

# 从编目到元数据管理：图书馆知识组织的发展路径

贾君枝

**摘要** 以传统编目为核心的知识组织活动面临着诸多挑战,不能有效地支持智慧图书馆所开展的知识服务,如何适应当前知识服务的发展是图书馆需要深入思考的问题。本文梳理国内外文献,寻找编目与元数据之间的有机联系,从编目规则、编目系统、编目业务流程角度对编目环境发展变化进行客观分析,旨在对元数据管理发展的必要性进行探讨,在此基础上提出图书馆元数据管理的发展思路。基于 FAIR 原则,本文提出了元数据统一管理框架、语义化及结构化的知识表示及语义增强元数据的实现,以期充分利用语义网相关技术实现对元数据全生命周期的管理,推动图书馆元数据质量提升,深入开展一系列围绕元数据的服务以带动知识服务。图 2。表 1。参考文献 24。

**关键词** 图书馆 知识组织 元数据 编目

**分类号** G254

## From Cataloging to Metadata Management: The Development of Library Knowledge Organization

JIA Junzhi

### ABSTRACT

The development of smart libraries puts forward more in-depth requirements for knowledge services, which not only must adapt to the ubiquitous services of multiple application scenarios but also require closeness and accuracy of service content. As far as the current service situation is concerned, libraries are limited to the digital management and provision of collection resources, and lack the disclosure, reorganization and reproduction of knowledge units within the resources, and its business processes and system platforms are limited by traditional information organization methods. With the continuous growth of the number, scope and type of collection resources, it is unable to effectively meet users' demand for knowledge contained in the resources. How to adapt the knowledge organization activities with traditional cataloging as the core to the current development of knowledge services will be a problem that libraries need to consider deeply. This paper finds that the weakening of cataloging and the prominent role of metadata make it particularly necessary to re-examine the knowledge organization activities of the library.

This paper analyzes the development and change of the cataloging environment from three perspectives: cataloging rules, cataloging systems and cataloging processes. The change in cataloging rules has brought about changes in cataloging ideas. From closeness to openness, from isolated ones to interconnected ones,

通信作者:贾君枝,Email:Junzhij@163.com,ORCID:0000-0003-1486-673X( Correspondence should be addressed to JIA Junzhi,Email:Junzhij@163.com,ORCID:0000-0003-1486-673X)

and from resources descriptions to content descriptions in resources, these changes make cataloging not just a process of revealing and recording resource characteristics, but an important business process supporting the user-centered knowledge service. The function expansion of the cataloging system and its support for new cataloging rules have made the cataloging organization become the center for metadata management and control, which is responsible for the full life cycle management of metadata and a series of applications around metadata. These changes also put forward higher requirements for the optimization of cataloging process and the improvement of professional skill of catalogers.

Based on the principle of FAIR, this paper proposes the construction of a unified metadata management framework, semantic and structured knowledge representation and the realization of semantic enhanced metadata. The unified metadata management framework includes resource layer, business layer, system layer and service layer. The resource layer includes various forms of resource objects. The business layer aims to form a standardized process of metadata processing, and realize the management of different life cycle stages of metadata. The system layer builds a metadata repository, establishes the mapping of different metadata standards, and realizes the conversion between metadata records. The service layer provides various basic and value-added metadata services inside and outside the library. The semantic and structuring of metadata is embodied by refining the elements in the metadata standard, using the classes and attributes of ontology to describe a series of specifications for its elements, and fully defining the relationship types between elements to reflect the relationship between different data units, so as to facilitate data reprocessing, analysis and reuse. Semantic enhancement needs to rely on a wide range of internal and external resources to achieve multi-party cooperation with publishing houses, database vendors, museums, archives and other institutions. Through mapping, integration, association and other means, the content of metadata can be enriched continuously and achieve the purpose of expanding the internal metadata of the library. 2 figs. 1 tab. 24 refs.

#### KEY WORDS

Library. Knowledge organization. Metadata. Cataloging.

## 0 引言

图书馆从信息服务到知识服务已历经二十多年的发展,在学科、主题及个性化服务方面取得了长足的进步。智慧图书馆的发展对知识服务提出更深层次的要求,既要有适应多种应用情境的泛在化服务,还要求服务内容的贴身化及精准性。知识服务能力提升一方面依赖于馆员素质及现代化技术手段,另一方面取决于资源的加工深度及建设水平。就服务现状看,图书馆局限于馆藏资源的数字化管理及提供,缺乏对资源内部知识单元的揭示及重组再现,业

务流程及系统平台受限于传统的信息组织手段及方法,随着馆藏资源数量、范围及类型的不断增长,已无法有效地满足用户对文献中所包含知识的需求。

2009年发布的美国图书馆协会政策指南明确提出图书馆的核心能力之一是知识及信息的组织能力<sup>[1]</sup>。随着云计算、人工智能技术的发展及其在新一代图书馆系统中的应用,作为知识服务机构,图书馆如何有机地整合内外部资源,将开放存取资源、网络资源统一纳入馆藏资源管理中,实现与出版机构、数据库商、博物馆、档案馆等机构之间的资源合作与共享,形成自己独特的资源优势,对于提升知识服务规模及

水平具有重要价值,而这很大程度上取决于相应的知识组织理念、方法及手段。知识组织以知识的获取、表示、关联及整合为核心,通过应用一系列知识组织工具实现对知识的有序化管理,并提升知识处理效率。从满足用户知识需求角度而言,知识组织是图书馆提升服务的根本,随着与出版机构、数据库商所提供服务的日益趋同,知识组织也是图书馆获得未来竞争优势的基础,是满足智慧图书馆知识服务需求的必备条件。

以传统编目为核心的知识组织活动面临着诸多挑战,为更好地适应未来知识服务需求,图书馆需进一步考虑如何有效地开展知识组织以推动资源的深层开发与利用。本文通过梳理国内外文献,发现从编目到元数据管理这一变化的内在逻辑;对图书馆编目环境发展变化进行客观分析,明确其对知识组织所提出的要求,尤其对元数据管理发展的必要性进行探讨;在此基础上对图书馆元数据管理的核心内容进行阐

述,以期推动图书馆知识组织活动从传统编目到元数据管理的转换。

## 1 研究进展

随着图书馆编目活动逐步弱化, MARC、XML、MODS、DC 等不同元数据标准的引入,馆内外不同资源之间集成需求的提出,元数据在编目活动中的重要作用日渐显现。因此,许多学者围绕编目活动,从元数据角度探讨两者之间的关系与转换,以客观分析图书馆编目环境的变化及未来可能的应对。对 Web of Science 中 1980 年后出版的关于图书馆编目和元数据的研究文献进行调研可以发现,关于图书馆编目的研究一直处于较平稳的增长态势(见图 1 灰色线),1995 年以后研究文献中编目与元数据的共现频次增加(见图 1 绿色线),2001 年至今关于图书馆元数据的研究呈现出快速上升趋势(见图 1 蓝色线)。

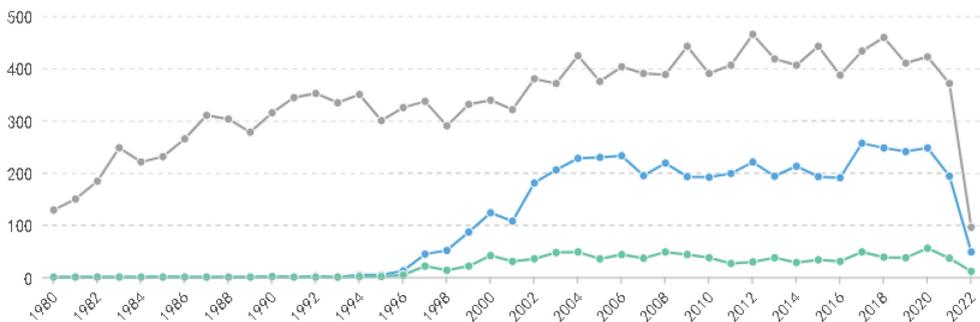


图 1 Web of Science 中关于图书馆编目与元数据的研究文献分布

自 1980 年以来,学者开始意识到编目业务的萎缩,编目自动化及大规模编目外包致使图书馆对专业编目馆员需求减少,从事编目的人员数量持续下降<sup>[2]</sup>。研究发现,近 5 到 10 年,编目馆员人数减少的图书馆占调研馆的 41.8%<sup>[3]</sup>。另外,Portia 等指出信息技术的变化致使编目馆员角色发生变化,他们需要更新编目知识及技能,而且这种改变影响到信息组织课程的教学内容<sup>[4]</sup>。康奈尔大学图书馆于 1997

年设立了元数据馆员,以体现更广泛的编目馆员范围,自此一些图书馆纷纷设置元数据馆员职位。Han 等分析了 86 个元数据馆员职位广告,发现与编目馆员职位相比,元数据职位需要传统编目知识,编目职位则将元数据标准知识(除 MARC 标准外)列为优先考虑条件<sup>[5]</sup>。可以看出,图书馆对编目馆员的职责要求也在发生变化。相关研究对 2016 年 8 月至 2018 年 8 月的招聘广告进行分析,发现关于图书馆编目馆

员及元数据馆员职位的广告有 192 个,其中需要编目馆员的岗位有 103 个,需要元数据馆员的岗位有 50 个,兼具元数据及编目的岗位有 39 个<sup>[6]</sup>,且二者之间的能力要求逐渐趋同。尽管图书馆对编目馆员的需求仍占主流,但是对编目馆员的元数据能力提升要求已见端倪。

与此同时,图书馆内部多个独立的系统以及异构的元数据催生了对图书馆系统的变革要求,元数据管理进一步得到重视。Bowen 认为如何让传统书目数据在下一代 Web 发现环境中发挥作用已成为图书馆所面临的挑战,把握技术愿景、了解用户需求、加深对元数据的理解并从中获取可能的价值成为应对挑战的必要条件<sup>[7]</sup>。以书目数据建设为核心的编目工作一度是图书情报机构的核心能力之一,然而正是固守传统编目业务使得图书情报机构当前面临元数据发展的困境<sup>[8]</sup>。美国国家信息标准协会(NISO)明确指出当前软件包大多数是元数据驱动的,图书馆、档案馆、博物馆等文化机构在创建、共享元数据方面具有悠久历史,可以帮助用户快速查找、定位所需资源<sup>[9]</sup>。图书馆下一代集成系统的开发,需涉及不同资源的元数据导入、规范、合并、共享等管理活动,用户通过检索界面才能检索到这些资源<sup>[10]</sup>,这对元数据管理提出了新的要求,即完整的生命周期管理、资源类型兼容、标准规范兼容、互联网环境下的书目与规范控制、知识组织与知识融合<sup>[11]</sup>。

元数据作为图书馆重要的资产,其战略性地位日渐突出。英国图书馆自 2015 年就意识到图书馆馆藏元数据的重要性,先后提出了主题为“释放价值”(2015—2018 年)<sup>[12]</sup>与“未来基础”(2019—2023 年)<sup>[13]</sup>的元数据发展战略,将元数据视为重要资产,并作为基础设施加以建设,以更好地提供访问、协作和开放式重用。美国联合计算机图书馆中心(OCLC)提出了下一代图书馆元数据转变的目标,旨在将 MARC 记录转变为编码良好、可共享、可链接的记录,以实现机器理解<sup>[14]</sup>。随着数据科学发展,数据增

值的实现不仅依赖于数据,更大程度上取决于数据的正确描述,即元数据<sup>[15]</sup>。

从国内外研究可以看出,编目业务的弱化,元数据作用的突显,使得重新审视图书馆知识组织活动尤为必要,如何通过元数据管理带动知识组织发展成为图书馆适应未来发展需要亟待解决的问题。

## 2 图书馆编目环境变化对知识组织发展的要求

与知识组织所处理的知识单元相比,信息组织局限于以“文献单元”为处理对象,主要围绕文献资源的数字化开展,属于知识组织的早期阶段。当前,图书馆实际上更多地停留在信息组织阶段。与信息组织相关联的主要业务活动是编目,即将采访到的馆藏资源按照一定的著录规则进行著录及标引后存入资源库,供用户检索使用。资源处理早期以实体资源为主,后发展到电子资源。随着编目业务进一步外包,信息组织活动在图书馆机构内部逐渐弱化,但是随着图书馆对知识服务需求的日益增加,图书馆信息组织活动迫切需要转向知识组织。图书馆编目环境的变化,不仅使编目理念发生了改变,而且使传统的信息组织手段及方法面临着极大的挑战。编目规则用于规范编目活动,编目系统是支持编目活动开展的技术工具,编目业务流程用以组织人员开展编目活动,这三者共同作用形成了图书馆内部的编目环境,因此本文从这三方面分析图书馆信息组织发展变革的必要性。

### 2.1 编目规则

编目规则实际是元数据标准,规定了基本的描述项目及编目活动应遵循的原则,用以指导编目人员,保证生成的元数据记录的标准化与规范化,推动图书馆机构之间的数据交换。编目规则随着资源及用户检索需求的变化而不断更新发展,从 ISBD、AACR2 到 2010 年 RDA 的

提出,国际编目规则发生了重要变化,我国于2021年也相应制订了《资源描述》标准以适应发展需求。

传统的编目规则服务于早期的手工编目,以单个资源的外在特征为识别对象,注重检索点的选取。新编目规则 RDA 重在资源描述而不是显示<sup>[16]</sup>,描述对象的确定及其属性特征的选取取决于用户的关注点,尤其对文献单元中知识内容及不同描述对象之间关系的揭示给予重点考虑,旨在使图书馆内外部数据不再孤立,以数据网络的形式建立链接,以适应未来资源整合需要。RDA 是实施知识组织的关键,应用 URI、本体、RDF、关联数据等语义网技术,实现对资源内容的揭示与标引,由此带来编目系统及业务流程的变革,比如实体资源和电子资源合并描述,描述内容、描述单元间的关系及描述格式的变化需要在新系统中得到体现。另外,RDA 并非是图书馆内部的规则,也适用于其他领域的资源描述,其应用范围的扩大,一定程度上推动着图书馆的开放性,使图书馆数据融入网络环境中,在更大范围内被用户使用。

RDA 编目规则基于实体关系模型<sup>[17]</sup>,从作品层、内容表达层、载体表现层、单件层角度分别对资源进行描述,建立了不同层次之间的语义关系,使描述的颗粒度由粗变细,描述的关系由扁平化向网络状发展。各个机构配合 RDA 规则发布了相应的描述词表,以实现资源的规范化描述,使规则具有可操作性。RDA 词表及美国国会图书馆发布的 BIBFRAME 词表定义了丰富的类及属性集合,有助于揭示不同资源的各种特征,也便于图书馆及其他机构根据资源特点选用,适用于智慧图书馆环境下全媒体资源的描述。

## 2.2 编目系统

编目系统是图书馆业务人员开展编目活动所使用的信息系统。编目系统的功能随着编目业务活动自动化要求的深化而不断地拓展,从

单纯的书刊著录发展到集采购、编目、典藏、流通一体化的集成系统。由于电子资源的编目不能有机地纳入集成系统中,因而需要有额外的 ERM 系统、数字资产管理系统等提供支持。此外,不同类型资源编目活动需要不同的编目系统,多个系统林立导致管理维护成本提高,数据分散式管理使得系统间数据交换难度增加,阻碍了编目流程的优化及集成服务的开展。2011 年以来新一代图书馆服务平台的出现,贯通了采购、编目、流通、电子资源管理、链接解析、资源发现等全业务流程,推动了编目系统及编目业务流程的新变革。与图书馆集成系统相比,新的图书馆系统支持用户对跨媒体资源内容的发现、识别、选择、获取与导航,体现了开放性、共享性、互操作性等优势<sup>[18]</sup>。

编目系统在新一代图书馆服务平台中已转变为元数据编辑系统,通过对元数据的查询、创建、更新、服务等全生命周期的管理,形成元数据仓储库,以实现元数据的统一管理。其支持的功能包括:①支持实体资源和电子资源统一编目,容纳包括 MARC、DC、RDA 等不同类型的元数据编目标准;②建立元数据与数字对象之间的链接<sup>[19]</sup>,以实现电子资源的全文获取;③建立不同元数据标准的映射,提供不同元数据记录之间的转换功能;④提供数据清洗、质量校验功能,确保元数据的准确性、规范性、一致性;⑤以 MARC、XML、RDF 多种格式展现编目生成的元数据记录,以满足不同用户的需求;⑥支持元数据收割,基于标准协议或者标准接口能从出版社、数据库商、图书馆联盟等机构获取不同来源的数据,如电子资源使用统计信息 COUNTER 报告<sup>[20]</sup>;⑦提供开放系统接口 API,允许元数据导出,支持元数据重用,提升元数据价值;⑧支持关联数据的发布,建立与外部资源的链接,提供对外元数据服务。

## 2.3 编目业务流程

为适应编目外包需求,图书馆形成了采编藏一体化体系。编目业务流程实际是对不同类

型资源特征的描述及揭示以生成元数据记录的过程。编目流程标准化及规范化水平不仅制约编目活动的开展水平,还会影响后期数据再加工及利用水平。当前编目业务规范以规定实体资源为主,对于电子资源、开放资源、自建资源等缺乏详细而明确的规定<sup>[21]</sup>,对于外包的编目数据缺乏明确的质量控制程序及有效手段,这些在很大程度上制约着编目数据的质量及其管理水平。

随着下一代图书馆服务平台的应用,实体与电子资源的合并编目、元数据统一管理,从数据层面实现了资源集成与共享,从而对现有编目的内涵及业务流程提出了变革要求,需要构建合适的业务流、组织结构、规章制度以及岗位职责等业务框架。编目作为元数据编辑过程,仅仅是业务流程的一个环节,由编到管——由编目到元数据管理是当前业务流程转变的关键。需要建立与系统平台相适应的业务流程,整合原有按照资源类型分割的编目活动,将编目人员纳入元数据管理团队,负责全馆内外元数据管理。元数据管理人员的责任是:①规范编目流程,熟悉不同元数据标准,尤其是新型编目规则;②熟悉元数据表示的各种形式化语言,实现不同语言格式的数据转换;③明确在服务平台中对应的角色分工及操作规范;④管理控制元数据质量;⑤对生成的元数据记录进行有效管理(包括导入、导出、收割、映射、集成)等。

可以看出,编目规则的变化带来了编目理念的变化。从封闭到开放,从孤立到互联,从对文献的描述到对文献内容的描述,编目已不单纯是资源特征揭示及记录过程,更成为支持以用户为中心的知识服务理念的重要业务流程。编目系统的功能扩展及其对新的编目规则的支持,已使编目机构成为元数据管理及控制中心,负责管理元数据全生命周期及围绕元数据所开展的系列应用。这些变化同时对编目流程的优化及编目人员专业素质水平的提升提出了更高要求。

### 3 图书馆元数据管理的发展思路

图书馆元数据不仅包括图书馆自身创建的数据,还包括从其他机构中获取的数据。与出版社、数据库商等其他机构建立广泛合作,不仅可以使图书馆充分利用外部元数据以丰富图书馆资源,而且可以将图书馆元数据广泛应用于互联网环境中,提升图书馆元数据的开放度、能见度及影响力,推动元数据的深层次开发利用。FORCE11 社区提出的 FAIR 原则,即可发现 (Findable)、可访问 (Accessible)、可互操作 (Interoperable)、可重用 (Reusable) 四大原则,已成为实现数据科学管理的重要依据。图书馆在遵循这些原则的基础上,可实现对图书馆元数据的统一管理,以提供高质量的元数据,保证元数据开放及共享,提升元数据服务效能。

#### 3.1 建立元数据统一管理框架

元数据统一管理框架包括资源层、业务层、系统层及服务层,如图 2 所示。资源层是元数据处理的对象,包括多种形式的资源对象,如图书馆需自建的馆藏实体资源、电子资源等的元数据资源,出版社、集成商等第三方机构的元数据资源。业务层形成元数据处理的标准化流程,实现对元数据全生命周期的管理。系统层构建元数据仓储库,建立不同元数据标准的映射,实现元数据记录之间的转换。服务层向图书馆内外部提供各种元数据基本服务及增值服务。各个层次之间彼此关联,共同作用,支持图书馆用户更好地获取及发现资源。

##### (1) 业务层

业务层是指将图书馆多元化资源对象纳入统一的元数据处理流程中,从元数据导入、编辑、质量检验到导出等系列环节,实现对元数据的管理。导入是将来自于图书馆内外部已有的元数据导入到系统中等待处理,这一环节能够接收不同格式的元数据,并将其转化成待处理的标准化格式。编辑是基于不同的元数据标准,

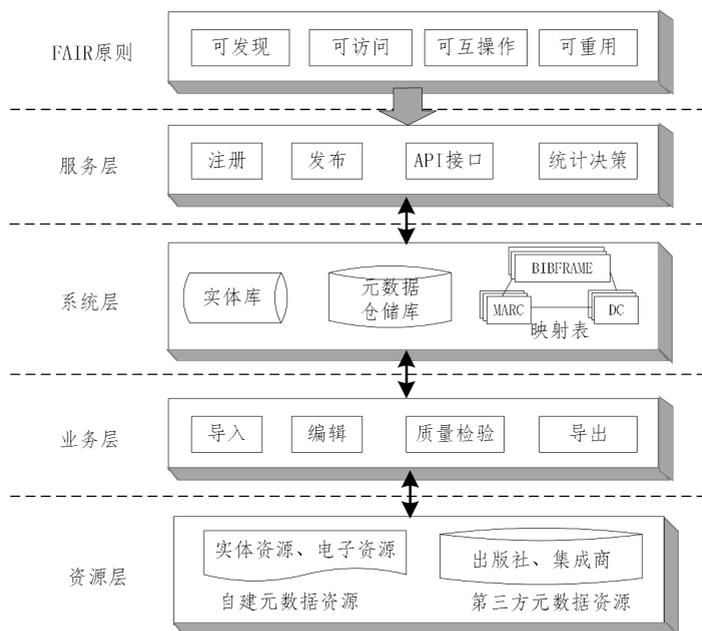


图2 元数据统一管理框架

对资源对象进行描述生成元数据记录,并进行增删改等操作,建立元数据记录与名称规范档关联,通常采用计算机辅助人工标引实现,一般包括新记录的生成、删除、修改、合并、链接、更新等。质量检验环节要对编辑好的数据进行质量评估,提供计算机自动检验及人工辅助功能,并将达到一定质量水平的元数据存入到系统。导出是将系统中的元数据按照一定的格式输出,供显示、打印或下载,提供人机可读的多种格式。

### (2) 系统层

图书馆资源的多元化导致其采用的元数据标准不一,而使用不同的元数据又会使图书馆内部资源整合工作具有难度。为提高资源之间的互操作性,需要建立统一元数据仓储库。元数据仓储库应设计统一的数据模型,定义其中的类及关系,在图书馆所使用的各个元数据标准的基础上,尽可能抽取出共同元素,基于关联性及重用性原则,形成具有灵活性、表现性的语义框架。元数据仓储库服务于元数据统一管理,用于元数据记录的统一存储。映射表在BIBFRAME、MARC、DC等不同的元数据标准与

数据模型之间建立对照关系,通过映射转换模块,将不同格式的元数据记录转换成统一格式。实体库是对分类表、主题词表、名称规范档等受控词表采用实体管理方式,使用URI命名,建立与图书馆外部词表的链接,进行统一的创建、发布和维护,不仅对资源取值进行规范性描述,还可以用来对不同来源的元数据进行合并且建立关联,以便对这些元数据进行调用及集成应用。

### (3) 服务层

图书馆元数据的服务对象包括馆员、读者、出版机构及数据提供商等数据使用者。不断积累的高质量元数据,逐步成为图书馆重要的数据资产,以元数据为中心开展各类对内对外服务,有助于提升图书馆外部影响力。服务层包括元数据注册、发布、提供API接口、统计决策等。元数据注册通过注册平台对图书馆资源描述所使用的元数据标准、受控词表及本体实现统一管理及有效存取。该平台支持元数据的注册、上传、编辑、维护、内容校验、格式转换、发布功能。注册内容包括元数据模式的元数据信息、各个元素的定义、取值、元素间关系、版本信

息、表示语言等。元数据发布是将元数据记录及受控词表转换成关联数据形式,建立与外部资源的链接,实现数据的开放关联,将图书馆资源融入 web 服务中,通过 API 或 SPARQL 端口发布以被图书馆内外部机构共享及再利用。提供多个 API 接口,允许应用程序调用,提供元数据查询服务,方便图书馆内外其他平台可以无缝地整合图书馆元数据内容。统计决策能够实现对不同资源类型元数据的多维度展示,包括与用户使用数据相结合的资源利用统计,提供可视化馆藏元数据展示,并在此基础上形成馆藏评估、资源利用及采购决策分析报告。

### 3.2 元数据的语义化及结构化表示

知识服务是围绕资源开展的,需将不同类型的资源及其知识内容整合成各种数据形式以便于将其嵌入服务之中,被不同服务所调用。资源整合及其服务调用依赖于知识的表示方式。将文献单元划分为更细粒度的知识单元,并且建立不同描述单元之间的联系,通过不同内容单元的组合及其关联,才能更好地适应未来知识服务的需求。知识表示主要体现在资源描述及标引过程。资源描述的维度、描述所形成的数据之间的关联都将作为图书馆知识组织所面临的关键问题,而资源描述依赖于描述的元数据标准。元数据标准的语义化及结构化是实现知识表示的关键。元数据语义化体现为细化元数据标准中的元素,使用本体的类与属性,对其元素进行一系列规范说明,明确取值类型,限定取值来源于一定的概念体系,通过与概念链接,使数据拥有语义。元数据结构化体现为各个元素之间不再是线性排列,而是通过充分定义元素之间的关系类型,以体现不同数据单元之间的关系,便于对数据进行再加工分析及重用。元数据包含描述、管理、结构三类,从元数据类型看,语义化更多地体现在描述性元数据层面,结构化体现在结构性元数据层面。

描述性元数据揭示描述对象的内外属性特征。描述对象从文献对象拓展到文献中的内

容信息,抽象出多种实体对象,如作品、代理、事件、机构、时间、空间等,并定义相应的属性,使资源描述颗粒度逐步细化。为保证不同颗粒度下资源描述的标准化及规范化,通常采用多个不同类型的元数据标准(也可称之为词表)共同加以描述,包括定义文献层特征的属性词表如 MARC、DC,内容层的属性词表如 RDA、BIBFRAME、CIDOC-CRM,属性值词表如中国分类主题词表、DDC、LCSH、SKOS 等。当前元数据标准为适应此需要,进一步将元素细化为丰富多样的类或属性,DC 词表从 15 个核心元素发展为包含 22 个类、55 个属性、12 个数据类型和 21 个编码体系的集合<sup>[22]</sup>;RDA 词表包含了 13 个类、3 013 个属性<sup>[17]</sup>;BIBFRAME 包含了 186 个类、195 个属性<sup>[23]</sup>。

结构化元数据揭示不同元数据之间的联系,是元数据关联化的重要体现,其价值在于以元数据的关联建立所描述资源之间的联系,形成有机的数据网络,便于实现资源的聚合、导航、推荐及发现。结构化元数据的增加意味着所定义的关系类型增多,呈现多样化特征,表现为不同实体对象之间的关联,包含资源层内部、资源层与内容层、内容层内部的关系。资源层的关联重在揭示图书馆内外部不同资源对象之间如单件与实例、单件与单件、实例与实例的关系。资源层与内容层的关系体现在载体形态与作品、单件与作品、单件与代理等的关系上。内容层关系丰富,不同内容对象之间包含作品与作品、作品与代理、作品与时空等多样化关系。除此之外,丰富的数据关系打破了数据孤立状态,展现了数据的多维度、多层次特性,可借助数据挖掘方法发现新的知识以服务于用户。关系类型需借助于词表中的对象属性定义其语义关系,不同词表所定义的结构化元数据侧重点不同。DC 没有明确定义对象属性,仅定义了元素的值域范围,元素值域来自所定义类的属性,共有 24 个,其中资源与代理之间的关系占到 7 个。RDA 对象属性有 359 个,文献层与内容层的关系类型数量分别是:代理(包含个人)与单

件 31 个、代理(包含个人)与内容表达 120 个、代理(包含个人)与载体表现 60 个;内容层之间的关系类型数量分布是:代理(包含个人)与作品 123 个,代理(包含个人)与代理(包含个人) 13 个,代理与地点 6 个,代理与时间 6 个。BIBFRAME 对象属性有 142 个,但实际上具有明确语义类型的属性仅 42 个,其中作品与作品、事件等内容层属性有 12 个。RDA 与 BIBFRAME 词表内容层的类及对象属性分布情况如表 1 所

示,其中类包含了作品、事件、主题、评论、代理、角色、类号、摘要、目录等,可以实现对资源对象内容的充分表达。对象属性中最核心的关系是代理与作品的关系,尤其在 RDA 表现突出。通过资源描述充分揭示这些关系,不仅为实现资源间的关联提供了基础,并且可以在已有数据的基础上,通过数据整合建立不同数据对象之间的关系。

表 1 内容层的类及对象属性

类型	RDA	BIBFRAME
内容层的类	Work、Agent、Place、Person、Family、Timespan、Corporatebody、CollectiveAgent	Person、Work、Agent、Place、Event、Review、Summary、Role、Classification、Temporal、Note、Topic、Title、Notation、Content、Polarity、Ensemble TableOf-Contents、Scale
内容层的对象属性	is film director agent of、is radio director agent of、is television director agent of、is issuing agent of、is dedicator agent of、is honouree agent of work of、is programmer agent of、is designer agent of、is cartographer agent of、of work of、is composer agent of work of 等(代理—作品关系类型共 123 个) has related agent of person、has colleague、has assistant、is assistant to、has coworker、has partner、has fellow student、has friend、has student、has teacher、has related person of person、has place of residence、has place of death has place of birth、has country associated with person、has related place of person、has period of activity of person、has date of death、has date of birth、has related timespan of person	arrangementOf、content、role、expressionOf、eventContent、ensemble、instrument、music-Medium、voice

### 3.3 元数据语义增强功能实现

语义增强是指采用各种技术和方法向内容添加语义元数据以有效增强数据价值的策略,可达到提高元数据可发现、可重用、可用性目的<sup>[24]</sup>。图书馆资源描述能力有限,生成的元数据表达的语义内容有限,因此需要借助广泛的内外部资源,实现与出版社、数据库商、博物馆、档案馆等机构的多方合作,通过映射、整合、关联等手段不断丰富元数据内容,以提供更多的情境信息。语义增强实现需借助以下手段:①建立以实体(主题、人、机构、地点)为中心的资源描述方式,运用

URI 对其进行唯一标识,通过构建实体之间的关系以明晰实体的语义含义,便于对实体进行浏览、搜索及发现;②充分运用名称规范档、受控词表、本体等多种类型的知识组织系统,对实体的取值进行限定,以提升资源描述的规范性,同时对地理名称、人名、机构名、主题名进行精确化语义增强,以实现多语种、多来源、多维度的资源展示;③对图书馆实体与外部资源实体之间进行对齐、映射、链接,将图书馆外部多种资源作为语境化信息链接到图书馆元数据,使图书馆资源不再是数据孤岛,有机地与外部资源建立联系,以便

用户可以访问更丰富的资源;④抽取外部资源中有价值的信息,作为实体值或者实体关系添加到资源描述中,扩充图书馆内部元数据,进一步丰富其内容。

#### 4 结论与展望

随着知识组织理论及方法体系的逐步完善,图书馆知识服务需求日渐提升,图书馆传统的编目业务流程面临着极大的挑战。从编目到元数据管理,是适应编目环境变化、顺应当前知

识组织发展趋势的必经途径。统一元数据管理框架的构建,语义化、结构化的知识表示及语义增强元数据的实现,旨在充分利用语义网相关技术推动图书馆元数据质量提升,形成具有竞争力的图书馆元数据资产,助力图书馆深入开展一系列围绕元数据的服务,带动知识服务能力的提升。元数据管理作为数据管理的重要组成部分,具有一定的复杂性,就当前我国图书馆发展来看,元数据管理面临着诸多困境,未来研究需对此进行系统梳理,以针对性地提出适合我国图书馆发展的对策。

致谢:本文系中国人民大学科学研究基金项目“数据驱动下的知识组织发展研究”(编号:202230018)的研究成果。

#### 参考文献

- [ 1 ] ALA. B. 1 core values, ethics, and core competencies[ EB/OL]. [ 2020-10-24]. <https://www.ala.org/about-ala/governance/policymanual/updatedpolicymanual/section2/40corevalues>.
- [ 2 ] MARIE B. Automation of cataloguing: effects on use of staff, efficiency, and service to patrons[ J]. *Journal of Academic Librarianship*, 1988, 14( 3 ): 145-149.
- [ 3 ] KENNETH F. The impact of automation on professional catalogers[ J]. *Information Technology and Libraries*, 1990, 9( 3 ): 242-251.
- [ 4 ] PORTIA B-C, HOSEIN Y. Map cataloguing training at the University of the West Indies, St. Augustine[ J]. *Library Review*, 2014, 58( 1-2 ): 97-108.
- [ 5 ] HAN M-J, HSWE P. The evolving role of the metadata librarian[ J]. *Library Resources & Technical Services*, 2010, 54( 3 ): 129-141.
- [ 6 ] TURNER R. Analyzing cataloging job descriptions: are cataloging jobs disappearing, changing, or merging[ J]. *Cataloging & Classification Quarterly*, 2020, 58( 6 ): 1-12.
- [ 7 ] BOWEN J. Metadata to support next-generation library resource discovery: lessons from the eXtensible Catalog, phase 1[ J]. *Information Technology & Libraries*, 2008, 27( 2 ): 5-19.
- [ 8 ] 丁迺劲, 曾建勋. 文献元数据集管理研究[ J]. *情报学报*, 2019, 38( 6 ): 568-577. ( DING Q J, ZENG J X. Literature metadata integration[ J]. *Journal of the China Society for Scientific and Technical Information*, 2019, 38( 6 ): 568-577. )
- [ 9 ] RILEY J. Understanding metadata[ EB/OL]. [ 2020-11-24]. [https://groups.niso.org/apps/group\\_public/download.php/17446/understanding%20metadata.pdf](https://groups.niso.org/apps/group_public/download.php/17446/understanding%20metadata.pdf).
- [ 10 ] 白雪, 李广利, 牛爱菊. Alma 环境下电子资源元数据的管理及对用户的影响——以北京师范大学图书馆为例[ J]. *新世纪图书馆*, 2019( 5 ): 49-52. ( BAI X, LI G L, NIU A J. Electronic resources' metadata management in Alma and its impact on users: taking Beijing Normal University Library for example[ J]. *New Century Library*, 2019( 5 ): 49-52. )

- [11] 许磊,夏翠娟. 第三代图书馆服务平台的元数据管理——以 FOLIO 的 Codex 方案为例[J]. 中国图书馆学报,2020,46(1):99-113. (XU L, XIA C J. Metadata management of the third generation library service platform;a case study of FOLIO Codex[J]. Journal of Library Science in China,2020,46(1):99-113.)
- [12] British Library. Unlocking the value; the British Library's collection metadata strategy 2015-2018 [EB/OL]. [2020-11-24]. <https://www.bl.uk/bibliographic/pdfs/british-library-collection-metadata-strategy-2015-2018.pdf>.
- [13] British Library. Foundations for the future; the British Library's collection metadata strategy 2019-2023 [EB/OL]. [2020-11-24]. <https://www.bl.uk/bibliographic/pdfs/british-library-collection-metadata-strategy-2019-2023.pdf>.
- [14] KAREN S-Y. Transitioning to the next generation of metadata [EB/OL]. [2020-11-24]. <https://doi.org/10.25333/rqgd-b343>.
- [15] GREENBERG J. Big metadata, smart metadata, and metadata capital; toward greater synergy between data science and metadata [J]. Journal of Data and Information Science, 2017, 2(3):19-36.
- [16] 吴贝贝,宋文. 从 MARC 走向 Bibframe——后 MARC 时代的书目记录 [J]. 图书情报工作, 2014, 58(9):85-90. (WU B B, SONG W. From MARC to Bibframe; bibliographic descriptions of post-MARC era [J]. Library and Information Service, 2014, 58(9):85-90.)
- [17] RDA element sets; RDA entity properties [EB/OL]. [2020-11-24]. <https://www.rdaregistry.info/Elements/x/>.
- [18] 许磊. 图书馆系统演变及其元数据管理 [J]. 图书馆论坛, 2021, 41(10):118-126. (XU L. A brief history of library systems and their metadata management [J]. Library Tribune, 2021, 41(10):118-126.)
- [19] 窦天芳,杨慧. 清华大学图书馆一体化资源管理平台建设——以 ALMA 系统实施为例 [J]. 数字图书馆论坛, 2020(5):2-7. (DOU T F, YANG H. Unified resource management platform construction of Tsinghua University Library; taking the implementation of ALAM as an example [J]. Digital Library Forum, 2020(5):2-7.)
- [20] 陈大庆,叶兰,丁培. 电子资源使用统计收割标准 SUSHI 的实施与应用研究 [J]. 中国图书馆学报, 2018, 44(2):46-60. (CHEN D Q, YE L, DING P. Implementing the standardized usage statistics harvesting initiative (SUSHI) protocol [J]. Journal of Library Science in China, 2018, 44(2):46-60.)
- [21] 李园贞,冯彩芬,郑建瑜. 大学图书馆编目业务外包的实践和思考——以中山大学图书馆为例 [J]. 大学图书馆学报, 2010, 28(3):71-74, 84. (LI Y Z, FENG C F, ZHENG J Y. Cataloging outsourcing; case study of Sun Yat-sen University Libraries [J]. Journal of Academic Libraries, 2010, 28(3):71-74, 84.)
- [22] DCMI Usage Board. DCMI metadata terms [EB/OL]. (2020-01-20) [2020-11-24]. <https://www.dublincore.org/>.
- [23] Bibliographic framework initiative [EB/OL]. [2020-11-24]. <https://www.loc.gov/bibframe>.
- [24] 曾蕾,谭旭. 数据的语义增强——解读图档博支持数字人文的新动向 [J]. 数字人文研究, 2021, 1(1):65-86. (ZENG L, TAN X. Semantic enrichment of data-interpreting the new trend of LAM data in supporting digital humanities [J]. Digital Humanities Research, 2021, 1(1):65-86.)

贾君枝 中国人民大学信息资源管理学院教授。北京 100872。

(收稿日期:2022-04-11;修回日期:2022-05-21)