



长江航道要素数据资源元数据研究^{*}

李学祥^{1,2}

(1. 武汉理工大学, 湖北 武汉 430063; 2. 长江航道规划设计研究院, 湖北 武汉 430011)

摘要: 在国内外相关行业元数据标准研究的基础上, 综合考虑长江航道监测与测量数据采集的实际情况、航道维护管理业务决策和航道综合信息服务的需要, 将长江航道要素数据资源划分为 7 大类、38 小类。针对长江航道要素数据的查询、检索、定位、服务的需求, 确立了长江航道要素数据资源元数据的内容框架, 包括标识信息、限制信息、数据质量信息、数据维护信息、空间表示信息、参照系信息、内容信息、覆盖范围信息、分发信息等 9 个方面, 并运用面向对象的方法进行了表达, 给出了核心元数据, 为推进长江航道要素数据资源的标准化和高效共享提供技术支撑。

关键词: 航道要素; 数据资源; 分类编码; 元数据

中图分类号: U 612. 26 文献标志码: A 文章编号: 1002-4972(2016)01-0027-05

Research on metadata of waterway elements' data resource of the Yangtze River

LI Xue-xiang^{1,2}

(1. Wuhan University of Technology, Wuhan 430063, China;
2. Changjiang Waterway Planning, Design and Research Institute, Wuhan 430011, China)

Abstract: Based on the study of the metadata standard about related areas at home and abroad, we divide the data resource of the elements of the Yangtze waterway into 7 categories, 38 small classes, with a comprehensive consideration of the waterway monitoring and data collection and the need of maintenance decisions and information service of waterway. We then establish the content framework of the metadata of the data resource of the Yangtze waterway's elements, which includes nine aspects, such as identification information, limited information, data quality, data maintenance information, spatial information, frame of reference information, main content information, coverage information, distribution information, all of which are expressed by the object-oriented method. The core metadata is also shown. All will help promote the standardization of data resources of the elements of the Yangtze waterway and provide a technical support for highly efficient sharing.

Keywords: waterway elements; data resource; classification and coding; metadata

近年来, 随着信息技术的飞速发展, 内河航道信息化已经成为变革传统的航道养护管理模式、提升内河航道综合信息公共服务能力的重要支撑手段, 同时也是综合交通运输信息化的重要组成部分。由于内河航运发展对内河航道的服务需求

以及我国内河航道通航条件变化的复杂性, 需要对各类航道要素进行采集、处理与加工、形成实时完备的数据源, 以构建完善的航道综合信息服务体系为基础, 推动实现航道实时动态监测、联网便捷服务, 才能更科学高效地支撑航道养护管

收稿日期: 2015-10-25
^{*}基金项目: 交通运输部 2013 年信息化重大专项 (2013-364-548-200)
作者简介: 李学祥 (1981—), 男, 博士后, 高级工程师, 从事水路交通信息化、数字航道和智能航道研究。

理决策、更便捷地满足内河航运等涉水活动业务应用。然而作为内河水运主通道的长江航道,仍然存在数据资源分散、标准不统一等问题,与长江航运等涉水业务应用对长江航道要素数据的查询、检索、定位、服务的需求不适应。以长江航道为样本,研究航道要素数据资源元数据的定义问题,制定满足交通运输元数据标准的长江航道要素数据元数据标准,支撑长江航道要素数据资源的“组织-发布-发现-定位-访问”,为推进长江航道要素数据资源的标准化和高效共享提供技术支撑。

1 国内外研究现状分析

长江航道要素数据资源是长江航道数据资源的核心组成部分,属于内河航道基础地理空间信息的范畴。元数据为信息系统各个层次的内容提供规范的定义、描述、交换和解析机制,为数据的整合与分布、异构系统之间的互操作提供服务。

1.1 国内研究现状

在航道数据资源标准方面,我国已有 JT/T 697-10—2009《交通统计信息基础数据元》、JT/T 697.4—2013《交通信息基础数据元 第4部分:航道信息基础数据元》^[1]、《内河航运综合信息服务电子报文数据元》(征求意见稿)等,分别从不同的侧面对航道数据资源的模型进行了规范化。在元数据方面,文献[2]系统探讨了元数据的理念,给出了多种不同的定义表述。文献[3]认为元数据就是一种对数据资源的内容、结构、价值、访问、使用等各个方面进行描述的工具,也是一种有效的资源组织方法,并对资源集合、对象元数据以及元数据统一组织方案的设计思路与方法等进行了分析研究。

国家标准 GB/T 19710—2005 是结合我国国情修改国际标准 ISO19115:2003 制定的元数据标准。SDS/T 2112—2004 是科技部的《科学数据共享工程》的建设成果,旨在提高科学数据的建库质量,规范数据加工的标准化,促进科学数据的共享与交

换,该标准适用于科学数据集元数据的管理,但是抽象层次较 GB/T 19710—2005 和 ISO19115:2003 更高,这些都是基础性、上层次的标准,均不宜直接用于具体行业应用。

国家标准委于 2007 年 11 月发布了《政务信息资源目录体系与交换体系》标准。文献[4]针对政务信息化建设中出现的“信息孤岛”等问题,在分析政务信息资源目录体系架构的基础上,提出了基于元数据的目录服务,对政务信息资源目录服务体系的框架进行了分析研究,并提出了基于元数据的政务信息资源的目录服务系统实现的关键技术和方法。文献[5]紧密结合水利信息资源目录体系,研究了水利信息资源目录服务元数据的定义问题,在充分考虑水利信息资源特点的基础上,定义了一套支持资源“发布-发现-访问”的元数据标准。国家海洋信息中心构建了基于元数据的渤海海洋信息资源目录服务系统^[6]。

1.2 国外研究现状

元数据(Metadata)一词最早出现在 1988 年美国国家航空航天局(NASA)发布的《目录交换格式》(DFI)手册中,是为了解决网络资源无序化而提出来的^[7]。目前国际上存在多种适用于不同领域、不同类型资源的元数据。其对元数据的研究主要集中在数字图书馆、数字博物馆、网络资源描述、学习资源描述、地理数据资源描述、音频视频资料描述等领域。都柏林核心(Dublin core, Dc)元数据主要用于资源的发现,机读目录格式(Marc)和美国联邦地理数据委员会的数字化地理元数据内容标准(FGDC/CSDGM)主要用于资源的著录,RSLP 集合描述(RSLPCD)、都柏林核心集合描述应用规范(DCCo)和数字资源集合访问 239.50 规范(Zcollection)等主要用于对资源集合的描述与组织等。

1.3 研究现状分析

从已有的研究现状来看,数据资源相关的标准主要有数据标准、编码标准、元数据标准、分类标准。元数据一般由标识信息、类型信息、管理信息、业务信息等 4 部分构成。数据资源目录

服务与数据交互技术平台相关的标准主要有通信协议、业务模式、业务模型、业务流程。目前的研究主要针对空间地理信息数据，抽象层次较高，与交通行业耦合度不够。虽然支持对非地理数据的描述，但是在描述航道要素信息资源的获取途径、分类体系和编码等方面显得不足，而且不能很好地描述航道要素数据，特别是未定关系数据库的结构、服务信息等元数据项。另外，我国内河特别是长江，涉及的航道要素众多且复杂，不能直接利用欧洲内河相关的数据标准规范。

2 长江航道要素数据资源的内容组织与分类编码研究

通过对长江航道数据资源的系统梳理，并从长江航道监测与测量数据采集的实际情况、航道维护管理业务决策和航道综合信息服务等方面综合考虑，确立长江航道要素数据资源的核心内容。

2.1 内容组织

从构建服务目录、便于检索应用的角度出发，长江航道要素数据可整合为航道信息、航道维护信息、航道整治建筑物信息、航道助航设施信息、航道通航建筑物信息、航道水文地形信息、航道气象信息等 7 个方面（表 1）。

表 1 整合后长江航道要素数据的内容组织	
整合理由	核心内容
便于人们了解航道概况	航道信息
便于管理部门和航道维护人员检索信息、业务决策	航道维护信息
便于航道系统治理及工程人员检索信息、工程决策	航道整治建筑物信息
便于港航企业、船舶船东检索信息、船舶装载及航行决策	航道助航设施信息
跨临河信息的整合	航道通航建筑物信息
与河流紧密相关信息的整合	航道水文地形信息
与气象紧密相关信息的整合	航道气象信息

2.2 分类编码

在长江航道要素数据的分类编码方面，针对长江航道要素数据的核心内容，划分为 7 个大类、38 个小类（表 2）。

表 2 长江航道要素数据的分类	
核心内容（大类）	小类名称及数量
航道信息	航道地理位置、航道基本信息、航道辅助信息、航道技术状况信息，共计 4 类
航道维护信息	航道水道信息、浅滩信息、碍航物信息、航道维护站点信息、航道计划维护信息、航道实际维护信息，共计 6 类
航道整治建筑物信息	护岸建筑物信息、护底建筑物信息、护滩建筑物信息、坝体信息，共计 4 类
航道助航设施信息	航标信息、航行基面信息、航行水尺及水位站信息、信号台信息、航道监控站点信息、航道雾情监测站点信息、航道测量控制点信息、航道助航设施维护站点信息、锚地信息，共计 9 类
航道通航建筑物信息	桥梁信息、隧道信息、船闸信息、升船机信息、过江（河）线缆信息，共计 5 类
航道水文地形信息	航道地形信息、航道水位信息、航道水深信息、航道流速流向信息、航道断面流量信息、航道断面含沙量信息，共计 6 类
航道气象信息	温度信息、风力风向信息、天气信息、能见度信息，共计 4 类

航道代码采用 HD，内河航道代码采用 HD01（兼顾其他非内河航道），内河航道要素采用 HD0101（兼顾内河航道其他方面），一级分类代码为 HD0101XX，二级分类代码为 HD0101XXXXX。

3 长江航道要素数据资源的元数据研究

基于对长江航道要素数据资源内容组织与分类编码的理解，开展长江航道要素数据资源的元数据研究。

3.1 内容框架

长江航道要素数据资源元数据应当是航道要素信息资源的全面描述，应紧密结合内河航道要素数据资源的特点，遵循全面性、模块化、可扩展性、多粒度性、兼容性的原则。鉴于长江航道要素数据资源属于地理空间信息的范畴，包括属性信息和空间信息两个方面，而地理空间信息的生产与更新维护需要国家测绘部门的批准，在生产与应用时均有特殊的精度要求、可靠性要求、时效性要求，且服务对象既有航道管理部门、海事部门、航运管理等行业内的用户，又有船公司、航行船舶等社会公众。因此，其内容框架可参考

水利地理空间信息元数据标准，应包括标识信息、限制信息、数据质量信息、数据维护信息、空间表示信息、参照系信息、内容信息、覆盖范围信息、分发信息等 9 个方面（表 3），图示表达类目信息可以参照内河电子航道图的相关标准。

表 3 长江航道要素数据资源元数据的内容框架

内容	内容说明
标识信息	描述唯一标识数据的基本信息，应包括有关资源的引用，数据集的摘要、目的、可信度、状况和联系方式等信息
限制信息	应包含对访问和使用资源所施加的限制内容
数据质量信息	数据质量评价信息，包括与数据生产有关的数据质量信息的一般说明
数据维护信息	应包括有关数据资源的更新频率及更新范围的信息
空间表示信息	数据集中用于表示空间信息的机制信息
参照系信息	包括数据集使用的空间和时间参照系的说明
内容信息	应提供数据内容特征的描述信息，并表明数据集所在的资源范围
覆盖范围信息	为描述有关实体的空间、时间覆盖范围信息，应包括地理覆盖范围、时间覆盖范围、垂向覆盖范围、要素覆盖范围的信息
分发信息	有关数据资源发布方的信息及用户获取数据资源的途径

3.2 表达

从长江航道要素数据资源元数据的内容框架可以看出，长江航道要素数据资源元数据具有一定的层次结构，各层级内部具有某一方面描述信息内容，层次之间又呈现出一定的关系，包括关联关系、聚合关系、组合关系、继承派生关系等。为了准确表达长江航道要素数据资源元数据的层次结构和各层次结构之间的关系，区别于传统采用 UML（unified modeling language，统一建模语言）表达方法，采用面向对象的分析方法对各层级元数据实体进行表达（图 1）。

3.3 核心元数据

长江航道要素数据资源核心元数据是表示唯一标识长江航道要素数据资源单元文档所需的基本的、最少数量的元数据元素和元数据实体，为长江航道要素数据资源元数据元素集的子集（表 4）。

表 4 长江航道要素数据资源的核心元数据

结构名称	元素名称	元数据项
实体信息	元数据文件标识符	MD_元数据. 文件标识符
	元数据标准名称	MD_元数据. 元数据标准名称
	元数据标准版本	MD_元数据. 元数据标准版本
	元数据采用的语种	MD_元数据. 语种
	元数据采用的字符集	MD_元数据. 字符集
标识信息	元数据创建日期	MD_元数据. 元数据创建日期
	数据集名称	MD_元数据 > MD_标识. 名称
	数据集创建日期	MD_元数据 > MD_标识. 创建日期
	数据集版本	MD_元数据 > MD_标识. 版本
限制信息	数据集负责单位	MD_元数据 > MD_标识. 负责单位. 单位名称
	访问限制	MD_元数据 > MD_限制. 法律限制. 访问限制
	使用限制	MD_元数据 > MD_限制. 法律限制. 使用限制
质量信息	比例尺	MD_元数据 > MD_质量. 比例尺分母
	内容完整性评价	MD_元数据 > MD_质量. 内容完整性评价. 说明
	逻辑完整性评价	MD_元数据 > MD_质量. 逻辑完整性评价. 说明
	空间完整性评价	MD_元数据 > MD_质量. 空间完整性评价. 说明
	时间完整性评价	MD_元数据 > MD_质量. 时间完整性评价. 说明
维护信息	生产进展	MD_元数据 > MD_维护. 数据生产进展
	更新频率	MD_元数据 > MD_维护. 数据更新进展
空间表示信息	空间表示	MD_元数据 > MD_空间表示. 矢量空间表示. 拓扑等级
空间参考信息	时间基准	MD_元数据 > MD_参考系. 时间基准
	高程基准	MD_元数据 > MD_参考系. 高程基准
	平面基准	MD_元数据 > MD_参考系. 平面基准
	投影方式	MD_元数据 > MD_参考系. 投影方式
内容信息	内容名称	MD_元数据 > MD_内容. 内容名称
	要素类型	MD_元数据 > MD_内容. 要素子类型
	数据格式	MD_元数据 > MD_内容. 数据格式子类型
范围信息	河流名称	MD_元数据 > MD_覆盖. 河流名称
	时间覆盖范围	MD_元数据 > MD_覆盖. 时间覆盖范围
	航道覆盖范围	MD_元数据 > MD_覆盖. 航道里程覆盖范围
分发信息	分发格式	MD_元数据 > MD_分发. 离线分发
	资源 URL	MD_元数据 > MD_分发. 在线分发. 链接地址 URL
	在线功能	MD_元数据 > MD_分发. 在线分发. 在线功能
	分发单位名称	MD_元数据 > MD_分发. 分发单位. 单位名称
	分发单位地址	MD_元数据 > MD_分发. 分发单位. 单位地址
	分发单位联系电话	MD_元数据 > MD_分发. 分发单位. 联系电话

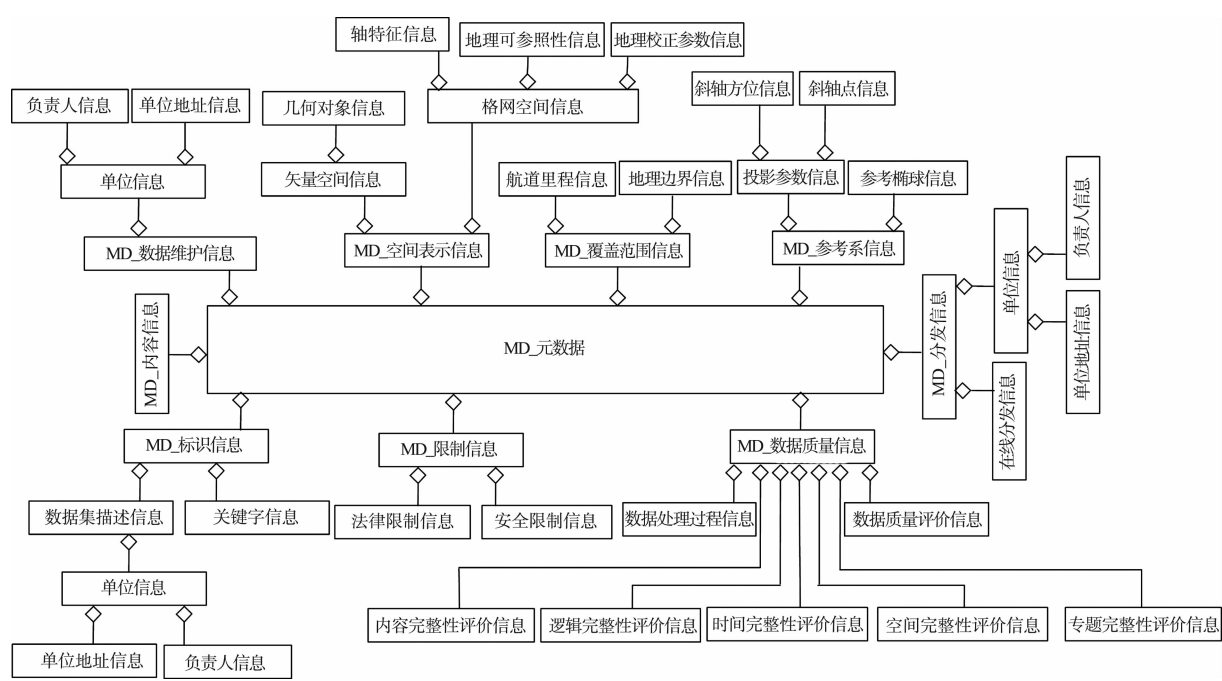


图 1 基于面向对象方法的长江航道要素数据资源元数据的表达

4 结语

- 1) 完成长江航道要素数据资源分类编码、给出元数据描述，是实现长江航道要素数据资源描述、检索、定位、服务的一项基础性工作，对长江航道要素数据资源整合和综合服务体系构建具有重要的指导意义；
- 2) 从构建服务目录、便于检索应用的角度，长江航道要素数据资源整合为 7 个一级分类、38 个二级分类，一级分类代码可采用 HD0101XX，二级分类代码可采用 HD0101XXXXX；
- 3) 按照全面性、模块化、可扩展性、多粒度性、兼容性的原则，长江航道要素数据资源元数据内容框架可划分为标识信息、限制信息、数据质量信息、数据维护信息、空间表示信息、参照系信息、内容信息、覆盖范围信息、分发信息等 9 个方面，其核心元数据由这 9 个方面的 36 个元数据项构成；
- 4) 随着长江数字航道工程建设的推进，在运行实践中也需要进一步检验长江航道要素数据资源

元数据的适应性和可扩展性。

参考文献：

[1] 全淑贞,段成民,黄莉莉,等. 行业标准《航道信息基础数据元》修订情况介绍[J]. 交通标准化, 2014(16): 8-11.

[2] 刘嘉. 元数据: 理念与应用[J]. 中国图书馆学报, 2001(5): 32-36.

[3] 赵军. 数据资源描述与组织的元数据方法[J]. 天津: 天津大学, 2005.

[4] 潘光辉. 基于元数据的政务信息资源目录体系应用研究[J]. 信息管理与信息学, 2009(10): 42-45.

[5] 冯钧,唐志贤,朱跃龙,等. 水利信息资源目录服务元数据定义研究[J]. 水利信息化, 2011(9): 19-22.

[6] 张峰,李四海,王伟. 基于元数据的渤海海洋信息资源目录服务系统[J]. 环境保护与循环经济, 2009(9): 31-34.

[7] NASA. Directory of Inter-change Format Manual [M]. USA: NSSDC/WDC-A-R&S, 1998: 88-89.

(本文编辑 武亚庆)