

GLAM领域资源描述参考模型及应用场景研究*

陈人语
(首都图书馆, 北京 100021)

摘要: 规范描述、开放共享人类文化遗产信息资源是语义网时代GLAM的重要使命。资源描述参考模型是信息资源语义化、关联化的底层逻辑。面对GLAM领域不同应用场景的众多模型, 研究典型模型的内容与特点, 对人类文化遗产信息资源的知识融合、本体构建与映射、知识发现与服务创新等有重要意义。本文通过整理研究CIDOC CRM、RiC-CM、IFLA LRM、EDM和CDWA等5种模型的类别、属性、关系, 对比分析各模型语义揭示的应用过程, 总结资源描述应用场景的适用特点, 挖掘资源描述参考模型及应用场景研究的应用价值, 为文化遗产领域知识组织本体构建提供参考。

关键词: 资源描述; 参考模型; GLAM; CIDOC CRM; RiC-CM; IFLA LRM; EDM; CDWA

中图分类号: G254 **DOI:** 10.3772/j.issn.1673-2286.2023.01.001

引文格式: 陈人语. GLAM领域资源描述参考模型及应用场景研究[J]. 数字图书馆论坛, 2023 (1) : 1-9.

美术馆 (Galleries)、图书馆 (Libraries)、档案馆 (Archives)、博物馆 (Museums) 是保存人类文明及文化遗产的重要组织机构, 通常使用英文缩略词GLAM来表示这些机构的集合。组织揭示、开发利用馆藏资源, 传承、发扬人类文明成果, 一直以来是GLAM领域各机构共同承担的社会职能。随着信息技术的发展和语义网的诞生, GLAM馆藏资源的组织揭示逐步向语义化、关联化发展。资源描述参考模型是资源揭示和语义关联的底层逻辑, 是资源描述本体、核心词汇及语义关系的基础。为更好地揭示和组织各种信息资源, 针对不同应用场景的资源描述参考模型相继产生。CIDOC CRM (Conceptual Reference Model)、RiC-CM (Records in Contexts-Conceptual Model)、IFLA LRM (Library Reference Model)、EDM (Europeana Data Model)、CDWA (Categories for the Description of Works of Art) 是欧美等国家的博物馆、档案馆、图书馆

等机构广泛应用的资源描述参考模型, 对其进行分析研究, 对我国文化遗产保护机构信息资源的知识挖掘、数据融合、数字人文发现等领域的研究具有重要意义。

1 GLAM资源描述参考模型及相关研究

资源描述参考模型是资源描述过程中通过提取所描述资源中隐性和显性的核心概念而建立的、揭示概念间属性和关系的模型。这种参考模型可以理解为资源描述的形式本体, 包括类目、属性和关系。在语义网背景下, 参考模型是资源描述的底层框架。通过参考模型能够对所描述资源进行细粒度知识组织, 集成不同来源的信息, 揭示资源之间的关系, 建立深度广泛的语义关联, 实现资源的知识融合与发现。

1996年, 国际文献工作委员会 (International Committee for Documentation, CIDOC) 推出了促进多样化文化

收稿日期: 2022-12-17

* 本研究得到国家社会科学基金后期资助暨优秀博士学位论文项目“我国公共文化数字资源融合的标准体系建设研究” (编号: 20FTQA003) 资助。

遗产信息源的集成、转接和相互交换的概念参考模型CIDOC CRM。这一参考模型的诞生虽然不是最早的,但其属性关系的复杂程度是最高的,并且衍生开发了与之兼容的其他模型,包括FRBRoo、PRESSoo、CRMinf、CRMarchaeo、CRMsci、CRMgeo、CRMdig、CRMba、CRMtex、CRMsoc、CRMact。这些模型涵盖图书目录、连续出版物目录、生物多样性、地质学、地理学、考古学、文化遗产保护等各个领域^[1],CIDOC CRM成为GLAM领域具备兼容性且涵盖领域最广的一种资源描述参考模型。由于文化遗产信息资源具特点,不同的机构组织开发了适用于各自场景的资源描述参考模型。例如,国际档案理事会开发了档案记录参考模型RiC-CM、国际图联开发了图书馆参考模型IFLA LRM、欧洲数字图书馆(Europeana)开发了文化遗产参考模型EDM、Getty博物馆开发了艺术作品参考模型CDWA等。这些参考模型涵盖了GLAM领域各种资源类型,突出了不同资源的内涵和特点。

我国对资源描述概念模型的研究主要集中在单一模型介绍、模型应用、本体构建、元数据标准、词表映射等方面。李文^[2]、蔡丹等^[3]分别从档案资源和图书馆资源的角度,对RiC-CM和IFLA LRM模型进行了介绍;翟琼等^[4]一直关注CIDOC CRM模型在数字博物馆领域语义框架、语义检索等方面的应用问题;吴琼等^[5-6]、王艺茹等^[7]在CIDOC CRM模型的基础上,探讨了不可移动文物信息参考模型、建筑领域非遗知识本体等的构建问题;董焱^[8]、贾君枝等^[9]、张俊娥等^[10]从博物馆元数据标准的角度,对相关模型进行研究;陈艳等^[11]进行了CIDOC CRM和DC元数据的映射研究。对概念模型的综合研究不多,主要有范炜等^[12]、左丹等^[13]、黄永欣^[14]对文化遗产数据资源关联集成模型、人文信息资源语义描述、文化遗产资讯领域的参考模型进行的介绍和研究,涉及CIDOC CRM、CDWA、EDM等模型。国外对概念模型的研究主要集中在模型的应用和模型的映射上。Mäkelä等^[15]研究了CIDOC CRM在第一次世界大战关联数据中的应用;Bountouri等^[16]对CIDOC CRM和编码档案描述(EAD)进行了映射研究。

由于知识组织方法不断向细粒度、语义化深化,资源描述和揭示已经突破行业、介质、载体等界限,形成以资源中蕴含的知识关联为主的趋势。在知识关联方便、机器可读等方面,资源描述参考模型起到了重要作用。但是,资源描述参考模型的发展需要一个过程。在早期阶段,不同行业根据各自需求设计各种模型并进

行应用,这给语义网背景下的资源融合带来了阻碍。因此,研究资源描述参考模型,了解模型的实体、属性、关系,有利于通过映射建立不同领域资源描述的关联,实现知识融合与发现。同时,研究不同模型的适用特点和应用场景,有利于在知识组织、知识挖掘过程中,借鉴不同模型的本体,构建适合特定领域应用场景的自定义参考模型,实现知识服务创新。目前,很多领域在知识组织系统建设中,因为对资源描述模型缺乏了解,导致本体构建不科学、不规范。笔者根据模型研发机构和世界各大主流文化遗产网站语义模型使用情况,选择国际上GLAM领域影响力比较广泛的5种资源描述参考模型(CIDOC CRM、RiC-CM、IFLA LRM、EDM和CDWA)进行比较研究,希望通过对参考模型及应用场景的研究为领域知识组织建设提供参考。

2 GLAM资源描述参考模型对比

2.1 GLAM资源描述参考模型具体内容

CIDOC CRM、RiC-CM、IFLA LRM、EDM和CDWA这5种GLAM资源描述参考模型都属于E-R(实体-关系)模型,因此从类目、属性及关系等具体内容入手,对各种模型进行对比分析。

CIDOC CRM于1996年推出以来,经历了多次修订。根据2022年公布的CIDOC CRM 7.2.2版^[17],CIDOC CRM主要有81个类目,类目层级最多的有9级。其中,11个核心类目为E1 CRM Entity(CRM实体)、E2 Temporal Entity(暂存实体)、E77 Persistent Item(持久性项目)、E52 Time-Span(时间段)、E53 Place(位置)、E54 Dimension(维度、尺寸)、E92 Spacetime Volume(时空容积)、E59 Primitive Value(原始值)、E70 Thing(事物)、E5 Event(事件)、E39 Actor(行动者)。CIDOC CRM模型一共有161种属性,按照实体的类别层级,以“实体-事物-事件-时间-空间-尺寸-时空容积-原始值”的逻辑主线来编制属性关系,主要语义关系包括事件关系、时间关系、空间关系、时空关系等。

2016年,国际档案理事会发布RiC-CM模型。根据2021年7月修订版^[18],RiC-CM 0.2定义了22个类目和4个层级。其中,8个核心类目为RiC-E01 Thing(事物)、RiC-E02 Record Resource(记录资源)、RiC-E06 Instantiation(实例)、RiC-E07 Agent(代理)、RiC-E14

Event (事件)、RiC-E16 Rule (规则)、RiC-E18 Date (日期)、RiC-E22 Place (地点)。RiC-CM具有40个属性,79种关系。RiC-CM模型描述逻辑主要围绕资源、实例、代理、事件、规则展开,涉及时间、空间,主要语义关系包括“整体-部分”关系、顺序关系、主题关系、种源关系、管理层关系、事件关系等。

2017年,国际图联发布IFLA LRM模型^[19],该模型定义了11个类目和3个层级。其中,9个核心类目为LRM-E1 Res (资源)、LRM-E2 Work (作品)、LRM-E3 Expression (内容表达)、LRM-E4 Manifestation (载体表现)、LRM-E5 Item (单件)、LRM-E6 Agent (代理)、LRM-E9 Nomen (命名)、LRM-E10 Place (地点)、LRM-E11 Time-span (时间段)。IFLA LRM模型总共具有37个属性,主要语义关系包括:作品、内容表达、载体表现和单件之间的关系;行为者与作品、内容表达、载体表现和单件之间的责任关系;主题关系;称谓关系;行为者之间的关系。

EDM是2013年欧洲数字图书馆在ESE (欧洲语义元素)和DC元数据的基础上推出的参考模型^[20]。因此,EDM的类目主要分为复用已有命名空间元素和自建元素。复用元素主要来源于RDFS、OAI、SKOS、DC、DCAT、CC、SIOC,一共7类;自建类目为12类。其中,核心类目主要有NonInformationResource (非信息资源)、InformationResource (信息资源)、Webresource (网络资源)、ProvidedCHO (CHO提供)、Collection (收藏品)、Dataset (数据集)、Event (事件)、Agent (代理)、Place (地点)、Physical Thing (物理事件)、Time Span (时间段)。对于不同来源的类目,通过SKOS:Concept进行术语规范,通过ORE:Aggregation进行各种资源聚合。EDM的属性种类有79种,也分为复用和自建。其中,来自ORE (开放档案协议复用和交换元素)的属性有3种,来自DC的属性有37种,EDM自有属性有39种。兼容性是EDM最大的特点。主要语义关系包括资源来源关系、事件情境关

系、资源对象间关系、资源对象主题关联关系、聚合结构关系等。

CDWA是1990年由美国国家人文基金会 (National Endowment for the Humanities)、保罗·盖蒂信托基金 (J. Paul Getty Trust) 等联合制定的用于描述艺术品、建筑以及艺术品的视觉或文字替代品的元数据方案^[10]。CDWA最初不是一个数据模型,但它被安排在一个概念框架中,可用于设计数据模型和数据库以及访问信息。经过多次修订,CDWA配合其他的Getty词表 (AAT、CONA、IA、TGN、ULAN) 形成一套概念模型。这套概念模型主要反映7个实体 (对象/作品、地名规范、人名规范、主题规范、通用规范、相关可视化文档、相关文本文档) 的语义关系。语义关系是指对象/作品与对象/作品、对象/作品与相关可视化文档、对象/作品与相关文本文档、对象/作品与4个规范词表之间的关系。根据2022年6月CDWA的最新修订版,CDWA定义了31个主类目、501个属性。

2.2 GLAM资源描述参考模型对比

通过对5种资源描述参考模型的内容对比发现,所要描述的资源特点和应用场景是资源描述参考模型建模首要考虑的两个方面。这两个方面决定了模型的元素量级、描述维度、语义关系规模等。整理5个模型的类目、层级、属性等得出如表1所示的参考模型对比表。下面分别从模型内容、场景适用性、语义表达侧重点三方面对这5个资源描述参考模型进行对比分析。

2.2.1 模型内容对比

从各模型产生历史及修订时间来看,CDWA最早诞生,但它起初以词表形式产生,逐步形成关系体系和概念模型;从产生过程来看,RiC-CM和IFLA LRM是在CIDOC的倡议下研发的,与CIDOC CRM的理念一

表1 资源描述参考模型对比分析

	主要适用领域	诞生时间/年	最近修订时间/年	类目数量/个	层级数量/级	属性数量/种	应用场景特点
CIDOC CRM	博物馆	1996	2022	81	9	161	以事件为中心
RiC-CM	档案馆	2016	2021	22	4	40	以人物为中心
IFLA LRM	图书馆	2017	2018	11	3	37	以作品为中心
EDM	文化遗产领域	2011	2017	19	3	79	以对象为中心
CDWA	美术馆、艺术品收藏领域	1990	2022	31	6	501	以细节描述为中心

致。EDM是欧盟数字图书馆发展的产物,它是为了聚合不同来源的资源,既兼顾CIDOC CRM的建模原则,又兼顾DC元数据以及W3C语义网框架等国际标准;从类别和属性来看,CIDOC CRM是类目最多、最全面的,CDWA是描述最细致的,EDM的类别和层级相对较少,但属性数量相对较多,它是最灵活和轻量级的。下面从各参考模型的场景适用性和语义表达侧重点两方面进一步对比分析。

2.2.2 场景适用性对比

虽然上述5个模型涵盖GLAM系统的各种场景,支持文化遗产领域的资源描述的各种应用,但是由于各文化遗产机构馆藏资源之间的交叉,资源揭示完全采用一种模型可能在某些属性揭示方面受到限制。因此,选择何种参考模型来描述资源,不仅要考虑行业场景,而且要了解各描述模型场景适用的优缺点。

CIDOC CRM针对博物馆领域文物资源承载信息丰富的特点,对所描述资源的材质、形态、归属、存放地点、时间跨度等形成了类目最多、属性最多的综合性模型,具备对资源内容各个维度细粒度揭示的能力。但由于类目细致、属性丰富,导致关系繁冗成为该模型的一个弊端。RiC-CM、IFLA LRM作为档案馆和图书馆的专业模型,主要依据档案资源和文献资源的著录特点,揭示了档案资源和文献资源不同实体之间的关联。但对于其他资源如文物、艺术品等内容进行语义揭示,就会存在一些维度的缺失。EDM是欧盟文化遗产数字平台设计的参考模型,以ESE和DC等元数据为基础。EDM模型多数的大类和属性复用了已有词表,并采用聚类的方式将不同元素聚合,是一个兼容性极高的模型。但随着兼容性的增高,EDM模型对资源特点的揭示会缺乏深度。CDWA最初是著录艺术作品的一套词表,后来和Getty的其他规范词表形成一整套的概念模型。但是,CDWA在类别和属性的定义上是分类词表的编制原则,通过下位类的方式表示类目属性。CDWA的概念模型在其自身的系统中比较完善,但与外部映射时,由于系统过于复杂,导致与外界交互不太“友好”。

2.2.3 语意表达侧重点对比

CIDOC CRM能从事件关系、时间关系、空间关系、时空关系的维度来揭示资源内容。由于博物馆资

源承载了丰富的历史信息,因此,CIDOC CRM擅长“以事件为中心”的语义表达,能通过对事件发生的时间、空间关系的揭示,展示事件的发生发展关系。RiC-CM主要强调档案中人物、活动以及所属,档案资源的特点来自档案文献中“活动”“所属”“授权”,“Agent”“Event”“Rule”是模型表达的语义主体。因此,语义主体之间的主要关系成为RiC-CM“以人物为中心”语义表达的侧重点。IFLA LRM的语义关系主要体现的是文献资源的作品、内容表达、载体表现等维度之间的关系,始终围绕“作品”来进行描述,通过“以作品为中心”的语义关系来聚合不同的资源形态。EDM由于高兼容性和高聚合度,侧重“以对象为中心”的语义表达,在聚合不同来源的资源上能凸显优势。CDWA的编制是为了满足对艺术品描述与揭示的需求,其最大的特点就是语义揭示的“细致”,语义表达“以细节描述为中心”。艺术品的每个细节、每个维度,CDWA都进行了归类,并建立了人名、地名、机构名、主题四个维度的规范,与CDWA的核心类目形成语义模型。

3 GLAM资源描述参考模型应用场景实例分析

以上对比分析了5个资源描述参考模型的应用场景及特点,下面通过实例分别进行验证和解释。

3.1 “以事件为中心”的应用场景

CIDOC CRM模型能够提供事件、时间、地点、时空等维度的关系揭示,但“以事件为中心”的语义关系在事物活动、资源内容的揭示方面更有深度。在Volume A: Definition of CIDOC Conceptual Reference Model^[17]中,通过对图1左右两张图片的关系解读,展示了CIDOC CRM模型如何进行“以事件为中心”的语义描述。这两张图片讲述了一个与艺术史相关的历史事件——温克尔曼看见拉奥孔(E5 Event-Winkelmann sees Laocoön)。模型左边是拉奥孔群雕(Laocoön Group)图片,右边是德国艺术史学家约翰·约阿希姆·温克尔曼(Johann-Joachim Winkelmann)的图片。1755年,约翰·约阿希姆·温克尔曼在梵蒂冈看到了拉奥孔的群雕组图。1764年,他在《古代艺术史》(*History of the Art of Antiquity*)中描述看过拉奥孔群雕之后的印象,认为希腊艺术“高贵的单纯,静穆的

伟大”。但是，图片中的这座希腊风格 (hellenistic) 的拉奥孔群雕却被学者广泛地认为是来自帕拉马的青铜器雕塑 (Pergamene Baroque) 的复制品，这件复制品根据罗马委托的原作复制而成 (Roman-commissioned

copy of the Laocoön Group)。图1还显示，约翰·约阿希姆·温克尔曼出生于1717年，是安娜·玛丽亚·迈耶 (Anna-Maria Meyer) 的孩子，1768年在的里雅斯特去世。

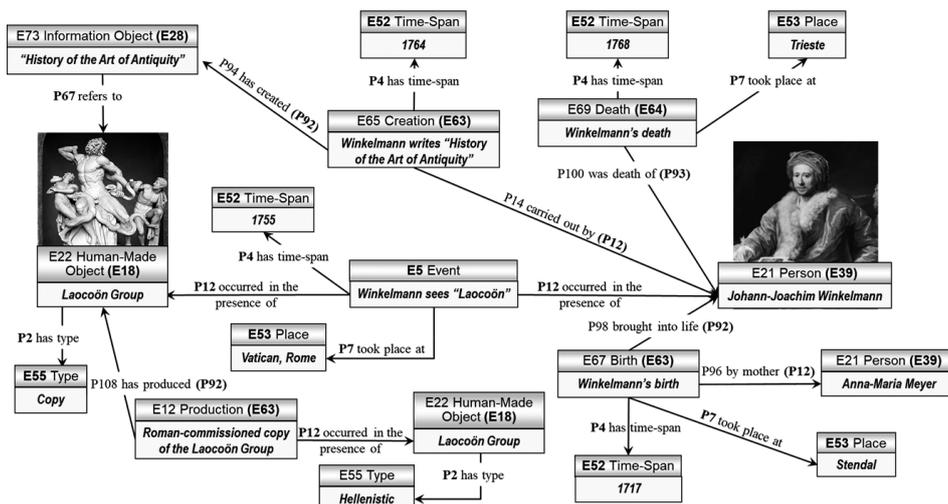


图1 “温克尔曼看见拉奥孔”语义关系

在整个事件的语义关系图中，每个方框的上方是类目名称，下方是实例。属性表示为带有属性标签的箭头。通过类目标签和属性标签，可以看到整个事件的发展历程。即信息从希腊青铜雕塑原件，传递到拉奥孔群雕副本和约翰·约阿希姆·温克尔曼的头脑，再传递到他的作品《古代艺术史》中，并且每个事件发生的时间点（约翰·约阿希姆·温克尔曼的出生时间和去世时间），以及他和母亲的关系都在整个语义关系中表现得非常清楚。正是CIDOC CRM“以事件为中心”的特点，让两部作品承载的相关历史事件得到了体现。

3.2 以“人物为中心”的应用场景

RiC-CM的4个重要的实体为资源、实例、代理和活动，四者体现了档案记录的“以人物为中心”的特点。在RiC-CM 0.1^[21]中，通过“19世纪法国公证人P.G.F.Léveau”的档案信息资源描述，展示了RiC-CM模型在揭示人物的活动、社会关系等方面的语义关联。

图2中数据来自法国国家档案馆 (Archives nationales de France, ANF) 和当地档案机构Cher省的档案部门。在图示右边浅色部分“Cher省档案部门”的关系图中，主要记录的语义关系为Mechun的公证人办公室创造了Mechun档案办公室，Mechun档案办公室拥有公证

人P.Léveau的档案卷宗；公证人P.Léveau档案卷宗有目录、文档、复本等形式；Mechun的档案办公室属于Cher省的档案部门。在图示左边深色部分法国国家档案馆的关系图中，主要记录的语义关系为P.G.F.Léveau在1820—1842年担任第一公证人职务；Mechun的公证人办公室创建了第一公证人职务；Mechun的公证人办公室位于Mechun；Mechun位于Cher省；Cher省的档案部门也位于Cher省。此外，P.G.F.Léveau是E.J.Léveau的孩子，也是他的继任者；E.J.Léveau在1804—1820年也担任第一公证人职务；E.J.Léveau和P.G.F.Léveau在法国国家档案馆有各自的档案，并且他们的档案还合并成了一套卷宗号为ANF2003056/6的系列卷宗。

在这个实例中，法国国家档案馆和Cher省档案部门因为都拥有“P.G.F.Léveau”这位公证人的档案而被关联在一起。并且，“P.G.F.Léveau”与父亲的社会关系也得到了体现。通过时间关系的揭示，还发现了P.G.F.Léveau原来是他父亲的继任者。所以，RiC-CM参考模型在“以人为中心”的应用场景中具备优势。

3.3 “以作品为中心”的应用场景

IFLA LRM是在图书馆系统的FRBRoo、FRAD、FRSAD模型基础上形成的。该模型适用于图书馆基于

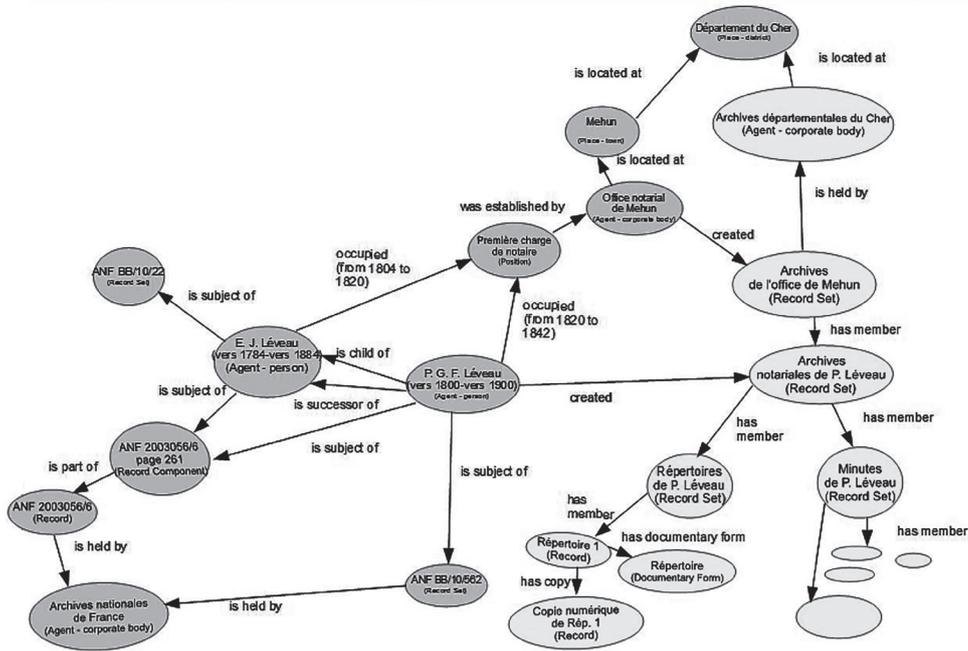


图2 “19世纪法国公证人P.G.F.Léveau” 语义关系

多种“内容表达”和“载体表现”的“作品”的聚类，更好地满足了用户“以作品为中心”的应用场景。

通过纽约市立公共图书馆 (New York Public Library) 的OPAC, 以题名为“Little women”进行检索, 对IFLA LRM的模型进行检测。当输入“Little women”后, 检索得到7 750条记录^[22]。本文选取了5条题材、载体形态、责任者、出版者、出版年等各不相同的记录。在检索结果中, 内容表达 (Expression) 有小说、电影、歌剧、音乐剧等; 载体形态 (Manifestation) 有图书、有声书、DVD、CD等; 单件 (Item) 是每一条记录的收藏单位, 图3的图例中给出了第一个检索结果的单件示例。纽约市立公共图书馆还提供了精确检索, 包括体裁 (Genre)、语种 (Language)、格式 (Format)、出版地 (Place) 等。

通过图3所示的检索实例, 说明以IFLA LRM模型能够实现“以作品为中心”的不同内容表达、载体形态、单件、主题等核心元素的书目资源的聚类, 能够满足用户查找不同作品的检索需求, 是一个立体的、关联的书目描述概念模型。

3.4 “以对象为中心”的应用场景

EDM作为欧盟文化遗产数字平台的资源描述参考模型, 更希望让不同平台、不同来源的资源集成在一起。因此, 实现不同馆藏、网络及数据库等资源的聚合, 成为EDM模型考虑的重点。复用DC、RDFS、SKOS等其他元数据词表, 能更方便地适用“以对象为中心”的应用场景, 同时让EDM的体量更加轻便。

本文选取一个关于“巨石阵” (Stone-Henge) 的实例, 来展示EDM模型如何实现不同资源类型之间的关联。如图4所示, 左边是一幅1800年的巨石阵水彩画, 来自英国图书馆Online Gallery的数字资源^[23]; 右边是1987年拍摄的巨石阵照片, 来自英国文化遗产网站 (<http://viewfinder.english-heritage.org.uk>)。这两张图片作为edm:ProvidedCHO的实例, 因为具有相同的空间覆盖范围——巨石阵, 所以通过“dcterms:spatial”产生了联系。在这种情况下, 可以直接将两张图片链接在一起, 允许用户从一个对象浏览到



图3 “Little women” 检索实例展示

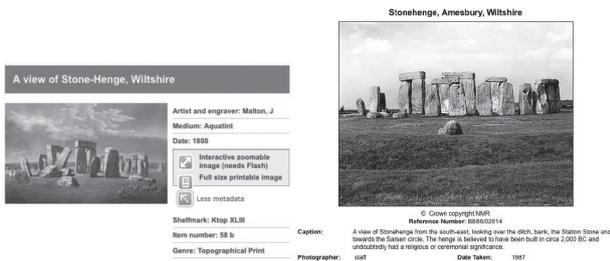


图4 “巨石阵”图片实例展示

另一个对象，实现不同对象之间的聚合和关联。EDM模型在“以对象为中心”的应用场景中，能够很好地完

成任务。

3.5 “以细节描述为中心”的应用场景

CDWA是全面揭示美术作品、视觉作品、艺术藏品等的重要参考模型。详尽的分类体系和规范控制体系，使得Getty博物馆实现对藏品各个维度的揭示，包括对藏品细节特征的描述。图5展示的是以凡·高 (Van Gogh) 为检索词的实例。

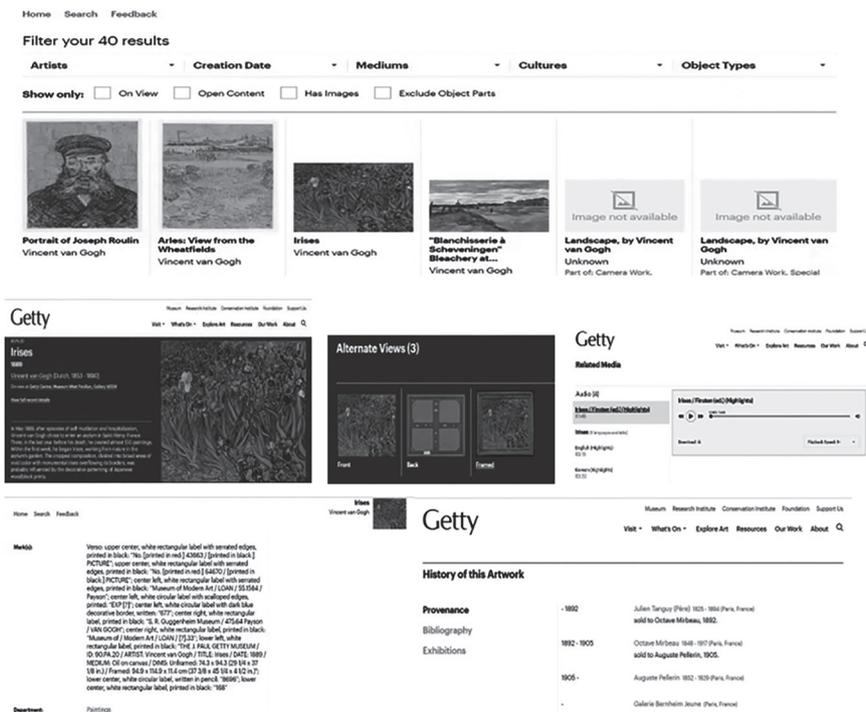


图5 凡·高作品实例展示图

通过输入“Van Gogh”，检索到40条记录。Getty博物馆给出了5个维度的聚类，分别为“Artists”（艺术家）、“Creation Date”（创作日期）、“Medium”（博物馆）、“Culture”（文化）、“Object Type”（资料类型）。值得说明的是，“Culture”（文化）维度的下级类目为国别。因此，这5个维度概括为“人”“时”“地”“机构”“类型”，这是很多模型中本体的通用类别。查看凡·高最著名作品《鸢尾花》(Iris)的检索界面，可以看见对这个作品的详细描述，包括作品创作过程、作品各角度展示、作品相关音频、作品印记描述、作品流传历史等极其丰富的内容。无论是载体形态，还是作品流转过程都能得到清楚的揭示。特别是“Mark(s)”这种类别的揭示，对艺术藏

品的考证更具价值。查看其他检索结果，还发现有的款目对画作的左边、右边等都有充分描述和揭示。可以说，CDWA模型对艺术作品的细节性描述进行了全面设计。

4 GLAM资源描述概念模型及场景研究的应用价值

GLAM资源描述需要揭示资源的物理信息和语义信息，并且不同的用户根据应用场景对资源描述有不同的需求。5种模型涵盖了GLAM领域的不同用户需求，通过不同应用场景的实例分析，验证了5种资源描述参考模型在语义揭示等方面的功用。作为语义网背景下发

展起来的资源描述底层逻辑,作为人机交互、资源开放共享的标准框架,资源描述概念模型及场景的应用价值主要体现在以下3个方面。

4.1 实现关联与映射,加强资源共享与融合

Web 3.0通过RDF数据模型、规范描述方法及相关技术体系,实现了网络资源的语义关联与共享。要打破GLAM领域界限,实现文化遗产资源的开放共享、关联融合,就需要掌握各资源描述概念模型的实体、属性内容,建立各参考模型核心元素的相互映射。只有实现了概念模型之间的映射,才能真正实现资源内容的互联融合。目前,CIDOC CRM已经建立了与上述4个模型之间的映射;EDM是一种兼容性特别强的模型,作为以DC为基础的模型,它与其他4种模型的映射也都有建立;CDWA在Getty网站上将与其他模型、词表及元数据的映射也做了发布。因此,资源描述参考模型之间的相互映射使文化遗产资源的关联与融合变成现实。

4.2 适用不同应用场景的本体构建,助力知识挖掘与发现

对资源的细粒度描述和语义关联是资源描述参考模型的主要功能,也是资源知识挖掘与发现的基础。为了更好地开发文化遗产资源,针对不同资源集成系统和应用场景,选择适用的参考模型非常关键。只有充分了解每种参考模型的描述特点、语义关系,才能科学构建资源描述领域本体,形成丰富的语义关系,实现资源的知识挖掘和语义推理,从而达到知识增值和知识服务的目的。本研究对不同应用场景进行分析,给出了知识挖掘时的具体应用场景。因此,研究者可以根据资源所属领域、资源描述的核心主体,选择适用的资源描述参考模型,或者复用其中的元素,构建适合资源特色的本体及关系,实现知识挖掘与发现。随着各领域知识挖掘的不断深入,行业之间界限逐渐被打破。图书馆在对古籍文献揭示时,如果要展现古籍的所属与流转等情况,就可以借鉴CIDOC CRM中“以事件为中心”的相关实体和属性;博物馆如果要集成相关文物的网络资源、数据库资源等不同来源的资源,就可以借鉴EDM中关于“聚合”的相关属性;如果做相关专题的知识图谱,在本体构建时,可以从实际需要出发,综合考虑相关模型中的实体和属性,构建符合需求的本体和概念

模型。

4.3 加强数字人文研究,创新文化服务

数字人文是信息化浪潮下,人文研究与信息技术结合的创新方法,近些年在文化遗产的保护与挖掘方面得到了广泛应用,如欧洲时光机、中国历代人物传记资料库、中国家谱知识服务平台、数字敦煌等数字人文项目均产生了深远影响。每一个数字人文项目的成功,除了相关技术和工具之外,资源描述的概念模型也至关重要。不同的数字人文项目,都根据各自研究重点和应用场景构建了概念模型,是对经典通用参考模型的深化和创新。分析研究国际通行的GLAM领域资源描述参考模型,能为我国本地数字人文建设的标准化、规范化以及文化创新贡献力量。

5 总结与展望

本文通过对GLAM系统中5种模型内容和应用场景区的对比分析研究,总结了各自模型的特点及应用价值,发现不同资源描述参考模型的规范元素的映射是文化遗产资源开放共享、资源融合的关键;通过相关模型及元素的复用,构建适合资源特点的语义模型,能为文化遗产资源的知识挖掘与开发提供逻辑支持。GLAM领域及W3C创造的资源描述参考模型远不止文中提到5种,希望以后在参考模型逻辑关联规则、语义推理规则、映射融合规则等方面有更多发现,更好地推动文化遗产领域资源融合、知识增值与智慧服务。

参考文献

- [1] Compatible models & Collaborations [EB/OL]. [2022-12-16]. <https://cidoc-crm.org/collaborations>.
- [2] 李文. 档案资源的语义化组织标准: ICA RiC-CM记录概念模型及RiC-O记录本体分析 [J]. 档案, 2021 (5): 49-52.
- [3] 蔡丹, 罗琳. 从FRBR-LRM到IFLA LRM [J]. 山东图书馆学刊, 2018 (5): 52-58.
- [4] 翟琼, 刘宏哲. CIDOC CRM在数字博物馆中的应用 [J]. 中国博物馆, 2015, 32 (2): 26-30.
- [5] 吴琼. 基于本体的不可移动文物数字博物馆模型研究与设计 [D]. 北京: 北京交通大学, 2008.
- [6] 吴琼, 鲍泓. 基于地理本体的不可移动文物信息参考模型研究

- 及应用[J]. 计算机科学, 2008, 35(3): 113-116, 212.
- [7] 王艺茹, 史东辉. 使用CIDOC CRM构建建筑领域非遗知识本体[J/OL]. 计算机工程与应用: 1-13 [2022-12-24]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20220323.2236.013.html>.
- [8] 董焱. 数字博物馆元数据标准初探[J]. 北京联合大学学报(自然科学版), 2005, 19(2): 61-65.
- [9] 贾君枝, 史璇. 数字博物馆元数据标准构建研究[J]. 山西大学学报(哲学社会科学版), 2015, 38(1): 114-119.
- [10] 张俊娥, 王亚林. 博物馆元数据标准构建研究: 以盖蒂研究所元数据标准为例[J]. 大学图书馆学报, 2018, 36(6): 55-63.
- [11] 陈艳, 周馨. 基于CIDOC CRM的文化遗产资源的元数据集集成——以DC元数据的映射为例[J]. 现代情报, 2010, 30(5): 60-63, 84.
- [12] 范炜, 林君松. 文化遗产数据资源关联集成模型研究[J]. 信息资源管理学报, 2022, 12(6): 111-122.
- [13] 左丹, 欧石燕. 人文信息资源语义描述、语义组织研究与实践述评[J]. 图书馆论坛, 2019, 39(8): 21-31.
- [14] 黄永欣. 文化遗产资讯领域中的参考模型[J]. 图书馆学研究, 2012(11): 57-61, 41.
- [15] MÄKELÄ E, TÖRNROOS J, LINDQUIST T, et al. WWILOD: an application of CIDOC-CRM to World War I linked data[J]. International Journal on Digital Libraries, 2017, 18(4): 333-343.
- [16] BOUNTOURI L, GERGATSOU LIS M. The semantic mapping of archival metadata to the CIDOC CRM ontology[J]. Journal of Archival Organization, 2011, 9(3/4): 174-207.
- [17] Volume A: Definition of the CIDOC Conceptual Reference Model [EB/OL]. [2022-12-16]. <https://cidoc-crm.org/Versio n/version-7.2.2>.
- [18] Records in Contexts-Conceptual Model [EB/OL]. [2022-12-16]. <https://www.ica.org/en/records-in-contexts-conceptual-model>.
- [19] IFLA Library Reference Model [EB/OL]. [2022-12-16]. <https://repository.ifla.org/handle/123456789/40>.
- [20] Europeana Data Model Primer [EB/OL]. [2022-12-16]. https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf.
- [21] Records in Contexts Conceptual Model Version 0.1, September 2016 [EB/OL]. [2022-12-16]. <https://www.ica.org/sites/default/files/RiC-CM-0.1.pdf>.
- [22] New York public library [EB/OL]. [2022-05-16]. https://browse.nypl.org/iii/encore/search/C__Rb21311863__SLittle%20women__P0%2C1__Orighresult__U__X7?lang=eng&suite=def#resultRecord-b21311863.
- [23] A view of Stone-Henge, Wiltshire [EB/OL]. [2022-05-14]. <http://www.bl.uk/onlinegallery/onlineex/kinggeorge/a/003ktop00000043u058b0000.html>.

作者简介

陈人语, 女, 1980年生, 硕士, 副研究馆员, 研究方向: 知识组织、图书馆学, E-mail: chenrenyu@163.com。

Research on GLAM Domain Resource Description Reference Model
and Application Scenarios

CHEN RenYu
(Capital Library of China, Beijing 100021, P. R. China)

Abstract: Standardized description and open sharing of human cultural heritage information resources are important mission of GLAM in the semantic web era. The resource description reference model is the underlying logic of information resource semantics and association. Facing many models in different application scenarios in the GLAM field, studying the content and characteristics of typical models is of great significance to the knowledge fusion, ontology construction and mapping, knowledge discovery and service innovation of human cultural heritage information resources. In this paper, the classification, attribute and relationship of CIDOC CRM, RiC-CM, IFLA LRM, EDM and CDWA are collated and studied, and the application process of semantic disclosure of each model is compared and analyzed. The applicable characteristics of resource description application scenarios are summarized, and the application value of resource description reference model and application scenario research is excavated, which provides a reference for the construction of knowledge organization ontology in the field of cultural heritage.

Keywords: Resource Description; Reference Model; GLAM; CIDOC CRM; RiC-CM; IFLA LRM; EDM; CDWA

(责任编辑: 王玮)