大模型赋能可信数据空间：数据安全治理与信任机制构建*

裴 雷[1] 陈晓宇[2]
（1. 南京大学数据智能与交叉创新实验室，南京 210023；
2. 上海大学文化遗产与信息管理学院，上海 200444）

摘 要：[**目的/意义**]探讨可信数据空间的关键构成要素及其实现路径，重点分析大模型技术在数据安全治理和信任机制框架构建中的作用和影响。[**方法/过程**]采用理论分析与实践案例相结合的方法，围绕数据安全治理、大模型技术路径和信任机制构建三个方面展开论述。[**结果/结论**]可信数据空间的建设需要在制度、技术和多主体协作之间实现深度融合。大模型为数据安全治理提供了强大支撑，尤其在实时监测、异常识别、智能决策及自动化合规管理等方面展现出显著优势。数据安全治理的完善依赖于信任评估体系的动态优化，通过分类分级保护、跨主体协作机制和可解释性增强，提升可信数据流通的透明度与可控性。可信数据空间的建设不仅是技术创新的过程，更是数据治理范式的变革，其实施路径需兼顾合规性、互操作性与国际化标准对接，推动数据要素市场的高效配置。

关键词：可信数据空间 大语言模型 数据安全治理 信任机制 数据治理框架
分类号：G250.7
DOI：10.31193/SSAP.J.ISSN.2096-6695.2025.02.02

0 引 言

数据安全治理与信任机制构建已成为全球数字化发展的核心议题[1]。数据作为数智时代的关键生产要素，其可信流通和有效治理不仅是推动数字经济高质量发展的关键，也是国家竞争力提升的重要支撑[2]。党的二十届三中全会明确提出：“加快建设全国统一大市场”“健全数据要素市场化配置制度”[3]，为我国数据要素市场的规范化发展提供了政策指引，同时对数据安全治理体系提出了更高要求。在此背景下，如何在保障数据安全的同时，构建跨行业、跨领域、多主体

* 本文系国家社会科学基金重大项目“复杂信息环境下的数据要素流通政策仿真与评价监测研究”（项目编号：24&ZD190）的研究成果之一。
[作者简介]裴雷，男，教授，研究方向为政策计算与政策扩散、数据治理与智能决策等，Email：plei@nju.edu.cn；陈晓宇，男，副教授，研究方向为数字文化资源管理、用户信息行为、社会信息学等，Email：xiaoyu-chen@shu.edu.cn。

协同的可信数据流通机制，成为当前数据治理的核心议题。

随着大模型技术的飞速发展，数据流通与治理的范式正在发生深刻变革。大模型凭借其强大的数据处理能力、知识推理能力和智能决策能力[4]，在复杂信息环境下为数据安全治理和信任机制构建提供了新的技术支撑[5]。特别是在数据可信共享、隐私保护计算和跨域协作等领域，大模型能够提升数据使用透明度，降低数据流通中的安全隐患与合规风险。然而，数据安全是可信数据空间的基础，技术的不透明性、数据聚合的合规性以及数据权属的复杂性，依然是大模型技术在数据治理实践中必须解决的问题。在充分发挥大模型智能化优势的同时，要确保数据要素的安全可信流通[6]。

《可信数据空间发展行动计划（2024—2028 年）》（以下简称《行动计划》）提出，到 2028 年我国将建成 100 个以上可信数据空间，形成覆盖广泛、互联互通、价值共创的数据流通生态网络[7]。这一政策目标不仅强调了数据要素市场化配置的重要性，也对可信数据空间的技术架构和治理体系提出了更高要求。可信数据空间的建设不仅是技术层面的探索，更是制度与治理模式的创新。通过技术路径与治理框架的融合，构建一个高效、安全、透明的信任机制，是推动数据要素市场健康发展的重要任务[8]。然而，在大规模数据流通、跨行业数据共享的安全性和信任机制方面，现有政策框架尚缺乏针对性方案。

本文围绕可信数据空间的建设路径展开讨论，着重分析大模型在数据安全治理与信任机制构建中的作用，并系统探讨数据可信共享的制度设计与技术架构。首先剖析可信数据空间建设中数据安全治理的需求；其次探讨大模型如何在隐私保护、数据治理和跨域协作中发挥作用；最后构建相关信任机制框架，提出未来可信数据空间的发展路径，并展望其对数据要素市场和数字经济高质量发展的作用。

1　可信数据空间建设中的数据安全治理

1.1　可信数据空间的内涵特征与现状

可信数据空间是一个基于共识机制、多主体协同参与的数据治理生态系统，其关键特征是通过技术手段和制度设计，保障数据在流通、共享和交易中的安全性、可信性和可控性[9]。从本质上讲，可信数据空间不仅是一种技术架构，更是数据要素市场化配置的基础设施，其核心在于解决数据流通中“数据主权”、“数据安全”和“信任机制”问题。

可信数据空间的建设具有多重目标。第一，强调数据主权与自治性，通过技术和制度赋予数据提供方对数据使用过程的自主权和控制权，确保数据共享和流通过程中所有权和隐私不受侵犯[10]。第二，可信数据空间注重多方协作与互操作性，通过统一的技术标准和接口规范，打破“数据孤岛”，推动不同平台和系统间的数据资源高效流通与利用[11]。第三，安全性和合规性是可信数据空间的基本要求，通过区块链、数字签名、智能合约等前沿技术，保障数据传输过程的完整性、不可篡改性和可追溯性，同时确保数据使用符合不同国家和行业的法律法规。第四，在经济性和高效性方面，可信数据空间通过设计合理的激励机制和优化数据流通路径，显著提升数据要素交易效率，释放数据经济的潜在价值[12]。

目前，国内外可信数据空间建设已取得初步进展，正成为推动全球数字经济发展的关键支点[13]。在国际层面，欧洲较早展开了系统性探索，依托国际数据空间协会（International Data Spaces Association，IDSA）等组织，构建起一套具有代表性的可信数据空间通用架构与标准体系，如 IDSA 4.0 版参考架构模型，并在汽车制造（如 Catena–X 数据空间）与供应链管理［如智能连接供应网络（smart connected supplier network，SCSN）］等领域推进多项示范项目。相比之下，美国主要通过谷歌云（Google Cloud）等云服务商推动跨企业数据共享与智能分析，其市场机制较成熟，平台化能力较强，但在跨行业互操作性和数据安全保障层面仍存在一定挑战。亚洲地区，日本和韩国也在数据空间建设中展现出积极态势。日本发布《互联工业开放架构》（*Connected Industries Open Framework*，CIOF），强调工业数据跨主体可信流通的标准化与协作治理。韩国则以 MyData 个人数据服务为基础，探索数据主权导向下的个性化数据空间构建路径。

我国可信数据空间建设也在快速推进。从政策层面看，《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等一系列法律法规的出台，为数据流通与治理提供了坚实的法律保障。在实践方面，以区块链、隐私计算等技术为基础，积极探索数据安全框架，逐步实现数据可信流通的全生命周期管理[10]。上海、深圳等地已开始试点构建了多个行业数据空间，涵盖金融、医疗、工业等重点领域，为数据流通和价值挖掘提供了创新的应用场景。

这些实践为构建统一的数据治理模式奠定了重要基础[14]，主要实践案例及现状见表 1。

表 1　数据空间实践案例及现状

国家 / 地区	实践案例	主要特点	现实挑战
欧盟	• 国际数据空间协会（IDSA） • Catena–X（汽车制造） • SCSN（供应链管理）	• 制定可信数据空间通用框架和标准（如IDSA4.0版参考架构模型） • 行业示范项目促进多领域协作 • 数据资源高效流通与共享	• 跨行业互操作性不足 • 跨境数据合规难题 • 数据主权争议
美国	Google Cloud 等云服务平台	• 基于成熟商业模式推动数据共享与价值挖掘 • 强调技术驱动的数据经济发展模式	• 跨行业互操作性存在局限 • 数据安全保障不足
日本	《互联工业开放架构》（CIOF）	• 多领域协同治理 • 通过标准化接口实现跨行业协作	• 行业间数据整合与协调性仍需完善
韩国	MyData 个人数据服务	• 以用户为中心的个性化数据空间 • 强调隐私保护与数据自主权	• 数据治理能力需进一步加强 • 国际规则适配能力不足
中国	上海、深圳等地行业数据空间试点	• 通过《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等提供政策支持 • 技术创新（区块链、隐私计算等）推动全生命周期管理 • 应用场景涵盖金融、医疗、工业等领域	• 全球数据治理框架尚未协调 • 部分行业数据治理实践能力不足 • 技术优化与扩展性问题

目前尽管国内外可信数据空间建设已取得一定进展，但其发展过程中仍然存在一些亟待解决的问题。首先，全球范围内的数据治理框架尚未实现协调统一，不同国家和地区在数据保护标准、技术规范以及合规要求上的分歧，给跨境数据流通和可信数据空间的全球化推广带来了障碍。其次，数据空间内部的技术实践仍需优化，现有技术手段在应对大规模、多场景、多主体的数据协作中，尚未完全解决性能、可靠性与扩展性之间的平衡问题。再次，部分行业在数据治理意识和实践能力上的不足，也对可信数据空间的广泛应用形成了一定阻力。最后，可信数据空间建设还需应对多主体利益协调的复杂性，数据提供方、使用方以及中介服务方之间的信任机制和协作模式有待进一步完善，而这些利益相关方的互动关系往往受到不同治理结构、法律环境及商业模式的深刻影响。

1.2 数据安全治理面临的关键挑战

数据安全是数字经济发展的重要基石，也是数据要素市场化配置的核心保障。在全球数字化转型的浪潮中，数据要素的安全可信流通，不仅关系到资源的有效整合和利用，更关系到社会经济运行的稳定性与可持续性[15]。近年来，随着数据规模的快速增长和使用场景的高度复杂化[16]，数据安全治理的内涵不断拓展，其重要性已远超传统的技术防护层面，逐渐成为影响国家治理、经济发展和社会信任的重要因素[17]。

数据安全直接关系到多主体协作的信任基础。数据流通的本质在于多方之间的数据共享与协同利用，在这一过程中，信任机制是实现高效流通的关键。然而，现实中数据泄露、隐私侵害和非法滥用等问题频发，数据的提供方担忧隐私信息的泄露与误用，而数据的使用方则面临数据质量难以保证和合法性存疑的困境。这种信任缺失不仅削弱了数据交易的流动性，也增加了数据使用的成本与风险，从而阻碍了数据要素在经济活动中的价值释放[12]。尤其是在涉及敏感数据的领域，数据安全问题更是带来了难以忽视的连锁反应。在金融领域，数据泄露不仅会导致客户隐私信息的外泄，还可能诱发大规模金融欺诈，威胁金融体系的稳定性。在医疗领域，患者健康数据的滥用或泄露，不仅侵害个人隐私，还会动摇公众对医疗服务的信任，影响医疗数据的共享意愿和研究价值的实现。在政务领域，政府数据作为国家治理的重要资源，若因安全问题引发数据丢失或不当传播，不仅可能影响公共服务的质量，还会削弱公众对政府机构的信任，甚至引发社会稳定问题[11]。从更深层次来看，信任缺失往往形成一种“数据孤岛”效应。由于缺乏统一的安全保障与信任机制，各数据主体倾向于保护自身数据资产而非共享，这不仅限制了数据价值的进一步挖掘，也造成了数据资源的浪费和冗余。这种割裂状态阻碍了跨行业、跨区域的协同创新，使得数据要素的流通效率和经济效益远低于预期。

国际数据治理格局的复杂性和不确定性在很大程度上加剧了数据安全问题的迫切性，数据跨境流通已成为经济和技术合作的核心环节。国际上对数据主权、隐私保护和合规性要求方面的分歧，使全球数据治理呈现出碎片化状态[9]。这种局面不仅增加了数据安全的潜在风险，也推高了跨境数据流通合规成本和操作难度。各国对数据主权的解读差异是数据流通规则碎片化的主要原因之一。以欧盟为例，《通用数据保护条例》（*General Data Protection Regulation*, GDPR）确立了严格的数据保护与跨境传输规范，强调隐私保护的核心地位，但其严苛的要求使其在其他地区的适用性相对有限。相较之下，美国倾向于采用以市场为导向的数据流通方式，

对隐私保护的监管相对宽松。而我国则以数据的国家安全属性为核心，构建了基于《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》的法律框架。这种基于不同国家利益和治理理念构建的规则体系，缺乏统一的标准，不仅导致安全规范和技术合规要求的差异化，还使跨境流通的合法性审核更加复杂[6]。

技术层面的互操作性不足与标准化体系的缺失同样是当前跨境数据流通的主要瓶颈。在不同国家和地区，监管要求的多样性导致数据加密、匿名化处理及访问控制机制的技术实现路径各异。实施细节上的不兼容进一步阻碍了数据的高效流通，也为安全隐患埋下了伏笔。尤其是在不同技术体系之间，标准割裂容易形成对接漏洞，成为恶意攻击的潜在入口，从而进一步放大了跨境数据流通的安全风险[18]。

解决数据安全问题的核心在于构建一个可持续、可信赖的安全生态。安全生态不仅是对数据资源本身的保护，更是释放数据要素价值、提升其社会经济效能的关键路径[19]。在安全生态框架下，数据资源的流通不再局限于保护隐私和防范风险，而是为多方协作和价值共创提供了坚实的信任基础。唯有在安全的前提下，数据的规模化可信流通才能得以实现，并通过推动数据要素在经济社会系统中的高效整合，促进协同创新和动态平衡。这一安全生态需要构建覆盖数据全生命周期的动态防护体系，包括数据采集的合法性审核、存储的加密与防护、传输的路径控制，以及使用过程中的实时监测与风险响应。同时，安全生态建设必须以多主体协作为基石，通过完善的制度保障机制和透明的治理模式，推动企业、政府、社会组织之间的协同合作，化解信任壁垒，提升数据要素的共享效率和治理效能。更为重要的是，安全生态的建立不应仅仅着眼于当下的技术问题，而需要立足长远，以低成本高效益的方式不断激发数据资源的潜能，为数字经济的可持续发展注入强劲动能。

2　大模型在数据安全治理中的技术路径

2.1　大模型驱动的数据安全治理优势及挑战

大模型作为人工智能技术的重要突破，正成为数据安全治理的重要支撑力量，并在数据分析、异常检测和风险预测等领域展现出显著优势。其强大的计算能力和深度学习模型，使其能够高效处理复杂数据环境下的安全问题，不仅提升了数据安全监测与响应能力，也在推动数据要素市场化和可信数据流通方面发挥了关键作用[18]。

在数据安全治理方面，大模型的优势主要体现在三个方面。第一，在动态安全监测与异常检测中，大模型能够通过深度学习算法分析海量数据，精准识别潜在的安全威胁。例如，在网络安全领域，大模型可通过实时分析数据流量，检测异常访问、恶意攻击及数据泄露风险，并生成自动化预警，提高安全管理的响应效率。第二，大模型的自学习能力使其能够持续优化安全策略，动态适应新的威胁环境。通过对攻击模式、数据流通规律的深度训练，大模型可以不断提升检测精度，使其在动态风险环境中具备更强的适应性[4]。第三，大模型在数据保护与隐私计算方面也展现出积极作用，如基于差分隐私、联邦学习等方法，在保障数据安全的同时，实现数据共享的可控性，为可信数据空间建设奠定技术基础。

此外，大模型广泛应用也可能带来新的安全挑战。其训练过程依赖于大量、多样化的数据，这些数据往往涉及个人隐私、商业机密甚至国家战略信息[10]。在数据高度聚合的背景下，如果缺乏严密的安全管理措施，不仅可能成为攻击者的主要目标，还可能因数据链条中的薄弱环节引发大规模信息泄露，造成难以估量的后果。在跨行业、跨领域的数据使用场景中，这种风险被进一步放大，对数据安全治理提出了更高的技术和管理要求[11]。

大模型的算法复杂性及“黑箱”特性增加了数据安全监管的复杂性。由于其推理过程难以解释，数据使用方和监管机构难以判断其决策逻辑是否符合要求，进而影响对模型预测结果的信任度。在隐私保护要求较高的场景，如医疗、金融或跨境数据流通，算法的不透明性可能引发数据合规性争议，甚至影响数据跨境共享的合法性[4]。此外，大模型需要持续迭代优化，而新数据的合法性、可信度和来源控制仍然存在漏洞，如何在模型优化过程中保证数据使用的合规性，成为数据监管与治理的关键难题。

在分布式计算环境下，数据安全治理更具挑战性。多节点协作模式虽然提升了计算效率，但也增加了数据传输和计算过程中的安全风险。计算节点的权限管理、漏洞修复与访问控制成为需要解决的问题，任何一个节点的安全失效，都可能通过链式反应影响整个系统的稳定性，甚至引发系统性安全事件。在跨境数据流通和多行业协作的应用场景中，数据需要在不同法律和技术标准下流转，各国对于数据隐私保护、网络安全合规的标准各异，进一步加剧了数据治理的复杂性。既要确保计算节点的安全性与合法性，又要兼顾跨境流通的数据合规要求，还要保证大模型的计算效率和响应速度，这种多目标优化问题对可信数据空间的技术设计和治理体系提出了更高的要求。

2.2 隐私保护、数据治理与算法透明性

大模型技术在数据安全治理中的广泛应用，使得隐私保护、数据治理和算法透明性成为可信数据空间建设的核心要素。这三者共同构成了数据流通过程中的信任基石，不仅影响数据安全和隐私保护机制的有效性，也直接关系到多主体协作的稳定性与合规性。

在可信数据空间中，隐私保护是确保数据要素安全流通的核心任务，特别是在跨主体协作和跨境数据流通的场景下，对敏感信息的保护需求尤为突出。大模型在训练和推理过程中需要依赖大量数据，而这些数据往往涉及个人隐私、商业机密，甚至国家安全信息。数据的集中存储和计算模式使其容易成为攻击目标，信息泄露的风险也随之增加。为了降低隐私风险，当前广泛采用差分隐私（differential privacy）和联邦学习（federated learning）等技术手段，实现数据“可用不可见”的可信计算。

差分隐私技术通过在数据分析过程中引入随机噪声，确保单个数据点的不可识别性，从而在保护个体隐私的同时，不影响整体数据分析的有效性。这种方法已被广泛用于个性化推荐、统计分析和公共数据共享领域。相比之下，联邦学习技术通过去中心化的分布式训练方式，使数据始终存储在本地，仅传输经过加密的模型参数，避免了数据的集中暴露风险。在跨机构数据协作场景中，如医疗、金融和智能制造等行业，联邦学习技术的应用可以实现不同主体间的数据协同计算，同时确保隐私保护的合规性。这些技术手段为可信数据空间中的隐私保护提供了坚实的技术支撑，使数据在保证安全的前提下实现价值最大化。

数据治理在可信数据空间中扮演着规则设定与安全管理的双重角色，既是保障数据合规性的重要机制，也是优化数据资源流通效率的关键环节。可信数据空间中的数据治理主要涉及数据存储、权限管理、访问控制和合规监管四个方面，以确保数据的合法使用与有效流通。

在数据权限管理和访问控制方面，基于角色的访问控制（role–based access control, RBAC）和基于属性的访问控制（attribute–based access control, ABAC）已成为主流治理模式。RBAC 通过定义不同用户角色及其权限，确保数据访问行为符合既定规则，而 ABAC 进一步结合用户行为、环境条件等因素，实现更精细化的权限管理。这些机制不仅能够提高数据访问的安全性，还能在不同主体间建立动态授权机制，提高可信数据空间的灵活性和适应性。

此外，数据治理还涉及跨境数据合规性问题。由于各国在数据主权、隐私保护和数据流通规则上存在较大差异，可信数据空间需要在技术层面提供合规保障，如数据来源合法性审查、跨境数据加密传输，以及基于智能合约的数据使用用途监管。例如，在欧盟《通用数据保护条例》的约束下，企业需确保用户数据在跨境流通过程中符合“数据最小化”原则，而《中华人民共和国数据安全法》则强调数据分类分级保护，对特定类型的数据出境设定了更严格的审查标准。因此，可信数据空间的治理体系需要适配多层次的数据安全与合规要求，以确保数据在不同法律体系下的安全流通。

大模型推理过程中的“黑箱”特性是对算法透明性的显著挑战。传统的深度学习模型往往难以解释其决策逻辑，导致使用方、监管机构乃至社会公众对其可信度产生疑虑。在数据治理场景中，这种不可解释性可能引发数据滥用、偏见放大和决策不透明等问题，特别是在涉及公共决策、医疗诊断、金融风控等高敏感领域，算法透明性的缺失可能影响数据空间的可信性。

可解释人工智能（explainable artificial intelligence, XAI）成为提升算法透明度的重要解决方案。XAI 通过特征重要性分析、反事实推理、模型可视化等方法，使模型的推理逻辑更加清晰，帮助用户理解其决策过程。例如，在金融风控中，XAI 可用于解释贷款审批的评分标准，使数据使用方能够理解拒贷决策的合理性，并确保算法的公平性和合规性。此外，算法审计和第三方评估也是提升可信数据空间透明度的重要手段。通过引入独立机构进行算法审查，可有效减少模型偏见、提升算法可信度，并增强各主体对数据流通过程的信任感。

在技术实践中，隐私保护、数据治理与算法透明性的协同作用主要体现在确保数据在可信环境中共享，降低数据泄露风险，通过分类分级管理、访问控制和合规监管，实现数据资源的有效配置，以及通过可解释性方法和审计体系，使数据决策的合理性、透明度和可追溯性得以提升。构建可信数据空间，需要通过这三大技术维度的紧密结合，在保障数据安全的同时，优化数据价值释放，使数据治理体系更具稳定性和可持续性。

2.3　跨境数据流通与分布式计算的安全策略

随着国际数据协作的日益增多和分布式计算的深度发展，数据流通与处理已不仅仅局限于技术层面的问题，更涉及复杂的法律法规差异、技术标准不一以及多主体协作中的信任机制缺失。目前，《行动计划》已提出构建全国统一的数据流通与治理框架，但仍缺乏细化举措。跨境数据流通与分布式计算在可信数据空间中的应用，不仅需要技术路径的创新，还需要制度设计与协同

治理的有机结合。

跨境数据流通的核心挑战在于如何在多法律体系间实现安全与合规的平衡。在国际规则未完全统一的背景下，各国对数据主权和隐私保护的要求差异显著，跨境数据流通需要同时满足不同地区的法规要求[1]。以欧盟《通用数据保护条例》为例，其对数据出境的严格限制要求提供详细的合规说明，而《中华人民共和国数据安全法》则更加强调数据流通的国家安全属性。在这种高度分散的法律环境下，区块链技术和零知识证明（zero-knowledge proof，ZKP）为跨境数据流通提供了创新性的解决方案。区块链的分布式账本特性通过不可篡改的记录保障了数据流通过程的透明性和安全性，同时结合智能合约实现自动化的权限控制和合规验证[20]。零知识证明则能够在不泄露数据本身的前提下完成合法性验证，为敏感数据的跨境使用提供了技术保障。这些技术手段的结合，不仅为国际数据流通奠定了技术基础，也为不同法律体系间的互操作性提供了可能[21]。

然而，技术手段的应用仍需政策和制度的保障。可信数据空间的跨境治理需要信任评估体系与智能合约机制的双重支撑，以实现多主体间的数据可信流通。建立基于信誉积分、数据使用记录和合规性审查的多维信任评估模型，确保跨境数据流通的透明度和安全性。数据提供方和使用方需定期接受独立机构的合规评估，并根据历史数据流通记录调整访问权限，以增强国际数据共享中的信任机制。智能合约可以在不同法律体系下实现跨主体的数据流通规则自动化执行，确保数据的使用范围、访问权限及用途监管满足合规要求。在国际供应链数据共享场景中，智能合约可以预设访问权限，仅允许符合特定条件的企业访问关键数据，并在数据流通过程中进行动态调整，以确保数据使用的合法性。

分布式计算环境下的数据安全问题则更加聚焦于技术实施的复杂性。在大模型的分布式训练或推理中，数据需要在多个计算节点间传输和处理，这种多节点协作模式虽然提升了计算效率，但也增加了安全风险。一方面，分布式计算中的单点故障可能导致整个系统的脆弱性；另一方面，不可信节点的参与可能引发数据泄露或计算结果的篡改。为了解决这些问题，可以通过细粒度权限控制和动态风险监测技术构建一个更加安全的分布式计算环境。细粒度权限控制通过对计算节点的访问权限进行精细划分，确保每个节点的操作范围受到严格限制。动态风险监测则利用实时分析技术快速识别异常行为，并对潜在风险进行及时干预。此外，分布式共识算法的引入能够提高容错能力，在部分节点失效或出现恶意行为时依然保障系统的稳定性和数据的完整性。

在跨境数据流通和分布式计算环境中，实现数据“可用不可见”是保障数据安全与隐私的核心目标[10,14]。这一理念要求在数据价值被充分利用的同时，确保数据的原始内容不被泄露或滥用。大模型在这一过程中发挥了重要作用，特别是在实时数据监测与异常行为识别方面，其强大的分析和推理能力能够对数据流通中的异常情况进行动态预警，并通过溯源机制快速定位安全事件的来源与责任主体。此外，基于大模型的多方数据权限控制与动态认证技术，可以通过对数据访问权限的精细化管理和实时调整，确保数据使用的合法性与透明性，其技术路径见图 1。

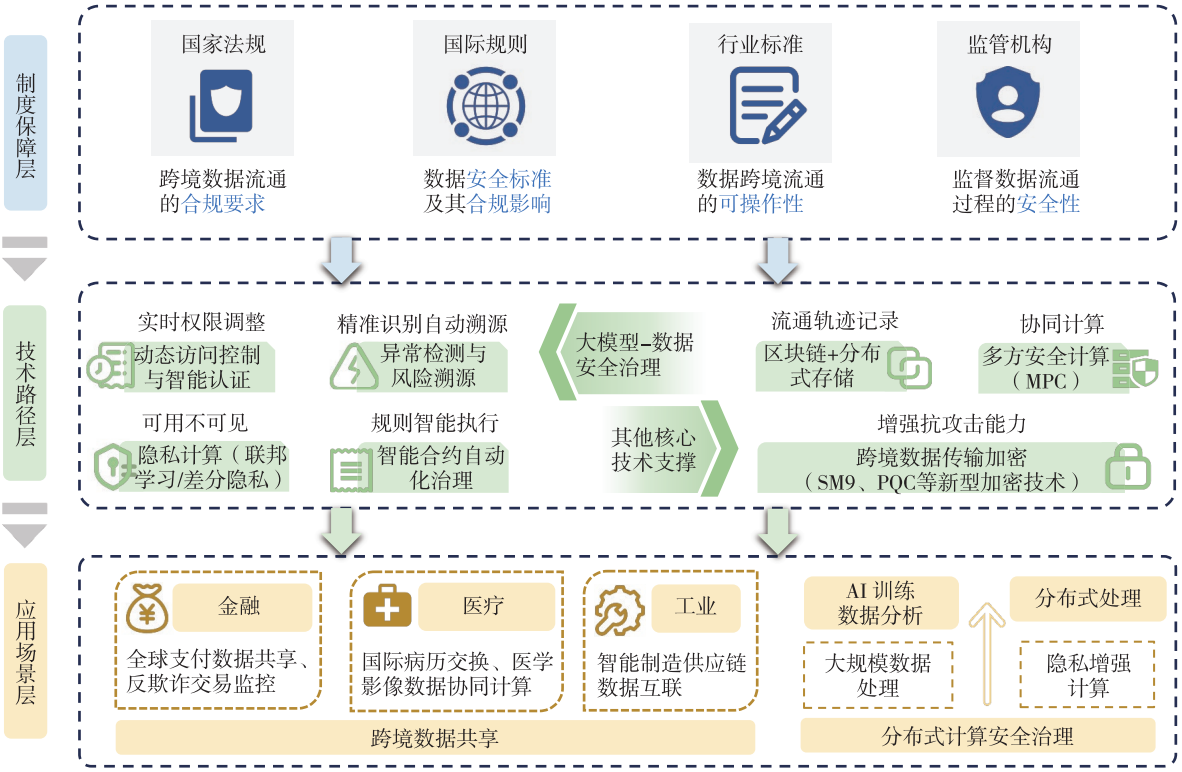

图 1　跨境数据流通与分布式计算在数据安全治理中的技术路径

3　可信数据空间的信任机制构建

可信数据空间有效运行依赖于完善的信任机制构建，需要通过制度保障、技术支撑和多主体协同，构建透明、安全、可追溯的数据流通体系。在流通过程中，信任不仅关涉数据提供方与使用方之间的安全交互，还涉及政府监管机构、行业联盟以及技术平台的协同治理。如何在制度、技术和协作层面构建稳固的信任框架，是实现可信数据流通的关键。

信任的基础首先来源于法律法规与政策框架的规范约束。数据要素的跨主体流通和交易，需要在法律层面界定数据权属、流通边界及合规要求。例如，《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》对数据分类分级保护、数据跨境流通以及用途监管提出了明确要求，确保数据合法合规流转。此外，欧盟《通用数据保护条例》等法规为数据跨境流通提供了不同模式的合规路径。在此背景下，可信数据空间的信任机制需要适配多层次的数据治理规则，并建立跨区域、跨行业的标准化认证体系，确保数据在不同法律体系间的安全互通。

技术信任作为可信数据空间构建的核心支撑，大模型在其中发挥着重要作用。基于大模型的动态访问控制与智能认证技术，使得数据访问权限可以根据用户行为、历史记录和风险评估动态调整，提高数据使用的透明度与安全性。在数据安全监测与异常检测方面，大模型能够通过实时分析数据流，精准识别异常访问和潜在攻击，并结合溯源机制快速定位责任主体。此外，可解释人工智能技术的引入，有助于提升数据处理和共享的透明度，使数据提供方、使用方及监管机构

能够理解模型决策逻辑，减少因算法“黑箱”效应带来的信任障碍。智能合约技术则为数据交易提供了自动化执行保障，通过预设条件触发机制，确保数据流通的安全性和不可篡改性，同时减少人工干预，提高数据治理的效率。

可信数据空间的信任机制需要兼顾多主体协同治理，构建基于规则约束、行为监督和激励机制的动态信任体系。政府在数据治理体系中发挥监管和标准制定的作用，行业联盟和数据交易平台则负责具体规则的落地实施。为了增强信任的可操作性，可以通过建立多主体数据流通的信任评估体系，引入信誉积分、数据流通记录、合规性审核等指标，对数据使用者进行信用评级，并基于信任等级调整数据访问权限。在智能合约的支撑下，这种信任评估可以被嵌入数据交易规则中，实现数据共享协议的自动执行，提高跨主体协作的可信度。此外，监管机构可以通过区块链技术构建透明可追溯的监管链，对数据使用情况进行实时审查，确保数据流通的合规性。

可信数据空间的治理框架需要制度、技术和多主体协同的融合推进。在制度层面，数据分类分级、合规监管和标准化认证体系构成了信任机制的基础约束；在技术层面，大模型、隐私计算和智能合约等技术为数据安全与流通提供了核心支撑；在协同治理层面，政府、行业机构和数据平台共同参与，构建动态信任评估和激励机制，以确保数据空间的稳定运行。这一信任机制的建立，将为数据要素市场化配置提供更加稳健的支撑，为数据共享与交易营造透明、安全、高效的运行环境。可信数据空间的信任机制框架及对应层次见图 2。

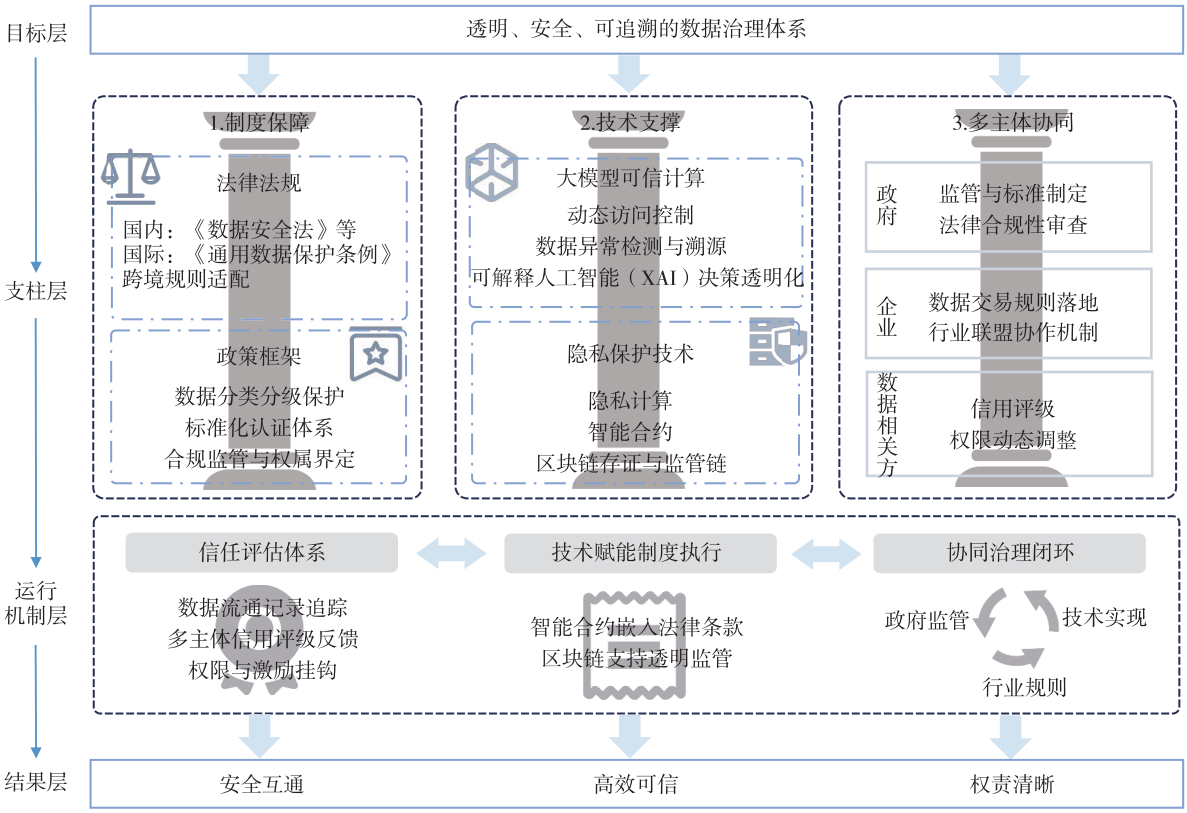

图 2　可信数据空间的信任机制框架

4 讨 论

可信数据空间建设是数据要素市场化配置和数字经济高质量发展的关键所在。在技术、制度和多主体协作的深度交织下，数据安全治理和信任机制构建成为这一领域的重要议题。基于《行动计划》政策框架，本文提出了一套以大模型为核心的数据安全治理体系及信任机制框架，弥补了当前政策在技术适配性、隐私保护、数据可信流通等方面的不足。未来，随着政策执行的深入推进，本文提出的智能合约监管、数据安全可控流通模式，有望为《行动计划》的落地提供技术与治理支撑。

当前，大模型技术为可信数据空间的构建提供了强有力的技术支撑，特别是在实时监测、权限管理和智能合约执行等方面，为数据“可用不可见”的实现奠定了基础。同时，数据安全治理在分类分级保护、动态风险管理与规则体系设计中发挥了核心作用，提升了数据流通的安全性和多主体协作的效率。然而，全球数据治理规则的不统一、技术标准的缺失和跨境流通的合规难题，依然制约着可信数据空间的广泛推广与应用。这些问题的解决，需要从技术创新、法律制度完善、多方合作模式等方面进行全面应对。

各国对可信数据空间建设采取了不同的发展路径。相比之下，我国可信数据空间建设具有明显的政策引导特征。近年来，以《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》为核心的法律体系逐步完善，各地方政府试点建设可信数据空间，形成多行业、多层次的数据治理模式。中国的路径更加强调“数据分类分级 + 技术创新 + 政府监管”的综合治理框架，在政府主导下，构建跨部门、跨区域的数据治理体系，如上海数据交易所、深圳数据要素流通试点等。通过政企协同，探索“数据权属确认—合规流通—价值变现”的新型数据交易机制，增强数据要素市场的流动性与可控性。

5 结 语

本文从可信数据空间内涵出发，聚焦数据安全治理与信任机制构建的核心问题，系统分析了当前建设过程中的关键需求与技术挑战。在此基础上，提出以大模型为代表的智能技术路径，探讨其在隐私保护、实时监测、智能合约执行等方面的适配机制，并进一步构建涵盖制度信任、技术信任与关系信任的多维框架，旨在为可信数据空间的可持续发展提供系统性解决方案与理论支撑。

展望未来，可信数据空间建设需在技术体系与制度规范的深度融合中不断演进，尤其要加快推进国际规则的协调与标准化进程，逐步建立兼容多国法律体系和行业标准的数据流通架构。下一步，应在加强技术创新能力的同时，构建多主体参与的协同治理机制，推动可信数据空间从局部试点走向系统落地、标准化推广。随着全球数据生态持续演化，可信数据空间有望成为支撑数字经济高质量发展和全球数据协作的重要基础设施。

【参考文献】

［1］付少雄，孙建军. 数据流通与安全：标准与保障体系［J］. 图书与情报，2023（4）：20–28.

［2］王艳，杨达. 中国式管理会计体系变革：从数据要素到数据资产［J］. 管理世界，2024，40（10）：171–189.

［3］坚持推进数据要素市场化配置改革——国家数据局介绍数据领域改革进展和成效［EB/OL］.［2025–04–22］. https://www.gov.cn/lianbo/bumen/202407/content_6964034.htm.

［4］刘江峰，张冉，张君冬，等. 以生成式人工智能赋能思想史计算研究：模型构建与应用探索［J］. 图书馆杂志，2025，44（3）：113–127.

［5］唐星龙，张昱，曾文. 复杂信息环境下科技情报技术基础的体系建设研究［J］. 情报学报，2024，43（7）：761–772.

［6］张衡. “数据二十条”下探析数据资源持有权的内涵及框架构建［J］. 信息资源管理学报，2024，14（2）：54–67.

［7］国家数据局关于印发《可信数据空间发展行动计划（2024–2028年）》的通知［EB/OL］.［2025–05–06］. https://www.gov.cn/zhengce/zhengceku/202411/content_6996363.htm.

［8］夏义堃，蒋洁，张夏恒，等. 发展新质生产力的信息资源管理学科回应与思考［J］. 农业图书情报学报，2024，36（1）：4–32.

［9］王雪，夏义堃，裴雷. 国内外数据要素市场研究进展：系统性文献综述［J］. 图书情报知识，2023，40（6）：117–128.

［10］臧国全，肖洋，张凯亮. 政府数据的隐私风险计量与分级保护机制研究［J/OL］. 中国图书馆学报，2024：1–14［2024–12–25］. http：//kns.cnki.net/kcms/detail/11.2746.G2.20241118.1338.002.html.

［11］范如国. 平台技术赋能、公共博弈与复杂适应性治理［J］. 中国社会科学，2021（12）：131–152，202.

［12］刘涛雄，戎珂，张亚迪. 数据资本估算及对中国经济增长的贡献——基于数据价值链的视角［J］. 中国社会科学，2023（10）：44–64，205.

［13］崔文波，张涛，马海群，等. 欧盟数据与算法安全治理：特征与启示［J］. 信息资源管理学报，2023，13（2）：30–41.

［14］孙建军，裴雷，付少雄. 兼收并蓄：信息资源管理学科建设背景下的数据管理［J］. 信息资源管理学报，2023，13（1）：9–17.

［15］王艳，杨达. 中国式管理会计体系变革：从数据要素到数据资产［J］. 管理世界，2024，40（10）：171–189.

［16］杨新涯，王莹，尹伟宏. 数据驱动的新型情报服务研究［J］. 文献与数据学报，2019，1（1）：32–41，117.

［17］李玉海，王蕊. 政府数据开放十年实践与未来展望［J］. 文献与数据学报，2022，4（4）：12–14.

［18］马海群，张涛. 数据生产力的理据阐释与“新质”力量［J/OL］. 中国图书馆学报，2024：1–16［2024–12–25］. http：//kns.cnki.net/kcms/detail/11.2746.g2.20241126.1531.002.html.

［19］刘昭阁，李向阳，乔立民，等. 案例支持下城市灾害风险应对的大数据治理模式分析方法［J］. 情报学报，2024，43（6）：672–684.

［20］刘明达，陈左宁，拾以娟，等. 区块链在数据安全领域的研究进展［J］. 计算机学报，2021，44（1）：1–27.

［21］吴一楷，李国安. 科技治理与治理科技：以区块链数字资产的规制研究为视角［J］. 中国科学院院刊，2024，39（8）：1375–1388.

裴雷，陈晓宇 . 大模型赋能可信数据空间：数据安全治理与信任机制构建［J］. 文献与数据学报，2025，7（2）：015–027.

Large Language Model Enabling Trusted Data Spaces: Data Security Governance and Trust Mechanisms Construction

Pei Lei[1]　Chen Xiaoyu[2]

(1. Data Intelligence and Interdisciplinary Innovation Laboratory, Nanjing University, Nanjing 210023, China; 2. School of Cultural Heritage and Information Management, Shanghai University, Shanghai 200444, China)

Abstract: ［**Purpose/Significance**］ This article explores the key components and implementation pathways of trustworthy data spaces, with a particular focus on the role and impact of large language model (LLM) technologies in data security governance and the construction of trust frameworks. ［**Method/Process**］A combination of theoretical analysis and practical case studies is adopted to examine three core areas: data security governance, technical pathways of LLM, and the establishment of trust mechanisms. ［**Result/Conclusion**］ The development of trustworthy data spaces requires deep integration across institutional frameworks, technological infrastructures, and multi-stakeholder collaboration. LLM offer powerful support for data security governance, particularly excelling in real-time monitoring, anomaly detection, intelligent decision-making, and automated compliance management. The improvement of data security governance hinges on the dynamic optimization of trust evaluation systems, which can enhance the transparency and controllability of trusted data circulation through tiered protection strategies, cross-actor collaboration mechanisms, and enhanced explainability. Ultimately, building trustworthy data spaces is not only a matter of technological innovation but also a paradigm shift in data governance. Its implementation must align with regulatory compliance, interoperability, and international standards to facilitate the efficient allocation of data as a production factor.

Keywords: Trusted data space; Large language model (LLM); Data security governance; Trust mechanism; Data governance framework

（本文责编：孔青青）