

# 健康医疗数据协同治理的现实困境与破解路径

张岚卿<sup>1\*</sup>, 王乐恒<sup>2\*</sup>, 何亦可<sup>1</sup>

<sup>1</sup>天津工业大学法学院, 天津

<sup>2</sup>天津工业大学电气工程学院, 天津

收稿日期: 2026年3月3日; 录用日期: 2026年4月13日; 发布日期: 2026年4月22日

## 摘要

随着医疗行业的数字化转型, 健康医疗数据加速资源化, 但实践中却深陷信息交互壁垒、合规风险不可控及协同低效的现实困境。面对兼具多重利益属性的海量诊疗信息, 传统的单一事后监管模式早已捉襟见肘。本文突破常规的制度修补思路, 探索“规则代码化”的创新实践: 首先通过构建分级治理体系纾解技术割裂下的数据碎片化分布难题, 进而将复杂的合规红线转化为底层智能合约代码, 实现自动化合规审查与权责明晰; 在此基础上, 引入Shapley值搭建价值量化反哺机制, 从利益根源上激发基层机构的协作意愿。这种多维协同方案, 为兼顾隐私保护与要素流通的健康医疗数据生态提供了现实可行的治理范式。

## 关键词

健康医疗数据, 协同治理, 规则代码化

# Practical Challenges and Solutions for the Collaborative Governance of Health and Medical Data

Lanqing Zhang<sup>1\*</sup>, Leheng Wang<sup>2\*</sup>, Yike He<sup>1</sup>

<sup>1</sup>School of Law, Tiangong University, Tianjin

<sup>2</sup>School of Electrical Engineering, Tiangong University, Tianjin

Received: March 3, 2026; accepted: April 13, 2026; published: April 22, 2026

## Abstract

With the digital transformation of the healthcare industry, the resource utilization of health and

\*共同第一作者。

文章引用: 张岚卿, 王乐恒, 何亦可. 健康医疗数据协同治理的现实困境与破解路径[J]. 社会科学前沿, 2026, 15(4): 163-171. DOI: 10.12677/ass.2026.154295

medical data is accelerating. However, in practice, it is deeply trapped in the realistic dilemmas of fragmented information exchange barriers, uncontrollable compliance risks, and inefficient coordination. Faced with massive diagnosis and treatment information possessing multiple interest attributes, the traditional single ex-post regulatory model has long been stretched to its limits. Breaking through the conventional approach of institutional patching, this paper explores the innovative practice of “Codification of rules”: first, by constructing a tiered governance system to alleviate fragmented data distribution caused by technological fragmentation; subsequently, by transforming complex compliance red lines into underlying smart contract code to achieve millisecond-level automated compliance review and clarify rights and responsibilities; on this basis, introducing the Shapley value to establish a value-quantified feedback mechanism, thereby stimulating the data-sharing motivation of grassroots institutions from the root of interests. This multi-dimensional collaborative scheme provides a practical governance paradigm for a medical data ecosystem that balances privacy protection and data element circulation.

## Keywords

Health and Medical Data, Collaborative Governance, Rule Codification

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

传统医疗行业正经历数字化转型, 健康医疗数据因此具备了重要的地位。数据的治理能力, 直接影响着公共卫生安全的保障水平与数字经济的发展质量。为此, 本文从健康医疗数据治理实践出发, 探讨将法律合规规则转化为底层代码的协同治理技术方案。

围绕健康医疗大数据的协同治理, 现有研究呈现出“技术驱动”与“制度规制”两条并行的演进脉络。在技术探索层面, 学者们多聚焦于打破信任壁垒的底层架构设计, 例如利用区块链智能合约构建防篡改的医疗档案管理系统[1], 或探讨联邦学习技术在实现数据“可用不可见”原则下的隐私保护潜力[2]。这类方案虽实现了数据层的安全隔离, 但因未能嵌入医疗场景的具体合规要求, 在实际应用中往往难以落地。与之相对, 制度层面的研究则侧重于宏观的权利分配与风险防范, 如通过比较域外立法模式反思现行医疗应用中法律框架的适用性缺失[3]。

法理层面的分析难以跟得上数据的实时交互, 现有方案在执行层面存在明显的局限。以数据产权理论为例, 其虽然为宏观制度层面“三权分置”提供了合法性基础, 但在具体的跨机构协作场景中, 由于难以精准拆解医疗数据在清洗、合并等过程的价值链条, 传统产权理论在指导实际利益分配时往往大打折扣。同时, 虽然合作博弈论已被广泛用于经济领域的利益清算, 但我们并不能将其直接套用到医疗系统的数据贡献度计算, 还需要额外考虑医疗机构体量不同带来的影响。若不加以修正则难以公正地反应真实贡献。

针对以上局限, 本研究试图在理论与应用层面对现有方案进行补充: 一方面, 提出“技术-法律”双头并进的方案, 将合规红线转化为底层代码以实现程序审查, 修补纯技术方案缺乏合规内核的短板; 另一方面, 针对健康医疗数据交互的非对称特征, 将数据确权逻辑与合作博弈论中的 Shapley 值算法相融合, 把数据贡献转化为可量化的指标。围绕上述构想, 本文首先剖析医疗协同中的数据连通与合规困境, 继而考察域外制度的成熟经验。在此基础上, 通过推演“规则代码化”的实际落地场景, 尝试为本土健康医疗数据治理提供一条可操作的协同路径。

## 2. 健康医疗数据协同治理面临的主要问题

### 2.1. 健康医疗数据的数据碎片化

所谓“数据碎片化”，是在信息产生来源、表现方式、储存位置、归属主体等不同而造成。长期以来，数据和信息被不同主体以不同的标准储存，忽略了同一主体所属信息之间的天然关联性与耦合性，呈现数据割裂的现象[4]。

#### 2.1.1. 传统医学数据处理办法难以适应技术发展

技术的高速迭代给医学行业提出了巨大的挑战，传统的健康医疗行业的数据技术处理难以适应其发展。基层卫生信息化涉及系统的开发、维护、升级等等，需要长期的资金投入。其建设资金主要来自于中央政府和地方财政，因此基层信息化的程度与当地的经济水平密切相关。各地的医疗实践虽各具特色，而对于基层医院以及欠发达地区，技术的应用未经普及。如上海市除了中央财政统一拨款之外，还有大量的经费支持来自于区级财政，对于经济相对不发达的远郊区，有市级财政通过转移支付加以补充；而安徽省基层卫生服务主要来自于中央财政，地方财政的投入有限，因此很多基层卫生信息化管理理念因缺乏资金支持而无法实施[5]。三甲医院拥有高标准数据中心与防护体系，已在使用人工智能辅助决策；基层医疗机构信息化投入不足，存在大量非结构化记录，防护薄弱，甚至对口的社区卫生服务中心还在处理手写的转诊单。

因此，技术设施的差距限制了知识图谱等治理技术在基层医疗机构的应用，导致统一的治理规则无法在基层落地，产生了技术割裂的现象，加剧了数据割裂。数据交互壁垒不仅是连接失效，更是技术应用割裂与生态失衡。

#### 2.1.2. 隐私泄露隐患阻碍底层数据流通

健康医疗数据包含大量敏感个人信息，如患者的既往病史、遗传数据及身份特征等，这些数据如果管理不慎导致泄露，会给患者以及医院带来不可挽回的损失。在跨机构的协作中，数据会流经各个机构，如何保护数据安全不泄露是个难题。基层医疗机构的投入有限，对于数据的保护措施比较松散，成为了数据流通网络防护最薄弱的环节。传输过程中，数据可能被黑客截获；存储环节，内部人员也可能违规调阅。

数据泄露带来的问题是多重的，不单是患者的隐私暴露、权益受损，医疗机构也会因此面临公众信任危机。最坏的情况下，公共卫生安全也会收到威胁。由于数据出域即面对泄露风险，医疗机构在是否共享数据的问题上需再三考虑。正是这种顾虑，使得医疗机构在参与数据共享的时候保持防御性心态，共享意味着潜在的责任风险，而收益又不明确。于是不敢共享成为了公共选择，数据碎片化进一步加深。

### 2.2. 制度与监管存在滞后性

2017年的政策红利期曾催生一批规范性文件，2020年后《数据安全法》《个人信息保护法》相继出台后，地方反而陷入观望，对健康医疗数据领域的管理未展现出应有的积极性。在人工智能，深度学习，联邦学习等技术高速发展的当下，能够针对性引领行业高质量发展的法规暂付阙如。

#### 2.2.1. 法律空白加剧合规难度

对于央地立法呈现出如此“空窗期”的情形，一方面，健康医疗数据治理高度依赖技术标准，而技术标准的制定权分散在工信、卫健、市场监管等多个部门，部门之间亦会存在协同的困难，加大立法难度；另一方面，跨部门规章对“重要数据”认定标准的碎片化与目录制定的裁量权冲突，显著提升了地方立法的合法性风险，客观上引发了规范制定的迟滞。此种制度供给的结构性短缺，并非简单的行政惰

性, 而是折射出现有成文法立法技术在应对高动态、强外部性的数据确权议题时的固有限制。

健康医疗数据存在权属之争。保护个人信息安全是健康医疗数据共享的基石,《个人信息保护法》中对个人信息的定义为“以电子或者其他方式记录的与已识别或者可识别的自然人有关的各种信息,不包括匿名化处理后的信息”,鉴于经法定标准匿名化处理的数据已阻断了个人身份识别链路,对其衍生出的大数据财产权益进行界定时,理应充分考量医疗机构在清洗、结构化等治理环节投入的实质性劳动。医疗机构是健康医疗数据的生产者,也是健康医疗数据的保护者和守护者,诊疗数据是医疗机构提供诊疗服务时产生的附属产物,医疗机构会对数据做清洗处理,在保证数据干净、安全的前提下挖掘数据的价值,赋予医疗机构一定的健康医疗数据财产权,能推动各方的发展。健康医疗数据产生于个人的诊疗过程,经过处理和汇聚后,会体现出公共属性。这种公共属性在传染病防控、医疗资源优化配置等场景中,体现得十分明显。健康医疗数据的公共利益,是政府通过数据治理落实“人民健康优先”原则的核心抓手。因此,需要确立“人格权益-财产权益-公共利益”的法益平衡框架[6]。

在当下的法律环境中,法律合规层面也面临巨大问题。关于“健康医疗数据”,我国法律尚未给出明确的定义。只能参考《信息安全技术健康医疗数据安全指南(GB/T 39725-2020)》和《国家健康医疗大数据标准、安全和服务管理办法(试行)》。法律规则的静态稳定性与技术创新的动态演进性矛盾,最终沉淀为实践中难以消解的合规风险。这种滞后性一旦投射到具体的新型医疗场景中,便会形成一个亟待填补的法律真空。以正在推行的疾病诊断相关分组(DRG)改革为例,它要求将医疗行为纳入标准化的支付框架,这本身就需要强大的数据支撑和精细化的管理工具。然而,当医院试图引入人工智能辅助决策系统来优化这一过程时,问题便接踵而至:AI给出的建议究竟处于何种法律地位?如果采纳了AI的建议导致医保支付偏差,责任该由谁来承担?是开处方的医生,是提供算法的厂商,还是引入系统的医院?这种关于“合规边界”与“责任归属”的模糊地带,使得许多医疗机构在拥抱智能技术时顾虑重重[7]。

不仅如此,健康医疗数据的多样性和复杂性也给数据安全保护带来了挑战,如个人属性数据、健康状况数据、医疗应用数据、医疗支付数据、卫生资源数据、公共卫生数据等,不同的数据采取的标准与技术不同,加大了健康医疗数据合规的难度。

### 2.2.2. 传统监管方法难以定分止争

2015年,全国最大的数据交易所——贵阳大数据交易所(下文简称“贵交所”)正式挂牌运营,但成立5年后,2020年时据前高管爆料称:“一整年贵交所连500万的项目都没有”。贵交所规模庞大、受到了政府与政策的大力支持,在政府的厚望下,贵交所在开办之初也展现出惊人潜力,但是最终也未能尽如人意。医疗健康数据交易作为数据交易的重要组成部分,却未能在市场中充分的流通,优化资源的配置。

当下执法模式依旧是传统监管模式。IDC市场研究公司预计,在2020年,全球42%的电子健康数据会处于无保护状态[8]。在面对海量、高频且深嵌于算法之中的诊疗数据时,政府主导的监管模式的风控能力有所不足。现有监管依赖人工抽查与事后追责,不及难以应对海量高频诊疗数据,亦无法深入算法内部拦截违规行为。尽管我国司法实践已逐步接纳区块链存证的法律效力,但在跨层级医疗机构组成的联盟链网络中,节点的异构性与时间戳时序的不一致性依然给底层证据的完整性认定带来技术挑战。

### 2.3. 价值量化标准缺位导致协同动力缺失

目前跨机构协作主要依靠行政指令推动,各机构缺乏主动合作的意愿,这导致真正意义上的协同难以落地。这种局面的形成,其实是因为利益分配不均。

基层医疗机构承担了大量基础性工作,比如初诊记录、慢病随访数据的采集。这些工作需要投入人力和时间,但数据采集完成后,通常被向上抽离,汇聚到三甲医院或科研平台。三甲医院凭借绝对的资

源与平台优势,天然成为了数据的加工方与利益的收割方[9]。而数据流出后,基层很难看到回报——既没有获得更先进的AI辅助诊疗工具,也没有换来优质专家资源的下沉。数据从基层流出后,很难再以资源或技术的形式回流。

这种现象的根源在于,现有的医疗协作模式基本都靠政府的行政指令推动,而不是因为市场需求自发地进行。政府要求基层共享数据,但贡献数据的多与少,质量如何,目前都无法评定。既然无法量化贡献,就无法按劳分配,只能平均分配。这种“大锅饭”现象,打击了基层机构的积极性[10]。

归根结底,协作动力不足不是因为基层不愿合作,而是因为其贡献无法被量化。基层数据对最终科研成果的贡献到底有多大?目前缺乏公允的计算方法。没有公允的计算方法,投入就无法与回报挂钩,基层的积极性就难以维持。

### 3. 健康医疗数据协同治理的域外经验与模式借鉴

#### 3.1. 以欧盟为代表的统一式制度治理模式

欧盟在数据治理体系上秉持强监管与高标准统一的政策导向,旨在依托《通用数据保护条例》(GDPR),自上而下稳步推进欧洲健康数据空间(EHDS)的建设进程。为宏观制度建设的完整性提供了坚实的制度保障——赋予患者对数据的全面控制权,并设立专门机构实行许可制度,确保数据在匿名化后用于公共利益[11]。

根据GDPR第1条规定的“保护基本权利与自由,特别是个人数据受保护的權利”与1995年《数据保护指令》第十条规定的“保护基本权利和自由,特别是隐私权”可知,欧盟将健康医疗数据涉及的隐私权作为基本权利加以保护,而不单视为传统隐私权客体进行保护。此基础上,欧洲健康数据空间(EHDS)赋予患者对数据的全面控制权——包括访问权、更正权、限制处理权等,但明确拒绝承认患者“拥有”数据。其系统性规划倾向于强化敏感健康数据与基本人格利益的深度锚定,在限制直接财产化交易以换取高信任环境的同时,亦通过许可制力图打通其科研二次利用的流通渠道。但是,由于制度对于个人信息的严格保护,数据利用的激励作用不足,各主体的科研或商业应用面临较高的合规门槛。

#### 3.2. 以美国为代表的分散式制度治理模式

美国采取截然不同的法理逻辑:采取市场驱动、政府引导的实用主义法理特征。法律框架以《健康保险流通与责任法案》为基础,通过21世纪治愈法案打通信息阻塞。核心抓手是可信交换框架与通用协议,通过认证合格健康信息网络构建网络的网络。该模式不建立中央数据库,而是签署通用协议遵循统一标准实现互联,并且,在政府的引导下,鼓励第三方服务,建立基于价值的分配机制,充分激发市场活力,培养内生动力。它充分发挥私营部门活力,通过市场竞争推动标准落地与服务优化,技术迭代快、商业价值释放充分。此时,健康医疗数据虽然涉及隐私,但经过清洗,达到匿名化标准后可以成为可交易的“准财产”。但缺点是碎片化严重,商业壁垒依然存在,各州也存在各具特色的规则使用且隐私法律差异增加了跨域治理复杂性。

#### 3.3. 对我国健康数据管理的借鉴与启示

横向对比欧美经验,我国应探索符合国情的治理体系。健康医疗数据跨境流动与隐私权保护的有效平衡之前提是遵循“合法、正当、必要”原则。在制度上,系统性规划的完整性不可或缺,但必须留出地方政府因地制宜的空间,立法上需要一定的灵活性;在管理上,政府主导不等于政府包办,可以引入市场化认证和第三方服务,当满足一定的条件,允许其参与市场竞争,可以激活参与者的积极性。

但中国现行法律体系尚未对前述问题予以具体化。所以,宏观制度建设需进一步法治化与标准化。

借鉴欧洲健康数据空间, 加快制定统一互操作性标准与分级分类细则, 打破行政与厂商造成的隔离, 明确数据确权与流转红线。此外, 引入市场机制与激励。路径上应坚持政府主导加市场运作, 依托国家平台构建信任底座, 利用隐私计算等技术实现数据可用不可见。

中国与欧美不同的是, “数据主体”并不单一: 个人的私权属性、机构的财产属性以及社会的公共属性。面对此种复合权利与多方利益深度交织的现状, 探寻依托“规则代码化”技术确立契合中国本土化医疗特征的数据权属配置与信任验证范式, 具有突出的现实迫切性。

## 4. 法律与技术融合的健康医疗数据协同治理路径探索

### 4.1. 建立健康医疗数据分级治理体系纾解数据交互壁垒

基层医院因资金有限, 往往没有建立统一的数据存储格式, 大量的病例由手写记录, 影像资料也格式不一; 相较之下, 三甲医院早已建立起高标准的数据中心, 对各种数据实现统一管理。两者之间的数据处理能力存在显著差异, 阻碍了数据的交流。

要解决数据流通时标准不一的问题, 可以从数据接入环节入手。在数据正式对接之前, 先对不同格式的数据进行清洗, 利用 NLP 技术识别手写病历, 用 OCR 技术标准化影像资料, 尽可能地把多源异构的数据统一成一个标准。

统一数据格式后, 就可以引入基于知识图谱的治理机制, 将《数据安全法》的分级分类要求落地。系统会根据数据的来源、敏感程度(是否属于核心隐私)以及未来的应用场景, 自动给数据打上标签。打上标签后, 防护策略也随之区分: 涉及个人核心信息的, 系统会自动施加强脱敏和加密; 而常规的诊疗数据, 则可以走简化流程实现数据共享。

为了解决机构间对于数据出域的隐私顾虑, 可考虑引入联邦学习机制。该机制将数据训练模型发往各个机构, 各机构在无需联网的情况下利用本地数据训练模型, 反馈时只需要上传运行结果, 不需要上传原始数据。数据不出库, 隐私泄露的风险便会大大降低。这种模式使得原本因隐私问题无法共享的数据进入流通网络, 释放了沉睡数据的价值。由于数据没有移动, 也就不会违反《个人信息保护法》第 23 条关于向第三方提供个人信息需取得单独同意的要求。这样的技术架构使数据不动的前提下, 完成价值的流通, 加快市场资源配置的效率[12]。

### 4.2. 推动规则与监管代码化降低合规风险

长期以来, 传统监管高度依赖人工抽查与事后追责, 这种治理的滞后性意味着当违规行为被发现时, 实质性的权益损害往往已经发生。而将抽象的法律规制转化为计算机可直接执行的底层代码, 即科技治理领域前沿的“规则代码化”思想, 可以帮助各方牢牢把握公共安全的主动权和主导权。

#### 4.2.1. 自动化合规审查机制的建立

将法律法规与各地、各机构因地制宜的合规要求写入自动执行的逻辑程序形成智能合约后, 监管的形态将被彻底重塑。以人工智能辅助处方开具为例, 当合规标准编入系统, 医生诊疗过程中, 在点击确认的瞬间, 底层代码将快速完成对患者病症、用药限制与合法性的多维比对; 诊疗结束后, 智能合约机制可在处方开具与数据调阅环节前置嵌入验证节点, 系统依据预设的合规指引进行比对, 实时阻断超适应症、越权操作等违规行为。

在这种前置干预机制的运转下, 确保证据的真实性、合法性与关联性, 确认信息权属、划分清晰的权责关系同样重要。依托区块链技术的防篡改属性, 将诊疗过程与系统拦截的关键核心节点实时加密并永久记录在案, 实现技术留痕, 形成一条能够达到严格司法采信标准的完整电子证据链。对于法院而言, 这种透明清晰的权责界定方法, 或许能在一定程度上缓解它们在面对数据违规争议时, 因缺乏细则标准

而犹豫的态度。

对于权属的界定上, 此类技术方案也能更好的回应《数据二十条》中的“三权分置”制度, 即数据持有权、数据使用权、数据经营权的分离。数据持有权的核心是实际控制数据, 同时保障数据安全。数据使用权的作用是在具体场景中利用数据, 挖掘数据的价值。数据经营权的重点是对数据开展市场化运营, 实现数据的流转和交易。这种权利划分结构, 能通过持有权明确承担数据安全责任的主体。同时, 它能够通过分离使用权和经营权, 让数据资源在不同主体之间有序流转。它能缓解数据私权保护和数据共享之间的矛盾。它能为激活数据要素市场提供制度支持, 为推动数据产品的交易和流通提供制度支持, 也为促进数字经济发展提供制度支持[13]。

通过数字技术的底层支撑与机制优化, 能够完成健康医疗数据的使用权分配; 通过“规则代码化”, 不同的主体在不同的场景可以得到不同程度的赋权; 此时, 多主体同时使用同一数据资源不仅不会产生权利冲突, 反而能够通过多重的应用场景释放出数据的多维价值, 有效实现数据从“静态存储”向“动态增值”的市场转化。

#### 4.2.2. 自动化规则适用的局限性与混合审查机制构建

将合规标准嵌入底层代码确实大幅提升了监管效能, 但这并不意味着算法能够完全取代人类的法律理性。底层逻辑程序的刚性与成文法适用的情境弹性之间存在着天然的矛盾。法律条文往往包含“合理”“必要”等不确定性概念, 为复杂的现实情境留出依情裁量的空间; 而智能合约执行的是确定的代码。在充满未知与人文伦理考量的医疗场景中, 并不能单纯依赖预设程序进行自动化裁判。试想在突发急救等极端情形下, 如果系统因患者授权期限刚好逾期而触发自动拦截信息获取, 将直接阻断主治医师获取过敏史等关键信息的渠道, 从而延误抢救时机, 构成对生命权和健康权的实质性侵害。

应对此类风险, 单一的代码规制需要转向“人机结合”的混合治理模式。一种可行的做法是, 在自动化系统之外设立独立的审查委员会, 成员涵盖医学伦理、法律和技术等专业背景。该委员会的职责贯穿整个周期: 在算法设计初期进行价值对齐与伦理评估, 在系统运行期间, 针对因自动化拦截产生的边缘争议案件启用人工复核, 必要时可赋予特殊场景豁免权。此外, 智能合约需要保留适当的更新机制, 以便在法规修订或突发公共卫生事件等情形下, 合法调整合规阈值。通过这种技术刚性与人文弹性的精细调和, 治理框架得以在保障数据流通效率的同时, 坚守医疗行为的伦理与法律底线。

“人机结合”混合裁决机制的建立, 是在代码效率和人文考量之间的平衡。不是过度追求用技术代替人类的判断, 而是用智能合约应对常规情景, 让人来决定特殊情况下的选择。二者各司其职。

#### 4.3. 搭建价值量化反哺机制破解单向虹吸难题

基层机构不愿共享数据, 表面上看是意愿问题, 根源上其实是利益分配问题。他们承担了数据采集的大部分基础工作, 但数据流向上层后, 换不回任何实质性的回报, 比如专家资源或新技术。要打破这种单向虹吸, 难点在于如何量化基层数据的贡献。没有一个公认的计算方法, 任何反哺都无从谈起。

在激励机制设计中, 可以引入合作博弈理论中的 Shapley 值。这一方法的核心思想是: 不按数据量多少平均分配收益, 而是根据每个参与方对最终成果的边际贡献来分配[14]。以医疗 AI 模型训练为例, Shapley 值可以测算某家基层医院的数据在多大程度上提升了模型的诊断精度。贡献越大, 获得的回报理应越多。

当然, Shapley 值只是理论算法, 落地需要配套的技术平台。这个平台需要记录每一次数据上传的时间、数量和质量, 并据此计算贡献积分。当基层机构积累到一定积分, 就可以公平兑换三甲医院的会诊配额、AI 辅助诊断工具使用权等资源。

这种机制可以将协作模式从行政指令驱动转向利益驱动。为了获取更多的医疗资源, 基层医院会主

动提升自身的数据质量, 而不是像先前那样被动完成任务。有了量化标准, 就可以进行按贡献分配, 基层就有了持续参与的动力。

#### 4.4. 价值量化机制的模拟验证: 基于 Shapley 值的算例分析

为验证 Shapley 值量化机制在评估医疗体系中各方贡献度的可行性, 本节构建一个简化的跨机构数据协同博弈模型进行数学推演。假设区域数据联盟内存在一家掌握海量常规诊疗数据的三甲医院  $A$ , 以及两家基层社区卫生服务中心  $B$  与  $C$ 。其中,  $B$  积累了少量但极具价值的糖尿病稀缺并发症随访记录,  $C$  则仅提供少量常规基础体征数据。以联合训练辅助决策模型的整体效用值  $v$  (涵盖模型精度与泛化能力, 满分为 100 单位) 作为分配基准, 单独建模时各方因数据维度受限效用均处于得分低位, 设定  $v(A) = 50$ ,  $v(B) = 10$ ,  $v(C) = 10$ 。

在多方协作场景中, 机构间数据维度的互补性开始显现。当三甲医院  $A$  与掌握稀缺数据的  $B$  融合时, 有效填补了模型在罕见病理上的认知盲区, 效用激增至  $v(A, B) = 85$ ; 相比之下,  $A$  与提供同质化数据的  $C$  合作带来的边际提升显著收窄,  $v(A, C) = 65$ ; 两家基层机构联手效用为  $v(B, C) = 25$ 。当三方全面开放协同, 模型达到全局最优  $v(A, B, C) = 100$ 。

为了打破粗放式的均分或按数据体积分配的积弊, 系统引入合作博弈论中的 Shapley 值公式测算各方真实贡献:

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!} (v(S \cup \{i\}) - v(S))$$

基于该算法对所有可能协作排列下各节点边际贡献的加权平均, 利益清算结果发生根本性扭转。最终, 三甲医院  $A$  的收益被精准界定为 63.3 单位; 掌握稀缺数据的基层机构  $B$  获得 23.3 单位的反哺资源; 仅提供常规数据的  $C$  获得 13.3 单位。

如果以数据量为唯一贡献度评价标准, 基层机构一般只能分得 5% 的残值。但在 Shapley 值算法之下, 我们可以看到数据提供方  $B$  没有因为数据量少而被边缘化, 反而精准锁定了其提供的稀缺数据, 为  $B$  分得了远高于按量分配机制下的分值。该例的计算结果提供了初步的经验证据, 证实了该量化机制可以识别数据质量与稀缺性对模型的贡献, 让拥有高价值数据的机构可以获得与之贡献相匹配的收益分配。本例只是简单的数学推演, 现实场景中需要考虑的数据重叠、模型效用精确度等问题, 这是我们后续的研究目标之一。

#### 4.5. 医联体糖尿病共病管理的场景推演

为了验证上述多维协同治理机制的可行性, 本节将对区域医联体内的糖尿病慢病管理进行推演。以某社区为例, 其档案室留存着近三年来的糖尿病患者血糖记录, 其中不少是家庭医生上门随访后手写的纸质单, 或是不同品牌血糖仪导出的格式各异的 Excel 表格。当三甲医院试图调用这些底层信息训练预测模型时, 首要工作就是数据清洗。这不仅涉及将手写病历电子化、统一化, 更关键的是要识别哪些是敏感信息。为此, 多维知识图谱需要先行介入, 依据《个人信息保护法》将数据划分为核心隐私、一般医疗数据和可公开数据, 从而确立后续流转的边界。

三甲医院发出联合训练请求时, 智能合约程序便会激活。这套程序可以在数据出库前检测数据调用请求是否合规, 一旦监测到患者家庭住址、联系方式等核心隐私信息试图越权出域, 底层代码便会即刻拦截传输并通过区块链技术保存操作日志。在此硬性的底层机制下, 联邦学习机制允许基层在不交出原始数据的前提下利用本地算力训练, 仅将加密参数回传主节点。这从操作层面消除了隐私泄露的顾虑。

仅靠消除隐私泄漏隐患并不足以长久维持协作机制, 影响基层参与医院的症结实际上在于数据单向

流出引发的利益失衡。为此, 系统在模型联合训练完毕后引入沙普利值(Shapley value)算法进行贡献度清算。该算法能够计算不同参与者对最终模型的贡献度。例如, 某家社区医院提供了一位糖尿病性大疱病(糖尿病中的边缘病症)患者的医疗数据, 尽管数据量很少, 但是对模型的实际贡献是巨大的。Shapley 值的巧妙之处即在此, 他不是只根据数据的量来判断贡献, 而是同时评估质和量。平台据此向高贡献度基层倾斜医疗资源。通过科学的贡献评估工具, 原本依赖行政指令的被动任务, 才能真正转化为利益驱动的自发行为, 一个兼顾隐私保护与要素价值流通的协同闭环由此落地。

## 5. 结语

解决健康医疗数据协同问题, 没有一招鲜的简单方案。本文尝试论证, 无论是单纯的制度修补, 还是孤立的技术应用, 都难以应对当前“数据割裂、合规风险、协同低效”这三重挑战。真正有效的路径, 可能是法律与技术的深度融合。本文正是对这种融合的初步探索: 它用知识图谱让数据标准对齐, 用智能合约让法律条文变成可执行的代码, 用区块链和联邦学习让协作变得可信且有激励。当然, 这套方案仍处于构想阶段。要实现数据“可用不可见”的愿景, 未来还需要跨学科持续研究, 以及在更多试点中的检验与磨合[15]。

## 基金项目

天津市大学生创新创业训练计划项目《健康算法盾——基于多维知识图谱的医疗智能合约协同治理中枢》, 项目编号 202510058090。

## 参考文献

- [1] Khatoun, A. (2020) A Blockchain-Based Smart Contract System for Healthcare Management. *Electronics*, 9, Article No. 94. <https://doi.org/10.3390/electronics9010094>
- [2] 刘辉, 陈晓云. 一种利用联邦学习基于区块链的医疗数据共享隐私保护方法[P]. 中国, CN114996762A. 2022-09-02.
- [3] 张宇清. 中国健康医疗数据政策法规现状及域外立法模式借鉴[J]. 医学信息学杂志, 2024, 45(2): 26-31.
- [4] 张畅, 李卫. 面向医疗健康领域的联邦学习综述: 应用、挑战及未来发展方向[J]. 工程科学学报, 2025, 47(9): 1825-1840.
- [5] 高星, 吕欣航, 徐晓敏, 等. 基层卫生信息化现状、问题与对策研究[J]. 中国数字医学, 2019, 14(9): 5-7.
- [6] 武翠丹. 健康医疗数据应用的法律规制[J]. 华东政法大学学报, 2025(5): 68-80.
- [7] 周俊, 方国英, 吴楠. 联邦学习安全与隐私保护研究综述[J]. 西华大学学报(自然科学版), 2020, 39(4): 9-17.
- [8] 陈锦, 徐秉楠. 健康大数据安全要闯哪些技术关[N]. 健康报, 2018-10-13(003).
- [9] 薛腾飞, 傅群超, 王枫, 等. 基于区块链的医疗数据共享模型研究[J]. 自动化学报, 2017, 43(9): 1555-1562.
- [10] 周正强, 陈玉玲, 李涛, 等. 基于联盟链的医疗数据安全共享方案[J]. 应用科学学报, 2021, 39(1): 123-134.
- [11] 李慧敏, 陈光. 论数据驱动创新与个人信息保护的冲突与平衡——基于对日本医疗数据规制经验的考察[J]. 中国科学院院刊, 2020, 35(9): 1143-1151.
- [12] 潘燕杰. 被遗忘权保护健康医疗数据的实践困境与对策[J]. 医学与社会, 2024, 37(9): 137-144.
- [13] 顾昕. “健康中国”战略中基本卫生保健的治理创新[J]. 中国社会科学, 2019(12): 121-138.
- [14] 黄丽华, 郭梦珂, 邵志清, 等. 关于构建全国统一的数据资产登记体系的思考[J]. 中国科学院院刊, 2022, 33(10): 1426-1434.
- [15] 殷明雪. 健康医疗大数据建设与隐私保护之冲突与协调[J]. 医学与社会, 2023, 36(2): 125-131.