# 教育数字化转型下可信数据空间的构建路径

## The Path of Building Trusted Data Space under Digital Transformation of Education

曾颖洁，刘佳音，刘毓，王涛*
[1]华中师范大学 人工智能教育学部 数字教育湖北省重点实验室 湖北 武汉 430079
* tmac@ccnu.edu.cn

**【摘要】** 教育数字化转型背景下，数据孤岛化与低可信度问题严重制约教育数据价值的释放。本文基于新质生产力视角，提出构建教育可信数据空间的系统性框架，旨在通过技术架构与治理机制的协同创新，推动教育数据的智能化转型。研究指出，当前国内教育数据共享面临数据主权归属模糊与跨系统兼容性不足等核心挑战。基于此，提出基于可信数据空间理论的教育数据治理框架，通过构建五层协同架构体系，融合区块链、身份治理等技术，旨在为解决教育数据流通多维困境提供系统性解决方案。通过场景推演验证，该架构支持区域教育资源智能调配与 AIGC 工具跨域训练，有效平衡数据隐私与效用。研究强调，可信数据空间需依托新型数字基础设施与教育主体素养提升，以“治理机制创新、技术工具嵌入、业务场景重构”推动教育生产力向数据智能范式转型，为教育资源的优化配置与教育服务提质增效提供理论支撑与实践路径。

**【关键词】** 可信数据空间；跨域数据共享；数据隐私保护；教育数字化转型

***Abstract:*** *In the context of education digital transformation, the problems of data siloing and low trustworthiness seriously constrain the release of education data value. Based on the perspective of new quality productivity, this paper proposes a systematic framework for constructing a credible data space in education, aiming to promote the intelligent transformation of education data through the synergistic innovation of technical architecture and governance mechanism. The study points out that the current domestic education data sharing faces the core challenges of fuzzy data sovereignty attribution and insufficient cross-system compatibility. Based on this, an education data governance framework based on the theory of trusted data space is proposed, which aims to provide a systematic solution to the multidimensional dilemma of education data circulation by constructing a five-layer synergistic architecture system and integrating blockchain, identity governance and other technologies. The architecture supports the intelligent deployment of regional educational resources and cross-domain training of AIGC tools, effectively balancing data privacy and utility, as verified by scenario projection. The study emphasizes that credible data space needs to rely on new digital infrastructure and education subject literacy enhancement to promote the transformation of education productivity to data intelligence paradigm through “innovation of governance mechanism, embedding of technology tools, and reconfiguration of business scenarios”, and to provide theoretical support and practical paths for optimal allocation of education resources and enhancement of the quality and efficiency of education services.*

**Keywords:** Trusted data space, cross-domain data sharing, data privacy protection, digital transformation of education

## 1. 前言

随着数字化时代的到来，教育领域的转型面临前所未有的机遇与挑战。然而，教育数据的价值释放面临诸多障碍，尤其是“孤岛化”现象和低可信度问题，这些问题阻碍了数据的有效利用。因此，突破这些制约因素，释放教育数据潜力，成为亟待解决的重要课题。为此，本文旨在提出可用的教育可信数据空间设计范式，探讨如何通过合理架构设计推动教育数据

的智能化转型，并揭示技术架构与教育业务场景之间的耦合机制，为教育领域的数字化转型提供理论支持与实践指导。

## 2. 相关研究

随着教育数据应用的不断深化，教育数据空间的研究成为学术界与实践领域关注的焦点。在国际层面，多个国家和地区已在探索如何通过构建教育数据空间来推动教育数字化转型。例如，欧盟通过打造数据空间(Data Spaces),为数据要素流通及服务提供底层基础设施，促进教育数据的流通与利用 (赵琳等人, 2024)，为教育政策制定和资源配置提供支持。美国则通过州级教育数据协作网络实现了不同州间教育数据的共享与协作，有效提升了教育数据的跨系统互通性和可操作性 (杨屿航等人, 2023)。这些国际实践展示了教育数据空间在推动教育现代化方面的重要作用，也为全球教育数据管理和利用提供了宝贵经验。

然而，国内在教育数据空间的建设方面仍面临一系列瓶颈。首先，数据主权归属的问题一直是制约教育数据共享的关键因素之一。不同部门和机构对于数据的控制权和使用权存在争议，导致数据资源配置低，供需之间难以进行有效匹配(杨艳等人, 2024)。其次，跨系统兼容性不足也是国内教育数据空间建设的一大障碍。教育领域内的数据管理系统众多且彼此独立，缺乏统一的标准和接口，使得数据的跨平台流通变得困难 (董晓辉等人, 2019)。国内教育数据共享的低效性主要体现在数据格式不统一和信息传递的滞后性等方面，这严重限制了教育数据的潜在价值释放。

教育数据空间通过数据智能化和可信协同作用，能够有效提升教育资源的配置效率和决策精准度，从而实现生产力的提升。相关跨学科的文献表明，数据空间的构建不仅仅是技术问题，更是制度创新与生产力提升的关键环节。在教育领域，通过建立高效的数据空间，能够突破传统教育生产力的瓶颈，推动教育事业向更加智能化、精准化的方向发展。因此，教育数据空间在提升教育全要素生产率和推动教育质量提升方面具有重要的潜力和现实意义。因此，如何突破这些瓶颈，构建适应中国教育发展的数据空间，成为当前研究的核心问题。

## 3. 教育可信数据空间架构设计

为更好地利用技术解决教育数据要素流通在实践中的困境，亟需构建全社会要素资源的网络化共享、集约化整合、协作化开发和高效化利用 (冯婷婷等人, 2024)。针对教育数据资源流通质量低、系统交互性差及多方主体权责失衡等问题，本文提出基于可信数据空间理论的教育数据治理框架，通过构建五层协同架构体系，融合区块链、身份治理等技术，旨在为解决教育数据流通多维困境提供系统性解决方案。

### *3.1. 数据源层*

数据源层作为教育数据的最初来源，主要涵盖 PC 端、移动端、智能设备及其他数字化教育工具等多个平台。其核心功能为数据收集、初步筛选及数据溯源，通过对数据的采集时间、地点、设备等元数据进行记录，并对所有数据采集终端进行身份认证和设备信任评估，确保数据的可追溯性，保证数据的来源可信。为确保数据的真实性与不可篡改性，本研究提出的功能实现主要依赖区块链技术与数字签名技术。具体而言，通过区块链的分布式账本特性，结合哈希算法和共识机制，同时，采用增强型身份治理（IAM）技术，为数据采集终端和用户分配唯一身份标识 (Esposito C et al., 2021)，并结合多因素认证（MFA）机制，进一步提升身份认证的可靠性和安全性 (Wang Q et al, 2023)。此外，利用可信代理（Trusted Proxy）对数据源进行加密处理，确保数据在采集端的安全性 (S. W er al., 2016)。这种结合区块链、

数字签名、增强型身份治理以及可信代理的技术方案，能够为数据流通生态提供全方位的安全保障，满足零信任架构下的动态访问控制需求。

### *3.2. 传输处理层*

传输处理层的核心功能是实现数据的加密传输与初步处理。在零信任架构下，该层需确保数据在传输过程中的严格加密与动态访问控制。为此，传输处理层需采用加密协议（如SSL/TLS）对数据进行加密传输，确保数据在传输过程中的保密性和防篡改性。同时，通过细粒度的访问控制策略，确保只有经过授权的人员能够访问敏感数据。无论是对原始数据进行去噪、格式化和标准化处理，还是进行数据分类、存档和生命周期管理，都需要依赖高效的算力设施和先进的数据模型，才能确保数据分析和机器学习模型能够在海量数据下高效运行，从而提升数据处理的性能和效率。同时，通过可信代理对数据流进行加密和解密 (U. D et al., 2023)，并动态调整数据传输路径，以应对潜在的安全威胁。

### *3.3. 中间服务层*

在零信任架构下，中间服务层通过微分段（Micro-Segmentation）和持续信任评估（Continuous Trust Assessment）技术增强其安全性 (C. E et al., 2018)，同时实现灵活的数据管理和高效的服务对接。通过将中间服务层划分为多个独立的安全区域，每个安全区域根据数据的敏感度和业务需求配置独立的访问控制策略，有效限制了潜在威胁的横向移动从而降低安全事件的扩散风险。同时，部署信任评估引擎（Trust Assessment Engine），实时监控数据访问行为，并结合用户行为分析（User and Entity Behavior Analytics, UEBA）技术，动态调整访问权限。该引擎能够基于用户行为模式、设备状态和上下文信息，持续评估访问请求的信任级别，确保每次访问请求均符合安全策略 (Martín A G et al., 2021)。通过支持自定义算法模型的运行与优化，例如个性化推荐算法、学习路径优化模型等。通过提供灵活的算法开发环境和高效的计算资源，中间服务层能够满足教育平台在智能化服务方面的多样化需求。中间服务层还需具备数据分级与分类功能，能够根据数据的隐私性和重要性将其划分为不同的敏感度级别。基于数据分类结果，实施差异化的管理策略，例如：个人隐私数据需要严格的加密和访问控制，而公开课程内容则可以开放访问。这不仅实现了数据的精细化管理和动态安全防护，还为教育平台的多样化服务提供了强大的技术支持。

### *3.4. 数据控制层*

为确保数据的访问与操作权限仅授予经过授权的用户或系统，防止未经授权的数据泄露或篡改，数据控制层承担着至关重要的角色。数据控制层通过全面的监管与管理机制，实现对数据访问的实时监控、审计与记录，涵盖每一次访问、修改和删除操作，从而保障数据使用过程的透明性。在数据的全生命周期管理中，数据控制层需依据相关法规和标准，对不再使用或达到存储期限的数据进行安全销毁。这一过程不仅符合数据保护法规的要求，还能有效防止信息泄露。以学校的学生信息管理系统为例，教师在操作学生成绩时，系统需通过严格的身份认证机制，结合多因素认证（MFA）技术 (Wang C et al., 2023)，确保用户身份的合法性。此外，系统基于身份、设备状态、上下文信息等多维度属性，采用基于属性的访问控制（ABAC）模型，对访问请求进行动态授权 (Li J et al., 2022)。所有成绩查询和修改操作均会被自动记录，形成完整的访问日志，以实现数据访问过程的可追溯性。当学生成绩数据超过保存期限时，系统将依据预设的安全策略进行数据的安全销毁，确保数据的完整性和保密性。通过上述措施，数据控制层不仅能够有效防止未经授权的数据访问和篡改，还能满足教育领域对数据安全与合规性的严格要求，为教育数据的可信流通提供坚实保障。。

### *3.5. 数据应用层*

数据应用层作为教育可信数据空间的实际应用层，主要涉及数据的可视化、汇聚与共享。在此层，需实现“可用不可见”的数据共享模式，通过隐私保护计算技术（如安全多方计算、联邦学习、可信执行环境等）确保数据在使用过程中的隐私性。同时，基于数据的隐私性的重要性，实施最小化访问控制策略，确保用户仅能访问其所需的最小数据集。该层整合来自不同平台和层级的教育数据，形成全面的教育数据集，支持数据的汇聚、共享与协同应用。通过数据治理和数据流通机制，促进不同学校之间共享教学资源，以及跨地区教育研究机构共享数据集进行联合研究。此外，数据应用层通过数据可视化技术（如统计图表、仪表盘、地图等），将复杂数据直观呈现，帮助教育管理者、教师、学生等不同角色快速理解数据背后的信息和趋势。同时，利用大数据分析技术和机器学习算法，对教育数据进行深入挖掘，发现潜在规律和趋势。例如，通过机器学习算法进行学生成绩预测、个性化学习路径推荐等，以支持教育的智能化应用。这种多层次的数据应用模式不仅提升了数据的利用效率，还为教育领域的数字化转型提供了坚实的技术支撑。

## 4. 场景推演与可行性分析

本文基于设计科学（Design Science）的架构评估方法，通过理论推演和技术可行性论证，提出了两种典型场景的替代案例验证。设计科学方法强调通过解决实际问题并结合创新技术，验证解决方案的有效性与可行性。在此框架下，我们不仅关注理论上的推导，还着重进行技术层面的可行性验证，以确保所提出方案的实施具有现实基础。

### *4.1. 区域教育资源动态调配*

在传统的区域教育资源调配模式中，往往依赖人工统计与手工调配，且信息流通滞后，导致资源分配不均和反应时间延长 (孟丽菊等人, 2024)。例如，在一些学校，特定学科或设备的资源利用率较高，但由于信息滞后，相关资源无法及时调配，造成教育资源的浪费或不足。为解决这一问题，可通过数据空间实时聚合各校的师资、设备与课程数据，并结合智能合约技术，实现资源调配的自动化与智能化。具体而言，当某一学校的实验室利用率超过90%时，智能合约将自动触发资源共享建议，通过系统推送相关信息，提示周边学校可以共享该实验室资源。此举能够显著提高资源使用效率，减少资源浪费。从技术角度看，区块链技术能够提供高度的透明性与安全性，确保数据的真实性与不可篡改性 (余胜泉等人, 2019)。在具体实现中，可以采用 Hyperledger 作为底层架构，其每秒交易处理量（TPS）足以满足区域教育资源的实时调配需求。此外，智能合约的执行可以保障各参与方的利益，实现自动化的资源共享协议。根据性能测试数据，Hyperledger 能够支持每秒上千笔交易，这对于大规模教育资源的动态调配来说是可行的。

### *4.2. AIGC 教育工具的数据供给*

在教育领域中，AIGC（人工智能生成内容）教育工具的研发确实需要大量的跨校数据进行训练，以提升模型的准确性与适应性。然而，隐私问题成为了一大挑战，尤其是在不同地区和国家之间的数据交流时，如何保障数据隐私，避免泄露敏感信息，成为不可回避的问题。

为应对这一挑战，可以结合联邦学习与数据沙箱技术，在不出域的前提下生成合成数据集进行训练。联邦学习允许多个机构在本地进行模型训练并共享模型参数，而不需要交换原始数据。数据沙箱进一步确保在处理过程中，敏感数据始终留在本地环境中，避免数据泄露。通过构建多层协同架构体系，融合区块链、身份治理等技术，建立跨区域数据共享机制与标

准化体系，旨在为解决教育数据流通多维困境提供系统性解决方案。这种方法不仅能够保护数据隐私，还能够促进数据要素资源的价值转换，提升数据流通效率，降低数据共享的门槛。

## 5.讨论与展望

在推动教育数字化转型的过程中，构建和完善教育数据空间是关键。首先是加强区域新型基础设施建设，推动 5G、大数据、云计算、人工智能等新一代信息技术的应用，持续建设信息网络、平台系统、数字资源、智慧校园、创新应用、可信安全等新型基础设施。打造智慧教育公共服务平台，统筹学习、教学和管理过程中的大数据，建立教育大数据仓，促进教育数据的贯通共享，组建教育大脑，统筹推进数据融合融通，面向学生、教师和学校建立数据应用和分析模式。其次，推进课堂教学过程数字化，教育改革的阵地在课堂，课堂教学是数字化转型的核心。探索基于各种生态的课堂教学过程数字化方式，从教学内容、学习资源、教学过程等方面进行数据采集、分析和应用，实现教学过程的数字化。并且，教育数字化转型的关键是做好教育大数据工作，这涉及到教育活动全过程中产生的海量教育大数据具有巨大的潜在价值。教育大数据可以促进教育从“用经验说话”变为“用数据说话”，实现数据创新、数据管理与数据决策的数字化新形态。通过这些措施，可以有效地构建和完善教育数据空间，为教育数字化转型提供坚实的基础和支持。

尽管上述方案具有显著优势，但在实际应用中仍面临一定的局限性。首先，初期部署的成本较高，尤其是在教育领域，很多学校或机构的数字化基础设施尚不完善，可能需要大量的投入来建设相关系统。其次，方案的实施依赖于教育主体的数字素养，教师和管理人员的数字化能力直接影响方案的实施效果。如果教育主体对新技术的接受度较低，可能会导致方案实施困难或效果不佳。

## 结论

通过“治理机制创新、技术工具嵌入、业务场景重构”三层作用，可信数据空间有效推动了教育生产力向数据智能范式转型。首先，治理机制创新为教育系统提供了规范化的管理框架，确保数据使用的透明性和安全性，提升了各方参与者的信任度。其次，技术工具嵌入为教育实践提供了强大的数据分析和决策支持能力，增强了教育活动的精准性和效率。最后，业务场景重构通过将数据智能与教育业务深度融合，促使传统教育模式向个性化、智能化发展，实现了教育资源的优化配置和教育过程的动态调整。

这种转型不仅提升了教育服务的质量和效率，也为政策制定者和技术开发者提供了新的理论参考。通过对可信数据空间的应用探索，可以更好地理解如何在现代教育环境中推动技术与治理的协同发展，进而促进教育领域的创新和发展。

## 参考文献

董晓辉,郑小斌&彭义平.(2019).高校教育大数据治理的框架设计与实施.中国电化教育(08),63-71.

冯婷婷,刘德建,黄璐璐,曹培杰&曾海军.(2024).数字教育：应用、共享、创新——2024 世界数字教育大会综述.中国电化教育,(03),20-36.

孟丽菊&张心誉.(2024).区域协调推进高质量教育体系建设：何为、难为与应为——兼评《我国教育综合发展水平区域差异研究》.教育理论与实践,44(31),30-35.

杨艳&林凌.(2024).数据要素高质量供给：内涵解析、困境挑战与规制设计.电子政务(11),15-

26.doi:10.16582/j.cnki.dzzw.2024.11.002.
杨屿航&马金晶.(2023).美国信息技术赋能教育的政策和实践演进.基础教育参考,(05),61-70.
余胜泉&李晓庆.(2019).区域性教育大数据总体架构与应用模型.中国电化教育,(01),18-27.
赵琳,钱雨秋&郑汉.(2024).欧盟数据要素市场培育政策、实践与模式.图书馆论坛 44(12),151-160..
Cayirci, E., & De Oliveira, A. S. (2018). Modelling trust and risk for cloud services. Journal of Cloud Computing, 7, 1-16.
Din, I. U., Bano, A., Awan, K. A., Almogren, A., Altameem, A., & Guizani, M. (2021). LightTrust: Lightweight trust management for edge devices in industrial Internet of Things. IEEE Internet of Things Journal, 10(4), 2776-2783.
Esposito, C., Ficco, M., & Gupta, B. B. (2021). Blockchain-based authentication and authorization for smart city applications. Information Processing & Management, 58(2), 102468.
Li, J., Zhang, Y., Ning, J., Huang, X., Poh, G. S., & Wang, D. (2020). Attribute based encryption with privacy protection and accountability for CloudIoT. IEEE Transactions on Cloud Computing, 10(2), 762-773.
Martín, A. G., Beltrán, M., Fernández-Isabel, A., & de Diego, I. M. (2021). An approach to detect user behaviour anomalies within identity federations. computers & security, 108, 102356.
Sun, W., Yu, S., Lou, W., Hou, Y. T., & Li, H. (2014). Protecting your right: Verifiable attribute-based keyword search with fine-grained owner-enforced search authorization in the cloud. IEEE Transactions on Parallel and Distributed Systems, 27(4), 1187-1198.
Wang, C., Wang, D., Duan, Y., & Tao, X. (2023). Secure and lightweight user authentication scheme for cloud-assisted Internet of Things. IEEE Transactions on Information Forensics and Security, 18, 2961-2976.
Wang, Q., & Wang, D. (2022). Understanding failures in security proofs of multi-factor authentication for mobile devices. IEEE Transactions on Information Forensics and Security, 18, 597-612.