

疾控数据分类分级实践探索^{*}

赵梓辰 傅 罡 张英杰

(中国疾病预防控制中心 北京 102206)

〔摘要〕 目的/意义 提高数据治理水平,保障数据安全,实现数据资源规范化管理。方法/过程 总结中国疾病预防控制中心数据分类分级工作实践经验,分析数据分类分级管理的目标和意义、实施过程、结果及其存在的问题。结果/结论 数据分类分级工作不仅提高了数据管理的规范化水平,还为保障公共卫生数据安全性、促进跨部门协作与信息共享提供了有力支持。

〔关键词〕 数据分类分级;公共卫生;数据管理;信息安全

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2025.04.005

Exploration on the Classification and Grading of Disease Prevention and Control Data

ZHAO Zichen, FU Gang, ZHANG Yingjie

Chinese Center for Disease Control and Prevention, Beijing 102206, China

〔Abstract〕 **Purpose/Significance** To improve data governance level, ensure data security, and achieve standardized management of data resources. **Method/Process** Practical experience of data classification and grading work in the Chinese Center for Disease Control and Prevention is summarized, goals and significance, implementation process, results and existing problems of data classification and grading management are analyzed. **Result/Conclusion** The implementation of data classification and grading not only elevates the standardization level of data management, but also provides strong support for ensuring the security of public health data, promoting cross departmental collaboration and information sharing.

〔Keywords〕 data classification and grading; public health; data management; information security

1 引言

在全球数字化转型浪潮推动下,大数据技术蓬勃发展,公共卫生领域数据量呈指数级增长。公共卫生数据是至关重要的公共数据资源,在疾病预防

控制、公共卫生决策制定以及社会健康保障体系中扮演着关键角色,也是提升公共卫生服务质量和效率的有力支撑。然而,随着数据规模的急剧膨胀,数据安全问题突显,数据治理成为关键。近年来,我国在卫生健康数据治理方面积极探索,其中数据分类分级工作已取得一定进展。如钱晨嗣等^[1]建设疾控数据资源管理共享平台,为数据统一管理、分类分级、开放共享提供了比较成熟、可复制的经验。张倩倩等^[2]针对临床数据中心提出医疗数据分类分级的方法和步骤,形成兼顾中医药大数据特点和安全管理要求的分类分级流程。路万里^[3]结合实践应用,为确定科学、合理、合法的健康医疗大数

〔修回日期〕 2025-02-17

〔作者简介〕 赵梓辰,研究实习员,发表论文2篇;通信作者:张英杰。

〔基金项目〕 中国疾病预防控制中心项目(项目编号:102393230020010000014)。

据分类分级规则、安全防护策略提供思路。陈烨等^[4]构建电子健康档案数据分类分级框架,明确原则与流程,并邀请专家评估其合理性和明确性。但是数据分类分级在具体实施、数据处理和共享等核心环节,仍然面临诸多严峻挑战^[5]。在实施方面,缺乏统一、科学的标准体系,导致不同机构对数据的分类分级存在差异,影响数据整合与利用;在数据处理环节,分析技术的精准性和效率不足,无法充分挖掘数据背后的潜在价值;在数据共享过程中,安全与隐私保护机制不完善,阻碍了数据在不同主体间的顺畅流通。

数据分类分级是一种行之有效的规范化管理手段,对公共卫生领域的数据治理具有不可估量的价值。通过科学合理的数据分类分级,能够更为精准地识别各类数据,制定并实施针对性的管理与保护策略,引导各机构充分认识数据的潜在价值,加速公共卫生行业的数字化转型进程^[6]。在公共卫生数据的流通与共享中,保障数据安全与合规,在确保数据可用的同时,实现价值的充分释放,为公共卫生事业的发展提供强大的数据驱动力^[7]。

2 数据分类分级的目标和意义

数据分类分级的主要目标是明确不同数据的管理责任和保护措施,确保数据在安全、合规和高效的环境中使用。具体而言,一是提升数据管理效率。通过系统化的数据分类和分级,明确不同类型数据的管理要求,从而提升数据使用的灵活性和效率。二是数据安全性与隐私保护。通过对敏感数据的分类和分级,采取相应的安全措施,保障个人隐私和数据安全。三是规范数据共享与应用。确保数据在安全合规的前提下发挥最大价值。

中国疾病预防控制中心(以下简称中国疾控中心)肩负着大规模健康数据的管理与应用重任,其开展数据分类分级工作具有重要的现实意义。一方面,这是信息安全管理的要求。通过构建完善的数据分类分级体系,能够有效提升数据的安全性与合规性,降低数据泄漏等安全风险。另一方面,也是优化数据使用、助力科学决策的关键举措。通

过精准的数据分类分级,能够提高数据处理与分析的效率和准确性,为公共卫生决策提供更为科学、可靠的数据支持。

3 实施过程

3.1 协同部署的实践策略

在数字化时代,数据分类分级工作因其系统性与复杂性,成为涉及全局的综合性业务。要从整体战略高度出发,对各部门在数据分类分级工作中的职责与分工进行界定。业务部门主要负责细致梳理业务流程,准确界定业务数据类型,并依据对业务的影响程度评估数据的重要性。这种基于业务实际的分析,为数据分类分级提供坚实的基础,确保数据的分类分级能够紧密贴合实际业务需求。信息和网络安全等技术部门负责搭建数据分类分级技术框架和提供技术支持。为确保数据分类分级工作有序推进,设立专门的工作小组——信息化建设工作专班,充分发挥统筹协调作用。

3.2 标准支撑

2024 年中国疾控中心依据《网络安全法》《数据安全法》,结合《卫生健康行业数据分类分级指南(试行)》、《信息安全技术 个人信息安全规范》(GB/T 35273—2020)、《信息安全技术 健康医疗数据安全指南》(GB/T 39725—2020)、《数据安全技术 数据分类分级规则》(GB/T 43697—2024)、《国家卫生信息资源分类与编码管理规范》(WS/T 787—2021)等法律法规和标准规范,制定《疾病预防控制数据分类分级实施细则(试行)》并组织实施,以此指导数据分类分级工作高效开展。

3.3 多维度数据资源梳理

为构建层次分明、逻辑清晰的数据资源框架,将数据来源划分为业务、科研、管理和外部 4 类。业务数据涵盖公共卫生、疾病监测、应急响应等日常运营和管理数据;科研数据涵盖各类科研项目中的数据资源,通常具有较高的学术价值;管理数据涵盖日常管理、内部调度和支持决策的数据;外部

数据涵盖来自其他机构或部门的数据资源，通过向外部机构采购或与外部信息系统对接获取。同时，数据资源的空间范畴包括现时数据和历史数据。现时数据反映组织当前的状态；历史数据记录不同时期的发展轨迹，通过对其深入分析，可以挖掘规律和趋势，为预测未来发展提供重要参考。此外，数据来源不仅包括信息系统中存储的数据，还包括调查、科研以及日常工作积累的数据文件。这些数据来源多样，为中国疾控中心的全面发展提供全方位的数据支持。

3.4 基于疾控领域特性的数据分类体系构建

基于业务梳理结果，采用“业务领域 - 关键业务 - 业务属性分类”的逻辑化、层次化分类方式，对数据进行分类。一级类目确定为基础资源、业务资源和主题资源，此为整个分类体系的顶层框架。二级类目进一步细化上级类目数据的范畴，使分类更具针对性。以基础资源为例，涵盖人口数据、地理数据、法律法规、人力资源、财务资源、设备资源和疾控信息化资源等。三级类目以人力资

源为例，包含人力资源规划、招聘与配置、培训与开发、绩效管理、薪酬福利管理、劳动关系管理和在职人员管理等。各级类目间逻辑关联紧密、层次分明。考虑到疾病预防控制业务的复杂性与动态性，四级、五级类目保留一定的灵活性，可依据具体业务属性进行合理扩展。这种设计既保证了分类体系的稳定性与规范性，又赋予其足够的弹性，以适应不断变化的业务需求，确保数据分类体系在实际应用中保持科学性与实用性。

3.5 科学严谨的数据分级

在数据分类基础上，依据时效性原则、从严原则和关联叠加原则，根据疾病预防控制数据重要程度以及泄漏后对国家安全、经济运行、社会秩序、公共利益、组织权益和个人利益造成的影响和危害程度，对数据资源进行分级。结合疾病预防控制机构属性和数据特点，按照数据重要程度从高到低分为核心数据、特别重要数据、比较重要数据、一般数据 4 个级别，见表 1。数据分级时应兼顾重要程度和危害程度等。

表 1 中国疾控中心数据级别确定规则

级别	影响对象和程度						与《数据安全法》
	国家安全	社会秩序	经济运行	公共利益	组织权益	个人权益	对应关系
4 级 核心数据	严重危害或特别严重危害	特别严重危害	特别严重危害	特别严重危害	—	—	核心数据
3 级 特别重要数据	一般危害	严重危害	严重危害	严重危害	—	—	重要数据
2 级 比较重要数据	—	一般危害	一般危害	一般危害	特别严重危害	特别严重危害	重要数据
1 级 一般数据	—	—	—	—	一般危害或严重危害	一般危害或严重危害	一般数据

3.6 以优化为目的的专家评审

对数据分类分级后，基于科学的管理理念，组织专家开展评审。精准识别数据分类分级过程中可能存在的遗漏、偏差或不合理之处，并提出针对性与建设性的改进意见，以推动数据分类分级工作不断优化与完善。

4 结果

4.1 数据资源的全面梳理与清晰分类

中国疾控中心对各来源数据进行全面梳理，完成初步分类，见表 2。再依据分类体系进一步对每项数据资源分类。

表 2 中国疾控中心数据资源分类统计（项）

数据来源	基础资源	业务资源	主题资源	合计
业务	2	217	9	228
科研	4	14	75	93
管理	58	0	0	58
外部	1	0	0	1
合计	65	231	84	380

4.2 科学分级和敏感数据的优先保护

中国疾控中心对数据进行科学严谨的分级，对于个人隐私和公共卫生敏感信息（如疾病监测数据、免疫接种数据、传染病疫情信息等）予以重点关注，将其纳入“3 级 特别重要数据”和“4 级 核心数据”范畴。其中特别重要及核心数据共 29 项，见表 3。

表 3 中国疾控中心数据资源分级统计（项）

数据类别	一般数据	比较重要数据	特别重要及核心数据	合计
基础资源	58	7	0	65
业务资源	76	128	27	231
主题资源	46	36	2	84
合计	180	171	29	380

4.3 数据共享与协作机制的规范化

数据分类分级工作的开展为内部以及与其他机构之间的数据共享和协作提供了保障。中国疾控中心印发《公共卫生大数据合作和共享审查委员会章程》，形成共享机制，在确保数据安全性和隐私性前提下，实现数据互通合作的进一步规范化。数据分类分级有助于更加精确地识别和挖掘各类数据的使用价值，不仅支持了中国疾控中心日常工作，还为科学决策提供了更加精准的依据。

4.4 全员数据安全意识得到增强

数据分类分级工作推动了中国疾控中心全员数据安全意识的提升。通过开展系列培训和实际工

作，向各直属单位和处室广泛普及了数据分类分级和保护相关知识。尤其在对核心数据和重要数据的处理方面，加深了工作人员对数据安全法律法规和操作规范的理解和掌握，与此同时也提高了全系统的数据管理能力。

5 讨论

5.1 数据管理者对数据分类分级工作体系性认知不足

在数据呈爆炸式增长的时代，数据治理工作已逐渐演变为一项既复杂又常态化的重要任务。数据分类分级，作为确保数据质量、提升数据价值的关键手段，其重要性不言而喻。然而，数据管理者对这一复杂任务普遍缺乏体系化认知。主要体现在多个方面，例如对数据治理目标的界定较模糊，未能充分理解数据治理是涵盖数据战略规划、数据标准制定、数据质量管理、数据安全保障等多维度的系统性工程；未能形成清晰的职责分工与协同机制，各环节之间缺乏有效衔接等。这种认知不足导致数据分类分级工作难以形成连贯、高效的推进路径，数据治理成果与预期之间常出现显著偏差，治理工作在反复调整与修正中进行，难以实现预期效果。这不仅造成了资源浪费，还极大地打击了技术人员的工作积极性与主动性^[8]。

5.2 应兼顾数据分类标准的统一性与灵活性

鉴于中国疾控中心各直属单位和处室业务的多样性和复杂性，在统一标准框架下实施数据分类分级面临诸多挑战。各直属单位和处室因其独特的业务职能，在数据生成、采集、存储和应用等方面存在显著差异。以科研数据与业务数据为例，不同单位的科研数据可能聚焦于特定疾病或特定应用场景，具有高度的专业性和独特性；而业务数据则围绕日常疾病防控、监测等常规工作，更注重实时性和全面性。这些差异使标准化的数据分类分级体系难以完全适配特殊需求。在追求标准化的过程中，应兼顾灵活性，充分满足不同业务场景下的数据管理需求。应进一步深入研究相关技术，结合实际工作中的复杂情况和需求，借鉴《数据管理能力成熟

度评估模型》^[9]等相关模型,进一步完善标准,增强其可操作性,确保数据分类分级工作的精准性。

5.3 应强化部门间沟通和协调机制

多部门协同参与是数据分类分级工作的显著特征。由于各部门在业务范畴、专业视角以及数据应用目的等方面存在差异,部门间沟通不畅极易阻碍工作顺利推进。例如,业务部门可能基于实际防控需求对数据分类,而技术部门则可能从数据存储和处理的便利性角度出发,如果两者缺乏有效沟通,就会造成数据分类分级的混乱。

为应对该挑战,一方面,应强化结果审核机制。通过建立科学、严谨的审核标准与流程,运用专业知识和数据分析工具,对数据分类分级的结果进行全面细致的审查。另一方面,应着重加强部门间的沟通与协作,定期组织跨部门交流会议,促进信息共享与传递。同时鼓励各部门及时反馈问题,共同探讨解决方案,形成高效的工作协同模式。此外,也可将数据分类分级工作流程进行规范化梳理,引入适宜的工具提高工作效率和质量,保障工作的有序推进^[10]。

5.4 确保数据安全是亟待解决的关键问题

随着数据量的增长,需要更多的技术支持和防护措施来确保数据在存储、传输和使用过程中的安全性。尤其是在大规模数据交换和共享的过程中,数据安全性容易受到威胁。尽管中国疾控中心已经部署加密技术和访问控制机制,但随着技术的进步和网络攻击方式的变化,如何保障数据安全仍是亟待解决的问题。

5.5 数据共享与跨部门合作依然存在挑战

在数据驱动决策的时代背景下,如何在严格保障安全的前提下,实现高效的数据共享与跨部门合作,已然成为公共卫生领域乃至整个数据管理范畴亟待攻克的关键问题。不同部门在数据生态系统中犹如相对独立的个体,各自遵循特定的数据标准、格式规范以及保密要求。数据格式的多样性同样加剧了跨部门合作的复杂性,增加了数据共享的技术

门槛。一些涉及患者隐私或国家安全的敏感数据在共享过程中要遵循严格的保密协议与安全措施,而过高的要求可能使数据共享受限,甚至无法共享。为此常采用匿名化处理,对数据中的个人标识信息进行删除或替换,以保护数据主体的隐私。然而,过度匿名化处理可能引发数据质量下降^[11]。要解决这些问题,需综合运用标准化制定、技术创新以及政策协调等多维度策略,以实现数据安全与数据共享的平衡。

6 结语

本研究构建了适用于中国疾控中心的数据分类分级体系,数据分类共包含 5 个层级,一级类目包括基础资源、业务资源和主题资源 3 类;数据等级则包括核心、特别重要、比较重要和一般数据 4 级;梳理 380 项数据资源,识别出其中 29 项为特别重要及核心数据,实现对敏感数据的重点保护。通过多维度数据梳理、贴合疾控特性的数据分类、严谨的数据分级,实现了公共卫生数据管理从无序到有序的关键转变,提高了数据管理的规范化水平,保障了数据的安全性,促进了跨部门协作与信息共享,在公共卫生数据治理领域具有开拓性意义。

目前数据分类分级工作仍面临标准化、安全性和技术更新等多方面挑战。未来中国疾控中心将通过持续完善技术平台、加强数据安全技术研究、深化跨部门合作,进一步挖掘公共卫生领域数据价值,助力精准防控、资源优化配置与科学决策,推动我国公共卫生事业迈向新高度。

作者贡献: 赵梓辰负责研究设计、论文撰写;傅罡负责提供指导、论文审核;张英杰负责论文审核与修订。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- 1 钱晨嗣,陈昕,夏天,等. 疾控数据管理共享平台的设计与实现 [J]. 中国卫生信息管理杂志, 2024, 21 (6): 889-894, 901.

(下转第 39 页)

研究预处理方法以及所获得经方医案数据的有效性, 可为下游智能诊疗模型训练提供可靠数据。

作者贡献: 文志华负责经方医案数据采集、处理与分析, 论文撰写; 张龙信负责经方医案校对工具构建; 刘青萍、肖莉、夏帅帅负责经方医案校对; 晏峻峰负责研究设计、提供指导。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

1 衣凯, 徐樱, 马云静, 等. 基于中医象思维的远程仿生复现脉诊仪的研发 [J]. 时珍国医国药, 2024, 35 (3): 747-750.

2 王玲, 林依凡, 李璐. 智能诊疗在舌象研究中的应用进展 [J]. 中华中医药杂志, 2021, 36 (1): 342-346.

3 姚帅君, 闫敬来, 杜彩凤, 等. 基于集成学习构建围绝经期综合征中医智能诊断模型 [J]. 中医杂志, 2023, 64 (6): 572-580.

4 DONG X, ZHENG Y, SHU Z, et al. TCMPR: TCM prescription recommendation based on subnetwork term mapping and deep learning [J]. Biomed research international, 2022 (2): 3776-3783.

5 JIN Y, JI W, SHI Y, et al. Meta-path guided graph attention network for explainable herb recommendation [J]. Health information science and systems, 2023, 11 (1): 5.

6 赵国桢, 郭诗琪, 庞华鑫, 等. 人工智能技术在辅助中医诊疗及诊疗标准化中的应用 [J]. 中医杂志, 2022,

63 (24): 2306-2310.

7 文志华, 夏帅帅, 刘东波, 等. 中医智能辨证诊断技术的演进与问题探讨 [J]. 世界科学技术-中医药现代化, 2021, 23 (11): 4298-4304.

8 孙宇衡, 王雨轩, 魏东升, 等. 基于数字孪生技术的中医智能辅助诊疗系统构建研究 [J]. 中华中医药学刊, 2024, 42 (9): 18-22.

9 陈玲玲, 刘诗发, 张晨. 中医医案类数据库的构建与数据处理研究 [J]. 中医文献杂志, 2014, 32 (6): 24-27.

10 朱彬, 朱垚, 陆明, 等. 基于 iOS 的多媒体医案采集移动平台的设计和实现 [J]. 计算机时代, 2017, 303 (9): 21-23.

11 章亚东, 胡孔法, 杨涛, 等. 名老中医医案数据采集与清理算法 DCCA 的研究 [J]. 计算机时代, 2018, 316 (10): 68-70.

12 韦昌法, 罗丽琴, 晏峻峰. 中医数字辨证配套医案智能采集与分析系统构建研究 [J]. 湖南中医药大学学报, 2020, 40 (1): 70-74.

13 马亚楠, 刘国华. 中医智能诊疗信息采集系统的设计与实现 [J]. 南开大学学报 (自然科学版), 2022, 55 (5): 67-71.

14 于琦, 王映辉, 李敬华, 等. 中医名医医案分析系统研究 [J]. 中国数字医学, 2015, 10 (3): 51-53.

15 于琦, 王映辉, 李宗友, 等. 基于语音识别技术的中医医案采集与应用研究 [J]. 中国数字医学, 2021, 16 (9): 80-83, 109.

(上接第 33 页)

2 张倩倩, 刘云, 彭远, 等. 医院中医药大数据分类分级研究与实践 [J]. 中国数字医学, 2025, 20 (2): 34-38.

3 路万里. 健康医疗大数据分类分级及安全策略 [J]. 网络安全技术与应用, 2024 (12): 65-67.

4 陈烨, 王阳, 徐亚兰, 等. 电子健康档案数据分类分级研究 [J]. 档案学研究, 2024 (3): 119-128.

5 许中缘, 何舒岑. 公共卫生领域大数据治理中个人信息的利用与保护 [J]. 中南大学学报 (社会科学版), 2022, 28 (3): 20-31.

6 毛丹, 夏天, 张诚, 等. 疾病预防控制数据治理主题设计与应用 [J]. 中国卫生信息管理杂志, 2022, 19 (2): 165-170, 177.

7 朱怡潇. 面向云端医疗数据流通的可验证计算研究 [D]. 西安: 西安电子科技大学, 2023.

8 李慧. 数据治理流程体系的建立与安全应用实践 [J]. 网络空间安全, 2024, 15 (4): 109-114.

9 林琳. 基于 DCMM 框架的医疗数据治理体系研究 [J]. 网络空间安全, 2024, 15 (4): 361-366.

10 祝守宇, 蔡春久. 数据治理: 工业企业数字化转型之道 [M]. 北京: 电子工业出版社, 2020.

11 唐明坤, 吴思竹, 周佳茵, 等. 医学数据共享隐私保护中基于聚类的匿名化算法关键技术研究 [J]. 医学信息学杂志, 2023, 44 (6): 65-71, 78.