从“探索”任务入手解读《IFLA 图书馆参考模型》*

吴诺曼　王瑞荣　黄毕惠

摘　要　《IFLA 图书馆参考模型》(IFLA-LRM)设有查找、识别、选择、获取、探索五项用户任务,其中“探索”是一个崭新的、重要的任务。结合信息搜索研究领域的热点理论“探索式搜索”,可以区分 IFLA-LRM 的两类用户任务“查找式搜索”和“探索式搜索”,从而理解“探索”这一用户任务之于 IFLA-LRM 的重要意义和驱动作用。以“探索”为钥匙,分析 IFLA-LRM 对实体、属性和关系的设计,有助于理解 IFLA-LRM 融入关联数据网的建模宗旨。IFLA-LRM 的落实,有赖于 RDA、BIBFRAME 的更新;在现有 MARC 格式下,应转换编目思维,重视“探索”任务从而透彻理解关联数据化字段,为 BIBFRAME 取代 MARC 做好准备,迎接书目从数字化走向数据化的趋势。图 1。表 4。参考文献 15。
关键词　IFLA 图书馆参考模型　用户任务　探索式搜索　关联数据

Interpretation of *IFLA Library Reference Model* from the Perspective of “Explore” Task

Wu Nuoman　Wang Ruirong　Huang Bihui

Abstract: *IFLA Library Reference Model* has five user tasks: Find, Identify, Select, Obtain and Explore, among which “Explore” is a new and important dimension. Combined with the hot theory “exploratory search” in the field of information search research, we can distinguish two types of user tasks “finding search” and “exploratory search” of IFLA-LRM, so as to understand the important significance and driving role of “Explore” to IFLA-LRM. Taking “Explore” as the key, analyzing IFLA-LRM's design of entities, attributes and relationships is helpful to understand IFLA-LRM's modeling purpose of integrating into the linked data communities. The implementation of IFLA-LRM depends on the updation of RDA and BIBFRAME; under the existing MARC format, catalogers should change their thoughts, pay attention to the task of “Explore”, so as to thoroughly understand the linked data fields, prepare for BIBFRAME to replace MARC, and meet the trend of bibliography from digitalization to datafication. 1 fig. 4 tabs. 15 refs.
Keywords: IFLA-LRM; User Task; Exploratory Search; Linked Data

国际图联于 2017 年正式推出《IFLA 图书馆参考模型》(*IFLA Library Reference Model*, IFLA-LRM,下文简称 LRM),中译版[1]由国家图书馆馆员翻译,于 2020 年 12 月发布。LRM 是 FR 家族(包含 FRBR、FRAD、FRSAD)的整合与升级版本,沿用“实体-关系”分析法,并对原有模型各要素进行合并、取舍和重新定义,形成更加严谨规范的统一体系。该模型解决了之前 FR 家族三模型存在的多头绪、冲突性问题带来的应用障碍,提升了模型的普遍性与通用性,并试图为实现书目开放关联提供基础架构。

相较于 FR 模型,LRM 更加突出用户任务的重要性。在章节安排上,用户任务的内容被提到实体、属性和关系之前,成为整个模型的焦点[1]。这意味着 LRM 模型的设计和决策取决于它所定义的用户任务,用户任务对建模起到定锚的作用。

* 本文系国家社科基金一般项目“开放共享背景下图书馆书目元数据组织模式的演变与发展研究”(项目编号:18BTQ026)的成果之一。

通讯作者:黄毕惠,电子邮箱:386182159@qq.com。

同时,LRM 增设的“探索”(Explore)是其相较于 FRBR 在用户任务方面最显著的变化。因此,本文首先解读用户任务,着重论述“探索”任务所体现的新维度、新趋势;其次,以“探索”为切入点,分析 LRM 对实体、属性和关系的相关设计;然后,围绕 LRM 模型的落实尤其是“探索”任务的实现,梳理 RDA、BIBFRAME 以及 MARC 21 近年来的一系列调整变化,从关联数据化的视角理解 LRM 的设计意图和应用前景。

1 “探索”任务释义

1.1 LRM 的“探索”任务

LRM 将“探索”任务定义为“利用资源之间的关系发现资源,并将它们放置在特定的情境中”,并在第三章、第六章中用较多篇幅予以说明。该定义强调“探索”任务的开放性,使用户的检索可以从一个资源关联到另一个资源,还可以帮助用户从呈现的关系中获得信息导航的功能,支持用户意外偶然性的资源发现。可见,资源与资源之间的关系(即资源元数据的关联化)是“检索”任务实现的基础,它将书目的检索功能进一步拓展为资源发现功能,这对用户实现“探索”目标具有十分重要的意义。

LRM 强调“探索”的特殊性,称其“是一个单独的维度,与其他任务不同”,是“用户任务中最具开放性的”。学者范炜指出,“探索”任务体现了探索式搜索(Exploratory Search)行为活动[2]。“探索式搜索”是信息搜索领域近年的热点,其研究主要围绕搜索的影响因素、结果优化和搜索系统构建展开,LRM 的“探索”任务从构建广义的书目元数据的角度对“探索式搜索”研究进行了丰富和细化。因此,本文借用“探索式搜索”理论研究“探索”任务之于 LRM 的重要意义和驱动作用,从而获得对书目在语义网时代的角色和作用更深入的理解,为编目人员领会模型并融会于实践工作提供参考;同时也从书目元数据角度,丰富关于“探索式搜索”的研究。

1.2 “探索式搜索”理论

“探索式搜索”理论由美国数字图书馆专家 Gary Marchionini 于 21 世纪初提出,他将信息搜索行为总结为查找、学习和调查三种[3]。其中,“查找”是最基本的搜索任务,在传统搜索系统中,查询表达式与系统储存的信息相匹配,即可为用户提供满意结果。而“学习”和“调查”则是更为复杂与高级的搜索任务,体现为无唯一答案、持久发生、多面性搜寻,在搜索过程中方向不确定、多次循环、多战术并用。这种用户行为越来越常见,Gary Marchionini 将“学习”和“调查”统称为“探索式搜索”,以区别于传统的查找式搜索。“探索式搜索”这一概念可以更好地体现信息需求的不确定性及动态变化性,客观地反映人机交互行为[4]。

表 1 搜索行为

查找式搜索	查找	事实检索、已知项搜索、导航、事务、验证、回答问题
探索式搜索	学习	知识获取、理解/解释、比较、聚合/整合、社会化
	调查	集聚、分析、排除/否定、综合、评价、发现、计划/预测、转换

1.3 “探索式搜索”理论对于理解 LRM “探索”任务的启发

自 FRBR 诞生以来,随着信息环境的变化,用户需求也逐渐升级,LRM 为适应开放的关联数据环境而诞生,尤其重视在语义网应用中的数据重用[1],以此打破图书馆资源的封闭性。因此,语义网背景下的用户搜索行为研究对于理解 LRM 的用户任务具有启发性。LRM 的五项用户任务中,前四项“查找”“识别”“选择”“获取”对应表 1 的“查找”,第五项“探索”则对应表 1 的“学习”和“调查”,也就是“探索式搜索”。将 Gary Marchionini 对搜索行为的分类与 LRM 的用户任务进行映射(见图 1),有助于理解 LRM 的建模出发点。

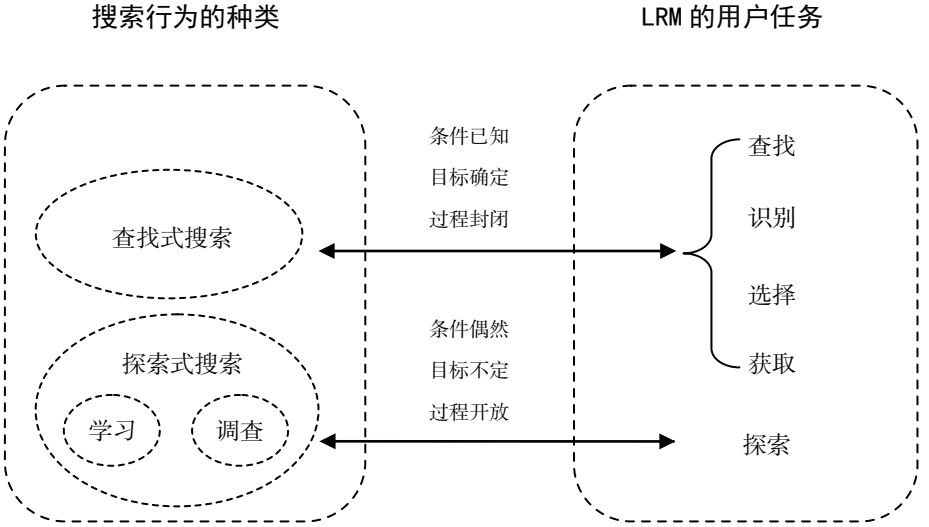


图 1　搜索行为的种类与 LRM 用户任务的映射

以书目用户为例,可以比较书单式搜索和迭代式搜索两种情形。书单式搜索,即用户的搜索条件已知、搜索目标明确,当用户进行“查找—识别—选择—获取”这一系列活动时,就完成了一次“查找式搜索”。迭代式搜索是指在大多数情形下,搜索是一个复杂的、不确定的过程,最初可能始于某个兴趣点或某个较为明确的目标,但在搜索时却从一个对象牵出另一个对象,层层递进和迭代。很多条件是在搜索过程中获得,其搜索目标也会产生变化,搜索路径充满各种随机性、可能性,是一个开放的探索式过程,这符合“探索式搜索”理论对于“学习”和“调查”行为的描述。除此之外,消费者浏览购物网站寻觅性价比最高的商品、游客浏览旅游网页以制定最佳路线、读者打开书评网站寻觅感兴趣的作品、科研工作者针对某项课题开展文献调研等活动,都具有学习和调查的性质,皆属于探索性的用户搜索行为。这种搜索行为是用户面对真实复杂的信息环境时更为常见的情形,用户对这种搜索功能的需求也越来越迫切。许多信息搜索理论都在试图解释这一搜索过程,如“采莓模型”[5]、“信息觅食模型”[6],而“探索式搜索”概念的提出,明确地将它与传统的“提问—应答”模式下的查找式搜索区别开来,并对信息组织工作提出了更高要求,指明了信息组织工作的发展方向。

“探索式搜索”理论有助于理解 LRM 新增的“探索”用户任务,即打破传统书目的封闭性,不再满足于“查找式搜索”,而是结合关联数据技术融入语义网,更好地帮助用户发现新知识,改善认知结构[7]。

2　以“探索”任务理解 LRM 模型

LRM 脱胎于 FR 家族的三个文件 FRBR、FRAD、FRSAD,其中 FRBR 的问世具有划时代意义。由 FRBR 首次构建的 WEMI 实体突破了传统编目以手边“一本书”为对象的扁平思维,形成以作品为中心的立体结构,着重体现资源间的关系,为绘制详尽的知识地图提供了可能。但是 FRBR 的目标主要在于以内容统领不同载体形态的资源,为用户提供尽可能全面且层次清晰的查找结果。而对于网络时代更加复杂的信息需求,比如对“学习”和“调查”的信息支持,FRBR 尽管较 AACR 时代有了很大突破,但仍然不能满足这一类探索性的搜索需求。因为 FRBR 主要解决了第一组资源实体关

系的揭示,对于第二组责任者实体、第三组主题实体的关系揭示则分别由 FRAD 和 FRSAD 继续深入,但三组实体之间的关系仍未得到普遍详尽的揭示,直至 LRM 统一整合和揭示三组实体之间的关系,将“探索”维度贯穿在实体和关系的设置中,为实现书目元素的普遍关联化提供了可能。

2.1 LRM 的实体分析

LRM 在保留 FRBR 核心实体的基础上,打破了按照作品、责任者、主题分组的设计,用一些更加泛化、概括力更强的实体,来覆盖更多可供提取、重组、量化分析的子类。LRM 基于“实体-关系”模型的设计,有利于揭示众多实体之间的复杂关系,为实现“探索”任务提供支持。

表 2 FR 家族与 LRM 的实体对比

<table>
<tr><th>模型</th><th colspan="2">FRBR</th><th colspan="2">FRAD</th><th>FRSAD</th><th colspan="3">IFLA-LRM</th></tr>
<tr><td rowspan="14">实体</td><td rowspan="4">第一组</td><td>作品 Work</td><td rowspan="14">第一组</td><td>个人 Person</td><td>主题 Thema</td><td rowspan="14">E1 资源 Res</td><td colspan="2">E2 作品 Work</td></tr>
<tr><td>内容表达 Expression</td><td>团体 Corparate body</td><td>主题表述 Nomen</td><td colspan="2">E3 内容表达 Expression</td></tr>
<tr><td>载体表现 Manifestation</td><td>家族 Family</td><td rowspan="13">—</td><td colspan="2">E4 载体表现 Manifestation</td></tr>
<tr><td>单件 Item</td><td>作品 Work</td><td colspan="2">E5 单件 Item</td></tr>
<tr><td rowspan="2">第二组</td><td>个人 Person</td><td>内容表达 Expression</td><td rowspan="2">E6 行为者 Agent</td><td>E7 个人 Person</td></tr>
<tr><td>团体 Corparate body</td><td>载体表现 Manifestation</td><td>E8 集体行为者 Collective agent</td></tr>
<tr><td rowspan="4">第三组</td><td>概念 Concept</td><td>单件 Item</td><td colspan="2">E9 命名 Nomen</td></tr>
<tr><td>实物 Object</td><td>概念 Concept</td><td colspan="2">E10 地点 Place</td></tr>
<tr><td>事件 Event</td><td>实物 Object</td><td colspan="2" rowspan="2">E11 时间段 Time-span</td></tr>
<tr><td>地点 Place</td><td>事件 Event</td></tr>
<tr><td colspan="2" rowspan="2">—</td><td>地点 Place</td><td colspan="3" rowspan="2">—</td></tr>
<tr><td rowspan="6">第二组</td><td>名称 Name</td></tr>
<tr><td>标识符 Identifier</td></tr>
<tr><td>受控检索点 Controlled access point</td></tr>
<tr><td>规则 Rules</td></tr>
<tr><td>机构 Agency</td></tr>
<tr><td>实体数量(个)</td><td colspan="2">10</td><td colspan="2">16</td><td>2</td><td colspan="3">11</td></tr>
</table>

注:“—”表示无对应项目

表 2 显示了从 FR 家族到 LRM 在实体设置上的整合,整体来看,IFLA 一系列资源描述标准的书目功能按照“重藏—重查找—重关联”的趋势发展。传统编目针对的是文献整体,不能对细粒度元数据进行提取、重组和分析,所以尚无必要区分一个对象的实体与命名(规范名称也只是命名的一种形式,而

不同于URI所标识的实体)。而在关联数据时代,为了保证计算机能够明确识别实体本身和对实体的命名,对两者分别进行定义成为必要。从FR家族到LRM的相关变化中可以看出这一走向:FRBR尚未区分实体和实体的命名;FRAD设置了名称、标识符和受控检索点三个实体,以便揭示实体与各种命名之间的关系;FRSAD仅有主题(Thema)和主题表述(Nomen)两个实体,Thema是一切可作为主题的事物,Nomen则是对主题的命名(FRAD所定义的名称、标识符和受控检索点隐藏在Nomen之中);LRM的E1资源(Res)和E9命名(Nomen)同样体现了"实体-命名"关系。其中,Res意为事物(Thing),是一个包罗万象的概念,是对Thema这一实体的泛化,凡LRM中明确定义或还未定义的任何其他实体都是资源(Res)的子类;E9 Nomen译为"命名",是对FRSAD中Nomen实体的泛化,其定义是"实体与指称该实体的名称之间的联系"。

区分实体与命名,对于元数据关联化(即"探索"任务的实现)必不可少。首先,将"命名"定义为实体而不是属性。在"实体-关系"模型中更便于表达名称之间的关系,如等同、派生、继承、整体/部分等,从而利于将资源之间的关系语义化。其次,可以为原本作为描述元素记录的内容提供检索点(如出版者),不再限于题名、责任者、分类主题等。新增检索点意味着新增关联,用户可由此挖掘相关资源,从而扩展新知识、促进利用。此外,通过赋予实体URI,编目界的规范控制更具操作性。不同机构可以根据各自惯例,采用不同的"首选形式",指向同一个URI代表的实体,而不必从不同名称形式中选择唯一的"规范形式"[8],从而优化用户搜索体验,消除语境差异造成的资源壁垒。

2.2 LRM的关系分析

"探索"意味着利用资源之间的关系发现资源,因此在信息组织中,对关系的揭示和表达尤为重要。LRM有一个特点,就是设置了更加抽象、概括力更强的实体,将FR家族中原有的大量属性转为关系,使原本不可拆分、不利于机器识别的描述性元素转为可供提取、重组、分析的实体。如FRBR中大量出现的"XX的地点""XX的日期"这类属性,在LRM中转为了R33(资源与地点相联系)和R35(资源与时间段相联系)这两种关系。同时,LRM还对书目中复杂的关系进行了梳理分类,抽象概括为36种关系,标识号为LRM-R1至LRM-R36。在实践中,可根据情境加以细化和扩展①。

LRM在第六章通过用例显示了实体、属性和关系与用户任务的匹配,本文特摘取"探索"任务的用例,结合实践中遇到的情况,来分析LRM所定义的关系如何支持"探索"的实现(见表3)。

表3 实体、属性和关系与"探索"任务的匹配

LRM第六章关于"探索"的用例		实例列举	涉及的关系
探索关系以理解主题领域的结构及其术语	例:浏览比起始主题范围更小的概念	逻辑学:辩证逻辑、形式逻辑、哲理逻辑、应用逻辑;模态逻辑、多值逻辑、认知逻辑、价值逻辑、时态逻辑、……	R1 资源与资源相联系 R31 行为者是集体行为者的一部分 R34 地点是地点的一部分 R36 时间段是时间段的一部分

① 参见LRM中译版表4.7。

续表

LRM 第六章关于“探索”的用例		实例列举	涉及的关系
探索一个实体不同实例间的关系	例 1:追踪原著作品与基于原著作品或改编自原著作品的派生关系	元代剧作家王实甫创作的的杂剧《西厢记》,灵感来源于唐代元稹的小说《莺莺传》。	R21 作品是作品的灵感
		F. J. 古尔德(F. J. Gould)的《写给青少年的希腊罗马名人传》(*The Children's Plutarch*)是普鲁塔克(Plutarch)的《希腊罗马名人传》(*The Lives of the Noble Grecians and Romans*)的转型作品。	R22 作品是作品的转型
	例 2:浏览与给定行为者相联系的作品和内容表达,及该行为者在创作和成果实现中的作用	巴金创作了《家》《春》《秋》等作品;翻译了《父与子》《快乐王子》等作品。	R5 作品由行为者创作 R6 内容表达由行为者创作
理解一个实体一个实例的各种命名间的关系	例 1:检查主题词汇表中论题性主题的变异名称	海洋化学 D 海水化学、化学海洋学	R15 命名等同于命名
	例 2:调查特定个人在不同应用情境中使用的变异名称(如宗教中使用的名称,官方名称)	第 42 任美国总统的官方名称是 William Jefferson Clinton,通常被称作 Bill Clinton。	
	例 3:查看国际团体在不同语言中使用的名称	Fédération Internationale de Football Association(法语)等同于“国际足球联合会”(汉语)。	
	例 4:探索在不同的受控词汇表中同一实体相同实例命名间的关联(如查找主题标题或术语对应的分类号)	“世界史”等同于 K1(《中国分类主题词表》分配的术语“世界史”与《中国图书馆分类法》中的一个有效分类号 K1 代表相同的概念)。	

从表 3 可见,LRM 列举的探索任务,在实践中均能通过实体与实体之间的关系揭示得以落实。LRM 用高度结构化的方式,将用户感兴趣的对象及其相互之间的联系加以提取、概括、命名,使传统编目的人工描述转为计算机可理解、可推理、可操作的结构化描述。这便于关系的自动发现与推荐,优化了信息组织,为实现“探索式搜索”提供了元数据建构层面的支撑。另外,LRM 虽然舍弃了阐明关系和提供依据这两项与规范数据相关的用户任务,但在模型的关系设计中,为这两项任务提供了支持(如 R15 命名等同于命名),其实质是把管理性任务隐藏在模型中,间接为终端用户任务服务。此外,“探索”任务发生的情景是复杂多变的,无法在模型中穷尽,但模型的关系设计具有很强的包容性、扩展性(如 R1 资源与资源相联系),为详细揭示千变万化的实体关联提供了空间,从而满足用户在信息世界中通过查找、学习和调查不断改善认知的需要①。

① 参见 LRM 中译版图 5.6。

3 LRM 的应用

LRM 对 FR 家族的整合体现了书目从数字化到数据化,也就是书目数据由机器可读走向机器可操作,进而融入互联网全球数据库的趋势[8]。LRM 在编目标准体系中属于概念模型层次,它的落实和应用,需要内容规则、格式规则的一系列跟进,而模型的调整与优化也必将带来国际编目标准的相应调整。新版 RDA 的出台、BIBFRAME 的修订以及 MARC21 近年来的变化,均体现了走向关联数据的趋势,即揭示实体之间的普遍关联,支持和提示用户执行“探索”任务,促进学术交叉融合和创新型知识成果的产出。

3.1 新版 RDA 对“探索”任务的支持

2004 年 RDA 开始编制,2010 年 RDA 工具套件正式发布,作为内容标准,RDA 忠实贯彻了 FR 功能模型,其元数据注册为书目走向关联数据环境做好了准备。2017 年启动的 3R 项目(RDA Toolkit Restructure and Redesign),是 RDA 指导委员会根据 LRM 对 RDA 实行的改造与更新(其成果即新版 RDA)。新版 RDA 从四个方面充分体现了对 LRM“探索”任务的支持。

首先,新版 RDA 的用户任务与 LRM 一致,同样在查找、识别、选择、获取的基础上新增了“探索”。

其次,新版 RDA 在实体的设置上根据 LRM 进行了优化,更利于“探索”任务的实现。LRM 将 FR 家族原有的 18 个实体整合为 11 个,新版 RDA 与此相呼应,定义了 13 个实体:RDA 实体、作品、内容表达、载体表现、单件、行为者、个人、集体行为者、团体、家族、命名、地点、时间段。其中 RDA 实体是其他实体的超类,对应 LRM 的资源(Res),是 Res 在应用中的一个实例。团体和家族在 LRM 中被整合至集体行为者,而新版 RDA 出于实践考虑,同时保留了这 3 个实体,除了 3 个实体的细微区别外,其他实体设置与 LRM 基本一致。命名(Nomen)实体源于 LRM,在新版 RDA 中将其定义为指代 RDA 实体的标识,包括名称、题名、检索点、标识符、分类号和主题标目。新版 RDA 设置的命名(Nomen)实体发挥区分真实世界对象及其名称形式的作用,如本文 2.1 节所述,此举对于 RDA 在关联数据环境中的运用意义重大,使大量传统编目中作为描述元素记录的内容可以升级为检索点,通过揭示命名与命名之间的关系可实现关联最大化,从而引导用户链接、跳转、意外发现,实现有价值的“探索”。

再次,LRM 取消了“国家书目记录基本需求”的相关内容,作为呼应新版 RDA 也不再强调核心元素,而是注册了更多元素。这意味着在新版 RDA 中,所有元素(即实体的属性和关系)是平等的,也都是可选的,在特定应用中可酌情忽略某些元素,也可根据需要呈现某些元素。因为新版 RDA 要处理的重点不再是传统编目观念中的“记录”,而是面向实体,细化到以三元组为单位的“数据”,数据的集成、复用、互联取决于用户的需要。新版 RDA 淡化核心元素而代之以增加更多元素,为“探索”提供了广泛的选项,因为“单一环境如图书馆目录中描述的一致性,与各种环境间进行连接的能力相比,正变得不那么重要”[9]。

最后,新版 RDA 在记录元素的方法上,提供了具有普适性的“四路径”,分别是非结构化描述(转录、修改后的转录、自由行文的附注等)、结构化描述(出版说明、论文说明、规范检索点等)、标识符(ISBN、ISSN、出版者编号等)和国际化资源标识符(IRI 等)[10]。其中,IRI 在新版 RDA 中被明确规定为路径之一,是融入关联数据的关键举措。IRI/URI 的使用使实体本身与实体的名称以及对实体的描述明确区分,完美解决了标目问题。实体的元素无论采取何种记录方式,只要有对应的 IRI/URI,都可以方便地进行关联,支持用户进行跨区域、跨学科、跨数据库的“探索”。

3.2 BIBFRAME 对“探索”任务的支持

书目要实现最终融入关联数据的目标,必须从底层元数据制作层面变革,但受制于 MARC 格式的封闭性、静态性、隐性语义揭示、粗粒度性等局限[11],书目数据难以彻底实现在更广泛的语义网中的共享和重用目标。因此,适用于新的网络数字环境的书目框架 BIBFRAME,自推出之日便

备受瞩目。BIBFRAME 的目标是取代 MARC 成为国际通用的元数据结构标准,用关联数据技术促进书目的交换、发布与共享。

BIBFRAME 的 1.0 和 2.0 版本分别于 2013 年和 2016 年正式推出,两者的最大区别是对核心类的设置,1.0 版定义了作品、实例、规范和注释 4 个核心类,2.0 版调整为作品、实例和单件 3 个核心类。2.0 版取消“规范”核心类的原因主要在于,BIBFRAME 运用关联数据技术提供的 URI/IRI 来标识实体,已经很好地实现了“规范”要解决的识别、消歧与汇集问题;其取消“注释”核心类的原因主要在于,随着关联数据环境的进步,外界已经有成熟的关联数据词表可供直接复用。2.0 版新增的“单件”核心类,原本位于 1.0 版注释类下的子类“馆藏资料”(Held Material)以及其下级子类“馆藏单件”(Held Item),两者被合并、升级成新增的“单件”核心类。BIBFRAME2.0 显著变化的动因来自对关联数据环境更深入的理解,数据环境的完善使 BIBFRAME1.0 的“规范”和“注释”两个核心类再无存在的必要,BIBFRAME2.0 顺应环境采纳外界成果、简化结构,使用 URI/IRI 和成熟的关联数据词表,以更高效的方式促进书目与网络环境的互联与交换,不但提高了书目的可见度,也支持用户“探索”发现更丰富的资源,使书目控制更加优化、普遍化,书目系统更加开放、通用。这不仅有利于用户从文献单元深入知识单元进行探索,而且在跨机构、跨数据库的融通和运用中数据不损失其语义,从而为用户实现知识挖掘、促进知识流动提供了良好的基础。

2021 年,美国国会图书馆(Library of Congress,LC)发布了 BIBFRAME2.1,新增了 Hub 类,位于作品、实例、单件之上,相当于 Share-VDE 的超级作品(Super Work),用以解决作品与内容表达、作品与作品之间的关系在 BIBFRAME2.0 中难以区分的问题[12]。这一调整被视为 BIBFRAME 对 LRM 以及 3R 项目的积极回应[13]。

BIBFRAME 经过数年试验,目前在欧美国家已进入推广应用阶段,在我国的应用案例有上海图书馆的家谱知识平台和华东师范大学的数字方志集成平台。上海图书馆是我国第一家应用 BIBFRAME 的图书馆,其开发的家谱知识平台是运用 BIBFRAME 揭示文献知识单元、丰富数据关联性从而帮助用户实现多面向“探索”任务的先例。该项目以 BIBFRAME2.0 为基础设计本体,复用 FOAF 中的术语来描述家谱中的先祖名人,并自定义“谱名、字、号、谥号”等中国历史人物特有的属性予以补充;复用 Geonames 的术语来描述家谱中的谱籍地;复用 Schema.org 和 W3C Organization 的术语来描述家谱资源涉及的收藏机构;复用 W3C Time Ontology 来描述时间信息,自定义了一些属性来描述家谱资源涉及的中国历史朝代信息[14]。该项目还将相关词表、数据集以关联数据形式发布,供其他机构共享重用。经过从 RDB 到 RDF 的数据转换,基于关联数据四原则的系统设计和基于语义技术框架的系统开发,以及利用可视化技术展示数据,利用 Web2.0 技术支持用户贡献知识,家谱知识平台实现了知识导航、知识发现和知识进化功能。以谱籍地名为例,地名不再作为一个以字符串存在的名称,而是一个与真实存在的地点相对应的实体。这个实体不仅有其名称,还有经纬度等 GIS 信息,可以在地图上准确定位,因而可以开发出基于地图的发现和探索功能。使用“地图画圈浏览”功能在地图上圈一个范围,系统会显示谱籍地在该地域范围内的所有家谱。又如,根据姓氏、谱籍、堂号、朝代的聚类展示,基于概念和实体间关联的关系发现,不同家谱中人与人之间的关系、人与地之间的关系的挖掘,对某一家族在地理空间上的迁徙路线、某地域范围内某姓氏的分布情况、某一宗族的散居地覆盖范围等的调查,这类过去需要专业学者皓首穷经方能实现的任务,都可在新技术的应用下实现“远距离”阅读,大大扩展了数据利用场景,释放了数据潜能[14]。

3.3 MARC21 字段变化对“探索”任务的支持

未来 BIBFRAME 取代 MARC 已是可见的趋

势,但当下毕竟尚未进入全面应用阶段。受制于技术和资金门槛、通用词表等数据基础设施薄弱、一线编目工作者对前沿技术缺乏了解等因素,国内图书馆界距离 BIBFRAME 的普及应用这一目标还有很长的路要走,当下大部分编目工作仍然在 MARC 格式下进行。正因如此,编目员深入学习 LRM、从概念模型层面理解书目的发展动向才尤为重要。

具体到对 MARC21 部分字段修订的理解,只有在关联数据的视角下,也就是从满足用户“探索”需求的角度出发,才能明白这些字段设置的意图和价值。为了应对关联数据的发展趋势,MARC21 已在结构与内容方面进行了调整,主要包括 6 个子字段(即 $0、$1、$2、$4、$e 与 $i)与一个字段(即 758)[15]。以小说《傲慢与偏见》为例,参考学者高斌运用上述字段与子字段对 MARC 原始记录进行的丰富化作业[15](本文对部分字段的运用有所调整),可以帮助编目员理解在现有的 MARC 格式下怎样使书目与关联数据环境相适应,促进 LRM 中“探索”任务的达成,具体如下。

原始记录:

100 1# $aAusten,Jane, $d1775-1817.
240 10 $aPride and prejudice. $lChinese
245 10 $aAo man yu pian jian/ $cZhen,Aositingzhu;[Xia Yinghuiyi]
700 1# $aXia,Yinghui.

采用 BIBFRAME 实体进行丰富化作业:

100 1# $aAusten,Jane, $d1775-1817.
$4http://id.loc.gov/ontologies/bibframe/agent(bf:agent)
$0http://viaf.org/viaf/102333412
$2 viaf
$1http://dbpedia.org/page/Jane_Austen
$2DBpedia
240 10 $aPride and prejudice. $lChinese
$4http://id.loc.gov/ontologies/bibframe/instanceOf(bf:instanceOf)
$0http://share-vde.org/sharevde/docBibframe/Work/139617-12
$2share-vde
245 10 $aAo man yu pian jian/ $cZhen,Aositingzhu;[Xia Yinghuiyi]
700 1# $aXia,Yinghui. $etranslator
$4http://id.loc.gov/ontologies/bibframe/agent (bf:agent)
$0http://worldcat.org/identities/np-xia,%20yinghui/
$2worldcatidentities
758 ## $1http://worldcat.org/oclc/213888776

原始记录取自密歇根大学图书馆,著录范围限于中文版《傲慢与偏见》的书目相关信息;经丰富化作业后,各字段关联作用如表 4 所示。

表4　MARC中关联数据化字段与子字段的实例说明

字段	子字段	说明
100	$4	标示字段100与245$a书目实体之间的资源责任关系为“行为者”agent(对应LRM-R5)。
	$0	著录VIAF提供的名称规范URI,除聚合各国语言的作者名称外,还可经由德国国家图书馆提供的URI关联至简·奥斯汀的家族成员,如简·奥斯汀的侄女(Caroline Jane Knight)。
	$2	说明$0中URI的来源。
	$1	著录DBpedia提供的URI,标示真实世界对象(real world object),聚合了简·奥斯汀不同语言的著者名称,也包括简·奥斯汀的不同英文作品。
	$2	说明$1中URI的来源。
240	$4	标示字段240与245$a书目实体之间的关系为:后者是前者的一个实例(对应LRM-R2)。
	$0	著录SHARE-VDE提供的作品规范URI,聚合了美国杜克大学图书馆、纽约大学图书馆、斯坦福大学图书馆、芝加哥大学图书馆、密歇根大学图书馆、宾夕法尼亚大学图书馆、耶鲁大学图书馆以及加拿大阿尔伯塔大学图书馆等有关英文版《傲慢与偏见》作品馆藏。
	$2	说明$0中URI的来源。
700	$4	标示字段700与245$a书目实体之间的资源责任关系为“行为者”agent(对应LRM-R6)。
	$e	与$4搭配使用,进一步说明责任关系为译者(对应LRM-R6)。
	$0	著录WorldCat Identities提供的URI,聚合了译者“夏颖慧”有关“简·奥斯汀(Jane Austen)”的中文翻译作品等相关信息。
	$2	说明$0中URI的来源。
758	$1	记载WorldCat中另一版本书目作为相关载体表现(对应LRM-R3)。

上述案例利用关联数据化字段与子字段,在现有MARC格式下将关联数据资源导入记录,丰富了书目平台用户的探索收获。更重要的是该做法将书目延展向关联数据空间,聚合相同URI不同来源的关联数据环境的外部资源,并且用$4明确标示资源之间的语义关系,这有利于书目信息被其他平台用户发现、利用,提高关联数据检索系统的会话性能,支持获取更为丰富的相关资源,提升“探索式搜索”的用户体验。

4　结语

“探索”是IFLA-LRM中新增的一项任务,结合信息搜索领域“探索式搜索”理论,可以加深对这一用户任务的理解,从而理解LRM模型体现的通用性、关联性,并对模型的建模思路以及实体、属性、关系如何相互效力以更好地融入关联数据环境有更加深入的认识。编目人员从“探索”任务视角了解模型融入关联数据环境的设计意图,进而理解RDA和BIBFRAME历年的更新和调整,有利于其突破传统编目思维,改善实践工作,为编目工作真正从数字化走向数据化做好准备。

参考文献

1　Pat Riva,等. IFLA图书馆参考模型:书目信息的概念模型. 蔡丹,等,译[EB/OL]. [2022-06-18]. https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017_rev201712-ch.pdf.

2　范炜. 书目资源的统一概念模型LRM论纲[J]. 文献与数据学报,2019(4):46-61.

3　Gary Marchionini. Exploratory search from finding to understanding[J]. Communications of the ACM,2006(4):41-46.

4　夏立新,等. 探索式搜索研究进展[J]. 图书情

报工作,2020(4):103-112.
5 Bates M J. The design of browsing and berrypicking techniques for the online search interface [J]. Online Review,1989(5):407-424.
6 Pirolli P,Card S. Information foraging in information access environments[C]//Association for Computing Machinery,Proceedings of the CHI'95 ACM conference on human factors in computing systems. New York:ACM Press,1995:51-58.
7 邢玉艳,刘萍.探索式搜索前后用户认知结构变化研究[J].图书情报工作,2021(22):74-84.
8 胡小菁.文献编目:从数字化到数据化[J].中国图书馆学报,2019(3):49-61.
9 胡小菁.RDA:从内容标准到元数据标准[J].图书馆论坛,2014(7):1-7.
10 王薇,罗翀.3R项目进展及国家图书馆RDA本地化实施的应对策略研究[J].国家图书馆学刊,2018(5):106-112.
11 王景侠.书目格式的关联数据化发展及其启示:从MARC到BIBFRAME[J].图书馆杂志,2016(9):50-56.
12 Hub:BIBFRAME模型下的超级作品[EB/OL].[2022-06-18]. https://catwizard.net/posts/20200628160934.html.
13 王兴兰.BIBFRAME的发展及其在国内外图书馆的应用[J].图书馆杂志,2022(3):79-87.
14 夏翠娟,等.家谱关联数据服务平台的开发实践[J].中国图书馆学报,2016(3):27-38.
15 高斌.MARC21关联数据化的转变与应用[J].图书馆学研究,2021(24):52-68.

(吴诺曼 馆员 四川大学图书馆,王瑞荣 馆员 四川大学图书馆,黄毕惠 研究馆员 四川大学图书馆)

收稿日期:2022-08-09

澳大利亚将公共借阅权扩展到电子书和有声书领域

2023年1月27日,澳大利亚政府宣布将公共借阅权从印刷书扩展到电子书及有声书领域,并计划在4年内追加1290万美元预算。澳大利亚公共借阅权制度始建于1975年,相继于1985年、1996年出台《公共借阅权法案》《公共借阅权计划》,由澳大利亚通讯、信息技术与艺术部拨款,对出版社及著作权人因其作品在公共图书馆被免费出借而给予补偿金,并设置公共借阅权委员会负责计划和实施。

资料来源

1 Australian authors to receive compensation for e-book loans for first time[EB/OL].[2023-02-06]. https://www.brisbanetimes.com.au/culture/art-and-design/australian-authors-to-receive-compensation-for-e-book-loans-for-first-time-20230127-p5cfxk.html.
2 Public Lending Right Scheme 2016[EB/OL].[2023-02-13]. https://www.legislation.gov.au/Details/F2016L01549.
3 Public Lending Right Committee annual reports 2021-2022[EB/OL].[2023-02-13]. https://www.arts.gov.au/publications/public-lending-right-committee-annual-report-2021-22.

(国家图书馆研究院 提供)