

# 书目资源的统一概念模型 LRM 论纲<sup>\*</sup>

范 炜

(四川大学公共管理学院信息管理技术系, 成都 610064)

**摘 要:** [目的/意义] 书目资源需要坚实的数据模型为基础。图书馆参考模型 (LRM) 是建立在 FRBR 家族模型之上的统一概念模型, 是书目资源实现开放关联的数据基础架构要件之一。全面分析书目资源的统一概念模型 LRM, 旨在加强书目资源的数据化实践。[过程/方法] 以 FRBR 家族模型的巩固、继承、协调、集成与统一的思路开展。首先, 对 LRM 的用户任务、新增实体与属性进行了分析比较; 其次, 在逻辑数据模型层面, 使用 UML 类图集成化呈现 LRM 全景, 对实体、属性与关系进行了全面分析; 再次, 对聚合与关联两个重要问题予以点明; 最后, 强调 LRM 定位书目资源的通用化、一般意义的概念模型。LRM 适用于广泛意义上的各类信息资源概念建模。与此同时, LRM 需要结合内容标准 (如 RDA) 进行应用落地, 利用语义网与关联数据实践, 开展多元化应用。[结果/结论] LRM 在“需求—行为—资源—模型—系统—应用”链条上促进信息资源深度开发与利用, 强调编目规则与数据建模手段的结合, 重视 LRM 的概念建模方法, 加强信息资源概念模型的语义丰富化研究, 推广图书馆、博物馆与档案馆跨界联合的数字资源专藏。

**关键词:** LRM 书目资源 概念模型 信息组织

**分类号:** G254.0

**DOI:** 10.31193/SSAPJ.ISSN.2096-6695.2019.04.05

## 0 引言

书目域 (Bibliographic Universe) 是书目记录功能需求 (Functional Requirements for Bibliographic Records, FRBR) 家族模型中普遍提及的一个概念。但是, 何谓书目域, FRBR 的报告并没有做出详细解释。笔者认为, 书目域具有实指与意指两层含义: 实指信息资源机构 (如图书馆、博物馆、档案馆等) 所拥有的书目记录 (数据) 集合, 即书目资源; 意指 FRBR 从现实世界向信息世界转换的概念建模, 书目记录中所蕴涵的实体集合。

书目域是图书馆编目工作情境的抽象化, 并不能直接具象化定义, 其内涵和表述会随着图书

<sup>\*</sup> 本文系四川大学创新火花项目“巴蜀非物质文化遗产知识图谱构建研究” (项目编号: 2018hhf-18) 的阶段性研究成果之一。  
[作者简介] 范炜 (ORCID: 0000-0003-0402-9767), 男, 四川大学公共管理学院信息管理技术系, 硕士生导师, 副教授, 博士, 研究方向为元数据、信息检索与情报服务。

馆信息资源建设与信息技术应用发展而改变。书目域的核心是书目资源。书目资源, 又称书目记录、书目数据等, 将其外推到网络世界, 可视为广泛意义的可记录并编码的信息资源。围绕书目资源的描述、组织与管理一直都是信息组织核心的研究与实践内容。

在 FRBR 已有研究与实践基础上, 国际图联 (International Federation of Library Associations, IFLA) 的图书馆参考模型 (Library Reference Model, LRM) 是 2017 年 8 月发布的面向书目资源的概念参考模型<sup>[1]</sup>。LRM 在满足最终用户信息需求的数据和功能方面, 提供了一个高层次的、抽象的元数据分析框架, 主要面向图书馆书目资源及相关资源的非管理性元数据。LRM 并不试图穷尽书目资源的所有方面, 旨在为书目资源的语义一致化、丰富化与集成化提供数据基础架构。

LRM 虽然是 IFLA 提出的, 但其适用范围较广, 如博物馆、档案馆、美术馆等, 可以为文化遗产资源的数字化、开放化、网络化、关联化以及国际化开发提供指南。

以大数据为代表的新型信息资源环境中, 智慧数据要求高质量、结构化、关联化的数据基础。长期以来, 图书馆编目工作形成的书目资源是高质量的结构化数据, 是智慧数据的重要来源类型之一<sup>[2]</sup>。集成化统一的 LRM 是书目数据向智慧数据转变的一个核心要件。通过 FRBR 家族模型及 LRM 的努力, 为书目资源向智慧数据转化打下基础。

## 1 LRM 概念模型概述

### 1.1 FRBR 家族模型基础

FRBR 是书目资源的第一个概念模型, 奠定了下一代书目信息系统的数据架构, 是迈入 21 世纪图书馆编目工作的重要里程碑<sup>[3]</sup>。FRBR 自 1998 年发布至今已有 20 年, 积累了丰富的建模经验与应用实践。这得益于编目标准的开发、信息组织与知识组织的理论发展以及数据库设计与数据库应用等多方面的共同努力。2011 年, FRBR 工作组基于 2009 年的 FRBR 最终报告发布了 FRBR 的 RDF 版本。由于 FRBR 采用实体关系 (Entity-Relationship, E-R) 模型, 也被称为 FRBRer。经过 FRBR 与 CIDOC CRM 两种概念模型的面向对象方法建模合作研究, 产出了 FRBR 模型的面向对象版本 FRBRoo, 早期版本只包括 FRBR, 第二个版本加入了 FRAD 和 FRSAD。在合作框架下, 通过 FRBRoo 实现了 FRBR 与 CIDOC CRM 的一致化。在 FRBRoo 的基础上, PRESSoo 设计用来表示持续型资源 (如连续出版物) 的形式化本体。从成果上反推, PRESSoo 是 FRBRoo 的扩展, FRBRoo 是 FRBR 和 CIDOC CRM 的扩展。

众所周知, 规范控制是高质量结构化数据的保障。在 FRBR 奠定书目记录的概念模型基础之后, 在书目规范控制方面, 针对第二组实体和第三组实体分别开发了 2009 年的规范数据的功能需求 (Functional Requirements for Authority Data, FRAD)<sup>[4]</sup> 和 2010 年的主题规范数据的功能需求 (Functional Requirements for Subject Authority Data, FRSAD)<sup>[5]</sup>。FRAD 考虑了用户需求与图书馆编目工作两方面, 贯彻实施了编目工作的规则与过程, 具体表现在阐明关系和提供依据两个用户任务。FRSAD 则侧重主题概念的识别与语义关系的表达, FRBR 是主干基础模型, FRAD 和 FRSAD 可视为 FRBR 的两个扩展。图 1 是 FRBR 家族模型的主要构成。FRBR、FRAD 和 FRSAD 是通常家族模型的代表。

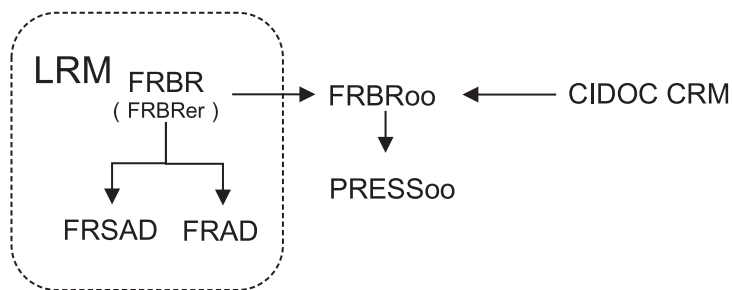


图 1 FRBR 模型家族示意图

FRBR、FRAD 和 FSAD 三个模型是由 IFLA 三个不同的工作小组开发。虽然三个模型的基础都是实体关系模型，但是三个工作小组各自有不同的侧重和解决问题的思路，在结构和概念化上存在一些差异。最早的 FRBR 定义相对自由，模型形式化定义相对较弱，这也为应用实践冲突埋下了伏笔。FRBR 家族模型在应用探索初期已经暴露出不一致性、错乱、冲突与复杂问题。

为了 FRBR 家族模型的巩固和持续发展，解决多模型的多头绪、复杂性、不一致性等问题，以及更好地适应开放互联的信息资源环境，如何对三个模型进行集成化统一，即是 LRM 的提出背景。LRM 的早期叫法是 FRBR LRM，现定名为 IFLA LRM。后文提到的 FRBR 家族模型主要 FRBR、FRAD 和 FRSAD 三个模型，这是 LRM 的构建基础，如图 1 左侧所示。

### 1.2 LRM 集成化统一考虑

LRM 的总体出发点是，一个通用的模型，一个入口，避免多个模型带来的多头绪问题，隐藏了多模型的诸多细节。在通用模型的一致化框架下再进行扩展，这样避免以后的冲突和错乱。从发展的眼光看，LRM 的出现是进化和整合的产物，是对 FRBR 家族模型的持续性发展，不是破坏性瓦解。LRM 的确立旨在巩固 FRBR 家族三个数据模型，对 FRBR、FRAD、FRSAD 三个模型进行集成化统一。

LRM 之所以能够对 FRBR 家族模型进行集成化统一，主要依靠 FRBR 家族模型的共通基础。FRBR 家族模型都定位为功能需求。具体体现为：以用户为中心、以 E-R 模型为基础以及 4 个核心实体 WEMI 的一致性，即作品 (Work)、内容表达 (Expression)、载体形式 (Manifestation)、单件 (Item)。另外，FRBR 家族模型报告的范围、定义、用户任务、实体、属性与关系都采用相同的内容体例。

在 LRM 集成化统一的定位考虑方面，力图揭示书目资源的共通性基础结构，提供概念模型框架，具备灵活的可扩展和互操作性，不与具体的标准、技术及系统绑定，为信息资源开发与利用而设计，面向跨界信息资源聚合与集成，为关联数据而优化。

### 1.3 LRM 概念模型的数据建模认识

总体而言，随着实践检验和理论深化，从 FRBR 家族模型到 LRM，概念模型越来越规范严谨，形式化定义得到重视与加强。

从数据库设计的数据建模角度理解 LRM，有助于明确 LRM 作为概念模型所处的建模阶段与转换方式。在数据库设计领域，数据建模包含三种数据模型：概念模型、逻辑模型与物理模

型<sup>[6]</sup>。概念模型是数据建模的第一步, 实现现实世界到信息世界的转换, 不依赖于计算机系统, 主要是为人所能理解, 作为设计者与用户之间的沟通工具。概念模型的建立需要洞察和理解现实世界中事物的存在及其运动状态, 在数据建模活动中反映为(事物)实体的识别和实体间关系的确立, 也就是最常见的 E-R 模型。LRM 即以 E-R 模型为基础, 以书目资源为目标, 构建信息资源概念模型, 提供信息资源描述的通用语义。

抽象概念模型构建好之后, 可以采用不同的逻辑数据模型来表达, 常见的逻辑数据模型有业界主流的关系型, 以及面向对象型、层次型等。新兴的各类 NoSQL 技术的数据模型也值得关注。逻辑数据模型可以被计算机系统和用户所理解, 是实现计算机系统存储与管理的重要基础。FRBRoo 对应的是逻辑数据模型层面的研究与实践探索。在数据建模中, E-R 模型与面向对象模型需要根据数据模型层次进行界定区分, 不能混为一谈。

## 2 LRM 以用户需求和用例为设计出发点

FRBR 家族模型始终坚持面向最终用户(End User)需要的数据与功能设计原则, 以用户需求和用例(Use Case)为设计出发点, 围绕信息搜寻行为, 明确用户任务框架, 开展模型设计。LRM 继承并集成了 FRBR 家族三个模型关于用户、用户需求与用户任务的阐释。这是信息组织以用户为中心的重大认知变化。FRBR 家族模型的报告都是将用户任务放在模型之后介绍, LRM 报告将用户和用户任务置于模型之前介绍, 内容结构上的调整体现了以用户为中心的建模思维导向。

### 2.1 LRM 的用户与资源

LRM 作为信息资源概念模型, 其面对的用户分为两类: 最终用户和模型使用者。面向的最终用户群体是广泛多元的, 涵盖各个行业领域和主题方面的信息资源场景和技术环境(如跨终端 Web)中各种类型的信息资源用户, 如学生、教师、医生、律师等, 最终用户类型根据具体情境进一步细分。LRM 使用者主要是信息链上的各种主体, 包括信息资源管理机构(图书馆、档案馆、情报所、博物馆等)、出版商、书店等。

由于面向最终用户, LRM 将模型的资源界定在与用户需求直接相关的描述性元数据方面, 不包括与最终用户需求间接相关的管理性元数据。

### 2.2 LRM 的用户任务分析

以已有的信息搜寻行为理论为指导, 对信息搜寻行为进行阶段划分, 将行为阶段定义为用户任务。由于信息搜寻行为过程不是单一线性的行进方式, 行为阶段存在往复、跳转、同时等现象, 所以行为阶段之间不存在直接的、严格的顺序关系。因此, 用户任务之间是相关的, 但用户任务之间的顺序需要根据具体情境和特定用户行为进一步分析。

LRM 在 FRBR 家族三个模型的用户任务(行为阶段)基础上, 进行了映射、综合与集成。LRM 定义 5 个用户任务: 查找(Find)、识别(Identity)、选择(Select)、获取(Obtain)和探究(Explore)。LRM 的用户任务命名与 FRBR 家族模型的用户任务名称保持一致, 没有使用新的用户任务词。LRM 前 4 个用户任务名称与 FRBR 的用户任务名称相同, 最后 1 个用户任务名称来

自 FRSAD。已发布的 FRBR 家族模型三份报告的中文版本中，用户任务名称的中文表达保持一致。

虽然 LRM 的任务名称与 FRBR 家族模型的任务名称相同，但对各个用户任务的内涵做出了新的阐释。从综合集成实质上讲，LRM 的用户任务定义是在已有 FRBR 家族模型用户任务基础上的泛化 (Generalization)。

根据用户任务的映射说明，图 2 显示了 LRM 对 FRBR 家族三个模型的用户任务进行综合集成的具体情况。FRBR 模型的用户任务是综合集成的主干，其中：查找和识别两个任务是三个模型的全内涵集成；选择任务是 FRBR 与 FRSAD 的全内涵集成；探究任务是 FRSAD 的探究任务与 FRAD 的部分阐明关系的集成；获取任务是 FRBR 的获取任务定义。FRAD 的提供依据任务面向图书馆编目工作，因此没有纳入 LRM 用户任务范畴。

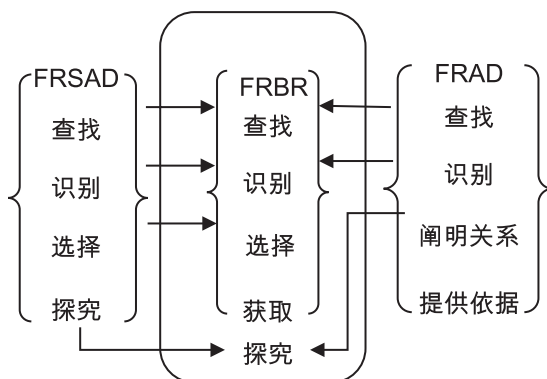


图 2 LRM 与 FRBR 家族模型的用户任务映射

表 1 在 LRM 用户任务定义的基础上，分析了各个任务的要点，并对应到信息检索系统设计中应考虑的功能要求。在用户任务中，用户所需的资源是具体的对象或实体，为确保表达的具体准确性，在用户任务定义中，对抽象的 Resource 使用资源对象一词来指代。

表 1 LRM 用户任务分析

任务名称	LRM 报告定义	任务要点	信息检索系统设计响应
查找 (Find)	通过任意相关条件的检索，将用户感兴趣的一个或多个资源对象汇集起来	相关条件：根据资源对象的属性与关系形成检索策略；汇集：资源对象的组织加工	搜索功能 搜索结果呈现
识别 (Identity)	准确理解已检索到的资源对象本质，区分开相似的资源对象	资源对象的本质即属性与特征	元数据描述
选择 (Select)	确定已检索到的资源对象的适用性，接受或拒绝特定资源对象	用户需求与资源对象的匹配	搜索交互 相关性判断、排序
获取 (Obtain)	访问资源对象的内容	从文档替代到文档本身的路径	获取途径 开放链接
探究 (Explore)	通过资源对象间的关系发现资源，将资源置于情境中看待	开放式、无明确结束点 关系驱动的联想与意外发现	情境化导航与浏览 资源发现、知识图谱



探究任务体现了探索性搜寻 (Exploratory Search) 行为活动。Gary Marchionini 将信息搜寻行为划分为界限模糊的三个类型: 查找 (Look up)、学习 (Learning) 和调研 (Investigation)。查找即知道要找什么 (Known-Item) 的搜索, 学习和调研两者被 Marchionini 归为探索性搜寻行为, 具体表现为关联式查找、意外发现等<sup>[7]</sup>。信息搜寻行为已有大量研究成果, 可用于 LRM 探究任务的理论分析, 为实践应用寻找理论支撑。

### 3 LRM 概念模型分析

#### 3.1 FRBR 家族模型的协调一致化

FRBR 家族三个模型为信息组织实践带来了挑战。在信息资源开发与利用实践中, 同时符合三个模型带来的工作复杂性和难度是显而易见的。在复杂性之下, 三个模型之间还存在不一致、冲突和矛盾。针对这一现象和问题, LRM 的努力目标是在三个模型基础上建立一个统一的模型框架, 集成化策略体现在: 以四个核心实体 WEMI 及其基本关系达成实体定义的一致性; 根据通用概念模型的抽象层次, 定义最常见的属性与关系, 将原先三个模型中的一些细节进行隐藏。

为协调一致性, LRM 工作组对 FRBR 家族三个模型的用户任务、实体、属性与关系进行了详细具体的转换映射。从某种意义上讲, LRM 模型是 FRBR 家族三个模型转换映射的结果。

在通用的、一般的模型抽象层次上, 根据用户需要和资源对象粒度要求, 对 FRBR 家族三个模型的已有元素的处理方式有以下几种: 保留、合并、泛化、另用或弃用。转换映射文件提供了 LRM 如何协调一致化 FRBR 家族三个模型的各种元素的依据, 随 LRM 报告一同发布。本文根据转换映射文件, 对 FRBR 家族三个模型之间的冲突与协调一致进行评析。

#### 3.2 FRBR 家族实体与属性的变动分析

在 LRM 的模型主干里, 以 FRBR 第一组实体为基础。在此基础上, 对实体与属性进行了调整。

##### (1) 实体的主要变化情况

(a) FRBR 第一组实体保留。

(b) 为 FRBR 第二组实体定义了一个新的超类实体 Agent, 保留 FRBR 的 Person 实体并改写定义, 新定义 Collective Agent, 包含之前的家族和团体组织这两个实体。

(c) FRBR 第三组实体弃用, 其中, 地点 (Place) 实体 (LRM-E10) 重新泛化定义, 新定义时间跨度 (Time-Span) 实体 (LRM-E11)。

(d) FRAD 的名称 (Name) 实体与 FRASD 的 Nomen 实体合并, 保留 Nomen 实体并泛化定义, FRAD 的识别符与受控检索点实体弃用, 可使用类型实体或 Nomen 实体的类别属性来表示。由于 FRAD 的机构 (Agency) 和规则 (Rules) 是受控检索点相关实体, 不在 LRM 模型定义范围之内, 因此被舍弃。

##### (2) 属性的主要变化

(a) 属性是实体特征的反映, 实体的变化会直接影响属性的变化调整。因此, 新增实体需要定义新的属性。例如, 为 Res、Nomen、地点、时间跨度等实体新增属性。

(b) 由于 LRM 定位更通用的概念模型, 属性数量大幅度减少, 将原有许多属性转为关系定义。这是一个重要的变化。例如, 作品的主题属性转为关系, 取代 FRAD 作品实体的主题属性 (subject) 和 FRBR 作品实体的坐标属性 (coordinates)。“……的地方” (place of) 和 “……的时间” (date of) 此类属性转为 Res 实体与地点、时间跨度实体之间的关系定义。

(c) 为作品实体新增代表性内容表达 (Representative expression) (LRM-E2-A2) 属性, 将之前既在作品实体又在内容表达实体中声明的属性集中在一处声明。代表性内容表达有助于提高作品识别的显著性。从理论上讲, 一部作品的所有表达是等价的, 定义代表性内容表达的好处在于, 编目员可以为作品选择一个优先题名。

(d) 对载体表现实体的众多载体形式属性进行了泛化定义, 新增载体表现状态 (Manifestation statement) (LRM-E4-A4) 属性, 根据需要进一步再细分类型。该属性的定义背后是载体表现实体的属性的大幅度合并, 或者说通用化抽象 (细节隐藏)。载体表现由原先的 38 个属性合并为 6 个属性。

(e) 将载体表现实体的有关访问属性合并成新的访问条件 (Access conditions) 属性, 访问限制 (Access restrictions) 属性更名为使用权限 (Use rights) 属性。

(f) 得益于实体层级结构与继承关系, 超类实体定义的属性无须在子类实体中再单独重复定义。例如, FRAD 个人、家族和团体组织的联系信息、语言、活动领域等属性合并为行为者实体的语言属性。虽然个人是行为者的子类, 但个人具备专有的一些属性, 如职业, 因此为个人实体重新定义了职业 (Profession/Occupation) 属性。除此之外, 在 FRAD 中个人实体拥有性别属性, 但 LRM 并未将其纳入定义范围。

(g) FRSAD Nomen 实体的属性与 FRAD 名称、识别符及受控检索点实体的属性合并, 成为 LRM Nomen 实体的属性。

(h) Res 实体的类别 (Category) 和注释 (Note) 两个新属性来自 FRSAD 的 Thema 实体属性。类别属性是对 Thema 实体的类型 (Type) 属性的泛化, 注释属性来自 Thema 实体的范围注释 (scope note), 表示内涵更加通用、宽泛的注释说明。

### 3.3 LRM 的重要实体分析

#### (1) Res (LRM-E1) 实体与 Nomen (LRM-E9) 实体

LRM 中的 Res 和 Nomen 实体定义来自 FRSAD。这两个实体的理解具有一定难度。首先, 要从 FRSAD 说起。FRSAD 是针对 FRBR 中第三组实体, 处理作品的主题、名称关系的概念模型。作品与主题之间的关系需要通过 aboutness 分析。aboutness 是知识组织的一个核心概念。FRSAD 的工作组主张现实主义, 从用户角度出发, aboutness 被视为作品与主题之间的关系, 而非属性, 不局限于作品“有什么”或“关于什么”。将 aboutness 视为一种关系, 避免哲学上唯名主义或理想主义的无休止争论。另外, 有关 aboutness 与 ofness 的区分, 主要体现在对艺术品 (照片、画作等) 的描述分析上。

FRSAD 定义了 Thema 与 Nomen 两个实体来处理作品的主题与称谓之间的关系。Thema 与 Nomen 是两个拉丁语单词。FRSAD 工作组认为, 这两个词在英语语境中没有预设含义, 是文化

中立的, 不需要翻译。Thema 相关的英文词有 subject、topic、concept 等, 不只有主题这一个含义。

在 FRASAD 中, Thema 是作为作品主题的任何实体, 也是所有 FRBR 实体的超类。Nomen 是 Thema 被知晓、被引用、被标记的任何符号 (sign) 或符号序列 (如字母、数字、字符、符号、声音等)。符号具有特定含义和指向, 令人产生联想。作品与 Thema 的主题对应关系是多对多, Thema 与 Nomen 之间的称谓关系也是多对多。

Nomen 对应的常见英文词是 name, 但又不只表示名称的含义。考虑到与 FRAD 中名称、识别符和受控检索点三者都属于 Nomen 的范畴, 同时, 又要与 FRAD 中名称实体区分开, 所以使用 Nomen 实体来表述。因此, Nomen 的内涵更广。

LRM 对 FRASAD 的 Thema 实体进行了泛化, 启用“Res”实体。Res 在拉丁语中是“thing”(事物)的含义。Res 实体作为 LRM 概念模型的顶层实体, 其定义为话语世界 (universe of discourse) 中的任何实体。这里的话语世界可以理解为人类可以理解和认识的语义空间, 通过语言和符号所能表达的, 与书目资源相关的所有实体。这部分的理论基础与符号学理论具有密切关系。

Nomen 是 Res 实体的称谓, 将 FRAD 的名称实体也纳入其中。因此, 从某种意义上讲, LRM 的 Nomen 比 FRASAD 定义的更广。将称谓作为实体的好处在于, 可为它们定义属性, 诸如语种、字顺、受控词表以细分同一实体不同称谓之间的关系等。举例来说, 区分个人的曾用名与现用名。值得注意的是, Res 实体是抽象的, 虽然可认知理解, 但其外显需要借助 Nomen。为避免文本表达混淆二者, 在 LRM 报告中, Res 实体用花括号表示, Nomen 实体用单引号表示, Nomen 实体示例的字符串属性则使用双引号表示。

在中文语境下, 为确保沟通交流的一致性, 建议直接使用 Res、Thema 和 Nomen 表述, 或使用芮斯、希玛、诺门三个中文音译词来指代, 当作专有术语来使用, 无须意译。

## (2) Agent (LRM-E6) 实体

LRM 对 FRBR 第二组实体进行了调整, 定义 Agent 作为第二组实体的超类, 整合家族和团体两个实体为 Collective Agent。Agent 是能够谨慎行动、被授予权利, 并能对其行为负责的实体。Agent 的上位类是 Res, 下位类包含 Person 与 Collective Agent。Agent 一般译为“代理、代理者”, 词典的一种解释是, 发生作用 (或产生某种结果) 的人 (或物)。综合 LRM 定义和中文语境, 笔者认为在信息资源概念模型中, “行为者”更能代表 Agent 的实体含义, 相应地, Collective Agent 可译为“集体行为者”。

关于 Person 的定义, LRM 界定为现实世界中的人, 包括健在的和已故的。因此, Person 译为“个人”较为恰当。文学创作、神话、民间传说中出现的虚拟人物或角色, LRM 不将其作为 Person 的实例。需要注意, 个人在 LRM 定义与 FRAD 定义之间存在一定冲突。

行为者实体拥有联系信息、活动领域与语言三个属性, 个人与集体行为者作为行为者的子类, 可继承父类行为者的属性。

## (3) 地点与时间跨度实体

LRM 增加了空间与时间营造的两个新实体: 地点与时间跨度。

在书目资源中, 地点实体主要体现为对地理区域或空间的社会与文化认同。地点实体可通过



地理名称、国家地区等确定。时间跨度包含开始时间、结束时间和持续时长。时间跨度很早就应用在编目中区分不同时期的人、作品和事件等，例如，第二次世界大战。

## 4 LRM 逻辑数据模型的集成化表示

LRM 基于实体关系 (E-R) 模型构建，一共定义了 11 个实体，37 个属性及 36 对 (72 个) 关系。LRM 为每个实体、属性和关系都赋予唯一的 ID。LRM 报告对实体层次、实体属性、关系分别做了详细介绍。

LRM 的实体是资源对象的概念化抽象，相当于面向对象和本体工程中类 (Class) 的概念。实体之间有层级结构 (isA 定义)，非平面列举式。超类 (Super class) 实体可以包含子类实体，子类可以继承超类的属性与关系。这样建模可以减少重复定义，支持自反和关系的推理，符合面向对象思想和语义网设计要求。实体之间包含 3 个层次，1 个顶层实体 (Res)、8 个二级实体，2 个三级实体。同一父类的实体之间是互斥的，如个人与集体行为者，即要么是个人，要么是集体行为者，不可能两者兼具。

LRM 的属性是描述实体的相关信息 (元数据)。LRM 对 FRBR 家族三个模型的实体的属性进行了较大的合并和精简，只保留了实体最常见的通用属性。LRM 报告列举了 37 个属性，每个属性与所属实体及实体层次相对应，其中包含了继承属性。例如，Res 实体的类别属性被作品、内容表达、载体表现、Nomen 及地点实体继承。集体行为者实体没有定义自己的属性，继承其父类行为者实体的属性。

LRM 定义的关系是一般意义上通用的。关系采用严谨的本体定义方法，为其指定 ID、定义域、值域及基数 (Cardinality) 限制等。每对关系都是互逆的 (Inversed)。在 LRM 定义的关系中，第一对关系是 Res 与 Res 的关联 (LRM-R1)，这是顶层关系，定义域与值域都是 Res 自身，正反关系名称相同 (is associated with)。不同资源对象情境下，顶层关系有具体所指。WEMI 实体间的关系依然是 LRM 的核心。

LRM 报告给出的实体关系图 (Entity-Relationship Diagram, ERD) 是一种自定义的图示表示法，直观易懂。为促进更广泛领域的模型设计、交流与实现，可以考虑在 LRM 建模应用中选用数据库业界常见的 ERD 图示法。ERD 图示有很多种，如 Peter Chen 的 E-R 图示法、信息工程 (Crow's Foot) 表示法、IDEF1X、ORM2、UML 等。ERD 的选择与模型的抽象程度及表达粒度有关。根据 LRM 的具体情况，综合考虑实体和属性定义、实体继承、关系定义及细分等要求，UML 类图是一种适合 LRM 的 ERD 表示方法。UML 是一种通用建模语言，其中 UML 类图可以用于 E-R 模型的形式化表示。LRM 实体的超类和子类的层级结构、属性和关系都可以通过 UML 类图展现。

LRM 报告首先给出了模型全局概览图，其次分别给出了局部图，有助于逐步了解和接受。这里使用 UML 类图对 LRM 模型的实体、属性与关系进行集成化呈现，一张总图涵盖所有实体、属性与关系，既鸟瞰全局又兼顾细节。图 3 是根据 LRM 报告的相关内容，使用 UML 类图绘制的 LRM 全景模型。

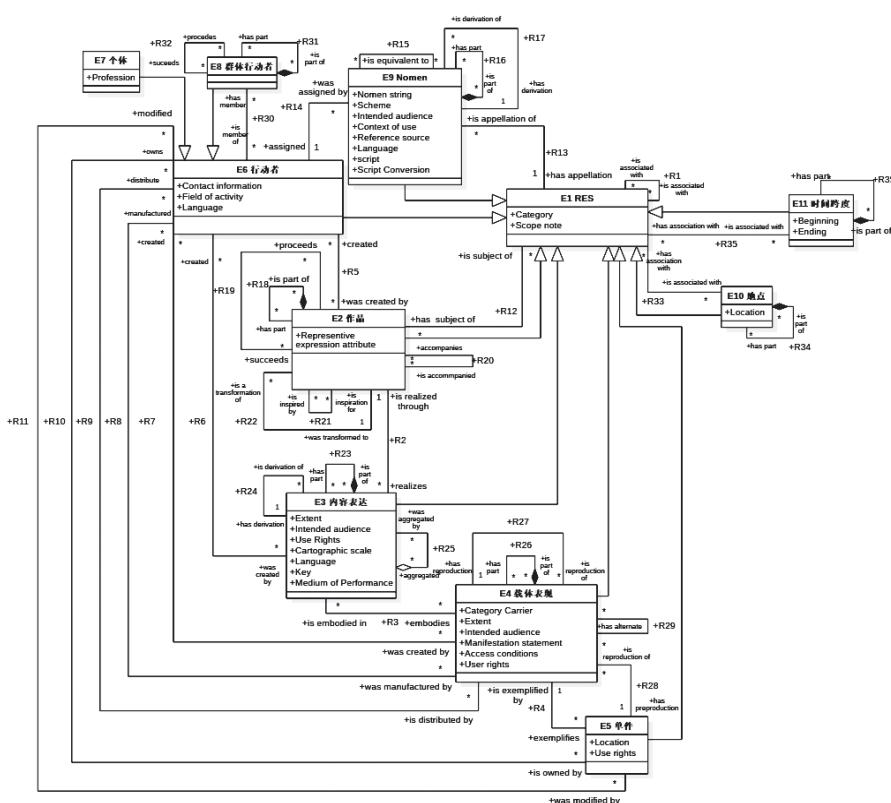


图 3 使用 UML 类图构造的 LRM 全景模型

在图 3 中，实体（类）用矩形表示，矩形分为三层，分别对应实体的名称、属性和方法。由于 LRM 概念模型仅用到类名和属性定义，所以方法一栏是空白的。方法留给后续应用程度来定义。因为 UML 的双向关联线条没有箭头，因此，将 LRM 关系做如下处理：LRM 关系的 ID 位于线条的中间，关系的名称位于线条的两端，根据定义域实体和值域实体分别靠近。关系的基数（1: 1、1: M、M: M）也写在线条的两端，与靠近的实体对应。图中基数 \* 代表多个，对应 LRM 的基数 M 和 FRBR 关系图示的双箭头。图中实心菱形表示组合关系，空心菱形表示聚合关系。图中空心三角箭头表示子类指向父类，子类实体可以继承父类实体的属性。

LRM 主要用到 UML 类图常见的集成、组合、聚合和关联四种关系。表 2 列出了 UML 类图关系用于 LRM 实体及关系定义的用法。

表 2 UML 类图关系与 LRM 的对比

UML 类图关系	LRM	说明
继承	E1 E2 E3 E4 E5 E6 E7 E8 E9 E10 E11	定义 LRM 的实体
组合	R16 R18 R23 R26 R34 R36	Nomen、WEM、地点、时间跨度自身
聚合	R25	内容表达之间的关系

续表

UML 类图关系	LRM	说明
关联	R1 R2 R3 R4	WEMI 核心关系
	R5 R6 R7 R8 R9 R10 R11	行为者与 WEMI 关系
	R12 R13 R14	作品、Res、Nomen、行为者的关系
	R15 R17	Nomen 自身关系
	R19 R20 R21 R22	作品自身关系
	R24 R27 R29	载体表现自身关系
	R28	单件自身关系
	R30 R31 R32	行为者关系
	R33 R35	Res 与地点、时间跨度的关系

这里将继承、组合、聚合之外的关系都归入关联关系，视为不同类型的关联关系。行为者与作品、内容表达各有 1 对关系，行为者与载体表现、单件各有 2 对关系。

在 LRM 报告中，列举了每个实体的关系定义，将实体作为定义域进行关系数量统计，得到表 3。关系数量的计算重复统计了互反关系的定义域与值域的实体互换，例如作品的包含关系 (has part/is part of) 算 2 次，依此类推。表 3 统计了 LRM 的 11 个实体的关系数量。值得注意的是，个人实体本身没有定义关系，然而这并不意味着个人实体是孤立的。个人实体通过其父类行为者实体与其他实体建立关系。

表 3 实体的关系数量统计

ID	实体 (定义域)	关系数量 (含互反关系)
E2	作品	13
E4	载体表现	11
E3	内容表达	9
E6	行为者	9
E9	Nomen	7
E1	Res	5
E8	集体行为者	5
E5	单件	4
E10	地点	3
E11	时间跨度	3
E7	个人	0

在 LRM 的 36 对关系中, 从关系类型的角度看, 有部分关系含义在不同实体上多次出现, 经过整理, 得出 24 对关系名称, 其中独有的关系名称有 20 对, 相同的关系名称有 4 对, 即不同定义域与值域的关系名称相同。相同关系名称的定义域与值域不同, 但基数都相同整体与部分 (has part/is part of) 关系出现次数最多, 达到 6 次, 另有 3 对关系各出现 2 次, 如表 4 所示。

表 4 LRM 相同关系名称整理

关系	定义域	值域	基数
整体 / 部分 (has part / is part of)	Nomen	Nomen	M: M
	作品	作品	
	内容表达	内容表达	
	载体形式	载体形式	
	地点	地点	
	时间跨度	时间跨度	
创建 / 由...创建 (was created by / created)	作品	行为者	M: M
	内容表达	行为者	
	载体表现	行为者	
先于 / 后于 (precedes/succeeds)	作品	作品	M: M
	集体行为者	集体行为者	
涉及 (has association with /is associated with)	Res	地点	M: M
	Res	时间跨度	M: M

对 LRM 关系的多维度分析, 有助于更好地理解 LRM, 为应用实践与扩展提供参考。

## 5 LRM 的聚合与关联分析

自 FRBR 发布以来, 聚合 (Aggregation) 是一个持续被讨论的开放问题。聚合的本质可归纳为: 根据特定情境与用户需求, 实现用户与资源对象的特定匹配。聚合不是简单的归集, 聚合依赖于关系的识别和确认, 需要在作品、内容表达与载体表现三者的关系转换框架下进行, 不同层次和不同粒度上应有不同的考虑。

根据聚合内涵的认识, LRM 的聚合主要处理多对多关系。内容表达与载体表现是 WEMI 中唯一一对多对多关系。作品与载体表现是一对多关系, 载体表现与单件是一对多关系。LRM 的聚合关系 (LRM-R25) 定义为载体表现包含多个内容表达。在这个定义中, 聚合操作以内容表达为中介对象, 聚合结果最终在载体表现上反映。

LRM 定义的聚合主要包括 3 种类型。一是独立作品内容表达的聚合。一般以特定类型为



聚合依据,聚合的内容表达具有相对独立性,可共处在一个载体表现上,聚合的结果是集合(Collection)。例如,学术期刊(多篇论文)、小说文集、诗歌选集、音乐CD专辑(多首歌曲)等。二是作品扩充内容表达的聚合。在不影响和改变原始作品和内容表达的前提下,对作品进行扩充的内容表达,例如,书籍的序言、插图、评论等。三是类似内容表达的聚合。载体表现可能包含同一作品的多个类似的内容表达。多语种作品是此类聚合的典型,例如,作品的英语版与中文版、双语政府文件等。除此之外,影视作品发行的DVD中声音对白和字幕也属于此类聚合。

需要注意,聚合(was aggregated by/aggregated)和组合(has part/is part of)是两种关系。从耦合度的强弱来看,聚合的耦合度较松。在LRM中,6个实体上定义了组合关系,如图3中的实心菱形表示。组合主要用于实体自身的整体-部分组成关系。在关系处理上要注意区分。连续出版物是同时具有聚合和组合关系的复杂结构,在新数字出版环境中基于LRM构建连续出版物的概念模型是一个待深入研究的重要问题。

要实现动态聚合,书目资源的关联基础非常重要。LRM除了提供概念模型框架之外,还将书目资源置于语义网与关联数据的开放数据生态。数据建模时考虑当前的技术环境对书目资源形态的影响,与语义网相关技术标准兼容,实现书目资源与外部相关数据的对接、集成与共享,如DBpedia、VIAF等。

未来,考虑到信息资源的开放集成性和LRM建模复用需求,需要将LRM的各要素进行声明发布,主要包括命名空间、URI和RDF Schema三个部分。Schema中每个元素、属性和关系都有唯一的URI,方便资源建模和描述时调用。信息资源对象使用LRM的Schema定义,根据一致性的描述机制,借助LRM数据中枢,在开放网络中实现RDF数据集之间的连接可能性。

LRM与FRBR家族三个模型的转换映射文件已经做了大量命名空间映射的准备工作。IFLA维护并管理FRBR家族模型的命名空间(<http://iflastandards.info/ns/fr/>),通过开放元数据登记(Open Metadata Registry, OMR)服务登记注册<sup>[8]</sup>,以RDF/XML形式发布各个模型的元素介绍及用户任务词汇说明。其中,FRBRer包含216个元素、FRBRoo包含188个元素、FRAD包含150个元素、FRSAD包含19个元素。每个FRBR家族模型的用户任务词汇都各有4个。未来LRM将会拥有唯一的命名空间,通过IFLA网站发布,以便于资源映射、关联与集成。

## 6 总结与展望

LRM是继FRBR之后信息组织的一个里程碑。LRM背靠FRBR家族的20年经验积累,将信息组织的理论与实践推进到新时代、新环境。LRM继续打破图书馆固守的资源大门,从传统的资源固守转向开放的资源融入。这是信息组织理念上的一次飞跃。

长期以来,信息组织和信息搜寻行为两部分的研究相对独立。从最终用户角度考虑,LRM在“需求—行为—资源—模型—系统—应用”链条上将起到重要的连接作用。LRM将信息组织与用户研究、信息检索联系在一起,置于信息资源管理整体框架之下。

本文全面梳理与分析了书目资源的统一概念模型LRM,以FRBR家族模型的巩固、继承、协调与统一思路展开分析,与语义一致化、丰富化与集成化目标形成对应。由于LRM是基于

FRBR 家族三个模型的整合, 体系庞大且内容丰富, 文章只是提纲挈领地进行了整体架构分析, 尽力梳理清楚 LRM 的内在关系, 构建出 LRM 的全景图, 提出一些研究问题。

相较于 LRM 报告文本和模型概览, 本文使用 UML 类图绘制的 LRM 全景模型, 对 LRM 的实体、属性与关系进行了集成化呈现, 更加直观、详尽, 兼具全局预览和局部放大。

除此之外, UML 类图的模型表示可进一步作为参考模型设计模板, 为特定情境的资源库概念建模定制与扩展。基于 UML 类图可正向工程生成代码, 与信息系统开发结合得更加紧密, 以提高应用开发与落地的效率。

强化 LRM 的理论研究与实践应用, 实质是回归信息组织这一核心领域的坚守与拓展。信息组织是图书情报专业知识体系的核心内容之一。LRM 的集成化、开放化、关联化特点为信息组织理论方法研究与应用实践打开了局面。强化以用户为中心和开放关联数据情境意识, 积极利用现代新技术理念与手段, 在巩固编目积累与信息组织经典理论方法的基础上开拓创新。从横向与纵向两个方面深化: 横向拓展, 鼓励跨界合作, 扩展到各个领域的信息资源建设; 纵向深入, 加强概念模型的稳固性、持续性与可扩展性。

在概念模型层次, 加强语言学、符号学、逻辑学、语义学、本体等方面的理论分析渗透。语言学和符号学中有关词汇、概念、含义、关系、所指、能指的理论, 例如作品、主题、表现与载体之间的关系分析深化。以 LRM 通用概念模型为基础, 加强数据建模的应用扩展研究, 主要体现在三个方面。

一是将编目规则的悠久传统与数据建模的现代手段结合起来研究。在新技术环境与开发语言中, 将编目规则人工语言转换成数机器可理解的规则(如, 本体规则定义), 便于关系的自动发现与推理。

二是开展语义丰富化研究。LRM 概念模型提供的只是通用的、一般意义的类、属性与关系, 需要根据具体信息资源开发场景中的用户需求, 进一步细化与关联扩展, 达到用户需求级别的语义丰富度。LRM 面向更广泛的信息资源, 可作为不同类型信息组织在数据架构设计时的重要参考。LRM 不仅关乎图书馆和编目人员, 还需要多学科多领域人士共同参与, 在开放关联的数据生态中实现各类数字资源的内容分发、检索、汇聚与利用。

在图书情报机构实践方面, 有意识加强图书馆编目工作与信息资源建设的深度整合。LRM 聚焦概念模型, 其使用需要与内容标准紧密结合。在 FRBR 家族模型的研究基础之上切实推进 LRM 的理论与方法研究, 强化与 RDA 实践研究的协同联动。在图书馆界, 资源描述与检索(Resource Description and Access, RDA)是内容类标准, 是 LRM 概念模型的实践支撑。RDA 在 LRM 还未正式发布之前, 就在积极探索 LRM 的适应性改造。LRM 为 RDA 提供了第一个稳固的聚合模型, 使得 RDA 的扩展性进一步增强。RDA 的 3R(RDA Redesign and Restructured) 重构项目会切实推动 LRM 的模型应用<sup>[9]</sup>。

三是开展从概念模型到逻辑数据模型的实现研究, 如关系型数据模型、面向对象数据模型、NoSQL(特别是图数据模型)及其之间的映射研究。在工具手段的利用方面, 充分利用 UML 建模语言的表达能力, 将 UML 的类图用于 LRM 概念建模研究与实践(扩展、定制与改造), 将 UML 的用例图用于 LRM 用户任务开发与分析。在信息资源开发与利用中尽可能地使用形式化、

规范化语言,提高信息资源相关系统设计与开发的效率。

目前,FRBR家族模型三个报告都已有中文翻译版<sup>[10-12]</sup>,LRM的中文版也即将在2020上半年翻译完成并发布。在行动方面,建议中国图书馆学会学术研究委员会信息组织分委会下组建LRM专题研究小组,联合图书馆编目工作者与高校教研工作者一起跟进LRM的最新发展,推动LRM的本地化研究。

在公共文化服务的大背景下,图书情报领域研究LRM的同时,应注意与博物馆、档案馆等一起开展文化类特色信息资源专题或特藏的LRM应用实践。加强LRM与CIDOC-CRM、CDWA等文化遗产资源数据模型的映射、兼容与互操作性研究。在文化大数据推进的过程中,推广LRM在信息资源专题数据库、知识库建模方面的通用方法论意义。

### 【注释】

collection在不同情境有不同译法,如馆藏、特藏、专题、选集等。

### 【参考文献】

[1] PAT R, PATRICK L B, and MAJA Ž. IFLA library reference model: A conceptual model for bibliographic information (LRM). Consolidation Editorial Group of the IFLA FRBR Review Group. [EB/OL]. [2019-12-20]. [https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017\\_rev201712.pdf](https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017_rev201712.pdf).

[2] 曾蕾, 王晓光, 范炜. 图档博领域的智慧数据及其在数字人文研究中的角色 [J]. 中国图书馆学报, 2018, 44 (1): 19-36.

[3] IFLA STUDY GROUP ON THE FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS. Functional requirements for bibliographic records: Final report (FRBR) [EB/OL]. [2019-12-20]. [https://www.ifla.org/files/assets/cataloguing/frbr/frbr\\_2008.pdf](https://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf).

[4] IFLA WORKING GROUP ON FUNCTIONAL REQUIREMENTS AND NUMBERING OF AUTHORITY RECORDS (FRANAR). Functional requirements for authority data: A conceptual model [EB/OL]. [2019-12-20]. [https://www.ifla.org/files/assets/cataloguing/frad/frad\\_2013.pdf](https://www.ifla.org/files/assets/cataloguing/frad/frad_2013.pdf).

[5] FUNCTIONAL REQUIREMENTS FOR SUBJECT AUTHORITY DATA, A CONCEPTUAL MODEL (FRSAD). 2011. IFLA Working Group on Functional Requirements for Subject Authority Records (FRSAR). Berlin/München: De Gruyter Saur. [EB/OL] [2019-12-20]. [https://www.ifla.org/files/assets/cataloguing/frbr/frbr\\_2008.pdf](https://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf).

[6] 王珊, 萨师焯. 数据库系统概论 (第5版). 北京: 高等教育出版社, 2014: 15.

[7] G Marchionini. Exploratory search from finding to understanding. Communications of the ACM, 2006, 49 (4): 41-46.

[8] Open Metadata Registry [EB/OL]. [2019-12-20]. <http://metadataregistry.org/>.

[9] RDA Redesign and Restructured [EB/OL]. [2019-12-20]. <https://www.rdatoolkit.org/3RProject>.

[10] 王绍平, 王静, 林明, 等译. 书目记录的功能需求 (FRBR): 最终报告 [EB/OL]. [2019-12-17]. <https://www.ifla.org/files/assets/cataloguing/frbr/frbr-zh.pdf>.

[11] 顾犇, 蔡成普, 杨熙, 等译. 规范数据的功能需求 (FRAD): 概念模型 [EB/OL]. [2019-12-17]. [https://www.ifla.org/files/assets/cataloguing/frad/frad\\_2009-zh.pdf](https://www.ifla.org/files/assets/cataloguing/frad/frad_2009-zh.pdf).

[12] 刘莎译. 主题规范数据的功能需求 (FRSAD): 概念模型 [EB/OL]. [2019-12-17]. <https://www.ifla.org/files/assets/cataloguing/frsad/frsad-final-report-zh.pdf>.

# The Outline of Unified Library Reference Model for Bibliography Resources

FAN Wei

(Department of Information Management, Sichuan University, Chengdu 610064, China)

---

**Abstract:** [ **Purpose/significance** ] Bibliographic resources need to be based on solid data models. The library reference model is an unified conceptual model based on the FRBR family models. It is one of the data infrastructure requirements for bibliographic resources to achieve open linked data. A comprehensive analysis of the unified library refernce model aims to strengthen the data practice of bibliographic resources. [ **Method/process** ] This paper carries out the thinking of consolidation, inheritance, coordination, integration and unification of the FRBR family model. First, it analyzes and compares the user tasks, newly-added entities, and attributes of LRM. Second, at the logical data model level, it uses UML class diagrams to integrate and present the LRM panorama, and comprehensively analyzes the entities, attributes, and relationships. The two important issues of aggregation and correlation are pointed out. Finally, it emphasizes the general utilization and general meaning conceptual model of LRM positioning bibliographic resources. LRM is applicable to the conceptual modeling of various types of information resources in a broad sense. At the same time, LRM needs to be combined with content standards (such as RDA) for application set-up, and use the Semantic Web and Linked Data Practices to develop diversified applications. [ **Result/conclusion** ] LRM promotes the in-depth development and utilization of information resources in the chain of “demand-behavior-resource-model-system-application”, emphasizes the combination of cataloging rules and data modeling methods, attaches importance to the conceptual modeling method of LRM, strengthens the semantic enrichment of conceptual model of information resources, and promotes the special collection of digital resources across libraries, museums and archives.

**Keywords:** LRM; Bibliography resources; Conceptual model; Organizing information

---

( 本文责编: 孔青青 )