·科学论坛·

隐私计算的关键理论与前沿应用*

郑志明[1] 何积丰[2] 唐立新[3] 童咏昕[1]
陈 厅[4] 谢 国[5] 王志衡[5] 吴国政[5]**

1. 北京航空航天大学，北京 100191
2. 同济大学，上海 200092
3. 东北大学，沈阳 110819
4. 电子科技大学，成都 610000
5. 国家自然科学基金委员会 信息科学部，北京 100085

[摘 要] 基于国家自然科学基金委员会第 352 期“双清论坛”，本文介绍了隐私计算在服务国家数字经济重大战略需求的重要作用，分析了国内外对该研究聚焦点的布局情况，阐述了隐私计算的内涵与外延，回顾了隐私计算研究历程的三个维度：隐私加密、隐私脱密与非规则博弈下新型隐私计算，凝练了隐私计算的非规则博弈、全生命周期保护与联邦计算范式等关键科学问题，研讨该领域前沿应用与未来研究方向，力争进一步推动我国隐私计算理论、方法和技术的突破性发展与应用。

[关键词] 隐私计算；加密与脱敏；非规则博弈；全生命周期；联邦计算

1 隐私计算研究概述

1.1 隐私计算发展背景

随着数字化与经济社会的深度融合，近年来隐私泄露事件频发，中小企业、普通民众都受到骚扰和侵害，严重者更直接威胁到国家安全和发展。2023 年 7 月，习总书记在南京考察时曾关切询问所提“信息安全的根问题”的解决进展。中国共产党中央委员会（以下简称“党中央”）和国务院在《“十四五”规划》与《数据二十条》中也多次强调加快培育数据要素市场，建立健全数据安全基础制度。解决数据要素流动与隐私保护的矛盾已成为服务国家数字经济战略的重要挑战。

隐私计算作为构建数字化社会信任底座、保障数据要素安全流通的关键技术，已成为当前研究的聚焦点。其技术起源实则由来已久，发展至今日其研究历程主要包含三个维度。维度一：隐私加密，主要结合密码学工具保护个人隐私敏感信息，维持传统数据加密的惯性，尽管安全性较高但性能表现难以落地应用。维度二：隐私脱敏，通过更改数据、增强信息不确定性来实现数据隐私保护，是一种平衡数据隐私性与计算可用性的折中方案，尽管缓解了计算与通信开销但损失了计算精度与效果。然而，以上技术难以满足数字化时代下的数据流通、社会治

郑志明 教授，中国科学院院士，万人计划国家教学名师，中关村国家实验室数据与应用领域首席科学家，中国通信学会副理事长，复杂关键软件环境全国重点实验室主任，国家区块链技术创新中心首席科学家，载人航天空间站软件专家组组长，未来区块链与隐私计算高精尖创新中心主任。

吴国政 博士，国家自然科学基金委员会信息科学部二处处长，主要研究方向为人工智能。

收稿日期：2024-02-26；修回日期：2024-07-15
* 本文根据国家自然科学基金委员会第 352 期“双清论坛”讨论的内容整理。
** 通信作者，Email：wugz@nsfc.gov.cn

理等新需求，郑志明院士等[1\]率先提出了第三个维度后纳什时代下面向非规则博弈场景下的新型隐私计算，针对群体面临数据价值驱动的复杂博弈与安全挑战，聚焦于打破理性假设下的隐私度量、价值激励与行为引导的新型隐私计算体系。

1.2 国内外布局情况

针对隐私计算技术日新月异的发展，目前世界多国争相从项目和政策方面战略性部署这一领域。在项目布局上，多个国家均先后设立了隐私计算相关项目，例如：2020 年，美国国防高级研究计划局设立了面向虚拟环境中的数据保护的项目（Data Protection in Virtual Environments，DPRIVE）；2022 年，欧盟委员会支持了构建隐私基础设施的 Nymtech 项目；日本数据保护机构个人信息保护委员会也推出 PIA（Privacy Impact Assessment）项目评估隐私影响风险。在政策法规上，欧美已经较早开展了相关法律法规的探索实践。美国从各州开始逐步自下而上推动了国家联邦层面立法：2018 年，加州较早颁布了《加州消费者隐私保护法案》，随后弗吉尼亚州和科罗拉多州等地积极出台相关草案并推动立法，直至 2022 年美国两院首次公布《美国数据隐私保护法案》（American Data Privacy and Protection Act，ADPPA）实现国家层面的数据保护立法。欧盟则从数据保护出发逐步完善了数据服务与流通交易等全方位的制度设计：在 2018 年与 2019 年分别正式实施《一般数据保护条例》（General Data Protection Regulation，GDPR）与《非个人数据自由流动条例》（Regulation on the Free Flow of Non-personal Data，RFND），是国际上最受关注的数据保护与流动法律文件，后续立法提案了《数字服务法》《数字市场法》《电子隐私条例》与《人工智能法案》，进一步完善数据资源的使用与监管体系。2022 年，欧盟因数据违规执行的罚款额已达 22 亿欧元。

我国与欧美相比，较早完成了相关政策法规的顶层设计，并着手布局相关项目。在政策法规上，2017 年我国正式实施《网络安全法》，并在 2021 年正式施行了《数据安全法》与《个人信息保护法》，为数据隐私安全提供了法律准绳。党中央和国务院在《“十四五”规划》中提出数字中国战略，并先后印发了《要素市场化配置综合改革试点总体方案》《“十四五”数字经济发展规划》《关于加快建设全国统一大市场的意见》与《关于构建数据基础制度更好发挥数据要素作用的意见》等，持续完善数据要素发展与管理体系。在项目布局上，国家自然科学基金委员会（以下简称“自然科学基金委”）在数据隐私保护方向资助国家杰出青年科学基金项目 1 项、优秀青年科学基金项目 6 项、青年科学基金项目 360 项、重点项目 16 项、面上项目 398 项、地区科学基金项目 43 项、联合基金项目 46 项、国际合作项目 19 项，总计资助 893 项，资助金额 5.2 亿元。此外，科学与技术部也在“十四五”重点研发计划“网络空间安全治理”中对相关技术进行了布局，如针对大数据平台隐私保护、多媒体数据隐私保护及分布式学习的隐私保护等方向资助了“隐私计算及安全保障基础理论研究”等项目。2023 年，国家数据局的正式挂牌成立恰逢其时、意义深远，这将进一步加快全国统一、辐射全球的数据大市场的建设，推动数字经济加速发展。

欧美在隐私计算技术的理论、硬件与软件等方面具有较大先发优势，形成了一定的技术壁垒。而我国在隐私计算领域研究和实践起步较晚，缺乏原始创新与基础理论突破，人才储备不足，缺少具有重大影响力的研究成果，成果转化率也偏低。因此，为服务数字中国重大战略需求，推进数字经济和社会治理现代化建设，自然科学基金委组织领域专家和学者深入讨论“隐私计算”中存在的科学问题，研究以“隐私计算”为主题的重大类项目的科学性，研究隐私计算的新理论、新方法和新平台，探索隐私计算新模式和新范式，增强数据隐私计算的安全性、可靠性、高效性，防范和化解我国信息安全领域面临的重大风险，充分发挥数据要素潜能，助力我国数字化建设。

1.3 隐私计算的内涵与外延

隐私计算是在保护敏感信息隐私安全的前提下完成数据处理与计算的一类技术，是泛在网络空间隐私信息保护的重要理论基础。根据场景与隐私保护力度不同，一般可分为隐私加密与隐私脱敏两类。隐私加密主要保证数据的机密性、完整性、不可否认性等，多使用安全多方计算与同态加密等隐私加密等手段进行处理，具有信息无损性的特点，但缺点是计算量大、扩展性差。隐私脱敏则主要是接收方难以获取发送方的完整信息，多使用匿名化与差分隐私等手段更改数据从而保护隐私，但缺点是会损失数据信息、影响计算精度。

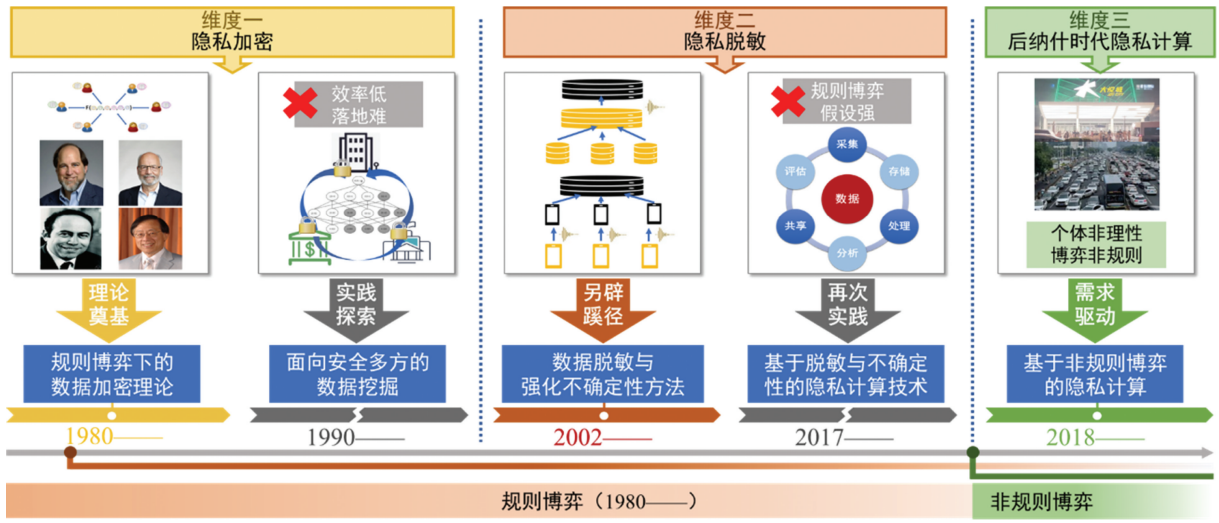

图 1　隐私计算发展历程脉络图

随着数据使用需求的迭代，不断拓展了隐私计算的外延。中国信息通信研究院在 2021 发布的《隐私保护计算与合规应用研究报告》与《隐私计算白皮书》中指出，隐私计算是面向数据采集、传输、存储、处理、共享、销毁等全生命周期中隐私保护的计算理论和方法，可以在不泄露数据提供方原始数据的前提下，完成对数据的分析计算，是一套在数据所有权、管理权和使用权分离时隐私度量、隐私泄露代价、隐私保护与隐私分析复杂性的可计算模型与公理化系统。不仅是对数据隐私与安全的有机融合，还需对区块链、联邦计算与可信硬件等多种技术有机结合。更进一步地针对群体非理性行为下的复杂博弈场景，探索突破原有隐私计算假设的基础理论创新，尝试与经济学、社会学及心理学等多个学科的交叉研究。

2　隐私计算研究历程

2.1　维度 1：隐私加密

隐私加密是基于数据保密的传统惯性，使用密码学工具对个人信息等敏感隐私加密，参与方遵守密码学协议规则完成计算任务。其代表性技术有安全多方计算与同态加密等。安全多方计算是指多个数据拥有方在不泄露原始输入数据的前提下共同完成目标函数计算。最早起源于 20 世纪 80 年代姚期智院士[2]提出的百万富翁问题，即两位百万富翁希望在不暴露自己财富值的前提下比较谁更为富有。针对这一问题发展出混淆电路、秘密共享与不经意传输等多种技术[3-5]。同态加密是一种特殊的加密协议，其同态性需满足在密文上进行特定运算的结果在解密后与明文直接进行运算结果相同[6, 7]。

隐私加密类技术尽管具有较高的安全性，然而其计算与通信开销高昂，难以适用于大数据时代的计算场景。20 世纪 90 年代，数据挖掘与隐私加密两个领域展开的交叉研究聚焦于保护多方联合数据挖掘过程中的隐私安全，涌现了基于安全多方计算的频繁模式挖掘、决策树分类与 K-Means 聚类等方法[8-11]。这一时期的研究主要受限于当时安全多方计算技术，而且算法效率评估以理论分析为主、缺乏实践验证。数据挖掘的全局性随机知识发现与隐私加密的局部性隐私安全保证间的根本矛盾日益凸显，二者难以折中。因此，该类技术后续研究逐渐聚焦于以落地应用为目标的通用工具库的构建[12-15]与高效专用协议的设计[16-18]。

2.2　维度 2：隐私脱敏

隐私脱敏是一种平衡数据隐私性与计算可用性的折中方案。为缓解隐私加密的高计算成本，数据脱敏方式旨在通过更改数据从而增强信息不确定性，保护数据隐私。其代表性技术有匿名化与差分隐私等。21 世纪初陆续出现 k-匿名化、l-多样性与 t-保密性等一系列隐私脱敏技术，其核心思想旨在缩小数据尺度、增强数据间的不可区分性[19, 20]。例如：k 匿名方法通过抑制或泛化等方式更改数据后，可以使数据集中任一条都难以和其余 k－1 条数据区分，并根据数据的发布次数与类型而衍生出不同变种。然而，此类技术抵御攻击类型有限，且均缺乏较为完善的理论保证。2006 年针对差分攻击的差分隐私技术[22]出现，其核心思想是通过添加随机噪声保护统计计算结果，可以保证隐私预算内的相邻数据集不可区分性。差分隐私具有更为严谨的数学模型、支持更为通用的计算函数并能抵御更为复杂

的攻击，因而得到了应用与发展。

尽管隐私脱敏类技术计算开销低，但在一定程度上破坏了数据信息，损失了计算精度与效果。因此，隐私脱敏类技术在功能上有一定局限，需在数据隐私性与计算精度间寻求平衡，难以适用于数据全生命周期的隐私保护。近年来此类技术也在持续发展，针对不同数据场景衍生出本地化差分隐私等多个变种[23-26]。

2.3 维度 3：后纳什时代隐私计算

21 世纪 10 年代后，随着信息化的快速发展，大数据的使用场景衍生出更为多样化的隐私保护技术。例如，面向云计算与外包数据库场景的密态数据管理、面向云边端一体化的协同计算与基于可信硬件的机密计算等[27, 28, 36]。为开展数据要素市场化建设、充分发挥数据要素市场潜能，国务院在《要素市场化配置综合改革试点总体方案》中提出“原始数据不出域，数据可用不可见”的数据要素流通新需求。联邦计算以其“数据不动计算动”的核心思想，实现了兼顾“自治与跨域”的大数据隐私计算新方法，成为隐私计算新型范式之一，在金融风控、医疗卫生与城市交通等多个领域均探索落地应用[29-35]。

然而随着数据要素化的深入推进，数据的价值愈发凸显。从而导致在数据要素流通场景中面临由价值驱动的多方复杂博弈，为数据隐私安全带来新的挑战。在数据交易与社会治理等真实场景中，参与成员在数据价值等利益驱动下，难以建立多方间的信任体系，可能采取不理性行为，从而违反纳什均衡引导下的规则博弈假设，破坏现有隐私计算工具的安全模型[1]。针对这一问题，郑志明院士率先提出了后纳什时代下的新型隐私计算理念，主要聚焦于建立非规则博弈场景下的隐私度量、价值激励与行为引导体系。亟需融合非规则与规则博弈下的隐私计算方法，进一步结合心理信息学与社会信息学等交叉学科，研究新型隐私计算的基础理论，为构建数字化社会的信任体系提供技术底座。

3 主要科学问题与研究方向

隐私计算是构建数字化社会信任底座的关键技术。当前隐私计算主要面临在高可信、高隐私与高可用间的平衡，在高可信方面存在群体非规则博弈复杂动态的挑战；在高隐私方面存在全流程隐私需求异质多样的挑战；在高可用方面存在自治大数据跨域流通低效的挑战。

为应对上述挑战，需实现面向复杂博弈场景下的数据全生命周期隐私保护，构建隐私计算的新型基础理论与关键技术体系，全面提升隐私计算的安全性、可用性和可信性，有以下科学问题亟待解决。

3.1 融合规则与非规则博弈的隐私计算

不同主体间隐私需求多样化，行为复杂化，且在价值驱动下呈现动态复杂博弈态势。为解决该科学问题，建议对以下三个方向开展重点研究：新型数字化安全信任根问题、后纳什时代隐私计算数学理论、非规则博弈下隐私度量激励理论。

科学问题一：新型数字化安全信任根问题

信任作为社会构建并维持运转的基础，是既古老又前沿的重大研究方向。该领域是多学科交叉研究的汇聚点，文史哲学科研究面向社会共识的建立与演变；而政经法学科研究面向社会规则的构建与完善。然而，在数字化趋势下，需要使用信息化手段构建新型信任体系。数字化建设依赖各社会主体间的信息流转，但个体组成的社会本质是非充分互信的，在信息流转过程中易产生熵增现象，导致数据隐私泄露风险高、危害大，甚至动摇数字化社会的安全底座。因此，亟需刻画新型数字化社会下的安全信任根问题，建立信息技术与心理学、社会学有机融合的新型信任体系，为数字化建设奠定基础。

科学问题二：后纳什时代隐私计算数学理论

后纳什时代下，数据处理的环境和任务复杂，个体的行为和决策动态变化，信息不对称性和不确定性增加，而现有隐私计算技术多立足于纳什均衡引导下的规则博弈理性假设，在处理后纳什时代下大规模数据共享和隐私计算时具有一定的局限性。针对日益复杂的数据环境和隐私保护新挑战，需要构建后纳什均衡引导下的新型隐私计算理论。研究融合信息论、密码学和概率论的隐私计算理论，探索针对复杂动态隐私计算场景的加密算法、不确定性理论和隐私计算协议等，支撑安全可信的隐私计算技术研究。研究融合行为经济学和社会科学的隐私动态博弈模型，描述和分析用户随时间变化的隐私互动行为，为隐私安全保障和多方激励引导提供科学可行的解决方案。

科学问题三：非规则博弈下隐私度量激励理论

数字经济与社会治理等真实场景应用中个体行为并非完全符合传统经济理论中的理性行为模型，增加了隐私数据共享行为激励引导的难度。针对隐私计算中存在个体成员不诚实、群体行为不理性、多方博弈不规则等问题，需要构建非规则博弈下隐私度量激励理论。研究数据价值评估理论，形成安全、

公平的隐私度量框架。研究非规则多方博弈下数据共享的动态博弈模型，实现精准的参与方行为描述。研究基于隐私度量和动态博弈的激励策略，优化个体和群体的隐私数据共享行为，确保非规则博弈下隐私计算的安全性、公平性和可持续性。

3.2　面向数据全生命周期的隐私保护

数据的隐私保护问题贯穿数据全生命周期，而数据流通致使所有权、管理权、使用权分离，各隐私保护环节需紧密耦合，为解决该科学问题，建议对以下三个方向开展重点研究：基于隐私抽象的软件形式化验证、平衡安全效能的数据加密与脱敏、全周期隐私保护的密态数据管理。

科学问题四：基于隐私抽象的软件形式化验证

智能软件作为智能算法载体，其软件过程与隐私安全息息相关，是构建隐私计算技术体系的重要着力点之一。随着智能化软件功能和结构日趋复杂化，其隐私泄露风险在泄露环节和泄露方式上均呈现多样化趋势。因此，需从软件工程角度结合数据抽象、概率模型、可信度量、形式化方法等手段，分析、设计并验证智能软件全生命周期中的隐私信息安全。刻画针对智能系统的不确定性，构建软件过程全生命周期隐私的定量度量与动态评估机制，增强隐私保护的自适应性。结合形式化描述、概率建模与自适应验证等理论方法，研究智能软件开发生命周期中的隐私保护，实现智能软件生命周期与隐私计算生命周期的有机统一。

科学问题五：平衡安全效能的数据加密与脱敏

数据加密与脱敏是隐私计算的两类重要技术路线，数据加密技术虽适配较强的安全假设但具有较高的计算代价，计算效率低；隐私脱敏技术虽具有较高的计算效率但仅支持特定的计算任务，数据效用低。因此，为充分发挥各类技术优势，有效解决数据应用和数据安全保护两难的问题，以高效用高效率为目标，设计融合加密和脱敏的新型隐私计算框架协议，包括基于安全混洗的差分隐私计算模型、融合新型轻量级密码工具的差分隐私计算机理研究。研究多方隐私计算协议的隐私安全风险分析方法，尤其是在复杂隐私计算环境下的隐私计算协议组合理论，以及针对脱敏与加密融合的新型攻击威胁识别方法。

科学问题六：全周期隐私保护的密态数据管理

数据的非排他性促进了数据要素的资产化与市场化。数据资产化是在数据被管理和处理过程中实现可用不可见，数据市场化是在数据被流通和交易过程中实现共用不共享。现有隐私计算技术主要关注数据市场化中数据价值可流通、数据交易可确权等问题，局限于数据分析阶段的数据价值转化与挖掘，难以覆盖数据全生命周期的隐私安全，忽略了在采集、存储、处理、流通等其他阶段均存在大量数据泄露隐患。因此，需针对数据资产化与市场化的隐私保护需求，研究面向全生命周期密态数据管理。构建全周期开放环境下的密态数据处理范式，设计隔离数据与用户的数据行为控制机制，实现覆盖多样化业务的完备性密态数据管理。研究降低业务性能损耗的软硬件协同密态数据高效处理方法，设计适应不同强度与维度下性能与安全性的平衡优化策略，实现面向密态数据系统生态的实用性、易用性与兼容性。为数据全生命周期提供统一水平的隐私性、机密性、完整性保障，端到端实现数据要素的保护与利用。

3.3　数据跨域流通高效协同计算

数据多孤立自治，多方数据协同是数字化建设的关键，而数据跨域共享隐私风险高，协同计算任务多样，针对这一挑战，建议对以下三个方向开展重点研究：大数据跨域安全流通的联邦计算范式、基于可信软硬件的云边端机密计算、面向工业互联网的隐私计算框架。

科学问题七：数据跨域流通的联邦计算新范式

联邦计算允许数据在不出本地的前提下进行多方联合计算，为“原始数据不出域、数据可用不可见”的大数据跨域需求提供了可行方案。然而，现有联邦计算技术基于规则博弈假设，忽略了个体的非理性、随机性行为，难以适用于开放、动态场景下的不可信环境。针对不可信环境中多方互信激励难、恶意行为检测难的挑战，需要构建大数据安全跨域流通的联邦计算范式。研究非规则博弈下的联邦激励理论，建立有限理性引导下的联邦激励模型，引导联邦参与方协作。研究针对恶意成员的联邦安全查询框架和复杂查询算法优化，实现抵御全流程恶意攻击的联邦多模式高效查询。研究面向异质数据的联邦鲁棒学习方法，有效协同多方非独立同分布数据进行均衡学习，充分挖掘数据价值。

科学问题八：基于可信硬件的云边端机密计算

云边端计算模型中存在硬件平台、算法模型、数据流通三方面的隐私威胁：硬件不可信，存在微架构信道漏洞的风险；模型易泄露，算法模型和数据被窃取导致隐私泄露问题；流通不合规，数据采集和流通的审计标准缺失。需要进行机密可信硬件自主化研

制，将可信执行环境管理任务与代码执行松耦合，采用物理解耦进一步保障管理任务自身安全性，拓展异构架构防护能力。研究基于存内计算的智能模型加解密和基于隐私预算的数据自销毁机制，实现智能模型与数据的隐私保护，包含基于内生指纹的数据集与模型版权保护和面向长时段多平台的时空数据隐私保护。研究数据流通的隐私合规性检测，对过度搜集行为进行检测并设计隐私政策审计机制。研究软硬件协同的机密计算平台设计模式，构建云边端隐私计算新范式。

科学问题九：面向工业互联网的隐私计算框架

工业互联网促进企业间形成有机循环，实现数据、资源、物流、能源和信息的存储、流通和使用，而该过程中跨企业数据的隐私保护是保障系统安全的重要环节。在制造循环工业系统中需要采集企业集群的数据，通过标识解析管理各企业的数据，并将各方制造数据协同进行系统优化。然而，当前工业互联网中仍然存在数据流通壁垒高、数据标识解析难、数据集成隐私弱的挑战，使得工业数据难以充分流通，限制了制造工业循环系统的效能。需建立工业设备的层级密码支撑保障体系，网络安全博弈与策略优化，尤其是基于加密理论实现工业企业集群中的分解和协调。研究多源异构的工业互联网数据分布式管理与协同优化技术，建立安全可追溯的工业制造隐私计算体系，促进工业集群高效循环畅通。

科学问题十：基于区块链的数据高效流通体系

区块链技术以其具有的透明性、不可篡改性和可追溯性等特点，成为隐私计算中的重要技术路线之一，是实现数据流通可信、可控、可计量的关键手段。然而现有区块链技术存在着实体间信任构建难、安全与效能平衡难、隐私与可信兼顾难的挑战。因此，需要在区块链的数据隐私性、计算结果真实性和计算效率的权衡设计方面展开研究，构建基于区块链的数据高效流通体系。研究有限信任环境下数据访问控制与可信共享机理，尤其是基于区块链的身份认证、访问控制、数据共享技术。研究数据流通、协同计算中安全与效能间的平衡机制，实现融合区块链的高效协同隐私计算技术。研究明文不可见下计算结果验真技术，尤其是基于区块链的数据计算过程追溯、隐私保护模型校验技术。

4 隐私计算的典型应用场景

党中央和国务院在《“十四五”规划》中强调应加快数字化建设进程，激发数据要素潜能，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。隐私计算为隐私保护下的数据要素价值发挥提供了解决方案，在数字化建设中具有丰富的应用场景和应用优势。基于隐私计算关键技术，融合人工智能打造可信人工智能应用，解决大模型等人工智能产业的隐私风险；面向数据要素市场构建合规数据交易平台，保障数据交易过程安全公平、透明合规；协同各政府主体隐私数据共享分析，实现跨部门政务管理和应急处置；打通医疗数据壁垒，促进医疗资源互联；建立数字人民币交易隐私保护机制，打造数字货币安全生态。本节对隐私计算未来可能的五个典型应用场景进行了展望。

4.1 典型应用场景一——可信人工智能

新一代人工智能技术迅猛发展，在给人们生活带来了深刻变化的同时，也逐渐展现出其应用风险，如大模型中的提示词泄漏、训练数据泄漏、使用隐私泄漏问题，引发了社会公众对人工智能的信任危机。可信人工智能以稳定性技术、可解释性技术、隐私计算技术、公平性技术等为基础，构建具备可控可靠、透明可释、隐私保护、明确责任、多元包容特性的可信人工智能体系，实现信息、物理与人类的安全融合交互，应对人工智能应用面临的信任危机。针对人工智能数据安全风险高、隐私保护程度弱的挑战，可信人工智能将重建原有训练、推理架构，设计物理层、算法层、应用层全层级数据隐私安全防护的人工智能产业体系[37]，如结合联邦学习等隐私计算技术设计新一代人工智能算法、基于可信硬件构造新型人工智能应用框架、恶意攻击防御的可信智能软件验证方法。结合隐私计算新技术手段，打造多方可信的人工智能产业体系，提升社会对人工智能的信任程度，加快构建健康可信的人工智能产业生态，促进人工智能技术和数字产业的长远发展。

4.2 典型应用场景二——合规数据交易

数据要素是数字经济的基础资源和核心引擎。近年我国高度重视数据要素市场建设，北京、上海等地相继成立了数据交易所，积极探索“原始数据不出域，数据可用不可见”的数据交易范式。合规数据交易以数据隐私保护技术为基础，旨在建立数据采集、存储、处理、流通、应用等全生命周期的安全管理机制，培育高效率流通、高质量发展的数据要素市场。针对数据交易市场中数据安全难保障、场内交易难持续的挑战，将结合隐私计算完善数据确权、跨域流通、公平定价[38, 39]、安全管理基础设施建设，构建可

计量可追溯的合规数据交易体系[40]。一方面，基于密态数据管理等隐私计算技术搭建安全透明的数据交易服务平台，实现交易流程的合规；另一方面，以数据隐私保护法律法规为准绳搭建起隐私计算技术与数据交易法规之间的桥梁，实现技术合规性的智能化监管。合规数据交易的建设将进一步促进数据要素的安全流通和价值共享，培育安全可信的数据要素市场生态体系，降低各类主体的数据获取门槛，加速推进数字经济战略的建设进程。

4.3　典型应用场景三——数字社会治理

“十四五”规划明确提出要构筑共治共享的数字社会治理体系，提高数字政府建设水平。数字社会治理的核心在于运用数字技术实现数据泛在融通共享和协同智能分析，然而政务信息等数据涉及大量敏感的隐私信息，限制了各部门数据的共享协同。针对上述挑战，将以隐私计算技术为突破点，建立数字社会治理信任底座，打通跨域数据应用价值链，探索多部门联动、资源共享、安全高效的社会治理新模式。以数字政府服务建设为例，通过多方安全计算等隐私计算技术，搭建起政府各职能部门、社会主体间的数据可信流通通道，实现多部门协同治理和应急处置。基于隐私计算的数字社会治理体系将有效提高社会运转和经济运行的泛在感知和智能决策能力，实现跨边界、跨领域的整体化社会治理，提升国家治理效能和社会治理现代化水平。

4.4　典型应用场景四——医疗数据挖掘

随着智能移动设备普及化、医疗设备数字化及电子病历结构化的推进，医疗数据呈现爆发增长以及多模态的特点，为依托新一代数字技术的数字医疗建设提供了庞大发展驱动力，医疗数据的协同挖掘已成必然趋势。然而，医疗数据作为医疗卫生领域的重要资产，包含大量个人隐私信息，阻碍了医疗数据的共享分析和价值挖掘。面对海量多模态医疗数据全生命周期，应用隐私计算技术，将建立医疗数据安全“存—查—算”技术体系，可望打破医疗数据壁垒，赋能智慧医疗建设。具体包括：隐私敏感数据分级安全存储，利用生存分析等算法挖掘隐私敏感数据，从数据层面上分离出不同粒度的隐私数据，实现分阶段分等级的隐私数据独立保护存储；隐私医疗数据联合查询，基于多方安全计算等技术支撑影像数据、病理数据、诊断数据、医保数据等多源多模态数据的安全高效复杂查询，提升医疗信息共享水平；医疗数据安全分析计算，通过联邦学习等技术对数据进行挖掘分析，形成医疗大模型等医疗辅助平台[41]，提高医疗普惠性和公平性。隐私计算在医疗领域融合应用将驱动医疗资源互联，重构就医诊疗模式，为全面实施健康中国战略、构建优质高效的医疗卫生服务体系提供强劲动力。

4.5　典型应用场景五——数字人民币

数字人民币近年在多个领域开放试点，已取得诸多实质性成效，数字人民币生态体系不断完善。数字人民币体系设计坚持“安全普惠”理念，必须保证数字人民币基础设施的安全性和可靠性。针对数字人民币可控匿名的需求，依托隐私计算有望从技术角度建立交易信息隔离机制，解决交易匿名与交易透明可追溯间的矛盾，实现个人隐私数据保护和安全治理。另外，数字人民币通过搭载不影响货币功能的智能合约[42]，使数字人民币在确保安全合规的前提下根据交易双方的约定规则自动强制执行合约内容，从而降低履约成本和违约风险。针对当事人身份信息与合约内容的隐私需求，依托区块链、同态加密与可信硬件等技术构建软硬一体化的隐私保护智能合约机制，实现合约执行流程的安全透明和可追溯性，打造数字人民币智能合约的信任基础。面向数字人民币生态建设，隐私计算等新技术将赋能安全数字货币业务升级，支撑数字经济降本提效，促进数字人民币服务创新发展，提升数字人民币的治理水平。

5　结　语

在全球数字经济快速发展的背景下，大数据和新型人工智能等技术在智慧金融、智慧城市、智慧医疗等领域的应用创新层出不穷，与此同时也为隐私计算提供了新的机遇和挑战。由于我国在隐私计算领域的研究和实践起步较晚、人才储备不足、原创成果不多、成果转化率偏低，亟需基础理论突破和具有重大影响力的研究成果。

因此，深入开展信息科学、数学、管理科学等多学科交叉研究是解决上述挑战的必要途径之一，建议加大国家自然科学基金对隐私计算基础理论与关键技术中多学科交叉研究的资助力度，围绕三个前沿方向与十个关键科学问题开展原创性研究，提升我国隐私计算基础理论、方法和技术研究水平，促进隐私计算相关科技成果的应用与推广。

参　考　文　献

[1]　吕卫锋，郑志明，童咏昕，等. 基于大数据的分布式社会治理智能系统. 软件学报，2022，33(3)：931—949.

[2] Yao ACC. How to generate and exchange secrets. 27th Annual Symposium on Foundations of Computer Science (sfcs 1986). Toronto, ON, Canada. IEEE, 1986: 162—167.
[3] Evans D, Kolesnikov V, Rosulek M. A pragmatic introduction to secure multi-party computation. Foundations and Trends? in Privacy and Security, 2018, 2(2/3): 70—246.
[4] Shamir A. How to share a secret. Communications of the ACM, 1979, 22(11): 612—613.
[5] Rabin MO. How to exchange secrets with oblivious transfer. IACR Cryptology EPrint Archive, 2005: 187.
[6] Acar A, Aksu H, Uluagac AS, et al. A survey on homomorphic encryption schemes. ACM Computing Surveys, 2019, 51(4): 1—35.
[7] Manulis M, Nguyen J. Fully homomorphic encryption beyond IND-CCA1 security: integrity through verifiability. Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2024: 63—93.
[8] Lindell Y, Pinkas B. Privacy preserving data mining. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000: 36—54.
[9] Vaidya J, Clifton C. Privacy-preserving k-means clustering over vertically partitioned dataProceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '03. Washington, D. C.. ACM, 2003: 206—215.
[10] Kantarcioglu M, Clifton C. Privacy-preserving distributed mining of association rules on horizontally partitioned data. IEEE Transactions on Knowledge and Data Engineering, 2004, 16(9): 1026—1037.
[11] Wu YC, Cai SF, Xiao XK, et al. Privacy preserving vertical federated learning for tree-based models. Proceedings of the VLDB Endowment, 2020, 13(12): 2090—2103.
[12] Mohassel P, Rindal P. ABY^3: A mixed protocol framework for machine learning proceedings of the 2018 ACM SIGSAC conference on computer and communications security. Toronto Canada. ACM, 2018: 35—52.
[13] Keller M. MP-SPDZ: a versatile framework for multi-Party computation proceedings of the 2020 ACM SIGSAC conference on computer and communications security. Virtual Event USA. ACM, 2020: 1575—1590.
[14] Zahur S, Evans D. Obliv-C: a language for extensible data-oblivious computation. Cryptology ePrint Archive, 2015: 1153.
[15] Li Y, Xu W. PrivPy: General and scalable privacy-preserving data mining proceedings of the 25th ACM SIGKDD international conference on knowledge Discovery & Data mining. Anchorage AK USA. ACM, 2019: 1299—1307.
[16] Zhang C, Chen Y, Liu W, et al. Linear private set union from Multi-Query reverse private membership test. USENIX Security Symposium, 2023: 337—354.
[17] Chen Y, Zhang M, Zhang C, et al. Private set operations from? Multi-query reverse private membership test. Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2024: 387—416.
[18] Yadav VK, Andola N, Verma S, et al. A survey of oblivious transfer protocol. ACM Computing Surveys, 2022, 54(10s): 1—37.
[19] Sweeney L. K-anonymity: a model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2002, 10(5): 557—570.
[20] Machanavajjhala A, Gehrke J, Kifer D, et al. L-diversity: privacy beyond k-anonymity. 22nd International Conference on Data Engineering (ICDE'06). Atlanta, GA, USA. IEEE, 2006: 24.
[21] Ye M, Shen W, Zhang JW, et al. SecureReID: privacy-preserving anonymization for person re-identification. IEEE Transactions on Information Forensics and Security, 2024, 19: 2840—2853.
[22] Dwork C. Differential privacy. International colloquium on automata, languages, and programming. Springer Berlin: Heidelberg, 2006;1—12.
[23] Mironov I. Rényi differential privacy. 2017 IEEE 30th Computer Security Foundations Symposium (CSF). Santa Barbara, CA, USA: IEEE, 2017: 263—275.
[24] Cormode G, Jha S, Kulkarni T, et al. Privacy at sacale: Local differential Privacy in PracticeProceedings of the 2018 International Conference on Management of Data. Houston TX USA: ACM, 2018: 1655—1658.
[25] Cormode G, Bharadwaj A. Sample-and-threshold differential privacy: histograms and applications. International Conference on Artificial Intelligence and Statistics, 2022: 1420—1431.
[26] Feldman V, McMillan A, Talwar K. Stronger privacy amplification by shuffling for renyi and approximate differential privacy. Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA). Philadelphia, PA: Society for Industrial and Applied Mathematics, 2023: 4966—4981.
[27] Luo QY, Wang YL, Yi K, et al. Secure sampling for approximate multi-party query processing. Proceedings of the ACM on Management of Data, 2023, 1(3): 1—27.
[28] 冯登国. 从可信计算到机密计算. 中国计算机学会通讯, 2022, 18(2): 18—23.
[29] Yang Q, Liu Y, Chen TJ, et al. Federated machine learning. ACM Transactions on Intelligent Systems and Technology, 2019, 10(2): 1—19.
[30] McMahan HB, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data. (2023-01-26)/[2024-02-22]. https://arxiv.org/abs/1602.05629v4.
[31] Tong Y, Pan X, Zeng Y, et al. Hu-fu: efficient and secure spatial queries over data federation. Proceedings of the VLDB Endowment, 2022, 15(6): 1159.
[32] Tong Y, Zeng Y, Zhou Z, et al. Federated computing: query, learning, and beyond. IEEE Data Engineering Bulletin, 2023, 46(1): 9—26.
[33] Liu FX, Zheng ZM, Shi YX, et al. A survey on federated learning: a perspective from multi-party computation. Frontiers of Computer Science, 2023, 18(1): 181336.

[34] Wang YS, Tong YX, Zhou ZM, et al. Distribution-regularized federated learning on non-IID data// 2023 IEEE 39th International Conference on Data Engineering (ICDE). Anaheim, CA, USA: IEEE, 2023: 2113—2125.

[35] Shi YX, Tong YX, Zeng YX, et al. Efficient approximate range aggregation over large-scale spatial data federation. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(1): 418—430.

[36] Paju A, Javed MO, Nurmi J, et al. SoK: A systematic review of TEE usage for developing trusted applications proceedings of the 18th international conference on availability, reliability and security. Benevento Italy: ACM, 2023: 1—15.

[37] Li B, Qi P, Liu B, et al. Trustworthy AI: from principles to practices. ACM Computing Surveys, 2023, 55(9): 1—46.

[38] Song TS, Tong YX, Wei SY. Profit allocation for federated learning// 2019 IEEE International Conference on Big Data (Big Data). Los Angeles, CA, USA: IEEE, 2019: 2577—2586.

[39] Pei J. A survey on data pricing: from economics to data science. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(10): 4586—4608.

[40] Chen F, Wang JH, Jiang CK, et al. Blockchain based non-repudiable IoT data trading: simpler, faster, and cheaper// IEEE INFOCOM 2022-IEEE Conference on Computer Communications. London, United Kingdom: IEEE, 2022: 1958—1967.

[41] Rauniyar A, Hagos DH, Jha D, et al. Federated learning for medical applications: a taxonomy, current trends, challenges, and future research directions. IEEE Internet of Things Journal, 2024, 11(5): 7374—7398.

[42] Sharma P, Jindal R, Borah MD. A review of smart contract-based platforms, applications, and challenges. Cluster Computing, 2023, 26(1): 395—421.

Innovative Theoretical Methods and Key Applications of Privacy Computing

Zhiming Zheng[1]　Jifeng He[2]　Lixin Tang[3]　Yongxin Tong[1]
Ting Chen[4]　Guo Xie[5]　Zhiheng Wang[5]　Guozheng Wu[5*]

1. *Beihang Univesity*, *Beijing* 100191
2. *Tongji University*, *Shanghai* 200092
3. *Northeastern University*, *Shenyang* 110819
4. *University of Electronic Science and Technology of China*, *Chengdu* 610000
5. *Department of Information Sciences*, *National Natural Science Foundation of China*, *Beijing* 100085

Abstract　Based on the 352nd Shuangqing Forum of National Natural Science Foundation of China, this article elaborates the importance of privacy computing in solving the major needs of the national digital economy. Besides, this article analyzes both domestic and international research focus in this area, elucidates the connotations and extensions of privacy computing, and revisits its research journey across three dimensions: privacy encryption, privacy anonymization, and new types of privacy computing under non-regular game theory. Moreover, this article succinctly addresses critical foundational scientific issues in privacy computing, such as non-standard game theory, full lifecycle privacy preserving, and the federated computing paradigm. Finally, this article discusses the important future research directions in this field, aiming to further advance the breakthrough development and application of privacy computing theories, methods, and technologies in China.

Keywords　privacy computing; encryption and anonymization; Non-Standard Game Theory; full lifecycle; federated computing

（责任编辑　张　强）

* Corresponding Author, Email: wugz@nsfc.gov.cn