

·专题:公共安全数据管理·

## 基于生命周期的公共安全数据管理模型研究\*

陆莉<sup>1,2</sup> 沙勇忠<sup>1,2</sup> 徐雪峰<sup>1,2</sup>

(1.兰州大学管理学院 甘肃兰州 730000)

(2.兰州大学应急管理研究中心 甘肃兰州 730000)

**摘要:**数据生命周期模型是链接数据管理策略理念与实践的关键模型。文章通过梳理典型数据生命周期模型的构建思路与阶段特征,结合公共安全数据的来源与类型,构建了全面的公共安全数据生命周期模型,并以该模型为基础,在分析公共安全治理内容的阶段特征基础上,提出了基于生命周期的公共安全数据管理三维模型,并就模型的主要功能进行了探讨。

**关键词:**数据生命周期;公共安全数据;公共安全治理;数据管理要素;数据管理模型

中图分类号:D63;G203 文献标识码:A DOI:10.11968/tsyqb.1003-6938.2019054

### Research about Public Safety Data Management Model Based on the Life Cycle

**Abstract** Data Life Cycle (DLC) model is the critical medium interlinking the data management strategies and practices. This paper analyzed the typical DLC models' construction and phases' characteristics as research basis. Combining the sources and features of public safety data, the overall public safety data life cycle model was constructed. Then basing on the public safety activities, the public safety data management model was put forward which combines the public safety data life cycle and the public safety activities life cycle. Meantime, the function of the public safety data management model was discussed as well.

**Key words** data life cycle; public safety data; public safety governance; data management elements; data management model

## 1 引言

公共安全数据管理是一个新兴的交叉学科研究领域,旨在探索物理空间和社会空间中的公共安全治理活动在信息空间中的映射——公共安全数据的规律与特征。通过数据科学与公共安全领域知识的融合,凝练公共安全治理活动中的数据问题,关注公共安全数据本身的管理进而促进数据在公共安全治理中的价值实现。有效的数据管理通常要做好长期的数据管理规划,并且在适当节点采取相应的管理活动,进行人员、机构、资源和关键事件之间的积极协同。在这一复杂的管理过程中,数据管理模型作为链接管理策略理念与管理实践的桥梁发挥着基础性作用。从科学数据管理与政府数据管理领域的相

关研究成果来看<sup>[1-2]</sup>,已提出了数据管理的生命周期模型、机构数据管理基础设施模型以及数据监管能力成熟度模型等典型模型。其中,数据生命周期模型能够以形象和可视的方式定义与阐明数据管理的复杂流程,将数据管理过程分解为不同的组成阶段,识别和阐释不同阶段的参与者角色、管理活动职责、流程重要事件和其他关键成分<sup>[3]</sup>,是具有基础意义的组织数据管理活动和优化数据服务的关键模型。

数据生命周期研究数据本身,考察其在生命周期中的阶段、状态和规律<sup>[4]</sup>。通过定义数据的每个存在阶段,以及阶段之间的连接关系,使整个数据生命周期成为一个不断演化和循环的连续统<sup>[3]</sup>。公共安全治理活动受领域特点和业务流程影响,其过程中生成或使用的数据及其生命周期有自身的规律。曾子明

\* 本文系教育部哲学社会科学研究重大课题攻关项目“大数据驱动的城市公共安全风险研究”(项目编号:16JZD023)与中央高校基本科研业务费专项资金资助重点项目“基于大数据的城市公共安全风险预警研究”(项目编号:17LZUJBWZD012)研究成果之一。

收稿日期:2019-08-20;责任编辑:魏志鹏

和杨倩雯<sup>[5]</sup>从总体上提出了城市公共安全数据的生命周期,包括数据采集、数据处理、数据存储、数据挖掘与分析、知识可视化、数据长期保存与再利用 6 个阶段。也有研究者对不同领域和不同类型数据的生命周期模型进行了探索,关注不同组织环境和业务情境下数据生命周期的差异性。如美国联邦地理数据委员会提出了针对地理空间数据的生命周期模型,该模型包括数据的定义、数据清单生成与数据资产评估、数据的获取与访问、数据维护、数据的利用与评估以及数据归档 6 个阶段<sup>[6]</sup>;针对传感器数据的生命周期模型,则对协作式生态学研究传感器数据的来源、情境、特征等进行了分析,构建了涵盖实验设计、校准和地面实况分析、数据采集、数据清洗、数值数据获取、多源数据集成、数据分析、出版、存储和长期保存 9 个阶段的传感器数据生命周期模型<sup>[7]</sup>。数据生命周期模型在不同领域中得到了扩展和细化,根据地球观测卫星委员会 2012 年的调研,已有 52 个不同学科领域提出了有针对性的数据生命周期模型<sup>[8]</sup>。

公共安全治理常面临复杂的跨领域、跨组织边界治理难题,其治理过程中生成或使用的公共安全数据既具有海量多源异构等大数据的普遍性特征,又受不同行业领域和多样化载体的影响,展现出细节上的差异性和特殊性。因此,公共安全数据生命周期的研究既要关注其整体性特征,也要对不同类型公共安全数据的特征和规律进行细化,发现其生命周期的多样性,丰富公共安全数据生命周期的内涵,使基于生命周期的公共安全数据管理模型既能够在理论上实现相对的统一,又能够贴近和指导管理实践。根据这一理念,沙勇忠等<sup>[9]</sup>提出了公共危机信息管理的双生命周期理论模型,即在公共危机管理生命周期的每一阶段嵌入信息管理全生命周期进行聚焦分析和有续统合,既从整体上实现了公共危机管理的减缓、准备、响应和恢复四阶段生命周期与信息管理的信生产、信息组织、信息分析和信息利用四阶段生命周期的有机结合,同时又能够根据公共危机管理不同阶段的差异化信息需求在信息管理内容上有所侧重。面对更为复杂的公共安全数据管理问题,如何实现数据管理的主体、客体、情境与活动等管理要素整体性和多样性的统一是研究的关键目

标。因此,本文从典型的数据生命周期模型分析出发,在梳理公共安全数据类型的基础上,探索公共安全数据生命周期的整体性特征和差异化规律,并基于此构建了公共安全数据管理的三维模型,以数据生命周期为基础结合公共安全数据的应用情境,探讨该模型在组织公共安全数据管理和支持公共安全治理活动中的功能与应用。

## 2 数据生命周期典型模型分析

数据生命周期(Data Life Cycle,DLC)定义了从数据生产阶段到消费阶段完整的数据生命框架,有助于组织从数据这一视角出发定义并构建业务情境中数据生产者和管理者的活动<sup>[10]</sup>。由于从最初生成到最终应用的完成,数据不断经历格式、应用方式和目标等方面的多重转变,相关参与者必须不同阶段通过一系列的数据管理活动使数据保持活力并能够发挥最大价值,以支持组织的决策制定和业务规划。构建数据生命周期模型的目的就在于通过有效的数据组织和清洗,实现最优的数据管理,进而为用户提供适当的、满足其质量需求的数据产品<sup>[11]</sup>。Sinaeepourfard A 等<sup>[12]</sup>认为,构建数据生命周期模型的作用可体现在以下几个方面:帮助管理者做好数据规划,以应对在数据生命各个阶段进行管理的复杂性问题;有助于在考虑限制条件与效率的情况下为终端用户提供符合需求的数据产品;帮助数据生成者和管理者进行数据的质量控制,明确在何种阶段移除干扰和噪音;阐明与数据生命相关的一系列基础活动和关键事件;帮助设计者开发能够可持续使用的数据管理标准或工具。在不同的数据密集型科学研究和实践中,由于跨领域或组织的差异,不同数据集在目标、来源、属性、所有权和传播发布限制等方面差异性显著,因而不同版本的数据生命周期模型在细节和数据活动上存在多样的展现方式。

本文梳理了部分典型的数据生命周期模型(见表 1)。发现不同类型的数据生命周期模型从开发目的和建构侧重点来看存在两种模式:一是偏向细节和差异性的阐释,重点面向特定场景与应用领域,强调具体的用户需求和数据挑战,如专门针对社会科学研究的 DDI 模型和强调个人数据管理的

APDL模型;二是侧重数据生命周期的总体性概括与抽象总结,使模型具有一定的普适性和灵活性,能够适应不同领域、多样化情境中的数据管理问题,如DCC模型和COSA-DLC模型。两种不同的模型构建思路是在数据生命周期整体性与多样性、普适性与特殊性之间的权衡。同时,不同的数据生命周期模型在数据生命周期阶段划分上也存在不同的侧重和表述。核心阶段如数据收集、数据保存、数据获取与使用等在大部分模型中均会出现。结合不同情境和管理需求,部分模型则会加入特定的生命周期阶段和管理元素,如研究设计与数据规划是科研数据管理过程中的初始阶段。数据销毁则仅出现在了个人数据生命周期模型中,以符合相关的数据管理和隐私保护政策。

### 3 公共安全数据的类型与生命周期模型

#### 3.1 公共安全数据类型

公共安全数据是指在公共安全治理活动中产生的或通过其他方式所获取的,能够反映公共安全治理活动的现象、特征和规律,进而支持管理决策的原始数据或经过加工整理的各类数据集,具有明显的海量、多源、异构和价值密度低等大数据特征。公共安全数据目前并未有统一的分类标准,海量的公共安全数据正存在于不同行业领域、不同组织机构以及不同的数据终端中。黄全义等<sup>[18]</sup>认为城市的公共安全数据可分为城市公共基础(市情)数据、部门业务数据、社情民意数据、物理环境与灾害监测数据、城市运行数据、人行为(活动)数据、突发事件应急处

表1 数据生命周期模型

名称	阶段	说明
DCC 数据监管生命周期模型 <sup>[13]</sup>	数据的产生与汇集、评估与选择、接入与融合、长期保存与存储、获取使用与重用,以及数据转换。	连续的循环模型,用户根据当前需求确定进入的阶段;提出涉及全生命周期的活动,描述与表征信息(元数据)、保存规划(监管活动与评估目标)、社区观察与参与(用户使用及参与标准工具等的开发)、监管与保存(数据回顾与连续性活动评估)。
DataONE 数据生命周期模型 <sup>[14]</sup>	数据的规划、收集、质量控制、描述、长期保存、发现与获取、集成与分析。	DataONE 开发工具、服务和教育素材的基础框架;部分研究活动既可以是只涉及部分阶段的线性路径,也可以是涉及全周期所有阶段的多重循环。
DDI 联合生命周期模型 <sup>[15]</sup>	研究设计与数据规划、数据收集、数据处理、数据归档、数据分配、数据发现、数据分析、数据重用。	该模型主要针对社会科学研究数据,从数据应用的视角出发,模型主要阶段为线性流程,同时包含数据处理阶段的一个二项选择路径与数据分析之后的反馈回路。
ANDS 数据共享活动模型 <sup>[15]</sup>	数据创建、存储、描述、识别、注册、发现、获取、开发。	该模型可作为一种结构化技术帮助数据生产者和消费者进行操作规划和使用说明制定,以支持数据的共享和重用。
COSA-DLC 数据生命周期模型 <sup>[12]</sup>	数据采集(收集、过滤、质量控制、描述)、数据处理(处理、质量管理、分析)、数据保存(分类、质量管理、归档、发布)。	该模型从数据6V特征出发,是全面的可适用于任何领域或学科的抽象模型,即可适用于任何不可知情境,但同时能够灵活满足具体情境下数据管理的需求与挑战。模型包括三个主要模块,每个模块下又包含了具体的阶段和活动。
BLM 数据生命周期模型 <sup>[11]</sup>	数据规划、获取、维护、使用、评估与归档。	该模型面向美国公共土地资源数据的管理,目的在于促进消费者之间的信息共享和提供高质量数据。模型以非线性的方式表示,数据质量管理是核心议题。
LOD2 关联数据生命周期模型 <sup>[16]</sup>	存储/查询、人工校订/程序编写、数据互联/融合、数据分类/扩展、质量分析、评估/纠正、搜索/浏览/探索、数据抽取。	该模型目的在于帮助搜索与抽取符合最终用户需求的有用数据,模型各个阶段不是单独处理,而是采用互相依赖的方法实现数据的语义关联,促进基于数据网络的知识抽取。
JISC 研究数据模型 <sup>[11]</sup>	数据规划、创建、使用、评估、出版、发现、重用。	该模型旨在促进英国高等教育和研究中的数据管理与共享,其设计的主要目的在于促进研究数据的共享与发现。
APDL 个人数据生命周期模型 <sup>[17]</sup>	构建概念模型、初始计划、数据收集、保留、存取、回顾与矫正、使用、发布、销毁。	该模型用于支持个人数据的管理和可追溯性,识别了不同生命周期阶段个人数据的状态、处理活动和参与者角色,促进数据活动符合法律标准和提升隐私保护。

注:①DCC:Digital Curation Centre,英国数字管理中心;DataONE:Data Observation Network for Earth,美国国家科学基金会地球数据观测网;DDI:Data Documentation Initiative,描述调研、问卷、统计数据资料等社会科学研究信息的国际标准;ANDS:The Australian National Data Service,澳大利亚国家数据服务中心;COSA-DLC:Comprehensive Scenario Agnostic Data LifeCycle,与场景无关的全面数据生命周期模型;BLM:Bureau of Land Management,美国土地管理局;LOD2:Linked Open Data,关联开放数据项目,由欧盟委员会FP7信息与通信技术工作计划资助;JISC:The Joint Information Systems Committee,英国联合信息系统委员会;APDL:Abstract Personal Data Lifecycle,个人数据生命周期抽象模型,由牛津大学Majed Alshammari等人提出;②6V:价值(Value)、体量(Volume)、多样性(Variety)、速度(Velocity)、可变性(Variability)、精确性(Veracity)。



置数据和公共安全知识数据八大类,每个大类下又涵盖了各种类型的具体数据。Thakuriah P等<sup>[19]</sup>则将城市大数据划分为传感器系统(基于基础设施或移动目标传感器)数据、用户生成内容(“社会”或“个人”传感器)数据、管理(政府)数据(开放的和保密的微观数据)、私营部门数据(消费者和交易记录)、艺术和人文数据以及混合数据(链接数据和综合数据)。

巴志超等<sup>[20]</sup>认为国家安全大数据是以各种复杂关联形式分布在人、机、物高度融合的信息空间、物

理世界和人类社会三元世界中。由于广义上的公共安全指所有事关社会生活安全稳定的外部环境和秩序,包括经济安全、环境安全、公共卫生安全、信息安全、食品安全、生产和工作场所安全等<sup>[21]</sup>,因此,上述安全领域中产生或使用的数据均可称之为公共安全数据。本研究从不同安全领域数据的共性出发,从数据来源和数据类型上对公共安全数据的分类进行总结(见表2)。

公共安全数据的来源是由通信网、互联网、传感

表2 公共安全数据的来源与类型

来源空间	数据类型	内容	应用举例
社会空间	政府行政管理数据(可公开数据与保密数据)	①政府开放数据,如经济建设、资源环境、教育科技、道路交通、社会发展等领域可作为开放数据提供给公众的数据资源。	公共安全治理中的公众参与,如企业与公众利用开放数据进行应用软件开发。
		②政府行政管理过程中不同安全领域的其他业务数据,如食品安全抽检数据、危化品管理数据、消防数据、监管对象的证照数据等。	支持各类风险的预警与防控;支持各类公共安全事件应对过程中的管理决策;科研机构的公共安全管理工具研发。
		③保密数据,如公安部门等负有特殊职能的机构数据。	支持国家安全、公共安全事务及执法司法行动中的风险分析与安全决策等。
	非政府组织管理数据(包括非营利性组织和盈利性组织的数据)	①第三方机构数据,如民间应急救援组织的志愿者人数、救援物资等数据。	应急救援过程中政府职能部门与第三方机构之间的协调和管理。
		②高校与图书馆保存的公共安全知识数据,如专利、文献、科研数据等。	科研机构进行公共安全管理工具研发;公共安全管理政策建议的提出。
		③企业等私营部门数据,如消费者个人情况数据、交易记录数据、公共事业和金融机构的使用数据。	把握个体行为并进行信用评估,支持金融风险防控;支持城市规划决策等。
	人口统计学数据	关于年龄、性别、收入、职业、健康、社会福利、教育记录等个人层面的保密数据;公开的统计年鉴数据。	受灾人群的人口特征表征和社会经济状况剖析;支持减灾政策制定等。
	个体或群体行为特征数据	个人或群体的个性特征、情绪表征、认知决策和压力应对等心理数据与行为数据,如谷歌街景(Google Street View)可提供的行人活动数据 <sup>[22]</sup> ,智能穿戴设备提供的情绪特征与生理特征数据。	突发公共安全事件中的公众情感分析与情绪预警;人群状态分析与异常行为检测等。
物理空间	公共安全基础数据	①静态基础数据,包括关键基础设施、重大危险源、道路交通设施、电力设施等相关行业、部门的静态数据。	公共安全风险识别与风险源监测;关键风险与危险因子分布地图构建等。
		②动态基础数据:社会日常运行中产生的物流信息、车流信息、金融流转信息等动态数据。	时空尺度可变的键关键风险呈现与分析;支持实时动态的风险评估与预警等。
	传感器数据	①基于位置感知传感器的时空网络大数据(Big Spatio-Temporal Network(BSTN)Data) <sup>[23]</sup> ,如路线图、GPS轨道、交通信号计时、车流量测量数据等。	智慧交通监管;突发事件情境下的应急救援路径优化;危险品道路运输管理与监测等。
		②基于空中、空间和地面传感器的卫星遥感数据。	自然灾害事件的灾情评估;支持救灾行动与决策,优化救灾物资配置。
		③自然环境监测数据,如气象、海洋、地质等与自然灾害相关的数据资料。	自然灾害风险预警;自然灾害突发事件应对与应急决策。
虚拟网络耦合空间	用户生成内容数据	①社交媒体数据,如推特、新浪微博等社交网站数据,微信等及时通信数据,贴吧、论坛等内容数据。	网络舆情分析与预警;突发公共安全事件中的信息发布、风险交流与危机沟通。
		②自发地理信息数据(Volunteered Geographic Information,VGI) <sup>[24]</sup> ,如用户上传至互联网的地理位置与行动轨迹等信息,或带有地理定位的图片和视频。	大规模人流监测;突发事件过程中实时辅助捕捉主要危害的演变与影响。
		③群体智慧数据,如众包模式下应急管理过程中受灾公众对事件的描述信息、社会公众对灾情信息的加工与标注。	应急响应与应急恢复的决策制定。
		④其他社会生成数据,如社会民意调查数据等。	公众的风险认知状况调查与风险态度评估。
	基础网络与应用系统数据	信息化建设中基础网络及各类应用系统处理的数据和相关信息安全风险数据。	应急管理系统的构建与完善;信息安全事件应急处置等。

网等相互融合链接所形成的混合环境,即从“物理空间-社会空间-虚拟网络耦合空间”中剥离出来的数据空间。物理空间中“物”是公共安全数据的载体,既包括关键基础设施、工业危险源、道路交通等结构化、半结构化基础数据;也包括通过人类生活空间中部署的大规模多种类传感设备所感知到的环境物理参数、图片、声音、视频等非结构化、多模态数据。社会空间中,个体或组织则是公共安全数据的生产者或持有者,既包括基础的人口统计学数据与个人或群体的个性特征、情绪表征、认知决策和压力应对等心理数据与行为数据,也包括各类组织的业务数据或公共安全知识数据等。虚拟网络耦合空间是指网络空间与物理空间和社会空间的重叠耦合,其数据载体是互联网或移动终端设备,既包括信息化建设中基础网络及各类应用系统处理的数据和相关信息安全风险数据(物理空间与网络空间的耦合),也包括新闻报道、论坛帖子、博客文章等社交媒体中的非结构化和半结构化数据(社会空间与网络空间的耦合)。

总体来说,公共安全数据具有以下明显特征:①连续性,公共安全治理活动的持续性决定了公共安全数据生成和获取的连续性,如政府部门的业务系统不断产生更新数据;②复杂多样性,公共安全数据涉及不同领域、不同来源、不同结构类型的数据,且数据规模大,增长速度快;③关联性,公共安全治理活动需要跨部门、跨行业、跨领域的关联,与之相应公共安全数据需要进行语义关联以支持数据增值与重用;④涉及主体的多元性,公共安全数据的生产、管理和使用过程中涉及政府组织、企业、公众、研究人员等不同参与主体。上述数据特征也决定了其管理活动的复杂性和系统性。

### 3.2 公共安全数据生命周期模型

#### 3.2.1 公共安全数据生命周期的阶段划分

考虑到公共安全数据的类型及其特征,本文借鉴典型数据生命周期模型的阶段划分,将公共安全数据生命周期划分为五个主要阶段,每个阶段又包含了详细的子阶段。

(1)数据规划(Data Planning):数据规划是对公共安全数据整体状况的描述,对数据在整个生命周

期阶段中如何被管理进行总体上的设计,主要包括两个子阶段:①数据情境分析。该阶段是数据规划的起点与基础工作,结合公共安全治理活动中公共安全数据产生与应用的具体情境,明确不同情境下公共安全治理任务的实现对数据的需求和管理目标;②管理规划设计。包括不同生命周期阶段,执行数据活动的基本原则、方法、标准、政策法规、基础设施、知识产权等关键问题的讨论和设计。如数据产生与收集的工具、数据评估和描述标准的制定、数据质量控制措施等。

(2)数据采集(Data Acquisition):数据采集是在数据规划阶段提出的设计方案指导下,实现数据的创建、收集、描述与评估,保证数据采集的全面性、准确性和相关性,主要包括以下几个子阶段:①数据产生与创建。通过业务信息化、社会调查、部署传感器与物联网建设、基于社交媒体等的用户内容生成以及仪器工具测量等方式,将潜在的数据转化为可存储和可获取的数字形式;②数据收集与描述。根据规划方案规定的收集范围,采集确定来源、载体、格式等属性要求的数据,规定数据收集的模板以指导多样的数据收集者,并用适当的元数据标准对数据进行全面精准的描述。同时提供与数据相关的情境信息,包括数据收集方法及过程的描述、收集人员信息、业务情境、参数说明、知识产权等;③数据评估与选择。在数据收集过程中执行数据质量控制标准的基础上,进一步通过机器清洗或人工检查保证数据质量,使其符合相关的任务要求和政策规定。

(3)数据处理(Data Processing):数据处理是对采集到的数据进行集成、分析等工作,为数据的增值和应用做好准备,以实现从数据到信息再到知识和智慧的有效流动,主要包括以下几个子阶段:①数据集成。从格式、编码方案、本体等方面转换不同来源获取的相关数据集,形成机器可读的同质数据集,以支持数据的分析和挖掘工作;②数据分析。运用统计分析、模拟仿真、自然语言处理等方法,从数据中抽取有价值的信息并发现知识,以可视化的方式将分析结果呈现给终端用户,这是数据增值的关键步骤。

(4)数据保存(Data Preservation):数据保存是将有价值的数据根据公共安全治理活动需要和领域类

型的差异进行分类存储和选择性的长期保存,主要包括以下几个阶段:①数据分类。根据数据应用情境、所在安全领域和数据来源与类型,按照统一的数据组织主题类目整理数据,为实现有效的数据存储做准备;②数据存储。在数据分类的基础上,针对不同类型与来源的数据选择相应的介质进行存储,并通过存储设备维护、数据备份、检查不变性、权限设置等措施保证数据的可获得性、准确性和安全性;③数据长期保存。由于公共安全数据的海量与持续更新的特征,需要识别具有长期保存价值的数据,并将相关数据提交专门机构进行适当的归档与长期保存,以减小数据管理成本、保证数据管理的安全性和持续性,支持数据的不断增值与重用。

(5)数据使用(Data Using):数据使用指用户发现并获取已发布的数据,运用数据及其分析结果来支持公共安全治理活动和决策,主要包括以下几个子阶段:①数据发布。根据相关管理政策,将经过分级分类整理的数据予以发布,为指定用户或公共终端用户提供数据获取和使用权限,实现数据的共享共用;②数据发现与获取。用户通过一定的方式获取到关于数据的基本信息,进而定位和获取到潜在有用的数据,如通过政府数据开放平台提供的API等方式;③数据使用。数据生产者或者其他使用者运用数据及其分析结果支持管理决策,进行数据产品的开发和数据服务提供等;④数据重用。已保存的数据可支持领域公共安全数据库的建设与知识库开发,支持数据回溯、检查与对比分析,并作为历史数据支持学术研究与报告撰写等工作。

公共安全数据生命周期的各个阶段在理想状态下形成线性的循环,实践中则多会产生非线性的回路,或仅聚焦于几个特定的生命周期阶段。详细的子阶段划分体现了较为全面的阶段任务,具有很大的灵活性和适应性,能够根据公共安全治理活动不同情境的差异性满足特定数据需求。

### 3.2.2 公共安全数据生命周期的主体层次

从参与主体角度来看,公共安全数据生命周期模型可分为基于个体的数据生命周期、基于组织的数据生命周期和基于领域群体的数据生命周期。

(1)基于个体的数据生命周期。公共安全治理活

动中的个体是最细粒度的参与主体,典型如食品安全治理活动中的每一个消费者,或者自然灾害发生时的受灾人群。个体在公共领域有意识的产生,并能够提供开放利用的数据被认为是公共治理的基础之一<sup>[25]</sup>。作为公共安全治理活动的利益相关者,其产生的如人口统计学数据、社交媒体数据、个体的心理特征与行为数据等,可为支持大规模人流监测、救灾物资优化配置、危机沟通与决策制定等公共安全治理活动提供重要依据。但由于数据质量、格式、产生目的、存储等受个体影响差异性最大,使得数据的收集、融合和长期保存难度也最大。在基于个体的公共安全数据生命周期中,数据活动往往集中在上游阶段,很难形成可以不断演化循环的回路。

(2)基于组织的数据生命周期。不同类型的组织在参与公共安全治理活动和提供公共安全服务时产生了基于组织的数据生命周期,其中最典型的即政府组织。作为公共安全治理活动的主导者和公共安全服务的核心提供者,政府组织也是公共安全数据的主要管理者 and 使用者。政府行政管理过程中产生的业务数据是公共安全数据的重要组成部分,在数据格式标准化和质量管理方面也具有行政优势。且随着政府数据开放运动的发展和大数据发展战略的制定,政府组织内部的数据基本实现了从产生到发布共享与利用的数据生命周期各个阶段的管理。但由于其数据的海量、持续性和部分数据的保密性等特征,在不同阶段的数据活动投入不同,也使得全生命周期的数据管理存在薄弱环节,如数据规划、数据评估与选择等。企业等非政府组织也注重组织内全生命周期的数据管理。

(3)基于领域群体的数据生命周期。不同安全领域的公共安全治理活动利益相关者整体上构成了一个特定的领域群体,如食品安全领域的“食品安全监管者-食品经营者-消费者”群体。基于领域群体的数据生命周期提供一个更为宏观的视角对实践中的数据活动进行全面的理解,如DCC数据生命周期模型就是典型代表。该类生命周期为领域群体提供一个共享的、能够被普遍接受的基础解释,可以帮助组织规划其数据活动,保证生命周期发展的正确顺序和管理过程中必须考虑在内的管理元素。相比于基于



个体和组织的生命周期,基于领域群体的生命周期更能够表征领域实践的数据所经历的共同阶段,在细节和精确性上有所弱化,更注重普遍性特征的高度概括与总结,其内涵和特征在基于个体和组织的生命周期中有不同的侧重和体现。

### 3.2.3 公共安全数据的生命周期模型

综合上述分析,本文提出公共安全数据的生命周期模型(见图1)。该模型从公共安全数据出发,充分考虑了公共安全治理中数据活动参与主体的差异性,形成了包括数据规划、数据采集、数据处理、数据保存和数据使用五个主要阶段在内的循环流程。其中,五个主要阶段又包含了具体的子阶段,涵盖了14个公共安全数据活动的主要任务阶段。不同数据阶段之间的关联通过数据格式的转换、数据活动参与者职责的确定以及数据生产者、管理者和使用者角色的明确得以体现,是保证数据质量和促进全生命周期数据管理的关键。

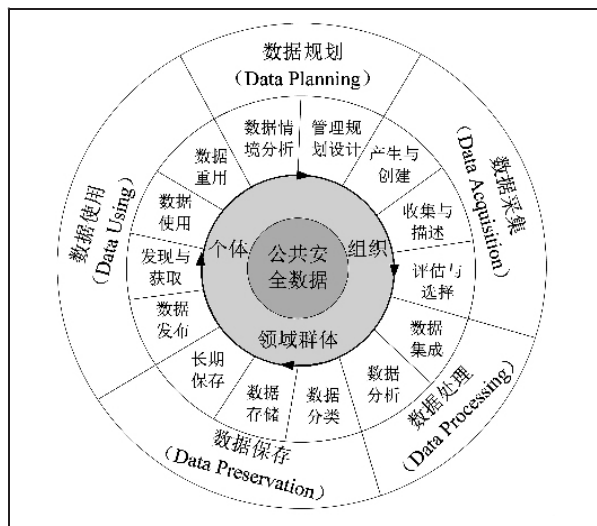


图1 公共安全数据生命周期模型

总体来看,该公共安全数据生命周期模型是一个全面的、概括的模型,涵盖了主要的数据生命周期阶段以保证其指导性和前瞻性。同时,该模型由于考虑了公共安全治理活动中参与主体与数据类型的差异性,能够迅速适应公共安全治理活动的不同场景,根据不同场景需求灵活地进行生命周期子阶段的选择、组合与扩展,满足不同主体在不同情境下的异质性数据需求,并保证较高水平的数据质量。其在应用中的潜在优势包括:可实现面向未来的数据发现、集

成和处理,管理和组织总体数据集;针对具体场景提供简单的个性化定制和采用方式;提高具体情境下的数据质量水平;对于具体应用者如数据、软件和系统设计者来说,能够消除额外的资源浪费和努力,根据自身需求灵活制定有效的数据架构。

## 4 基于生命周期的公共安全数据管理三维模型

### 4.1 公共安全数据管理三维模型的构建

公共安全数据管理必须紧密结合实践,体现公共安全治理活动的内容和需求。在明确公共安全数据生命周期模型的基础上,本文进一步结合公共安全治理内容与情境,构建公共安全数据管理的整体模型。从管理过程而言,风险管理、应急管理、危机管理组成了公共安全活动的主要内容<sup>[26]</sup>,也成为公共安全数据的主要应用情境。由于风险管理、应急管理、危机管理同样具有其生命周期的阶段特征,使得将数据生命周期嵌入公共安全治理活动,使数据活动响应不同阶段公共安全治理活动的实际需求成为可能。以此构建了基于生命周期的公共安全数据管理三维模型(见图2)。该模型将公共安全数据生命周期与公共安全治理内容的生命周期结合起来,帮助细化不同管理内容和情境下的数据需求,进而使数据管理活动能够切合公共安全治理活动和不同组织内部的业务管理活动,实现有效的、切合公共安全治理实践的数据管理。

公共安全治理从现实路径来看,以突发事件应急管理为切入点,延伸至公共危机管理,进而推动社会风险治理<sup>[27]</sup>。应急管理、危机管理与风险管理在其

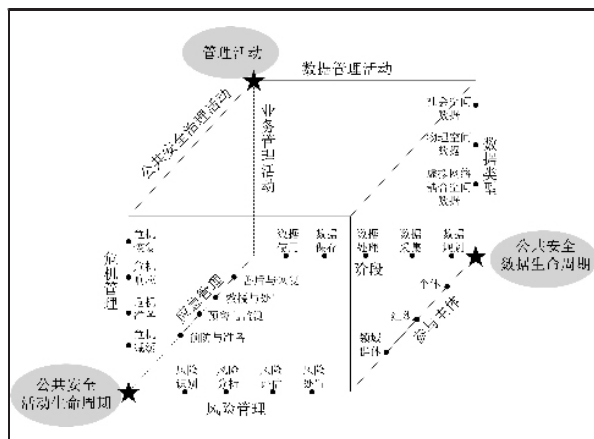


图2 基于生命周期的公共安全数据管理三维模型

运行机制上存在自身结构,不同管理要素的角色、功能和相互关系构成了其各自的生命周期循环。我国的应急管理循环圈包括预防与准备、预警与监测、救援与处置、善后与恢复四个阶段<sup>[26]</sup>;危机管理阶段论也有不同的分类,美国联邦应急管理局(FEMA)将危机发展过程分为减缓、准备、响应与恢复四个阶段<sup>[9]</sup>;风险管理则包括风险识别、风险分析、风险评估和风险处置四个阶段。

定义和构建公共安全数据活动框架的生命周期则可看成是一个嵌入到公共安全治理活动中的子集。不同的数据管理活动响应公共安全治理不同管理内容在不同阶段对数据的差异性需求。如应急预防与准备阶段、危机减缓和准备阶段、风险识别阶段均侧重于人口统计学数据、公共安全基础数据等反映自然与社会状态基础数据的收集与管理;应急预警与监测则注重对反映突发事件状况实时数据的动态收集和迅速处理,如自然环境监测数据、传感器数据和社会媒体数据等,以支持应急决策和救援资源配置工作。当然,不同管理内容与情境下的数据需求是一个复杂问题,涉及到不同安全领域又会加入更多的约束条件,需要结合实践进行深入的探索与分析。

在公共安全活动内容的生命周期与公共安全数据生命周期的嵌套融合之下,数据活动响应公共安全治理活动和不同类型组织的业务管理活动,使得数据管理活动既能涵盖数据管理过程中的各类要素,同时能够有效切合公共安全治理实践情境与组织的业务特征。

#### 4.2 公共安全数据管理三维模型的功能

基于生命周期的公共安全数据管理三维模型的功能体现在以下几个方面:

(1)有助于统筹公共安全数据管理过程中的各类要素。该模型将客体要素(不同来源和类型的公共安全数据)、主体要素(数据活动中不同的参与主体)和情境要素(公共安全活动内容)统一起来,明确了公共安全数据管理过程中主要要素的内涵和基本问题,并以生命周期为纽带形成了各要素之间的关联和作用机制,为实行有效的公共安全数据管理提供了基础条件。

(2)辅助公共安全数据管理规划的制定。公共安全数据管理规划从宏观上来说,是公共安全治理活动的重要组成部分;从中观上来说,涉及政府组织等核心主体的公共安全数据管理顶层设计与管理规划;从微观上来说,则是数据生命周期的第一阶段,需要考虑数据各个生命周期阶段的状态和活动。该模型为不同层面的数据规划制定提供一个统一的管理框架和灵活的适应机制,有助于不同层面数据管理规划的制定,构建并完善公共安全数据管理体系。

(3)指导不同参与主体的公共安全数据管理活动。将生命周期贯穿于整体管理模型的构建,实现不同参与主体的管理职能和角色职责随着生命周期阶段变化而转换,并在此基础上将公共安全数据管理活动统一集成在公共安全活动生命周期与公共安全数据生命周期的特征与规律基础之上,指导不同参与主体制定每个生命周期阶段内的管理策略、任务和活动。

(4)支持公共安全治理活动。基于以上功能的实现,该模型通过有效的公共安全数据管理,使数据在公共安全治理活动中充分发挥价值,实现数据驱动的公共安全治理的动态监测与预警、及时反馈与精准研判、决策支持与智慧管理目标。

## 5 结语

基于生命周期所构建的公共安全数据管理三维模型。既是一种理论概括,也是一种分析框架,对公共安全数据管理这一交叉学科领域具有理论认识和实践指导意义。一方面,该模型可作为关键基础模型,丰富公共安全数据管理的基础理论并为相关研究提供指导性框架;另一方面,该模型充分考虑了公共安全活动与数据管理要素、关系和过程的关联统一,有助于相关组织和领域群体进行数据管理规划和业务集成,开展全生命周期的公共安全数据管理活动。由于公共安全活动内容的领域多样性和参与主体的多元性,不同领域情境下对公共安全数据的异质性需求和数据管理模型的具体应用需要进一步探索,从数据类型、生命周期阶段、参与主体类型等多个角度对模型的合理性、解释性和适用范围进行验证,并结合领域实践进行不断修正与完善。



## 参考文献:

- [1] 王芳, 慎金花. 国外数据管护(Data Curation)研究与实践进展[J]. 中国图书馆学报, 2014, 40(4):116-128.
- [2] 黄如花, 赖彤. 数据生命周期视角下我国政府数据开放的障碍研究[J]. 情报理论与实践, 2018(2):7-13.
- [3] Carlson J. The use of life cycle models in developing and supporting data services[J]. Research data management: Practical strategies for information professionals, 2014:63-86.
- [4] 丁宁, 马浩琴. 国外高校科学数据生命周期管理模型比较研究及借鉴[J]. 图书情报工作, 2013, 57(6):18-22.
- [5] 曾子明, 杨倩雯. 面向第四范式的城市公共安全数据监管体系研究[J]. 情报理论与实践, 2018(2):82-87.
- [6] EPA IWGDD C. Harnessing the Power of Digital Data: Taking the Next Step[R]. Scientific Data Management (SDM) for Government Agencies: Report from the Workshop to Improve SDM (Report No. CENDI/2011-1). Washington D.C., 2011.
- [7] Wallis J C, Borgman C L, Mayernik M S, et al. Moving archival practices upstream: An exploration of the life cycle of ecological sensing data in collaborative field research[J]. International Journal of Digital Curation, 2008, 3(1): 114-126.
- [8] CEOS. Data lifecycle models and concepts[EB/OL]. [2018-11-18]. <http://wgiss.ceos.org/dsig/whitepapers/>.
- [9] 沙勇忠. 公共危机信息管理[M]. 北京: 中国社会科学出版社, 2016.
- [10] Bishop W, Grubestic T H. Data Lifecycle[M]. Geographic Information, 2016.
- [11] Sinaeepourfard A, Masip-Bruin X, Garcia J, et al. A Survey on Data Lifecycle Models: Discussions toward the 6Vs Challenges[DB/OL]. [2019-05-20]. <https://www.ac.upc.edu/app/research-reports/html/RR/2015/18.pdf>.
- [12] Sinaeepourfard A, Garcia J, Masip-Bruin X, et al. A comprehensive scenario agnostic Data LifeCycle model for an efficient data complexity management[C]. 2016 IEEE 12th International Conference on e-Science(e-Science). IEEE Computer Society, 2016.
- [13] Higgins S. The DCC curation lifecycle model[J]. International journal of digital curation, 2008, 3(1): 134-140.
- [14] DataONE. Data Life Cycle[EB/OL]. [2019-02-27]. <https://www.dataone.org/data-life-cycle>.
- [15] Ball A. Review of Data Management Lifecycle Models [DB/OL]. [2019-05-25]. <https://purehost.bath.ac.uk/ws/portalfiles/portal/206543/redm1rep120110ab10.pdf>.
- [16] Auer S, Bühmann L, Dirschl C, et al. Managing the life-cycle of linked data with the LOD2 stack[C]. International semantic Web conference. Springer, Berlin, Heidelberg, 2012: 1-16.
- [17] Alshammari M, Simpson A. Personal data management: an abstract personal data lifecycle model[C]. International Conference on Business Process Management. Springer, Cham, 2017: 685-697.
- [18] 黄全义, 夏金超, 杨秀中, 等. 城市公共安全大数据[J]. 地理空间信息, 2017, 15(7):1-5.
- [19] Thakuriah P, Tilahun N Y, Zellner M. Big Data and Urban Informatics: Innovations and Challenges to Urban Planning and Knowledge Discovery[M]. Seeing Cities Through Big Data. Springer International Publishing, 2017.
- [20] 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.
- [21] 刘莘. 公共安全与秩序行政法[J]. 江苏社会科学, 2004(6):49-50.
- [22] Yin L, Cheng Q, Shao Z, et al. 'Big Data': Pedestrian Volume Using Google Street View Images[M]. Seeing Cities Through Big Data. Springer, Cham, 2017: 461-469.
- [23] Gunturi V M V, Shekhar S. Big spatio-temporal network data analytics for smart cities: research needs[M]. Seeing Cities Through Big Data. Springer, Cham, 2017: 127-140.
- [24] Goodchild M F. Citizens as sensors: the world of volunteered geography[J]. GeoJournal, 2007, 69(4): 211-221.
- [25] Meijer A, Potjer S. Citizen-generated open data: An explorative analysis of 25 cases[J]. Government Information Quarterly, 2018, 35(4): 613-621.
- [26] 薛澜, 周玲, 朱琴. 风险治理: 完善与提升国家公共安全管理的基础[J]. 江苏社会科学, 2008(6):7-11.
- [27] 童星. 论风险灾害危机管理的跨学科研究[J]. 学海, 2016(2):96-101.

**作者简介:** 陆莉(1990-), 女, 兰州大学管理学院、兰州大学应急管理研究中心博士研究生; 沙勇忠(1968-), 男, 兰州大学管理学院、兰州大学应急管理研究中心教授, 博士生导师; 徐雪峰(1995-), 男, 兰州大学管理学院、兰州大学应急管理研究中心硕士研究生。