教育教学与管理 2026年第7期
ISSN: 3079-4439

数据活动图谱视角下数据安全风险评估机制研究

任鹏飞
中镪核信（上海）认证服务有限公司，上海，中国

【摘要】在数据安全治理实践中，制度规则的供给已较为密集，难点逐渐转向组织能力与评估工具的建设。现行分类分级、风险评估和合规管理虽然构成了治理的基本框架，但在落地过程中仍容易出现分类标签化、评估静态化、责任链条断裂以及安全投入收益难以说明等问题。传统信息安全风险评估通常围绕系统资产、威胁和脆弱性展开，对于数据在跨主体、跨场景、跨系统流动中产生的复合风险解释不足。基于对数据治理、风险治理和数据安全评估研究的梳理，本文提出“数据活动图谱”分析框架，将数据类型、处理主体、处理目的、流转路径、安全控制、风险后果和责任证据纳入统一评估视野，并据此构建“固有风险—控制有效性—残余风险—风险收益”的动态评估机制。研究指出，数据安全风险评估不能停留在数据是否敏感的判断上，更应考察特定活动链条中损害如何形成、控制能否生效以及责任如何归属。政务数据共享、企业用户画像和医疗健康数据委托处理三个场景的分析表明，该框架有助于提高数据安全治理的可解释性、可审计性和可问责性。

【关键词】数据安全治理；风险评估；数据活动图谱；分类分级；残余风险；数据要素流通

1.引言

随着数字经济向生产、治理和公共服务深处延伸，数据已经不只是被保存和调用的资源，而是在流动、关联、建模和复用中不断生成价值的要素。风险也因此呈现场景依赖性：同一组数据由不同主体基于不同目的处理，或者在不同技术条件下被组合使用，可能对应完全不同的风险水平；一些原本敏感度较低的数据，一旦与身份、位置、交易、健康或行为轨迹相连，也可能形成画像推断和差别化决策风险。

我国已形成以《网络安全法》《数据安全法》《个人信息保护法》《网络数据安全管理条例》及相关国家标准为支撑的数据安全制度体系。《数据安全法》确立了数据分类分级保护、风险监测、风险评估和应急处置等制度[1]；《个人信息保护法》强调个人信息处理的合法、正当、必要原则，并要求在特定情形下开展个人信息保护影响评估[2]；《网络数据安全管理条例》进一步细化了网络数据处理者义务、重要数据保护、个人信息保护和数据跨境安全要求[3]。但制度文本的丰富并不当然意味着风险被充分看见。实践中更突出的问题是：规则越细，组织越容易把治理压缩为材料留痕；流程越完整，责任链条反而可能在部门、平台和第三方之间被稀释。

目前不少数据安全风险评估仍沿用网络安全评估思路，重点放在系统资产识别、威胁识别和脆弱性分析上。ISO/IEC 27005 等标准提供了较成熟的信息安全风险管理路径[4]，但在数据跨部门共享、平台化运营、算法模型训练、第三方委托处理和跨境传输等场景中，风险并不总是稳定附着于某一系统资产，而是随着数据被调用、组合、推断和再利用而变化。本文关注的问题由此展开：在数据要素流通与安全治理并重的背景下，如何构建一种既能对接现行制度要求，又能呈现动态流动风险的数据安全评估机制？

本文据此主张，数据安全风险评估需要从“数据对象中心”转向“数据活动中心”。数据本身只能说明风险的起点，不能单独决定风险的强度和后果；真正需要评估的是数据类型、处理目的、主体关系、流转路径、技术控制和外部影响之间的组合关系。基于这一判断，本文提出“数据活动图谱”框架，并在此基础上建构动态风险评估机制。

2.文献综述与研究缺口

既有研究大体可以归纳为三条线索。第一条线索是数据治理研究。国外研究较早讨论数据治理中的组织结构、决策权配置和数据质量管理。Khatri 和 Brown 指出，数据治理本质上是围绕数据相关决策权和责任边界

展开的制度安排[5]。国内研究则更多与数据要素市场建设相结合，关注公共数据开放、企业数据合规、数据权益配置和数据流通机制[6,7]。这类研究为理解数据治理的制度属性提供了基础，但对安全风险如何在具体处理活动中生成，仍缺少足够细的分析工具。

第二条线索是风险治理研究。风险治理理论强调风险识别、评估、沟通和控制之间的闭环关系。NIST 隐私框架将隐私风险纳入企业风险管理过程，重视识别、治理、控制、沟通和保护之间的衔接[8]。ISO/IEC 27005 则以资产、威胁、脆弱性、影响和可能性为核心开展评估[4]。这一方法对传统信息系统安全管理具有较强适用性，但迁移到数据场景时会遇到边界问题：数据可以被低成本复制、快速流转和反复利用，风险随处理活动的展开而变化，并不完全等同于系统资产风险。

第三条线索是数据安全与个人信息保护评估研究。国内的实践应用多集中于数据分类分级、重要数据识别、数据安全风险评估、个人信息保护影响评估和数据出境安全评估等议题。国内相关标准对数据安全能力成熟度、个人信息处理规范、数据分类分级和数据安全风险评估方法提供了操作依据和时间方案[9-12]。这些研究推动了体系建设和合规实践，但仍有两个环节较为薄弱：其一，分类分级与风险评估之间缺少明确的衔接机制，这导致数据分类标签与后续安全控制策略脱节；其二，多主体、多场景流动中的责任证据因缺少取证依据而未充分纳入评估，风险发生后的可追溯和可问责的技术方式与制度保障仍然不足。

说到底，现在的研究就是缺一个能把“数据怎么流、风险在哪里、安全如何做、责任谁去扛”这四件事串在一起的框子。本文干脆换个路子：不聚焦死板的数据资产，把目标放在的数据活动。然后拿“数据活动图谱”当工具，把找风险、评措施、定责任这三步，全塞进一个分析结构里搞定。

3.数据活动图谱：数据安全风险评估的分析框架

本文所称数据活动图谱，是指围绕特定数据处理目的，对数据从收集、存储、使用、加工、传输、提供、公开到删除全过程中的主体、行为、路径、控制和责任关系进行结构化测绘。与传统数据目录相比，数据活动图谱不仅仅回答“有哪些数据”，还需要进一步了解数据为什么被处理、由谁处理、流向何处和可能造成何种损害以及责任最终如何落定。

数据活动图谱至少包括七个维度，具体见表 1。从数据类型影响初始敏感程度到处理主体影响责任复杂度，来认定处理目的划定正当性边界，另外，处理行为和流转路径决定风险扩散方式，安全控制反映风险削减能力，而责任证据则决定治理能否被审计和追溯。

表 1.数据活动图谱的构成维度

维度	主要内容	风险识别重点	证据要求
数据类型	一般数据、个人信息、敏感个人信息、重要数据、核心数据、商业秘密数据	敏感程度、规模、可识别性、可关联性	数据目录、分类分级记录
处理主体	数据处理者、受托处理者、共同处理者、接收方、云服务商、算法服务商	主体数量、权限边界、委托关系	合同、授权书、责任矩阵
处理目的	内部管理、公共服务、营销推荐、风控审核、科研分析、算法训练	正当性、必要性、目的变更	审批记录、影响评估报告
处理行为	采集、查询、融合、脱敏、导出、共享、公开、删除	高危操作、批量处理、二次利用	日志、工单、导出审批
流转路径	内部流转、跨部门共享、第三方委托、平台接口调用、跨境传输	路径长度、扩散范围、再共享风险	接口清单、调用日志、出境材料
安全控制	身份认证、最小授权、加密、脱敏、监测、告警、应急响应	控制覆盖率和有效性	配置记录、告警记录、演练记录
责任证据	审批、授权、合同、日志、整改、删除、审计材料	责任是否可追溯	证据链和留存期限

借助上述维度，数据安全风险评估可以从单点判断转向链条分析，避免把风险简单归因于某一数据表、某一系统或某一部门。更重要的是，该框架能够把分类分级结果与控制措施、残余风险接受和责任追溯连接起来，从而缓解分类制度与风险评估制度之间

常见的脱节。

4.动态风险评估模型：从固有风险到风险收益

基于数据活动图谱，本文构建“固有风险—控制有效性—残余风险—风险收益”的动态评估模型，见图 1。该模型与 GB/T 45577—2025 关于风险识别、风险分析、风险评价和风险处置的基本思路相衔接[12]，但其重点不在于复述一般风险管理流程，而在于把数据处理活动的场景差异、控制证据和责任边界纳入评估过程。

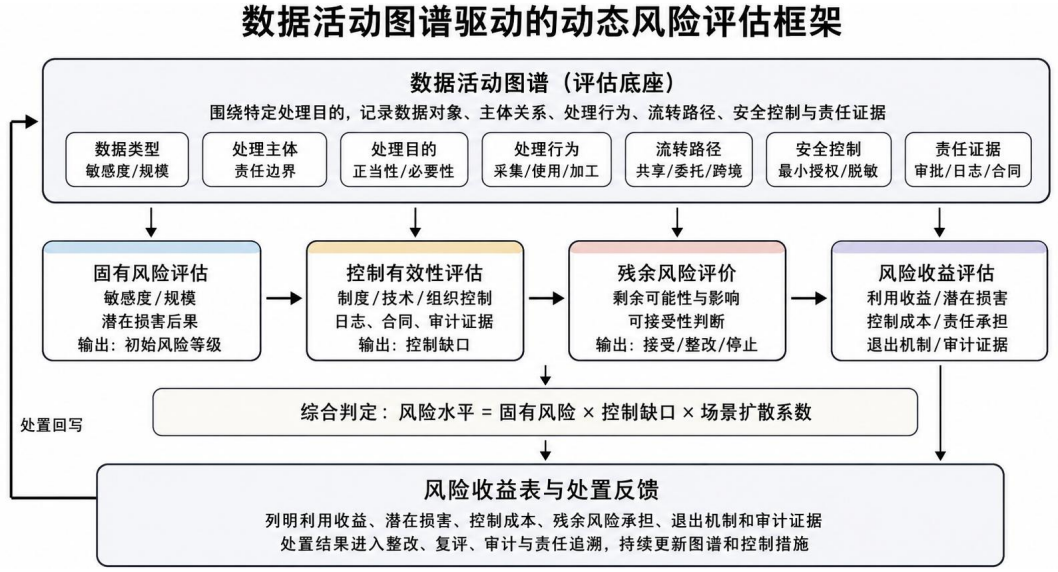

图 1.数据活动图谱驱动的动态风险评估框架

固有风险说的是：还没上任何安全措施之前，这个数据活动本身有多大风险。很多人做评估只看数据叫什么名字、是什么字段类型，这不够。你得同时看几件事：数据本身敏不敏感、体量有多大、涉及多少主体、在哪些范围流转、用来干什么、加工强度有多大，以及万一出事后果有多严重。举个简单的例子——同样是用户消费数据，拿来做内部统计，跟拿来做个性化推荐、信用评分、自动化决策，风险根本不在一个量级，后者显然要重得多。

控制有效性评估。固有风险说的是：还没上任何安全措施之前，这个数据活动本身有多大风险。很多人做评估只看数据叫什么名字、是什么字段类型，这不够。你得同时看几件事：数据本身敏不敏感、体量有多大、涉及多少主体、在哪些范围流转、用来干什么、加工强度有多大，以及万一出事后果有多严重。举个简单的例子——同样是用户消费数据，拿来做内部统计，跟拿来做个性化推荐、信用评分、自动化决策，风险根本不在一个量级，后者显然要重得多。

残余风险评估。残余风险就是上了控制措施之后还剩多少风险。数据安全治理的目标不是把风险降为零——这不现实，也没必要。关键是剩下的风险能不能跟处理目的、数据价值、责任承受能力匹配上。残余风险还偏高的，得加措施：比如强化审批、缩短留存期限、限制导出、做独立审计、定期复评。残余风险已经到不可接受程度的，该停就停，该禁就禁。

风险收益评估。做风险评估和整改，目的不是把数据流通一刀切地堵死，而是在可控的前提下把数据价值释放出来。倾向于用一张"风险收益表"来辅助决策：数据活动启动前，同步列清楚数据利用能带来什么收益、可能造成什么损害、控制成本是多少、残余风险由谁承担、退出机制怎么设、审计证据怎么留。这样可以避免两个极端：一是打着安全的旗子把合理利用空间压得太死；二是打着发展的旗子把安全成本和外部性甩给个人、社会或者公共部门。

该模型可概括为：风险水平 = 固有风险 × 控制缺口 × 场景扩散系数，见表 2。固有风险主要由数据敏感度、规模和潜在损害决定；控制缺口来自制度、技术、证据三个维度的不足；场景扩散系数则看涉及多少主体、流转路径有多复杂、数据会不会被再加工再利用。这个公式不是要算出一个精确数字，而是给组织做分级评估、给监管部门做差异化监管，提供一个可以讨论、可以校准的框架。

表 2.数据安全风险评估指标体系

一级指标	二级指标	评估问题	主要证据	输出结果
固有风险	敏感度、规模、损害后果	数据被泄露、篡改、滥用后会造成何种损害	分类分级记录、业务说明	初始风险等级
活动风险	主体、目的、行为、路径	处理目的是否必要，流转路径是否过长，主体是否过多	活动图谱、接口清单、审批记录	风险触发点
控制有效性	制度、技术、组织控制	控制措施是否覆盖关键节点并持续运行	日志、配置、演练、审计报告	控制缺口
残余风险	剩余可能性与影响	现有控制后风险是否仍可接受	复评记录、整改报告	接受/整改/停止
风险收益	收益、成本、责任、退出	数据利用收益是否足以支撑相应安全成本与外部性	风险收益表、责任矩阵	流通决策

5.典型场景分析

为检验上述框架的解释力，本文选取政务数据共享、企业用户画像与算法训练、医疗健康数据委托处理三个典型场景进行分析，见表 3。三类场景分别对应公共部门数据流通、平台企业商业化利用和高敏感行业数据处理，能够较集中地呈现数据安全风险在主体结构、处理目的和责任边界上的差异。

表 3.三类典型场景中风险识别与治理重点

场景	主要风险	图谱关注点	治理重点	残余风险处置
政务数据共享	共享目的泛化、授权边界模糊、二次共享失控	事项依据、字段清单、调用频率、使用部门	按事项授权、字段最小化、全量日志审计	高敏数据定期复评
企业用户画像	多源融合导致敏感推断、自动化决策不透明	采集来源、标签体系、模型用途、决策影响	用途限制、拒绝渠道、敏感标签治理	高风险画像禁用或降级
医疗健康委托处理	受托方权限过宽、脱敏不足、模型反推信息	委托协议、训练数据、模型输出、删除义务	数据隔离、强脱敏、受托方审计	独立审计与模型风险测试

5.1 政务数据共享场景

政务数据共享这事，公共价值高，但麻烦也多——涉及的主体层级多，社会影响大。大家平时谈风险，第一反应往往是数据泄露，但实际上政务数据共享的风险远不止这个。更常见的问题是：共享目的被泛化、授权边界模糊、数据给出去之后管不住。

举个现实中的例子。基层治理、公共服务、行政执法，这几个场景经常会同时调用户口、社保、医疗、住房、市场监管等各类数据。现在的情况是，监管部门能看到数据"已经被共享了"，但至于这次共享是不是有必要、有没有超范围、共享之后又引入了什么新风险——往往是一笔糊涂账，因为没有完整的数据活动图谱支撑。

所以在政务数据共享这个场景里，做风险评估要盯住几个关键点：共享目的是什么、有没有法定依据、字段是不是做到了最小化、调用频率是否合理、是哪个部门在用、日志能不能审计、二次共享有没有限制。

特别是高敏感政务数据，我觉得"按部门授权"这个思路有问题。更合适的做法是按事项授权——把授权边界绑定到具体的行政事项上，而不是笼统地给某个部门开一个长期访问权限。这样至少能知道每次调数据是为了办哪件事，边界清晰一些，后续追责也更容易。

5.2 企业用户画像与算法训练场景

平台企业通常会将用户注册信息、浏览记录、交易记录、位置数据和设备信息用于用户画像，并进一步服务于精准推荐、广告投放、风险控制或自动化决策。该场景的风险具有较强隐蔽性：单项数据看似敏感度不高，但多源融合后可能推断出收入水平、健康状况、消费能力、社会关系和行为偏好。Solove 关于隐私损害分类的研究提示，隐私风险并不限于信息泄露，也包括聚合、识别、二次使用和排除等多种损害形式[13]。

因此，对该场景的评估不能停留在隐私政策是否告知用户这一层面，还应进一步审查画像目的是否明确、数据融合是否必要、算法训练数据是否超出授权范围、自动化决策是否提供解释和拒绝渠道、敏感标签是否

被滥用。数据活动图谱可以呈现“采集—融合—建模—推断—决策”的完整链条，从而将个人权益风险纳入安全评估。

5.3 医疗健康数据委托处理场景

医疗健康数据兼具高度敏感性和公共价值。医疗机构可能将数据委托给技术公司，用于系统运维、科研分析、人工智能辅助诊断模型训练或区域健康平台建设。在这一场景中，风险主要来自受托方访问权限过宽、数据脱敏不足、科研目的与商业目的混同、模型训练数据难以删除以及成果反向识别等问题。

因此，评估应重点关注委托处理协议、访问控制、脱敏匿名化强度、训练数据隔离、模型输出风险、受托方删除义务和安全审计。尤其需要把“模型是否可能记忆或反推个人健康信息”纳入残余风险评估。由此可见，在人工智能应用背景下，数据安全风险已经从原始数据层延伸到模型层和输出层。

6.治理建议

第一，推动分类分级从“标签管理”转向“场景管理”。分类分级不能只依据字段名称和数据库归属作出判断，还应结合处理目的、流转范围、数据融合程度和损害后果动态调整。GB/T 43697—2024 确立的数据分类分级规则为组织建立基础目录提供了依据[11]，但分类结果只有嵌入具体数据活动链条，才能真正转化为控制要求[14]。

第二，建立数据活动图谱制度。对重要数据、敏感个人信息、大规模个人信息处理、公共数据共享和数据出境等高风险活动[15]，应要求形成数据活动图谱，并将其作为风险评估、合规审计和责任追溯的基础材料。

第三，强化控制证据而非形式合规。监管和内部审计不宜只检查制度文件是否齐备[16]，还应审查审批、授权、日志、合同、告警、整改和删除等证据是否形成闭环。缺少证据链支撑的控制措施，不能被简单认定为有效控制。

第四，建立残余风险接受机制。对于高风险数据活动，应明确残余风险由哪个管理层级批准[17]、由哪个部门承担、通过何种措施持续监测，避免风险责任在技术部门、业务部门和法务部门之间反复转移。

第五，将风险收益表嵌入数据流通决策。对于公共数据授权运营[18,19]、企业数据共享、算法训练和跨境传输等活动，应同步评估数据利用收益和安全外部性，使安全与发展的平衡具有可解释、可复核的依据。

7.结论

数据安全治理的核心问题，正在从“有没有制度”转向“制度能否转化为可执行、可验证、可问责的治理能力”。传统以静态资产为中心的风险评估方法，难以解释数据在多主体、多场景、多技术链条中形成的动态风险。本文提出的数据活动图谱框架，将数据类型、主体关系、处理目的、流转路径、安全控制和责任证据纳入统一分析，并进一步构建“固有风险—控制有效性—残余风险—风险收益”的评估机制。

本文的理论贡献在于，将数据安全风险评估从对象识别推进到活动链条分析；实践贡献在于，为组织开展分类分级、风险评估、合规审计和责任追溯提供一种可操作的工具框架。后续研究仍可结合行业案例和量化指标，对不同场景下的风险权重、控制有效性和残余风险阈值开展更细致的实证检验。

参考文献

[1] 数据安全法[M]//中华人民共和国法律汇编(2021).北京：人民出版社， 2022.
[2] 个人信息保护法[M]//中华人民共和国法律汇编(2021).北京：人民出版社， 2022.
[3] 网络数据安全管理条例[M]//中华人民共和国行政法规汇编(2024).北京：人民出版社， 2025.
[4] ISO/IEC. ISO/IEC 27005:2022 Information security, cybersecurity and privacy protection — Guidance on managing information security risks[S]. Geneva: ISO, 2022.
[5] Khatri V, Brown C V. Designing data governance [J]. Communications of the ACM, 2010, 53(1): 148-152.
[6] 王利明.论个人信息权益的法律保护[J].中国法学， 2021(4): 25-45.
[7] 周汉华.个人信息保护法的制度逻辑与时代展开[J].法学研究， 2021(5): 3-22.
[8] National Institute of Standards and Technology. NIST Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management， Version 1.0[R]. 2020.
[9] 国家市场监督管理总局，国家标准化管理

委员会. GB/T 37988—2019 信息安全技术 数据安全能力成熟度模型[S].北京：中国标准出版社， 2019.

[10] 国家市场监督管理总局，国家标准化管理委员会. GB/T 35273—2020 信息安全技术个人信息安全规范[S].北京：中国标准出版社， 2020.

[11] 国家市场监督管理总局，国家标准化管理委员会. GB/T 43697—2024 数据安全技术 数据分类分级规则[S].北京:中国标准出版社，2024.

[12] 国家市场监督管理总局，国家标准化管理委员会. GB/T 45577—2025 数据安全技术 数据安全风险评估方法[S].北京：中国标准出版社，2025.

[13] Solove D J. A taxonomy of privacy [J]. University of Pennsylvania Law Review, 2006， 154(3): 477-564.

[14] 中国信息通信研究院.数据安全治理实践指南[R].北京，2023.

[15] 高富平.数据要素市场化配置的法律机制[J].法学研究， 2023(1): 20-38.

[16] 张新宝.数据安全法的制度创新与规范适用[J].中国法学， 2023(3): 5-24.

[17] 龙卫球.数据分类分级管理的法律构造[J].法商研究， 2024(1): 3-17.

[18] 王锡锌.个人信息保护合规审计的制度逻辑[J].中外法学， 2024(2): 285-305.

[19] 梅夏英.数据安全风险评估的制度化路径[J].法律科学， 2024(3): 82-97.