Complexity of Interactive Proof Systems[J]. SIAM Journal on Computing, 1989.

[11] Kosba A , Miller A , Shi E , et al. Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts[C]// 2016 IEEE Symposium on Security and Privacy (SP). IEEE, 2016.

[12] Adhikari C . Secure Framework for Healthcare Data Management Using Ethereum-based Blockchain Technology. 2017.