

数据分类分级：研究趋势、政策标准与实践进展

严炜炜 谢顺欣 潘静 陆伟
(武汉大学信息管理学院, 武汉 430072)

摘要: 随着数据要素成为国家战略性资源, 数据分类分级主题研究呈现爆发式增长趋势, 相关政策和标准也陆续出台, 以推进数据安全保护工作。本文在对我国数据分类分级学术研究动态进行定量揭示的基础上, 提炼相关公开政策标准拟定趋势, 结合企业实践规律, 综合探究我国数据分类分级工作进展。研究发现, 科学研究层面, 相关研究聚焦于数据安全治理、数据治理、数据挖掘与应用等主题; 政策标准层面, 中央政策的提出经历了由点到面的过程, 地方标准多以政务数据和公共数据的分类分级为重点, 在突出地域特色的同时不断更新内容, 行业指南也已具备一定基础; 企业实践层面, 已形成以用户主体型、主题场景型和业务模块型为代表的分类分级实践模式。研究建议, 发挥图情学科优势推进数据分类分级理论和标准规范的跨学科研究合作, 提高地方标准的覆盖面和内容质量, 并推进行业指南完善和企业数据顶层规范建立。

关键词: 数据分类分级; 学术研究; 政策条例; 行业指南; 企业实践

中图分类号: G252 **DOI:** 10.3772/j.issn.1673-2286.2022.09.001

引文格式: 严炜炜, 谢顺欣, 潘静, 等. 数据分类分级: 研究趋势、政策标准与实践进展[J]. 数字图书馆论坛, 2022 (9) : 2-12.

数字经济与全球一体化背景下, 数据要素成为国家战略性资源^[1], 数字产业化和产业数字化进程的推进对数据安全保护和数字治理标准化提出更高的要求^[2]。2022年7月21日, 国家网信办公布对滴滴公司依法作出网络安全审查相关行政处罚^[3], 标志着国家进一步加大对数据安全的重视程度和保护力度。事实上, 近年来我国不断加强对数据安全的保护力度, 并先后颁布了《网络安全法》《数据安全法》《个人信息保护法》《关键信息基础设施安全保护条例》《网络安全审查办法》《数据出境安全评估办法》等法律法规, 并确立了“国家建立数据分类分级保护制度”^[4], 在“十四五”规划中亦进一步明确提出“要完善适用于大数据环境下的数据分类分级保护制度”^[5]。可见, 数据分类分级作为保障网络环境下数据安全的重要方法, 因符合不同类型和级别的数据属性对应着不同层级的安全防护需求这一客观要求, 得到了国家的高度重视。

近年来, 针对数据分类分级, 我国学者对相关主题的研究呈现爆发式增长。同时, 国家和地方政府出台了一系列数据分类分级相关的中央政策、地方标准和行业指南, 并指导企业实践。然而, 当前相关理论研究和实施制度尚处于摸索阶段, 并且存在范围局限和系统化程度不足等问题^[6]。系统梳理相关研究、政策与实践推进态势, 对于全面理解数据分类分级理论基础与科学规律, 掌握我国数据分类分级进展, 进一步完善数据分类分级行业指南, 切实促进业界安全有序落实数据分类分级标准具有重要意义。因此, 本文一方面利用信息计量、社会网络分析等方法对我国数据分类分级的学术研究趋势进行定量揭示; 另一方面, 利用政策文本分析方法, 归纳数据分类分级政策、标准等文件趋势, 并结合企业实践规律的提炼, 综合揭示我国数据分类分级现状, 为进一步深化该领域的研究和实践探索提供参考和指引。

1 数据分类分级科学研究趋势

为揭示我国数据分类分级相关研究趋势, 本文以中国知网 (CNKI) 作为数据来源, 以“数据分类分级”作为主题词进行检索, 将统计时间设为1990年1月—2022年7月, 在剔除重复或者不符合主题的文献后, 共获取313篇研究文献。

统计相关研究文献的时间分布可知, 1990—2019年发文量较少, 自2020年起发文量呈现爆发式增长, 2021年的发文量 (108篇) 约是2020年 (50篇) 的2倍多, 而2022年1—7月的发文数量 (67篇) 已与2020年全年发文量持平。这说明, 随着数据要素作为数字经济发展的核心受到重视^[7], 数据安全成为重要议题, 学界对数据分类分级的关注度在近3年显著提高, 并将持续增长。其中, 可以预见该热点主题后续发文量将会持续增长。

为深入了解和析我国在数据分类分级研究领域的学术热点, 本文进一步对研究文献进行关键词提取和聚类分析。首先提取313篇文献的所有关键词, 进行数据清洗, 将诸如“分类分级”和“分级分类”等语义相近关键词合并处理和清洗后, 共得到关键词208个。继而以关键词为节点运行CiteSpace, 并在可视化网络中, 选择“K聚类”和“LLR (Log-Likelihood Ratio) 算法”, 以关键词作为聚类标签来源, 可得到如图1所示的关键词聚类网络。其中, 节点代表关键词, 节点连线代表关键词共现情况, 结果呈现分类分级、数据安全、工信部、数据挖掘、数据共享、交易机制、数据分类、大数

据、电子政务、信用监管和数据管理等11个聚类主题, 反映出我国数据分类分级相关研究热点。

进一步通过对各聚类中的标签词及发文机构进行统计分析, 可发现研究主题分布于图书情报、法学、计算机、公共管理、遥感测绘和网络安全等多学科领域, 热门研究主题总结归纳为3个方面: ①数据安全治理, 包括#0、#1、#6, 关注数据安全和数据分类、数据迁移和分级存储等; ②数据治理, 包括#2、#4、#8、#9, 关注安全治理、数据共享、政务数据和社会信用体系等; ③数据挖掘与应用, 包括#3、#5、#7, 关注大数据和数据分类分级相关的方法与应用。

此外, 在关键词聚类的基础上, 本文进一步根据关键词时序排列进行研究热点演化趋势分析。结果显示, 1990—2000年, 数据分类分级通常与制图数据和层次模型同时出现, 主要服务于绘图数据。例如, 早在1992年, 鞠旭照^[8]采用层次模型和多边形邻接表结构, 对青岛一崂山旅游图街道数据的专题内容进行分类分级。华一新^[9]在总结地图要素分类分级知识的基础上, 提出地图要素表示等级的概念, 同时分析影响地图要素表示等级的因素。相关研究中的层次模型和地图要素分类分级概念, 对现阶段构建数据分类分级概念和模式具有一定的启示意义。在2002年前后, 随着计算机技术的进步, 数据挖掘成为因特网时代的研究热点^[10], 数据分类随之受到关注。2006年和2008年, 数据分类和数据分级分别成为热点。2014—2018年, 随着大数据的发展^[11], 学界开始关注数据安全的相关议题^[12]。2019年之后, 人工智能和云平台的研究热潮引发了学界对数据开放共享和国际监管的关注^[13-14], 数据分类分级的相关研究呈现爆发式增长。

综合来看, 在21世纪之前, 数据分类分级的相关研究与地图要素分类分级相关; 而随着大数据的发展, 数字经济和数据要素化的时代背景对数据安全提出了更高要求, 故近年来的数据分类分级研究侧重于从数据安全、数据治理、数据挖掘与应用等视角对数据分类或数据分级的方法、原则和路径等进行探究, 相关研究涉及多个学科, 正处于蓬勃发展阶段。

2 数据分类分级政策条例分析

数据分类分级作为保障网络安全的重要防护手段, 是国家安全战略的重要组成部分, 也是构建安全的数字要素市场不可或缺的保护制度。近年来, 我国政府和相



图1 数据分类分级研究文献关键词聚类结果

关机构已经开展许多探索：党中央和国务院出台的多项重要意见和法律法规，主要从大数据环境和国家数据安全等宏观视角对数据分类分级保护制度提出了要求；而地方出台的标准，多以政务数据和公共数据的分类分级为重点，聚焦数据全生命周期的分级安全防护。

2.1 中央政策

中央高度重视数据分类分级保护制度建设，为了在大数据环境下构建数据安全的保障制度，研究制定了一系列宏观政策文件，旨在关注数据配置管理体系，建立健全数据要素市场规则并营造良好的数字生态。

在国务院官网检索获取数据分类分级相关中央政策，归纳如图2所示。早在2018年3月，国务院办公厅就发布了《科学数据管理办法》，明确了科学数据分类分级的相关规定，例如法人单位要明确科学数据的密级

和保密期限、开放条件、开放对象和审核程序等^[15]，对科学数据管理有直接的帮助^[16]，但相关规定仅针对科学数据，未充分考虑其他类型的数据。在数据要素化的市场背景推动下，2020年4月，国务院出台《关于构建更加完善的要素市场化配置体制机制的意见》，特别指出要加强对政务数据、企业商业秘密和个人数据的保护^[17]，重点关注在市场中流通的关键数据的分类分级。2021年3月，《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》更是明确提出“要完善适用于大数据环境下的数据分类分级保护制度”^[5]，进一步拓宽了分类分级保护的数据范围。可见，从明确科学数据密级到提出数据分类分级等保护制度，相关政策的提出是由点到面逐渐宏观的，这与大数据和数据要素化等时代发展的背景密不可分，其目的始终在于保护核心数据安全，构建安全的数字生态。

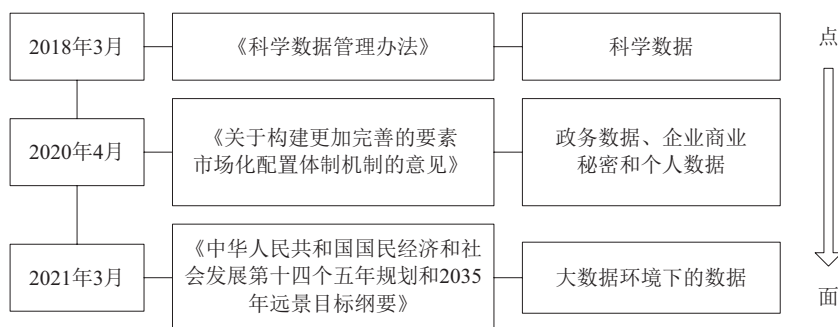


图2 数据分类分级主要中央政策

2.2 地方标准

进一步在全国标准信息公共服务平台检索数据分类分级相关地方标准文件，主要地方标准整理如表1所示。其中，贵州省大数据产业发展领导小组办公室于2016年6月发布的《政府数据 数据分类分级指南》是贵州省政府数据进行数据分类和分级顶层标准，也是全国范围内出台的最早一批数据分类分级标准指南。各地地方标准的分类分级框架结构整体较为接近，包括分类原则、分类方法或维度和分类特殊要求等版块，但在具体内容上存在一定差异。此外，地方出台的标准指南更多聚焦于政务数据和公共数据的分类分级，可以将其视为以政府数据为切入点所进行的初步尝试。

2.2.1 分类分级原则

各地方标准中的数据分类分级原则如图3所示。贵州省于2016年6月发布的地方标准中，“科学性，扩展性”分类原则和“自主定级，明确需求”分级原则在后续出台的各地方标准中得以保留和延续，有着重要的参考价值。不同的是，上海市地方标准中将分类和分级的原则合二为一，提出了“兼容性，安全性，科学性，需求导向，可操作性，可扩展性”的原则。福建省的地方标准一方面借鉴了上海市的分类分级原则，删去了“可操作性”并将其作为分类原则；另一方面借鉴保留了贵州省的分级原则。重庆市在数据分类原则中，专门提出了数据分类与数据分级之间的关系，认为“数据分类是数据分级的基础，数据分类与分级密不可分”，提炼补充“关联性”这一分类原则，并进一步增加了“分级管

控”等数据分级原则，进一步明确了数据分类和数据分级的关系。贵州省于2022年3月出台的新版地方标准在旧版基础上精简了分类分级原则的相关内容，并将原则和要求合并纳入“扩展分类”章节中。北京市地方标准中未对分类原则作详细介绍，其分级原则虽然与重庆市一致，但对原则的解释说明更加细致，如在解释“综合判定”这一原则时提出要“场景导向，兼顾内容”，强调“数据分级时应结合数据的应用场景、组合、取值、数据量的大小等”，力求数据分级准确合理的同时也深化了分级原则的内涵。

从时间维度看，地方标准中的分类原则经历了一个“简—繁—简”的过程。这可能是由于数据分类分级作为一种新的数据保护方法，在初步明确相关原则后，利用实践逐渐查漏补缺，有所补充；但过于烦琐的原则又限制了可操作性，故而近年来出台的地方标准中的分类原则有删繁就简的趋势。此外，各地方标准的分级原则在整体上基本一致，主要在原则各细目的级别权重上稍有不同。

2.2.2 分类分级方法

各地方标准中的分类分级方法归纳如图4所示。概括而言，各地方标准以多维度法和线分类法为主，个别省市也会采用面分类法进行标签构造。例如，除福建省是直接采用二级分类法之外，贵州省、浙江省、重庆市和上海市均在确定分类维度的基础上，再在各维度下

表1 数据分类分级主要地方标准

实施时间	发文机关	标准文件	框架结构
2016年6月	贵州省大数据产业发展领导小组办公室	《政府数据分类分级指南》 ^[18]	分类原则
			分级原则
			分类方法
			分级方法
2019年11月	上海市经济和信息化委员会	《上海市公共数据开放分级分类指南（试行）》 ^[19]	分级分类原则
			分级分类规则
			开放类别
			分级分类流程
2020年6月	福建省大数据管理局	《福建省公共数据资源开放分级分类指南（试行）》 ^[20]	分级原则
			分类分级方法
			分类原则
			分类编码说明
2021年7月	浙江省市场监督管理局	《数字化改革公共数据分类分级指南》 ^[21]	分类维度
			级别变更要求
			分级维度
2021年10月	重庆市大数据应用发展管理局	《重庆市公共数据分类分级指南（试行）》 ^[22]	分类分级原则
			分级方法
			分类方法
			分级管控要求
2022年3月	贵州省市场监督管理局	《政务数据 数据分类》 ^[23]	分类方法
			扩展分类
			分类与类别
2022年4月	北京市市场监督管理局	《政务数据 分级与安全保护规范》 ^[24]	分级原则
			分级流程
			分级方法
			分级安全保护要求

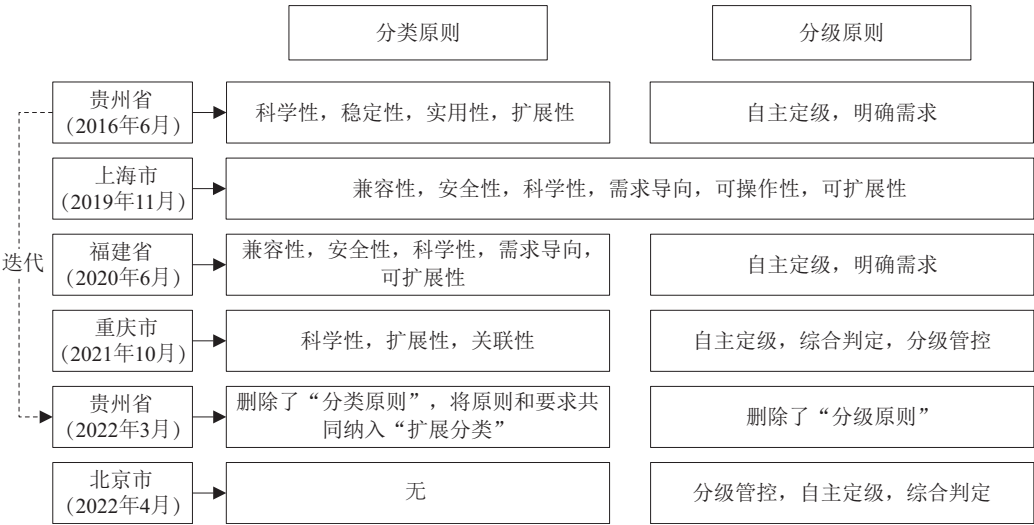


图3 地方标准分类分级原则

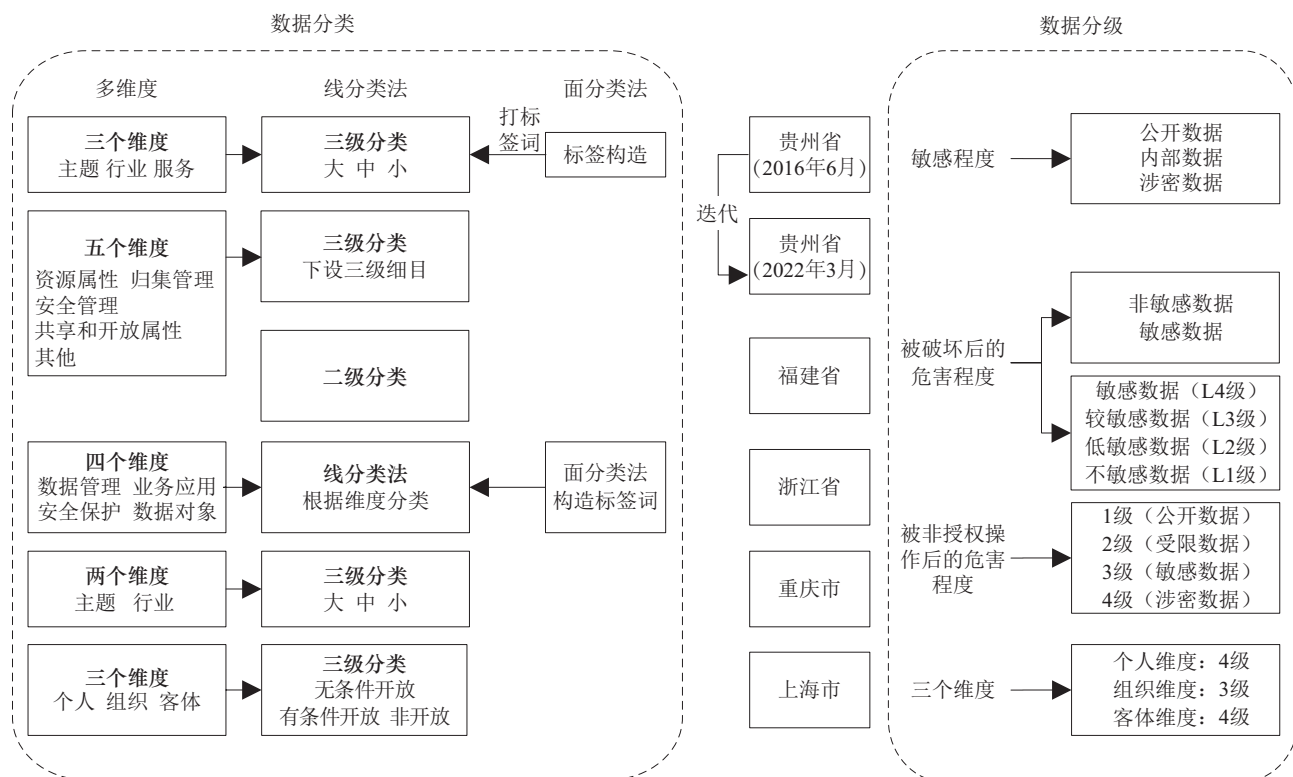


图4 主要地方标准中的数据分类分级方法

分别采用线分类法划分为包括大类、中类和小类的三级分类。在数据分级方法上，除上海市地方标准采用与数据分类相同的三个维度的视角以外，其他省市均以数据敏感程度或被违规操作后的危害程度为划分依据，按照数据敏感程度分为2级、3级或4级。特别地，2022年更新版的贵州省地方标准相较旧版，将数据分类的三个维度扩充成五个维度，并进一步细化了三级分类下的细目，删减了使用面分类法构造标签词和数据分级方法的相关内容。可见，考虑到政务数据和公共数据自身的类别和属性，各地方的数据分类方法大多从粗粒度的维度到细粒度的类别逐层划分，且大多依据数据敏感程度进行分级，存在一定共性。

2.2.3 分类分级要求

各省市在分类分级的一般要求和特殊要求上存在较大差异。在一般要求上，如浙江省和北京市的地方标准中，虽均对数据的采集、存储、加工、共享和使用等各个环节进行了详细要求，但划分数据生命周期的方式和各阶段的具体要求不尽相同，北京市还针对数据安全保护的通用要求、技术要求和管管理要求等进行了具体

规定。在特殊要求上，福建省拓展了“分类编码说明”章节，专门对公共数据资源分类进行了具体规定；浙江省考虑到因为业务需要变更公共数据级别的情形；重庆市根据数据生命周期，详细规定了各个环节的分级管控措施。此外，江西省出台的《生态文明数据分类及编码规范》、内蒙古自治区出台的《草原大数据：数据分类与编码》以及山东省出台的《农业大数据分类与编码规范》等标准，还针对具有地方特色的数据进行了专门规范，对于特殊数据的标准化管理具有重要意义。

3 数据分类分级行业实践分析

在政策条例的引导下，国内各行业已逐步开展关于数据分类分级指南的探索，其中，跨行业的指南注重从整体上进行规范，适用场景并不聚焦特定行业，涉及信息安全事件、大数据和工业数据等领域。而以通信、金融和医疗行业等为代表的公开行业指南则关注行业内部及用户数据，保护框架包括分类分级范围、原则和方法等方面。此外，企业也相继开始制订内部数据安全标准，开展数据分类分级实践。

3.1 行业指南分析

目前公开的行业指南可归纳如表2所示，其中涉及跨行业的数据分类分级指南均为国家标准，从对具体

的信息安全事件的分类分级，到大数据管理和工业数据分类分级，数据范围逐步扩大，数据分类和数据分级的方法也更细致具体。

表2 已公开的数据分类分级行业指南

适用行业	实施时间	发文机关	行业指南	框架结构
跨行业	2007年6月	国家质量监督检验检疫总局； 中国国家标准化管理委员会	《信息安全技术 信息安全事件 分类分级指南》 ^[25]	信息安全事件分类
				信息安全事件分级
	2019年8月	国家市场监督管理总局； 中国国家标准化管理委员会	《信息安全技术 大数据安全管理指南》 ^[26]	数据分级方法
				数据分类方法
				数据分级分类保护要求
	2020年2月	中华人民共和国工业和信息化部	《工业数据分级分类指南（试行）》 ^[27]	适用对象
				数据分级分类方法
				数据分级分类原则
				数据分级管理方法
通信	2014年12月	中华人民共和国工业和信息化部	《电信和互联网服务 用户个人信息保护 定义及分类》 ^[28]	用户个人信息保护范围
				用户个人信息内容和分类
				用户个人信息保护分级概述
				用户个人信息保护分级方法
	2020年12月	中华人民共和国工业和信息化部	《基础电信企业数据分级分类方法》 ^[29]	数据分级分类原则
				数据分级分类流程
金融	2018年9月	中国证券监督管理委员会	《证券期货业数据分类分级指引》 ^[30]	数据分级分类前提条件
				数据分级分类方法及原则
				数据分级分类中问题处理
	2020年2月	中国人民银行	《金融数据安全 数据安全分级指南》 ^[31]	数据安全定级
				重要数据识别
				数据安全定级
医疗	2021年7月	国家市场监督管理总局； 中国国家标准化管理委员会	《信息安全技术 健康医疗 数据安全指南》 ^[32]	数据分类范围
				数据分级方法
				数据安全分级原则

在跨行业的指南中，围绕信息安全事件，分类方法综合考虑信息安全事件的起因、表现、结果等，将信息安全事件分为有害程序事件、网络攻击事件、信息破坏事件、信息内容安全事件、设备设施故障、灾害性事件和其他信息安全事件7个基本分类，事件分级则根据信息系统的重要程度、系统损失和社会影响分为四级。围绕大数据安全管理，分类分级标准则更具有概括性，规定数据分类要按照数据主体、主题、业务等不同的属性进行；数据分级则需要组织的业务部门领导、业务专家、安全专家等共同确定，对各个行业数据分类分级标准规范的制定具有一定的参考价值，为其分类分级的原则和方法都提供了基础的指导思路。围绕工业数据，数

据分类需要工业企业结合生产制造模式、平台企业结合服务运营模式对工业数据进行分类梳理和标识，形成企业工业数据分类清单；数据分级则着重考虑工业数据遭篡改、破坏、泄露或非法利用后，可能对工业生产、经济效益等带来的潜在影响。概括而言，现阶段跨行业指南已围绕数据分类和数据分级的方法和原则进行了一定的规范。此外，从已公开的具体行业数据分类分级指南来看，包括部分国家标准和行业标准，涉及通信、金融和医疗等行业。

（1）通信行业。通信行业在2014年发布了《电信和互联网服务 用户个人信息保护 定义及分类》，规定了电信和互联网服务用户个人信息保护的术语和定义、

保护范围、信息内容和分类,初步提出了用户个人信息分类的方法。在2020年发布的《基础电信企业数据分类分级方法》中,进一步提出了针对电信企业的数据分类分级标准,数据分类根据基础电信业务运营和企业自身管理特点,按照树形结构,建立数据资源分类目录;数据分级则根据基础电信企业数据重要程度和敏感程度,确定数据资源的安全等级。概括而言,通信行业以个人信息保护为抓手,已经基本明确了个人信息和企业数据的分类分级方法,遵循“数据汇总-数据归类-数据定级”的步骤。

(2) 金融行业。金融行业在2018年发布的《证券期货业数据分类分级指引》中就已经对证券行业的数据分类分级制定出较为详细的标准规范,证券领域的数据采用业务条线标准进行数据分类,即先业务细分、再数据细分,继而形成从总到分的树形逻辑体系结构;数据分级则在数据分类的基础上将数据由高至低分为极高、高、中、低四级。在2020年发布的《金融数据安全 数据安全分级指南》将范围扩大至金融业,制定了详细的数据分级标准,但缺少数据分类维度的划分视角。在金融电子数据的分级依据方面,该指南将数据遭到破坏可能带来的影响作为重要标准,考虑对象和影响程度两个方面,将数据安全级别由高至低分为五级。此外,该指南还明确了金融数据安全分级的目标、原则、范围,以及数据安全定级的要素、规则和定级过程。整体上,金融行业数据分类分级指南主要针对证

券数据和金融电子数据,遵循“细分业务-数据归类-级别判定”的流程,已经基本形成了行业数据的分级标准,但相应的分类标准仍有待补充。

(3) 医疗行业。医疗行业在2020发布的《信息安全技术 健康医疗数据安全指南》以信息安全为出发点,对健康医疗数据进行分类分级,促进信息开发利用。在数据分类上,健康医疗数据可被分为个人属性数据、健康状况数据、医疗应用数据、医疗支付数据、卫生资源数据和公共卫生数据6类;在数据分级上,根据数据重要程度、风险级别以及对个人健康医疗数据主题可能造成的损害和影响,健康医疗数据则可由高至低被分为五级。概括而言,健康医疗领域作为医疗行业数据分类分级的初步探索,已形成初步框架,但数据分类分级的覆盖范围、具体措施和方法案例等内容还有待进一步扩充。

3.2 企业实践模式分析

随着国家、地方政策不断完善对数据分类分级的规定,以及行业指南对数据分类分级的要求深化,企业开始相继制定内部数据安全标准,开展数据分类分级实践。然而由于企业实践需要兼顾自身权益、数据使用和流通效率、符合相关的行业规范和保障用户的数据安全等多方面,导致数据分类分级需求及侧重点有所不同,继而产生了用户主体型、主题场景型和业务模块型3种主要企业实践模式,如表3所示。

表3 企业数据分类分级实践模式

实践模式	重 点	需 求
用户主体型	根据主体在数据生命周期产生的数据类型进行数据分类	提高企业数据管理规范性
	根据主体数据的敏感程度进行数据分级	
主题场景型	根据数据信息分区主题进行数据分类	提高企业数据使用的效率
	根据数据泄露在相应领域造成的负面影响进行数据分级	
业务模块型	制定数据分类分级模板并划分业务域	提高企业不同业务模块之间数据互相流动的能力
	结合实际业务和行业标准灵活细化数据分类分级方法	

3.2.1 用户主体型

用户主体型的数据分类分级模式围绕产生数据的主体,即用户展开,旨在提高数据管理的规范性,该实践模式归纳如图5所示。以中国移动为例,其在2016年发布了《中国移动大数据安全管控分级分类实施指南》和《中国移动IDC维护管理规定 数据安全管分

册(2016版)》等文件关注企业内部数据活动及数据分类分级,提出数据分类分级原则、方法、流程及示例,进行数据安全管控^[33]。具体而言,中国移动针对IDC系统中的数据,以用户为核心,将企业内部数据分为用户身份相关数据(A类)、用户服务相关数据(B类)、用户服务衍生数据(C类)和企业运营管理数据共四大类(D类),在与电信行业要求的分类相对应的同时,结

合企业自身管理需要, 将“其他信息”具体规定为“企业运营管理数据”。而数据分级则根据数据敏感程度不同, 以用户个人信息保护级别即敏感度进行数据级别划分, 主要强调用户数据这一主体自身的信息保护级别, 且遵循就高不就低原则, 将数据由高至低分为四

级。同时, 在数据管控中将依据客户所处行业、业务类型、业务范围等标准将客户分为四级, 不同子类数据等级依据客户等级不同进行细分, 相应采取不同的防护等级, 形成一一对应的关系。

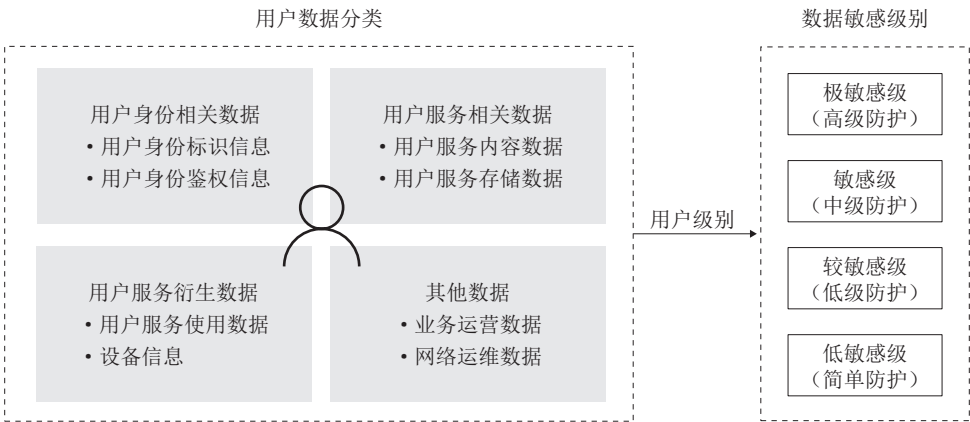


图5 用户主体型企业实践模式

可以看出, 用户主体型企业实践模式适用于以用户为核心的服务型企业, 如电信运营商和金融服务平台等, 其特点是, 数据类型集中在单一行业, 但在进行数据分类分级时需围绕不同层级的用户展开。该模式的优势是, 能在一定程度上提高企业数据分类分级的规范性, 且能够为不同等级的用户提供不同等级的数据保护方案; 但仍存在关注点集中于用户, 对企业其他数据的管控稍显不足等不足。

3.2.2 主题场景型

主题场景型的数据分类分级, 重点在于根据数据自身的来源属性和功能进行分类分级, 旨在提高数据使用的效率, 企业实践模式归纳如图6所示。例如, 云上贵州积极与政府合作, 搭建贵州省政府数据开放平台,

以主题、场景、部门三大标准进行数据分类, 其中, 主题分为城建住房、教育文化、工业农业、医疗卫生等21个不同功能的数据信息分区, 场景分为林业、交通、司法、教育等9个不同职能的分区, 为了明确分工, 部门则先分为省级部门和8个具体城市, 再具体划分到每个城市的不同政府部门^[34], 贴合平台的政务属性。在数据分级方面, 该平台将数据分为无条件开放数据和有条件开放数据, 无条件开放数据直接申请审批后即可下载或引用, 有条件开放数据则需要提交材料并通过2次审核后才可获取。此外, 平台还根据数据的敏感程度划分了3个敏感等级, 分别是公开数据、内部数据和涉密数据, 与开放属性存在一定的关联。

可见, 主题场景型企业实践模式适用于横跨多个行业的情形, 尤其是政务服务平台, 其特点在于: 数据横跨多个行业并且存在并列关系, 需要满足与企业外

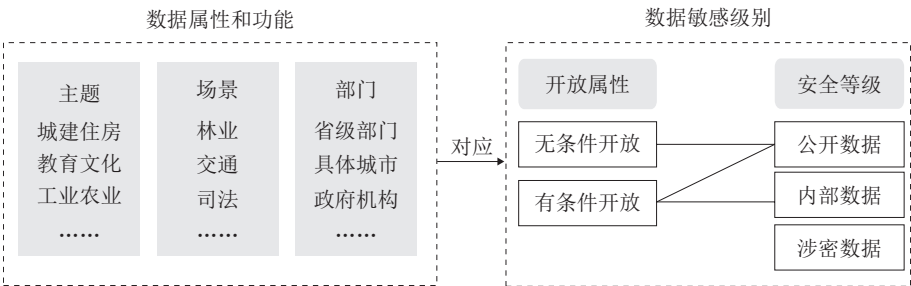


图6 主题场景型企业实践模式

部尤其是大众用户的数据互联互通需求。该模式的优势在于能够结合平台特点,分门别类建立具有主题特色的数据分类框架,充分发挥数据自身的属性;其劣势在于划分类别时的难度较高,可操作性不强,缺乏可供参考的划分方案。

3.2.3 业务模块型

业务模块型的数据分类分级,对不同业务模块的数据分类分级的方法进行了相应的细化,旨在提高企业不同业务模块之间数据互相流动的能力,相应的企业实践模式归纳如图7所示。以阿里巴巴为例,其数据安全分类分级模板支持自定义分类和分级,其数据安全分类分级

模板将数据分类为通用敏感信息、密钥敏感信息、敏感图片信息、位置敏感信息、个人敏感信息、设备敏感信息和企业敏感信息共7类52种敏感数据类型,并设三级敏感标准^[35]。在该模板的指导下,企业内部数据分类标准从业务板块出发,划分为不同的业务域,根据不同的业务流程和业务过程的不同阶段进行数据分类,其数据分类分级流程更具灵活性。同时,在标准分类分级规范的指引下,数据自动打标,同时给予人工调整的权限,更适合公司的各业务流程。此外,阿里巴巴不同业务板块数据分类分级标准也有所区分,例如金融板块分类分级标准会依据《金融数据安全 数据安全分级指南》行业标准,能源板块则会依据能源行业分类分级标准^[36],更有利于企业内部不同业务数据的运用与管理。

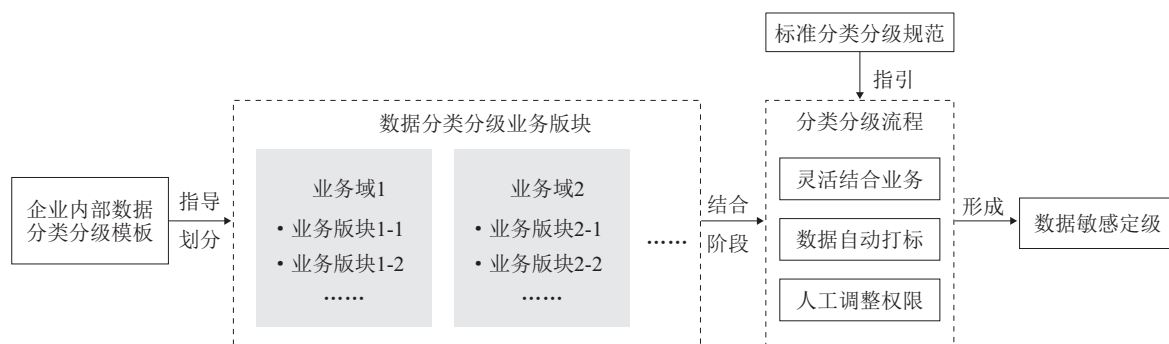


图7 业务模块型企业实践模式

业务模块型的企业实践模式适用于横跨多个业务板块的企业,其特点在于,重视数据的业务属性,以满足企业内部不同业务域之间的数据互联互通需求。该模式的优势在于:能够在模板和标准的指导下,以不同业务板块和业务流程为框架进行数据分类分级管理,操作灵活的同时保障了数据的规范流通性。然而,在以业务为导向的企业实践中,缺乏体系化的顶层规范,不同业务模块之间的管理差异可能为企业内外数据的流通造成一定困难。

4 结论与建议

日益复杂的大数据环境对保障数据安全提出了更高要求,数据分类分级也在此背景下成为重要的研究内容。本文利用信息计量、社会网络分析等方法对我国数据分类分级的学术研究动态进行定量揭示,并利用政策文本分析方法,归纳公开数据分类分级政策趋势,结合企业实践提炼应用模式,进而实现对我国数据分类分级

科学研究、政策条例、行业指南和企业实践等多维度现状、规律的综合揭示,为进一步深化该领域的研究和实践探索提供参考和指引。本文主要结论可概括如下。

(1) 科学研究层面,数据分类分级的概念演化与地图要素分类分级相关联,相关研究在近3年的关注度显著提高并呈现爆发式增长趋势,研究热点聚焦于数据安全治理、数据治理、数据挖掘与应用等主题,涉及图书情报、计算机、法学和公共管理等多个学科。

(2) 政策标准层面,中央政策的提出经历了由点到面的过程,重视大数据环境下的数据分类分级保护制度;地方标准多以政务数据和公共数据的分类分级为重点,聚焦数据全生命周期的安全防护,在突出地域特色的同时不断更新内容;同时,数据分类分级行业指南与标准制定也已具备一定跨行业基础,并在部分行业形成公开的指导框架。

(3) 企业实践层面,目前企业数据分类分级工作已展开了围绕产生数据的用户主体型,依据数据自身属性和功能的主题场景型,以及围绕企业不同版块业务

灵活展开的业务模块型等典型模式的实践探索。

在上述结论基础上, 本文针对数据分类分级研究与实践工作提出如下建议。

(1) 充分发挥图情学科优势, 拓展理论研究和跨学科合作。数据分类分级作为交叉学科的议题, 其概念演化、研究热点和趋势均具有明显的跨学科属性, 但当前的研究仅将其作为保障数据安全的方法或制度进行探讨, 缺少跨学科的视角。未来可将图情学科分类理论方法与数据分类分级的原则、路径和方法相结合, 并推进数据分类分级理论和标准规范的跨学科深度合作研究。

(2) 以点带面共同发文, 提升地方标准的覆盖度和内容质量。从已公开政策标准来看, 目前部分地区已经具备良好的数据分类分级制度基础, 但仍存在地方标准质量参差不齐的现实问题。建议未来在保持区域数据特色的基础上, 合理调配地区资源, 通过以点带面、共同发文等方式进一步提高地方标准的规范性和可操作性, 为数据安全跨省流动奠定良好的制度基础。

(3) 以推进传统行业数字化转型为契机, 完善行业数据分类分级指南。现阶段已形成数据分类分级框架的行业集中在数字化程度较好的互联网相关行业, 建议以传统企业的数字化服务转型为切入点, 拓展数据分类分级应用场景, 完善相关行业指南, 共促数据分类分级行业生态构筑。

(4) 建立企业数据分类分级顶层规范, 指引企业选择合适的实践模式。当前虽形成了部分企业实践模式, 但企业内部数据分类分级规章的透明度和时效性均有待提升。未来可在参照行业标准基础上, 从服务对象、业务规模、数据复杂程度和数据时效性等视角建立顶层规范, 根据企业自身业务范畴和数据特征选择合适的实践模式, 并进一步挖掘新模式, 将“完善适用于大数据环境下的数据分类分级保护制度”的指导思想落到实处。

参考文献

- [1] 夏义堃. 数据管理视角下的数据经济问题研究[J]. 中国图书馆学报, 2021, 47(6): 105-119.
- [2] 安小米, 许济沧, 王丽丽, 等. 国际标准中的数据治理: 概念、视角及其标准化协同路径[J]. 中国图书馆学报, 2021, 47(5): 59-79.
- [3] 国家互联网信息办公室对滴滴全球股份有限公司依法作出网络安全审查相关行政处罚的决定[EB/OL]. [2022-07-23]. http://www.cac.gov.cn/2022-07/21/c_1660021534306352.htm.
- [4] 中华人民共和国数据安全法[EB/OL]. [2022-07-30]. <http://www.npc.gov.cn/npc/c30834/202106/7c9af12f51334a73b56d7938f99a788a.shtml>.
- [5] 中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要[EB/OL]. [2022-07-20]. http://www.gov.cn/xinwen/2021-03/13/content_5592681.htm.
- [6] 卜学民. 论数据本地化模式的反思与制度构建[J]. 情报理论与实践, 2021, 44(12): 79, 80-87.
- [7] 陆瑶, 卢超, 董克, 等. 从幕后到台前: 数据要素化带来图情学科发展机遇与挑战[J]. 图书情报知识, 2021, 38(6): 123-133.
- [8] 鞠旭照. 机助旅游地图制图的一些结构特点——以青岛市交通游览图机助制图为例[J]. 地图, 1992(2): 27-33.
- [9] 华一新. 用专家系统技术确定地图要素的表示等级[J]. 测绘通报, 1993(3): 22-25.
- [10] 李勇, 桑艳艳. 网络文本数据分类技术与实现算法[J]. 情报学报, 2002(1): 21-26.
- [11] 王春华, 李维, 文庭孝. 我国图书情报领域大数据研究热点分析[J]. 图书情报知识, 2015(4): 82-89.
- [12] 巴志超, 李纲, 安璐, 等. 国家安全大数据综合信息集成: 应用架构与实现路径[J]. 中国软科学, 2018(7): 9-20.
- [13] 盛小平, 郭道胜. 科学数据开放共享中的数据安全治理研究[J]. 图书情报工作, 2020, 64(22): 25-36.
- [14] 毛子骏, 朱钰谦, 徐晓林. 中国省域政务数据安全政策文本量化研究[J]. 情报杂志, 2021, 40(12): 50, 72-79.
- [15] 国务院办公厅关于印发科学数据管理办法的通知[EB/OL]. [2022-07-20]. https://most.gov.cn/xxgk/xinxifenlei/fdzdgnr/fgzcgfxwj/gfxwj2018/201804/t20180404_139023.html.
- [16] 闫鑫, 黄国彬. 科学数据分类研究述评[J]. 图书馆论坛, 2020, 40(5): 45-54.
- [17] 关于构建更加完善的要素市场化配置体制机制的意见[EB/OL]. [2022-07-20]. http://www.gov.cn/zhengce/2020-04/09/content_5500622.htm.
- [18] 政府数据 数据分类分级指南: DB52/T 1123—2016[S/OL]. [2022-07-20]. <https://std.samr.gov.cn/db/search/stdDBDetaild?id=91D99E4D58EF2E24E05397BE0A0A3A10>.
- [19] 关于落实《上海市公共数据开放分级分类指南(试行)》相关要求的通知[EB/OL]. [2022-07-20]. <https://www.shqp.gov.cn/shqp/zwgk/zxgk/20210315/822898.html>.
- [20] 福建省公共数据资源开放分级分类指南(试行)[EB/OL]. [2022-07-20]. <https://data.fujian.gov.cn/oportal/news/f7fe8c7138714e669e4cdc6688f980ac/notice>.

- [21] 数字化改革 公共数据分类分级指南: DB33/T 2351—2021 [S/OL]. [2022-07-20]. <https://std.samr.gov.cn/db/search/stdDBDetailed?id=CDD3E8DBEE528602E05397BE0A0A25E9>.
- [22] 重庆市大数据应用发展管理局关于印发《重庆市公共数据分类分级指南(试行)》的通知 [EB/OL]. [2022-07-20]. http://dsjj.cq.gov.cn/zwgk_533/fdzdgknr/lzyj/xzgfxwj/202206/t20220609_10792615.html.
- [23] 政务数据 数据分类: DB52/T 1123—2021 [S/OL]. [2022-07-20]. <https://std.samr.gov.cn/db/search/stdDBDetailed?id=D7B6AF366D9AB0BFE05397BE0A0A41EF>.
- [24] 政务数据分级与安全保护规范: DB11/T 1918—2021 [S/OL]. [2022-07-20]. <https://std.samr.gov.cn/db/search/stdDBDetailed?id=D7B6AF366E66B0BFE05397BE0A0A41EF>.
- [25] 信息安全技术 信息安全事件分类分级指南: GB/Z 20986—2007 [S/OL]. [2022-07-14]. <https://std.samr.gov.cn/gb/search/gbDetailed?id=71F772D78E2D3A7E05397BE0A0AB82A>.
- [26] 信息安全技术 大数据安全管理指南: GB/T 37973—2019 [S/OL]. [2022-07-20]. <https://std.samr.gov.cn/gb/search/gbDetailed?id=91890A0DA63B80C6E05397BE0A0A065D>.
- [27] 工业和信息化部办公厅关于印发《工业数据分类分级指南(试行)》的通知 [EB/OL]. [2022-07-20]. http://www.gov.cn/zhengce/zhengceku/2020-03/07/content_5488251.htm.
- [28] 电信和互联网服务 用户个人信息保护 定义及分类: YD/T 2781—2014 [S/OL]. [2022-07-29]. <https://hbba.sacinfo.org.cn/stdDetail/a6a47f277260c7783e96d48ed5a659bf>.
- [29] 基础电信企业数据分级分类方法: YD/T 3813—2020 [S/OL]. [2022-07-20]. <https://std.samr.gov.cn/hb/search/stdHBDetailed?id=B80DCDBCBAF0CE3AE05397BE0A0A06D8>.
- [30] 证券期货业数据分类分级指引: JR/T 0158—2018 [S/OL]. [2022-07-20]. <http://www.csrc.gov.cn/csrc/c101862/c1022471/content.shtml>.
- [31] 金融数据安全 数据安全分级指南: JR/T 0197—2020 [S/OL]. [2022-07-20]. <https://std.samr.gov.cn/hb/search/stdHBDetailed?id=B081D125A6762DB8E05397BE0A0A5EA7>.
- [32] 信息安全技术 健康医疗数据安全指南: GB/T 39725—2020 [S/OL]. [2022-07-20]. <https://std.samr.gov.cn/gb/search/gbDetailed?id=B691BB7786CD126E05397BE0A0AF3B3>.
- [33] 中国移动IDC维护管理规定 数据安全分册(2016版) [EB/OL]. [2022-07-20]. <https://max.book118.com/html/2017/0903/131519772.shtm>.
- [34] 云上贵州系统平台项目情况 [EB/OL]. [2022-07-20]. http://dsj.guizhou.gov.cn/zwgk/xxgkml/zdlyxx/dsjyztg/201809/t20180927_10392519.html.
- [35] 阿里巴巴数据治理最佳实践 [EB/OL]. [2022-07-23]. <https://www.doc88.com/p-99259566525709.html>.
- [36] 阿里云数据安全和隐私保护白皮书 [EB/OL]. [2022-07-20]. <https://developer.aliyun.com/topic/download?id=8241>.

作者简介

严炜炜, 男, 1987年生, 博士, 副教授, 研究方向: 数据管理与用户行为, E-mail: yanww@whu.edu.cn。

谢顺欣, 女, 1999年生, 硕士研究生, 研究方向: 数据管理与用户行为。

潘静, 女, 1997年生, 硕士研究生, 研究方向: 数据管理与用户行为。

陆伟, 男, 1974年生, 博士, 教授, 研究方向: 知识管理与数据挖掘。

Data Classification: Research Progress, Policy Standards and Enterprise Practice

YAN WeiWei XIE ShunXin PAN Jing LU Wei

(School of Information Management, Wuhan University, Wuhan 430072, P. R. China)

Abstract: As data elements have become national strategic resources, research on data classification has shown an explosive growth trend, and relevant policies and standards have been introduced to build a data security protection system. This paper quantitatively reveals the academic research trends of data classification in China by using information measurement and social network analysis, refines the development trend of open data classification policy standards by using the method of policy text analysis, and comprehensively explores the progress trend of data classification in China in combination with the refining law of enterprise practice. The research found that at the level of scientific research, relevant research focused on topics such as data security management, data governance, data mining and application. As for policy standards, the proposal of the central policy has experienced a process from point to area. Local standards mostly focus on the classification of government data and public data, focus on the security protection of the whole life cycle of data, and constantly update the content while highlighting regional characteristics. For industrial standards, data classification and grading industry guidelines and standard formulation have also formed a basic framework. For enterprise practice, data classification and grading practice modes represented by user main body type, theme scene type and business module type has been formed. The research suggests that give full play to the advantages of library and information science, promote the interdisciplinary research cooperation of data classification and classification theory and standard specifications, improve the coverage and content quality of local standards, and promote the improvement of industry guidelines and the establishment of enterprise data top-level specifications.

Keywords: Data Classification; Academic Research; Policy Regulations; Industry Guide; Enterprise Practice

(收稿日期: 2022-08-29)