



# 关联型 MARC 的发展与应用研究\*

□陈辰 槐燕

**摘要** 在当前 MARC 和关联数据共存的复合环境下,研究 MARC 格式中标识符的部署和应用问题,对于简化从 MARC 到关联数据的转换具有重要意义。以 MARC21 和 UNIMARC 规范格式中设置的标识符为研究对象,识别和讨论文档网络和数据网络环境交叉融合所带来的新旧概念交替应用问题,具体对 MARC 规范格式中的记录控制号和永久记录标识符、本地和远程标识符、规范记录和真实世界对象标识符、RDF URI 和普通网页 URL 等概念和应用进行探讨。结果表明,MARC 规范格式为适应关联数据化发展做出了格式内容上的更新,各类标识符在 MARC 中承载的功能不同,应著录在不同的字段和子字段以示区分。未来仍需结合编目实践进一步研究标识符著录实施问题,包括 URI 与相关对象的对应协调,标识符的识别和选择,URI 链接与传统编目工作的权衡等问题。

**关键词** 关联数据 URI MARC 规范格式 唯一标识符 规范文档

**分类号** G254.36

**DOI** 10.16603/j.issn1002-1027.2024.01.010

## 1 引言

机读目录(Machine-Readable Cataloging, MARC)诞生于 20 世纪 60 年代,是图书馆在计算机环境下描述、存储、交换、控制和检索书目数据广泛遵循的规范<sup>[1]</sup>。尽管 MARC 被图书馆用以编制书目数据已有数十年的历史,但随着开放关联数据实践的开展,其局限性也越发凸显,早在本世纪初就有“MARC 必须死”的观点<sup>[2]</sup>,随后相关组织又提出了不同于 MARC 的 BIBFRAME 和 Schema.org 等关联数据格式。MARC 以文本字符串作为描述和控制对象的方式,不符合关联数据以实体或事物为中心的规范。MARC 记录中缺少可解析引用的 URI,导致机器难以理解字符串语义并进行语义操作,从而影响书目数据在语义 Web 环境下的共享。为应对开放关联数据环境带来的挑战,保证 MARC 继续在图书馆集成管理系统(Integrated Library Systems, ILS)和在线环境中发挥作用,相关组织对 MARC 格式和内容进行了一系列的调整。在 MARC 中添加和融入统一资源标识符(Uniform

Resource Identifier, URI)的研究和实践,即对于“关联型 MARC(Linky MARC)”的探索,推进了 MARC 向关联数据的过渡和转换<sup>[3]</sup>。

URI 构成关联数据生态系统的主干。作为语义网的一种轻量级实现方式,伯纳斯-李(Berners Lee)在 2006 年提出关联数据概念,旨在使用开放标准格式、接口和技术在网络上发布和关联数据,将万维网由文档网络转变为数据网络。数据网络是基于唯一标识符及其互链接的思想,以 URI 命名实体,且经由相同 URI 链接不同来源的数据<sup>[4]</sup>。在 MARC 记录中添加 URI 用以标识和链接实体,可以简化 MARC 到关联数据格式的转换,并将外部资源引入到书目系统中,从而开展知识增强和发现服务。

当前文档网络和数据网络环境产生了交叉融合,在 MARC 格式中调整和增加了融入 URI 的相关字段和子字段,由此带来了新旧概念应用的交替,那么如何在 MARC 字段或子字段中正确填充 URI,避免 URI 相关概念的混淆是摆在编目员面前的主要问题。本研究即关注 MARC 格式中的 URI 应用

\* 国家社会科学基金项目“名称规范控制向身份管理的转型研究”(项目编号:22BTQ038)的研究成果之一。  
通讯作者:陈辰,ORCID:0000-0002-0485-7525,邮箱:chenchensee@163.com。



问题,细致调研和评估 MARC 中承载各类标识符的字段和子字段,以期为编目员在 MARC 中添加和应用 URI 提供参考。

## 2 文献回顾

### 2.1 MARC 格式的调整

相关组织对 MARC 格式进行调整,以增加用于揭示 URI 的字段及子字段。2015 年,合作编目计划(Program for Cooperative Cataloging, PCC)政策委员会召集 MARC URI 任务组,识别和处理在 MARC 格式中部署关联数据标识符的问题,以提高 URI 应用的普及率。MARC 咨询委员会(MARC Advisory Committee, MAC)随后批准了在 MARC 的书目和规范数据中以 \$0 和 \$1 子字段著录 URI 的提案(MARC 提案号 2019-0316<sup>[5]</sup>和 2017-0817<sup>[6]</sup>)。这两个子字段用法的确定,是 MARC21 迈向关联数据的重要一步,实现了从规范控制向身份管理的重大转变。为进一步探索 \$0 和 \$1 子字段的最佳应用实践,PCC 政策委员会于 2019 年授权由 PCC NACO(Name Authority Cooperative Program)的身份管理任务组领导在 MARC 中添加 URI 的试点<sup>[7]</sup>。此外,MARC21 还调整和确定了 \$2、\$4、\$e、\$i 等子字段与 758 字段,均可用于著录关联数据 URI<sup>[8]</sup>。帕帕扎基斯(Papadakis)等还研究了能够著录关联数据信息的 MARC 字段,并对本地和远程定义的 URI 字段进行了区分<sup>[9]</sup>。

### 2.2 MARC 中融入 URI 的实践

在研究适合著录 URI 的 MARC 标准时,图书馆和相关组织还在考虑如何在其记录中融入 URI。PCC MARC URI 任务组开展了“MARC 中的 URI 使用”的专题调查,该调查于 2017 年 8 月至 9 月在相关的图书馆邮件列表中进行分发,最终收到 35 份回复,这些回复的图书馆都对在 ILS 环境中融入 URI 表现出浓厚兴趣<sup>[10]</sup>。美国乔治·华盛顿大学图书馆利用 MarcEdit 开源工具包将 URI 插入 MARC 记录中,外部标识符主要来自虚拟国际规范文档(Virtual International Authority File, VIAF)和美国国会图书馆的名称规范文档(LC-Name Authority File, NAF),URI 嵌入到 1xx、240、6xx、7xx 和 830 中的 \$0 子字段<sup>[11]</sup>。OCLC 参与了多项融入

URI 的 MARC 提案,且发布技术公告声明在其 MARC 记录中融入外部标识符,如在“技术公告 267”<sup>[12]</sup>中,对子字段 \$0 和 \$4 进行了重新定义,以包含 Web 检索协议形式的 URI,在“技术公告 268”<sup>[13]</sup>中,宣布提供 \$1 来容纳真实世界对象(Real World Object, RWO)的 URI。

在链接外部标识符的工具中, MarcEdit 根据新的 MARC 条款进行了调整,通过在 MarcNext<sup>①</sup> 中新增 Linked Identifiers 模块,支持对外部标识符的链接<sup>[14]</sup>。ILS 系统供应商已经或计划在其系统中包含必要的 URI 查找和插入功能,如 Ex Libris 的 Alma 利用其规范控制功能,通过链接已知权威词汇的 URI,自动丰富书目记录信息<sup>[15]</sup>。MARCIVE 和 Backstage Library Works 等供应商为 MARC 书目和规范数据提供 URI<sup>[16]</sup>。Casalini Libri 在其共享虚拟发现环境(Share Virtual Discovery Environment, SHARE-VDE)中,使用外部 URI 资源来丰富原始 MARC 记录,从而开展增强实体发现和关联数据服务。加利福尼亚大学戴维斯分校图书馆和 Zepheira 合作的 BIBFLOW 项目,其全称为“重塑编目:未来图书馆运作的模式”,成功修改了开源的 KUALI-OLE 描述模块,以在 LC、OCLC、VIAF 和 Getty 词汇表上执行关联数据查找,并将捕获的 URI 插入 MARC 记录<sup>[17]</sup>。

由上述分析可见,相关机构对“关联型 MARC”的理论和实践已经进行了积极探索,但目前只有少数图书馆使用了由 ILS 系统供应商提供的融入外部标识符的服务,图书馆在关联数据环境下的 URI 应用还并未普及。“MARC 与 LD(Linked Data)复合环境”<sup>[18]</sup>下 URI 应用实践的复杂性,如 MARC21 和 UNIMARC 在著录 URI 方面的差异,本地规范记录控制号和永久记录标识符的不同,规范记录和真实世界对象的区分,URI 和 URL(Uniform Resource Locator, 统一资源定位符)的比较,严重影响了图书馆应用 URI 的实践进程。鉴于此,本研究对 MARC 应用 URI 的相关问题进行讨论,阐述为何在 MARC 记录插入 URI,即产生“关联型 MARC”的内在机理,并在澄清相关概念的基础上,考察评估 MARC 中承载各类标识符的能力和范围,最后对有待进一步研究的问题进行探讨和展望。

① MarcNext 为集成在 MarcEdit 中的一个工具包。



### 3 关联型 MARC 产生的内在机理

“关联型 MARC”(Linky MARC)是由曾任职于 OCLC 的技术专家理查德·沃利斯(Richard Wallis)在 2018 年提出的概念,主要是为了区别于已有 MARC 语法中 URL<sup>[3]</sup>。它作为将 MARC 完全转换到基于 BIBFRAME 2.0 或 Schema.org 格式的临时过渡方案,为 MARC 记录中的相关实体(包括个人、组织、作品等)链接唯一、可识别、可操作、可共享的标识符,在 MARC 记录中使用 URI,直到完全转换到关联数据系统。以 URI 形式给出的全球唯一标识符是在语义 Web 上有效分发数据的关键要求<sup>[7]</sup>,这种受控标识符与文本字符串相比,具有语义清晰和机器可操作的优点,是资源描述框架(Resource Description Framework, RDF)建模方法的核心。

当前 MARC 记录中缺乏可解析引用的 URI,是其与关联数据不兼容的主要原因<sup>[19]</sup>。MARC 虽然对各字段有严格的规定,但其字段多由文本字符串控制,缺少 URI 标识,导致机器难以理解其丰富的语义,在进行关联数据解析和搜索时会丢失原有语义。基于特定 ISO 2709 格式产生的 MARC 数据,只能局限在图书馆系统内进行处理和使用,不能被关联数据信息检索系统采集和共享。国际编目界选择用关联数据取代 MARC 数据,其主要原因是关联数据使用 URI,比 MARC 文本字符串的传播范围更广。在 MARC 字段中添加机器可解析用的 URI,将简化从 MARC 到关联数据格式的转换,其优点如下:

一是,帮助维护规范信息,实现规范控制。使用 URI 可以帮助编目员快速确立标目信息,编目过程比手动更快,并可为规范描述带来更多的信息内容<sup>[20]</sup>;二是,添加 URI 可为 MARC 转换为 RDF 做准备。URI 是 RDF 模型的主要语法,在 MARC 记录中添加 URI 将有助于 MARC 向其他数据模型(如 BIBFRAME 或 Schema.org)的转换,并保留尽可能多的细节;三是,参引 URI 可将外部数据引入到目录环境中,以开展增强式目录发现服务。如 SHARE-VDE 通过应用外部 URI 资源,构建了驱动式联合书目发现环境。即使不构建关联数据,也可在原有目录系统中提供到外部数据的链接选项,以人类可读的方式在书目记录中显示,使目录数据更丰富<sup>[21]</sup>。再有,在书目数据中添加 URI 可以使转换工作普及化,编目员只需在当前工作环境中即可参与关联数据的创建<sup>[22]</sup>。

URI 是支持关联数据生态系统的重要标识和解析技术,在 MARC 中引入 URI 作为过渡方案,一方面保证了 MARC 继续在 ILS 中发挥信息组织功能(对自动化管理系统影响较小),另一方面又为 MARC 完全转换为关联数据做调整准备,这构成了“关联型 MARC”的内在需求或机理。

### 4 对 MARC 中的 URI 相关字段的调研

本研究的调研对象为 MARC21、UNIMARC 和 CNMARC 规范格式,旨在识别其新增或重新定义的著录 URI 的相关字段。由于 CNMARC 是等效采用 UNIMARC,所以本研究重点关注了 CNMARC 与 UNIMARC 相比尚未同步更新的字段。为了进行“MARC 和 LD”复合环境下 URI 相关概念的区分和比较,本研究除了关注承载关联数据 URI 的字段外,还将与 URI 相关的标识符都包含在调查范围内,包括对记录控制号和永久标识符,本地和远程标识符,规范记录和真实世界对象标识符,以及 RDF URI 和普通网页 URL 等概念和应用上的比较,希望为图书馆及相关机构正确使用 URI 提供参考。特定领域的标识符如国际标准书号(ISBN)、国际标准连续出版物号(ISSN)、国际标准文本代码(ISTC)、音乐导句等标识符不在本调查范围内。最终调查结果如表 1 和表 2 所示。

### 5 MARC 中的 URI 相关字段的应用

#### 5.1 记录唯一标识符

MARC 规范格式标识块有多项关于唯一标识符的字段,各字段的含义与功能相似但有所不同,在应用上容易引起混淆,本研究从如下三方面进行比较研究,以期对相关人员进行正确理解 MARC 字段并有效填充字段提供参考。

##### 5.1.1 记录控制号和永久记录标识符

MARC21 和 UNIMARC 的 001 字段,分别为控制号或记录标识符,虽名称略有差异,但表示的含义基本相同,为本地规范记录的唯一标识符,是制作、使用或发行记录的机构用作唯一标识的控制号。该号既可以自行构建,也可使用其他独立的标识符(如 ISBN 或国家书目号)作为本地规范记录标识符。如在 MARC21 中,可将 010(LC 控制号)或 016(除 LC 以外的国家书目机构控制号)的号码赋值本系统的记录标识符。



UNIMARC 的 003 字段表示永久记录标识符, MARC21 中没有与之相对应的含义相同的字段, 而是在 856 字段中新增子字段 \$g 来著录永久记录标识符, 表示制作记录的机构用作的永久记录标识符, 与 001 字段的记录标识符相比, 两个号码都具有唯

一性特点, 能够保持标识符和记录之间的一一对应关系。两个字段都表示规范记录(描述实体的记录)的唯一标识符, 而不是实体本身的标识符。编目员在填充应用这两个字段时, 还应注意其在含义、范围、特点和形式等方面的不同, 具体如表 3 所示。

表 1 MARC URI 相关的字段和分子字段

字段	MARC21	UNIMARC	CNMARC
001	控制号	记录控制号	记录控制号
003	控制号标识	永久记录标识符	永久记录标识符
010	美国国会图书馆(LC)控制号	国际标准名称识别码(ISNI)	——
015		国际标准规范数据号(已废)	国际标准规范数据号(已废)
016	国家书目机构控制号		
017		其他标准号	——
024	其他标准标识符		
033		其他系统永久记录标识符	其他系统永久记录标识符
035	系统控制号	其他系统控制号	其他系统控制号
810 \$a		参考数据源	参考数据源
856 \$g, \$h, \$u	电子资源定位与检索 \$g(永久标识符) \$h(失效 URI) \$u(URI)	电子资源定位与检索 \$u(URI)	电子资源定位与检索 \$u(URI)

表 2 MARC URI 相关的子字段

子字段	MARC21	应用的字段	UNIMARC	应用的字段
\$0	规范记录控制号或标识符	022, 024, 034, 043, 050, 052, 053, 055, 060, 065, 070, 075, 080, 087, 260, 360, 5xx, 7xx (除 788), 672, 673, 883, 885	说明语	4xx(除 431, 432)、5xx(除 531, 532)、640, 741, 742
\$3	——	——	规范记录控制号	241, 242, 4xx(除 431, 432 字段) 5xx, 6xx(除 610, 640 字段) 7xx (除 731, 732 字段)
\$1	真实世界对象(RWO)	022, 024, 034, 043, 050, 052, 053, 055, 060, 065, 070, 075, 080, 087, 260, 360, 5xx, 7xx (除 788), 672, 673, 883, 885	连接数据	240, 241, 242, 245, 430, 440, 441, 442, 445, 540, 541, 542, 545, 641, 642, 740, 741, 742, 745
\$u	统一资源标识符	031, 046, 670, 677, 678, 856, 883, 884	统一资源标识符	036, 334, 378, 608, 822, 856

表 3 记录控制号和永久记录标识符

	001(MARC21, UNIMARC)记录控制号	003(UNIMARC)、856 \$g(MARC21)永久记录标识符
含义	制作记录的机构用作唯一标识符的控制号	一个与数字对象关联的唯一、永久的字母数字字符串
范围	本地唯一识别号	Web 全局唯一性
特点	唯一性(本地)	唯一性、永久性、可描述、可互操作和可治理
形式	文本形式, 示例: 78-34279 FRBNF11911553 978-7-5013-5210-4 0115E5000F	互联网地址形式的 ARK, handle, DOI 等, 示例: <a href="http://catalogue.bnf.fr/ark:/12148/cb119115538">http://catalogue.bnf.fr/ark:/12148/cb119115538</a> <a href="http://errol.oclc.org/laf/nb2001-72552.html">http://errol.oclc.org/laf/nb2001-72552.html</a> <a href="https://orcid.org/0000-0003-0764-7889">https://orcid.org/0000-0003-0764-7889</a> <a href="https://isni.org/isni/0000000046506825">https://isni.org/isni/0000000046506825</a>



从含义上看,永久记录标识符是一个与数字对象关联的唯一、永久的字母数字字符串,是对数字资源的永久引用<sup>[23]</sup>,通过特定的协议可在网络上进行全局解析,将唯一标识符转换成相应的物理存放地址。而本地规范记录控制号,通常是由图书馆在创建记录中用作唯一标识符的控制号,在进行关联数据化转换之前通常是文本形式,并不支持全局解析功能。

从范围上看,记录控制号是制作记录的机构用作本地范围的唯一控制号,只在本地具有意义,不适用于在不同系统间进行交互使用。而永久记录标识符具有全局唯一性,是为记录定义的 Web 级唯一标识符。

从特点上看,相对于记录控制号,永久记录标识符最大特点是其“永久性”,该永久性是指以 HTTP URI 的方式支持内容协商,标识符不会随资源对象物理存储位置的改变而发生改变,这也是与普通的 URL 的区别(由于技术或人为原因,URL 可能会中断)。永久标识符可对规范记录进行参引,将 URI 加入到 Web 网站、博客、Email 等中参照引用<sup>[24]</sup>。永久标识符通常由专门机构提供管理和维护服务,从而保证其永久性,即使资源的位置发生了变化,也能随时对资源进行重新定位。从这个层面看,永久标识符不仅是数据基础设施的唯一标识符元素,作为一种服务,也能确保标识符解析到最新的资源访问位置<sup>[25]</sup>。

唯一性和永久性是标识符的重要属性,但并不能完全包含现有数据基础设施中对标识符所需的特性。为此,在永久标识符的基础上,ODIN(ORCID and DataCite Interoperability Network)联盟引入了“可信标识符”这个术语,将其定义为唯一的、永久的、可描述的、可互操作的和可治理的标识符<sup>[26]</sup>。可描述性,是指通过内容协商机制,将用户链接到一组描述实体的元数据,支持实体信息消歧或支持检索;可互操作,是指通过描述其关系的元数据元素与其他标识符关联。可治理,是指由一个组织将标识符分配、发布和管理作为其主要任务,有可持续的商业模式和一大批同意共同程序和政策的成员组织,有一个管理机构并致力于使用开放技术。这些特性代表了永久标识符的发展方向。

从形式上看,本地记录控制号通常以文本形式存在,表现为单个字段的的标准数字;永久标识符是一

种特定的互联网地址,允许在超文本链接的帮助下引用电子资源,同时确保这个链接不会改变。常见的永久标识符系统包括数字对象标识符(Digital Object Identifier, DOI)、档案资源标识(Archival Resource Key, ARK)、Handle、永久统一资源标识符(Persistent Uniform Resource Locator, PURL)、统一资源名称(Universal Resource Name, URN),以及开放研究者和贡献者标识符(Open Researcher and Contributor ID, ORCID)和 ISNI 等,其表现形式示例如表 3 所示。

需要指出的是,在记录控制号(或国家书目记录控制号)基础上可构建永久标识符,通过制定构建模式,将文本表示为永久的 HTTP URI,使其具有机器可操作性,支持内容协商。此类 URI 的创建模式一般使用标准路径作为前缀,使用记录控制号作为后缀。如 LC 的永久记录标识符构建就是通过前缀路径(<https://lcn.loc.gov/>) + LC 控制号构建,如 LC 名称标目为“Shakespeare, William, 1564-1616”的规范记录,其永久标识符为 <https://lcn.loc.gov/n78095332>,是在 LC 控制号(n 78095332)的基础上构建的“莎士比亚”规范记录的永久链接,可将其记录在新增的 856 \$g 子字段中。

需要注意的是,关联数据 URI 和永久记录标识符不完全等同。如 LC 在关联数据方面专门推出了 id.loc.gov 服务,其规范记录的 URI 构建通过前缀路径(<http://id.loc.gov/authorities/names/>(或 subject)) + LC 控制号的模式构建,如莎士比亚的关联数据 URI 为:<http://id.loc.gov/authorities/names/n78095332>。在关联数据应用于图书馆编目之前,永久标识符对于是标识记录还是实体,并未进行明确区分,这源于传统的编目规则未对实体和名称进行区分的做法,实体通常由名称或结合其他描述元素进行表达<sup>[27]</sup>。

### 5.1.2 本地标识符和远程标识符的区分

本地标识符是由使用标识符的图书馆自己创建,而远程标识符是由其他图书馆或机构创建。001(MARC21\UNIMARC)是本地机构定义的标识符,035(MARC21\UNIMARC)著录其他系统控制号,即远程定义的标识符。035 字段记录的控制号是除 001 字段(或 MARC21 中的 010 或 016 字段)以外的,但是存在于其他系统的 001 字段中。在 UNIMARC 格式中,003 和 033 字段,分别表示永久记录



标识符和其他系统永久记录标识符,即本地定义和远程定义标识符的区别。而 MARC21 的 856 \$ g 的永久记录标识符目前并未区分本地和远程。

### 5.1.3 其他系统永久记录标识符和其他系统控制号

033 (UNIMARC) 字段表示远程定义的其他系统的永久记录标识符,即从其他来源获得记录的永久标识符。035 (MARC21\UNIMARC) 其他系统控制号,表示远程的其他系统或数据来源定义的记录控制号。两者的不同,即为本地规范记录控制号和永久规范记录标识符的区分,但 033 字段中的永久记录标识符在万维网上等同于另一个数据库中记录的 035 字段中系统控制号<sup>[24]</sup>。

有关 001、033 和 035 字段的应用示例,如法国国家图书馆创建的记录,被整合到 Sudoc (法国国家学术联盟目录) 后,在 Sudoc 中会产生一个新的规范记录控制号(本地规范记录号),生成到 001 字段,那么原法国图书馆规范记录的控制号则著录在 035 字段中,而原法国图书馆永久记录标识符则著录于 033 字段中<sup>[24]</sup>,如:

001 099573598

033 ## \$a <http://catalogue.bnf.fr/ark:/12148/cb40133622z/PUBLIC>

035 ## \$a (FrPBN)FRBNF401336220000001

### 5.1.4 结论

在 MARC 标识块,记录控制号和永久记录标识符具有不同的表现,编目员可从含义、范围、特点和形式等方面进行区分。本地控制号为本地范围内的记录识别和管理而设置,而对远程标识符的著录可

提高本地系统与外界数据的交互。其他系统永久记录标识符和其他系统控制号都是外部系统构建的唯一识别号,不同在于是否具有永久性的特点。

## 5.2 规范记录和真实世界对象标识符

### 5.2.1 关联数据环境下的命名机制

语义网基础设施依赖于对实体的唯一识别,关联数据需要区分实体的规范记录和真实世界对象,如一个人和一条关于这个人的 MARC 规范记录是不同的,在 RDF 表示中,需要用不同的 URI 进行唯一识别,以获得清晰的语义。如果在 MARC 记录中著录 URI 信息时没有区分规范记录和真实世界对象,RDF 转换分析时则无法准确生成实体之间的关系<sup>[6]</sup>,如:

100 1# \$a Shakespeare, William, \$e author, \$d1564-1616,  
\$0 <<https://id.loc.gov/authorities/names/n78095332>>

\$0 子字段表示莎士比亚这条规范记录的 URI,当进行 RDF 解析时生成的是《哈姆雷特》这个著作是由威廉·莎士比亚这条规范记录编写,而不是指向莎士比亚这个人。因此,需要不同的 URI 区分实体的规范记录(实体的描述)和实体(真实世界对象)本身。

### 5.2.2 规范记录和真实世界对象的比较

为支持语义网环境下真实世界对象和规范记录分离的规则和趋势,MARC21 做了调整和区分,2019 年 MAC 批准通过了定义 \$0 和 \$1 来分别容纳规范记录和 RWO URI 的提案,其不同表现如表 4 所示。

表 4 MARC21 中的 \$0 和 \$1

	规范记录标识符(\$0)	真实世界对象标识符(\$1)
定义	表示规范记录控制号或标准编号,包含相关规范或分类记录的系统控制号,或标准标识符(如 ISNI)	标识实体的 URI,实体又称事物、真实世界对象或 RWO,无论是实际的还是概念的
范围	\$0 主要是关于图书馆社区对标目的管理及其标目来源的依据,并且是相对封闭的	在关联数据术语中,\$1 是身份管理的开放世界解决方案
结构	\$0 包含遗留和语义 Web 编码的混合,可以是文本或 URI 形式 \$aRoyal Society of Medicine Services (Great Britain) \$0(DLC)n ## 86108151 # \$aVine, Barbara, \$d1930 \$0(isni)0000000368647620 \$0 <a href="http://id.loc.gov/authorities/names/n88276221">http://id.loc.gov/authorities/names/n88276221</a>	\$1 包含符合关联数据规则的 URI \$aYeats, W. B. \$q(William Butler), \$d1865-1939 \$1 <a href="http://id.loc.gov/rwo/agents/n78095579">http://id.loc.gov/rwo/agents/n78095579</a> \$aLennon, John, \$d1940-1980 \$1 <a href="http://www.wikidata.org/entity/Q1203">http://www.wikidata.org/entity/Q1203</a> \$1 <a href="http://viaf.org/viaf/196844">http://viaf.org/viaf/196844</a>

从定义上看,\$0 子字段表示规范记录描述的控制号或标识符,其 URI 指向实体名称或标签的 URI,解析时重定向到实体记录的描述。记录是关于标目的特性信息,比如标目来源、修订历史,或者

用多种语言或文字表示,为权威标目来源提供依据。\$1 子字段指向真实世界对象,既可以是现实世界中的对象(如人、机构、地理位置等),也可以是概念上的抽象思想和不存在事物(如神话中的独角



兽)<sup>[28]</sup>。URI 指向实体或事物本身,解析重定向到实体本身信息的描述,如实体的传记细节、地理空间坐标、影响范围以及照片或其他图像等。

从用法范围上看,\$0 子字段主要针对图书馆社区对标目的管理,并且相对封闭。\$1 子字段则超出图书馆社区,成为身份管理的开放解决方案,专门用于关联数据应用程序,如聚合、事实提取、消除歧义和标识解析。在 MARC21 中应用的字段如表 2 所示,\$0 和 \$1 子字段可同时出现,也可以单独或组合出现在 MARC21 字段中,并且 \$0 和 \$1 子字段自身都可重复。

从结构上看,\$0 子字段包含遗留和语义 Web 编码的混合形式,既包括规范控制号文本形式,也包括 Web 检索协议的 URI。如果标识符是文本,则控制号或标识符前面会有适当的 MARC 组织代码或标准标识符源代码。\$1 子字段是符合关联数据规则表示的 Cool URI,通常要满足 W3C 对语义 Web URI 的定义,可利用 303 URI 和散列 URI 两种解决方案来标识真实世界对象。

目前 MARC21 通过 \$0 和 \$1 对规范记录和真实世界对象进行了区分,而在 UNIMARC 中并没有进行区分,只用 \$3 子字段连接规范记录的记录控制号(UNIMARC 中 \$3 子字段的应用如表 2 所示)。<\$3 子字段包含在标记 001 中定义的记录控制号<sup>[29]</sup>。

### 5.2.3 结论

传统规范控制的标目与实体对象是混淆的,记录描述的是标目内容,检索指向的是标目背后的对象。为了与语义 Web 社区保持一致,区分规范记录描述和真实世界对象,MARC21 利用 \$0 子字段指向引用描述事物记录的 URI 或者编号,利用 \$1 子字段指向直接引用事物(RWO)的 URI。\$0 子字

段关注来自图书馆社区相对封闭的名称标题,如“规范记录控制号或标准号”, \$1 子字段则以符合关联数据标准的 URI 方式,解决了开放世界身份管理问题,当进行解析时,URI 可指向该实体的描述。MARC21 中有多个字段可通过 \$0 和 \$1 子字段连接,而 UNIMARC 中只有 \$3 子字段表示规范记录描述。

## 5.3 RDF URI 和普通网页 URL

### 5.3.1 URI 相关概念

URI 是一个紧凑的字符序列,用于标识抽象或物理资源,通常以 URL 或 URN 的形式指定<sup>[30]</sup>。其中,URL 除了标识资源之外,还提供了一种访问和定位资源的方法,例如 http://或 ftp://或 mailto:等。URN 可以唯一地标识资源,但不一定指定其位置或明确如何访问资源。URI 旨在成为资源的永久标识符,与可能发生改变 URL 不同,URI 标识了信息资源的基础数据,该数据是保存在各种数据库中的特定二进制数据集<sup>[10]</sup>。

RDF URI 是语义网环境下的实体标识符表现形式,满足 W3C 对语义 Web URI 的定义,是可解析引用的 URI,支持机器可操作(Machine Actionable)或机器可遍历(Machine Traversable)。而国际化资源标识符(Internationalized Resource Identifiers, IRI)是 URI 方案的扩展,包含 Unicode 字符的全部范围<sup>[31]</sup>。IRIs 对于希望在各种文字中生成永久标识符的机构非常有用。

### 5.3.2 RDF URI 和网址 URL 的区分

对于编目员或其他用户而言,RDF URI 看似与用于浏览 Web 的 URL 是同一事物,然而两者在标识对象、功能、格式等方面存在差异,并记录在不同的 MARC 字段和子字段中,以便于数据处理和互操作。具体区分如表 5 所示。

表 5 RDF URI 和网址 URL 的区分

	RDF URI	网址 URL
对象	实体的 URI(真实世界对象)	文档(网站、网页)的 URL
功能	面向机器可读可操作,支持关联数据	面向人类可读,用以标识网站、文档或网页
格式	(a)http://id.loc.gov/rwo/agents/n95004729 (b)http://dbpedia.org/resource/Astronaut (c)http://www.wikidata.org/entity/Q488085	(a)http://id.loc.gov/rwo/agents/n95004729.html (b)http://dbpedia.org/page/Astronaut (c)https://www.wikidata.org/wiki/Q488085
相关 字段	024(MARC21)、017(UNIMARC)、033、822(UNIMARC)、7xx (UNIMARC)具有不同语言的外部资源的 URI	856 的 \$ u 810(UNIMARC)、670(MARC21)



从标识对象看,RDF URI 标识关联数据环境下的实体或真实世界对象,是为机器使用而设计的,支持在 Web 上公开和互连数据。网址 URL 是一种重要的 URI 类型,主要用以标识人类可读的网站、文档或网页,为其提供访问地址。

从功能上看,RDF URI 主要用于供机器使用,但是为了帮助人类用户理解 URI 标识的内容,RDF URI 经常参引到人类可读的网页 URL<sup>[32]</sup>。如试图从 Web 浏览器访问机器可读的 RDF URI: <http://id.loc.gov/rwo/agents/n95004729> 时,浏览器自动将请求重定向到 <http://id.loc.gov/rwo/agents/n95004729.html> 的 Web 页面,可见机器可读的 URI 和人类可读的 URL 之间的区别非常微小,很容易被忽略。如表 5 所示的格式示例,在 dbpedia 重定向中,机器可读 URI 中的单词“resource”被替换为“page”,在 Wikidata 中,“entity”一词被“wiki”取代。

### 5.3.3 相关字段分析

在容纳 RDF URI 和网址 URL 的 MARC 字段方面,由于两者在标识对象、功能等方面存在差异,应著录在不同的 MARC 字段。

#### (1) RDF URI 的 MARC 字段分析

##### ① 标识块

MARC21 的 024 字段用以著录 RDF URI,024 是与 1xx 功能块中命名的实体相关联的标准编号、代码或 URI。\$0 子字段标识规范记录的 URI(<http://id.loc.gov/authorities/names/n78095579>), \$1 子字段标识真实世界对象的 URI(<http://id.loc.gov/rwo/agents/n78095579>)。

UNIMARC 的 017 字段和 033 字段可用于著录 RDF URI,但目前并没有区分规范记录和真实世界对象,但本研究认为未来可利用 033 字段著录规范记录的 URI 形式(不排除著录标识符的文本形式)。017 字段用于著录真实世界对象的 URI 形式,它是与 2xx 功能块中命名的实体相关联的标识符,是用于标识实体的唯一的、永久的、国际公认的字母数字代码。

##### ② 7xx 功能块

7xx(MARC21)用于标目连接款目,如 700 字段表示标目连接款目一个人姓名,等同于著录在 1xx 功能块或相同记录中的标目,定义的 \$0 和 \$1 子字段都可著录 RDF URI,用于链接系统内的标题、来自不同词表或规范文档的标题以及不同关联数据源

的资源。

7xx(UNIMARC)表示 2-功能块中检索点的并行或替代语言和/或文字形式,如 700 字段表示其他语言和/或文字的规范检索点一个人名称。200 和 700 字段是有关同一实体的记录,但 700 比 200 字段多定义了一些子字段,如子字段 \$3 用以著录相应的远程资源,\$2 子字段定义远程资源的 MARC 机构代码,\$8 子字段著录相应资源的编目或基本检索点语种。与 MARC21 中的等效字段相比,UNIMARC 700 字段被明确地定义为有别于 200 字段使用的另一种语言的资源,因此,该字段可适用于著录具有不同语言的远程关联数据资源的 URI。

##### ③ 8xx 功能块

822(UNIMARC)用于容纳其他系统的相对应的映射数据,为 2019 年新增字段,此字段包含有关外部系统中相应数据的信息,其目的是实现系统之间的互操作性。定义了子字段 \$u(URI),用于识别与规范实体对应的外部系统 RDF URI。其范围是各类主题系统、地理数据库、分类表、规范文档等的关联数据源,如 LC 和 Geonames 的关联数据源,如下示例表示论题主题检索点 Fant? mes 与 LC 的 Ghosts 具有精确匹配关系,如:

```
250 # $aFant? mes
822 1# $aGhosts $2lc $uhttp://id.loc.gov/authorities/subjects/sh85054821$d2017-02-09
```

#### (2) 网址 URL 的 MARC 字段分析

为适应数字网络环境下 MARC 记录与外部资源描述链接的问题,20 世纪末在书目和规范记录中新增了 856 字段(电子资源定位与检索),记录所描述实体的电子文献获取信息<sup>[33]</sup>。在规范记录中,子字段 \$u,指向补充实体信息描述的 Web 地址,因此,856 \$u 子字段著录的 URL 主要用于位置访问,而不是 RDF URI 的标识目的,前者与文档网络兼容,后者与数据网络兼容。

MARC21 在 2022 年重新定义扩展了 \$u 子字段,以便在单个 856 字段中著录永久标识符和多个 URL 的记录<sup>[34]</sup>。永久标识符旨在解决 URL 的链接失效问题,在 856(MARC21)中不著录 RDF URI,但是可著录记录实体的永久标识符(在未涉及关联数据应用时,永久标识符并未区分记录和实体)和网页地址的 URL。





670(MARC21)和 810(UNIMARC)字段,表示与检索点或实体相关的参考数据源。其中,670(MARC21)设有 \$u 子字段,用以记录参考的数据来源或出处,使得所提供信息足以确认作品及其所在位置。但该字段应著录 URL,不用以著录远程关联数据资源的 URI。

810(UNIMARC)没有专门定义 \$u 子字段,但一些机构应用 \$a(源引用)来著录相关资源的 URL。无论是 670(MARC21)和 810(UNIMARC)都不应著录关联数据 URI,而用以著录网页地址 URL。需指出的是,810 和 822 字段都可以著录外部数据源,但前者是参考数据源,后者是实体映射数据。

#### 5.3.4 结论

根据对 MARC 字段的原有含义及最新调整的分析,同时结合 RDF URI 以及网址 URL 的标识对象和功能表现,表明两者应著录在不同的字段或子字段,以适应当前“MARC 和 LD”的过渡环境。网址 URL 主要与传统的文档网络兼容,而 RDF URI 与数据网络兼容,著录 URL 的目的是链接到外部文档,用以补充相关实体的描述信息,而著录 RDF URI 目的是链接到数据,支持实体识别和数据之间的关联和发现。在 MARC 中通过设置不同的字段容纳 URL 和 RDF URI,既可以满足当前 ILS 系统中的描述电子资源信息的需要,又可以为 MARC 转换为 RDF 做准备,进一步探索基于关联数据环境的目录发现服务。

## 6 总结与展望

### 6.1 总结

在 MARC 中融入 URI,创建“关联型 MARC”是将 MARC 转换为关联数据的过渡手段,利用 URI 控制、区分和关联实体资源,不仅可将外部资源引入图书馆目录中,还将简化 MARC 到关联数据格式的转换。由于当前图书馆正处于“MARC 和 LD”共存的复合环境,文档网络和数据网络环境产生了交叉融合,会带来新旧概念应用的交替,其中与 URI 相关的概念难免会给编目员带来应用上的混淆。本研究通过对记录控制号和永久记录标识符、本地和远程标识符、规范记录和真实世界对象标识符、RDF URI 和普通网页 URL 等概念和应用进行区分探讨,以期对相关编目机构理解和使用 URI 提供参

考。本研究只对 MARC 中的 URI 应用进行了概念上的探讨,但具体实施还面临多方面的问题,需要未来进一步研究。

### 6.2 未来研究问题

#### (1) URI 与相关对象的对应协调问题

研究 MARC 子字段通过单独或组合方式与 URI 对象对应。一是,URI 子字段的可重复性带来的实施问题,如同一字段 \$0 和 \$1 子字段可重复,可能造成无法与其他子字段或对象的精确对应问题。德国国家图书馆的做法是仅在指定同一实体时才可重复 \$0 子字段,此类问题可通过最佳实践来解决<sup>[35]</sup>。二是,URI 子字段与相同字段中的其他数据之间的关系,如 033 字段中的事件日期、事件和地点,\$0 子字段明确规定与 \$p 子字段(地点)相关。但多数情况下相关的子字段并没有指定,这将很难确定运用 URI 子字段的最佳实践<sup>[36]</sup>。为此 PCC 成立 MARC Object/URI Reconciliation subgroup 专题组,研究 URI 和 MARC 对象的对应关系<sup>[37]</sup>。关于哪些子字段标识的资源可以表示为 RDF 对象的问题,PCC MARC URI 任务组认为不可能制定适用于所有 MARC 字段的通用规则,而是需要利用编目员实践经验来提供分析,并枚举相关子字段,分析隐含的对应关系,以支持利用自动化手段准确地将 MARC 转换为 RDF<sup>[10]</sup>。

#### (2) 标识符的识别和选择问题

外部标识符来源的多样性给编目员带来识别和应用上的复杂性。一是,\$0 和 \$1 子字段的区分,需要编目员理解标识符是表示描述还是表示事物,明确每个标识符记录在哪个子字段中,但编目员由于经验和认知的不同,会产生不一致的做法。如对于 ISNI 是著录到 \$0 还是 \$1 子字段中,并未达成一致。MARC21 标准示例中,出于 ISNI 最初数据来源于图书馆规范记录的考虑,将 LC-NAF 和 ISNI URI 都著录在 \$0 中,但也有人认为 ISNI 与 VIAF、ORCID 等其他标识符相同,应著录到 \$1 中,\$0 只保存由图书馆或其他机构创建的规范记录标识符。二是,图书馆与其他社区对于真实世界对象的范围并未形成一致看法,如图书馆社区对于主题标目是否命名为 RWO 并未形成一致的、明确的看法<sup>[38]</sup>。

W3C 对 RWO 的定义确实包含“抽象的想法和不存在的东西,比如神话中的独角兽”,从这点看,主



题标目属于 RWO 范围。传统图书馆规范控制将人名规范文档和主题规范文档进行了区分,如同一个人名实体(如 Shakespeare),既可以作为责任者标目又可以作为主题标目,在对其进行关联数据化转换时,应将责任者标目和主题标目合并为一个面向对象实体的视图。如 LC 规范关联数据服务,将原来的责任者标目和主题标目只创建了一个 RWO URI,而不再单独为主题标目单独创建 URI。德国集成规范文档(Gemeinsame Normdatei, GND)也放弃以前对主题编目和描述性编目的规范控制之间的正式区分,即不再对人名是名称规范还是主题规范进行区分,而是将其合并为一个基于实体人的规范记录<sup>[39]</sup>。

再有,外部标识符来源的选择偏好问题。数据发布者使用和提供 URI 的模式千差万别,编目员并不能判断哪些 URI 源是首选,但为了帮助图书馆及相关机构使用关联数据,PCC 身份管理任务组编制了《创制和获取 URI:常用词表和参考源指南》报告<sup>[20]</sup>,收录了本领域及互联网上其他可用数据集,便于选择适用本地用例的 URI 源。

### (3) URI 链接与传统编目工作的权衡问题

URI 在关联数据环境中对于实体消除歧义和信息发现至关重要。在 MARC 中添加 URI,使得图书馆可以使用越来越多的资源,链接外部的 URI 越多,提供的信息就会越全面,本地所需的描述工作就越少。但是如何确定链接外部 URI 和本地描述工作的平衡点,以及在多大程度上使用非传统信息源(如图书销售商和维基百科<sup>[40]</sup>),是需要未来考虑的重要问题。编目描述规则在卡片或 MARC 编目环境中发挥了重要作用,在未来大多数实体都可以通过 URI 识别的环境中,如果可以从其他数据源(例如出版商或供应商)检索该信息,编目员还需要创建多少描述性数据?也需要结合相关实践经验进一步研究和确定。需要肯定的是,随着 URI 在 MARC 记录中的融入,图书馆未来规范控制工作的重心会有所调整,由重视对名称文本字符串的创建,到重视实体之间的区分和链接,使得编目员的工作重心不再是描述实体,而是链接和关联实体,依靠外部标识符进行数据的识别和增强。

### 参考文献

1 吴丹. 近五年国内外元数据研究进展[M]//陈传夫. 图书馆学研

究进展. 武汉:武汉大学出版社,2010:382.  
2 Tennant R. MARC must die[J]. Library Journal, 2002, 127(17): 26-28.  
3 Wallis R. The three linked data choices for libraries[EB/OL]. [2023-02-10]. <https://www.dataliberate.com/2018/05/22/the-three-linked-data-choices-for-libraries/>.  
4 Heath T, Bizer C. Linked data: evolving the Web into a global data space[EB/OL]. [2023-02-26]. <http://linkeddatabook.com/editions/1.0/#htoc45>.  
5 Library of Congress. MARC proposal no.2019-03: defining subfields \$0 and \$1 to capture URIs in field 024 of the MARC 21 authority format[EB/OL]. [2022-12-23]. <https://www.loc.gov/marc/mac/2019/2019-03.html>.  
6 Library of Congress. MARC proposal no.2017-08: use of subfields \$0 and \$1 to capture uniform resource identifiers(URIs) in the MARC 21 formats[EB/OL]. [2022-12-23]. <https://www.loc.gov/marc/mac/2017/2017-08.html>.  
7 Program for Cooperative Cataloging. PCC task group on URIs in MARC format[EB/OL]. [2023-01-23]. [www.loc.gov/aba/pcc/bibframe/TaskGroups/URI-TaskGroup.html](http://www.loc.gov/aba/pcc/bibframe/TaskGroups/URI-TaskGroup.html).  
8 陈亚宁, 温达茂. MARC21 链结资料化的转变与应用[J]. 教育资料与图书馆学, 2020, 57(1): 35-72.  
9 Papadakis I, Kyprianos K, Stefanidakis M. Linked data URIs and libraries: the story so far[J/OL]. D-Lib Magazine, 2015, 21(5/6): 10.1045/may2015-papadakis. [2023-01-23]. <http://dlib.org/dlib/may15/papadakis/05papadakis.html>.  
10 Shieh J. PCC's work on URIs in MARC[J]. Cataloging & Classification Quarterly, 2020, 58(3-4): 418-427.  
11 Shieh J, Reese T. The importance of identifiers in the new Web environment and using the uniform resource identifier (URI) in subfield zero (\$0): a small step that is actually a big step[J]. Journal of Library Metadata, 2015, 15(3/4): 208-226.  
12 OCLC. Technical bulletin 267 [EB/OL]. [2023-02-26]. <https://www.oclc.org/support/services/worldcat/documentation/tb/267.en.html>.  
13 OCLC. Technical bulletin 268 [EB/OL]. [2023-02-26]. [https://help.oclc.org/worldcat/cataloging\\_documentation/technical\\_bulletin\\_ins/268](https://help.oclc.org/worldcat/cataloging_documentation/technical_bulletin_ins/268).  
14 Shieh J. Embedded URI in MARC: an essential for linked data [EB/OL]. [2022-12-20]. [http://downloads.alcts.ala.org/ce/10262016\\_MARCtoBIBFRAMEseriesPart3\\_Embedded\\_URI\\_MARC\\_Slides.pdf](http://downloads.alcts.ala.org/ce/10262016_MARCtoBIBFRAMEseriesPart3_Embedded_URI_MARC_Slides.pdf).  
15 Ex libris. BIBFRAME support in Ex Libris Alma [EB/OL]. [2023-02-26]. [http://www.casalini.it/EBW2018/web\\_content/2018/docs/BIBFRAME\\_Support\\_in\\_Ex\\_Libris\\_Alma.pdf](http://www.casalini.it/EBW2018/web_content/2018/docs/BIBFRAME_Support_in_Ex_Libris_Alma.pdf).  
16 PCC. URI FAQs PCC URI task group on URIs in MARC[EB/OL]. [2022-12-24]. <https://www.loc.gov/aba/pcc/bibframe/TaskGroups/URI%20FAQs.pdf>.  
17 Smith M, Stahmer C, Li X, et al. BIBFLOW: a roadmap for library linked data transition[EB/OL]. [2023-02-02]. [https://www.library.ucdavis.edu/wp-content/uploads/2020/01/bibflow\\_roadmap\\_revised\\_3\\_14\\_2017.pdf](https://www.library.ucdavis.edu/wp-content/uploads/2020/01/bibflow_roadmap_revised_3_14_2017.pdf).  
18 Smith-Yoshimura K. What metadata managers expect from and value about the research library partnership[EB/OL]. [2023-02-02]. <http://hangingtogether.org/?p=6683>.  
19 Cannan J P, Frank P, Hawkins L. LC/NACO authority file in the library of congress BIBFRAME pilots[J]. Journal of Library Metadata, 2019, 19(1-2): 39-51.



- 20 Program for Cooperative Cataloging Task Group on URIs in MARC. Formulating and obtaining URIs: a guide to commonly used vocabularies and reference sources[EB/OL].[2022-10-10]. [https://www.loc.gov/aba/pcc/bibframe/TaskGroups/formulate\\_obtain\\_URI\\_guide.pdf](https://www.loc.gov/aba/pcc/bibframe/TaskGroups/formulate_obtain_URI_guide.pdf).
  - 21 Giusti Serra L, Schneider J A, Segundo E. Person identifiers in MARC 21 records in a semantic environment[J]. *Cataloging & Classification Quarterly*, 2020, 58(5): 505-519.
  - 22 Stalberg E, Riemer J, MacEwan A, et al. Exploring models for shared identity management at a global scale: the work of the PCC Task Group on identity management in NACO[J]. *Cataloging & Classification Quarterly*, 2020, 58(3/4): 428-437.
  - 23 ORCID Support. What are persistent identifiers (PIDs)? [EB/OL].[2023-02-26]. <https://support.orcid.org/hc/en-us/articles/360006971013-What-are-Persistent-identifiers-PIDs->.
  - 24 米尔娜·维勒. UNIMARC 手册:规范格式(第三版)[M].《中国机读规范格式》工作组,译.北京:国家图书馆出版社,2013.
  - 25 Digital Preservation Coalition. Persistent identifiers [EB/OL].[2023-03-01]. <https://www.dpconline.org/handbook/technical-solutions-and-tools/persistent-identifiers>.
  - 26 Fenner M, Thorisson G, Ruiz S, et al. D4.1 conceptual model of interoperability[EB/OL].[2022-02-18]. <https://doi.org/10.6084/m9.figshare.824314.v1>.
  - 27 胡小菁.文献编目:从数字化到数据化[J].*中国图书馆学报*, 2019(5): 49-61.
  - 28 W3C. Cool URIs for the semantic Web, W3C Interest Group note 03 december 2008 [EB/OL].[2023-02-26]. <https://www.w3.org/TR/cooluris/>.
  - 29 IFLA. UNIMARC authorities format manual: control subfields [EB/OL].[2023-02-26]. [https://cdn.ifla.org/wp-content/uploads/files/assets/uca/unimarc\\_updates/AUTHORITIES/u\\_a\\_control\\_subfields\\_update2020\\_online\\_final.pdf](https://cdn.ifla.org/wp-content/uploads/files/assets/uca/unimarc_updates/AUTHORITIES/u_a_control_subfields_update2020_online_final.pdf).
  - 30 Network Working Group. Uniform resource identifier (URI): generic syntax [EB/OL].[2023-02-26]. <https://www.ietf.org/rfc/rfc3986.txt>.
  - 31 W3C Internationalization. An introduction to multilingual Web addresses[EB/OL].[2023-02-26]. [www.w3.org/International/articles/idn-and-iri/](http://www.w3.org/International/articles/idn-and-iri/).
  - 32 Library of Congress. MARC discussion paper NO.2017-DP01 [EB/OL].[2023-02-26]. <https://www.loc.gov/marc/mac/2017/2017-dp01.html>.
  - 33 王雪原.高校图书馆学位论文 856 字段的作用[J].*图书馆建设*, 2015(4): 51-53.
  - 34 Library of Congress. 856-electronic location and access(R)[EB/OL].[2023.02-17]. <https://www.loc.gov/marc/authority/ad856.html>.
  - 35 PCC Task Group on URIs in MARC. Year 2 report to PoCo, october 2017 [EB/OL].[2023-02-26]. [https://www.loc.gov/aba/pcc/documents/PoCo-2017/PCC\\_URI\\_TG\\_20171015\\_Report.pdf](https://www.loc.gov/aba/pcc/documents/PoCo-2017/PCC_URI_TG_20171015_Report.pdf).
  - 36 Library of Congress. PCC task group on linked data best practices, final report.[EB/OL].[2023-02-26]. <https://www.loc.gov/aba/pcc/taskgroup/linked-data-best-practices-final-report.pdf>.
  - 37 Library of Congress. PCC URI task group report april 15, 2017 [EB/OL].[2023-02-26]. [https://www.loc.gov/aba/pcc/bibframe/TaskGroups/PCC\\_URI\\_TG\\_20170415\\_Report.pdf#page=4](https://www.loc.gov/aba/pcc/bibframe/TaskGroups/PCC_URI_TG_20170415_Report.pdf#page=4).
  - 38 PCC task group on URIs in MARC. Year 2 report to PoCo, april 2017 [EB/OL].[2022-12-26]. [https://www.loc.gov/aba/pcc/bibframe/TaskGroups/PCC\\_URI\\_TG\\_20170415\\_Report.pdf#page=4](https://www.loc.gov/aba/pcc/bibframe/TaskGroups/PCC_URI_TG_20170415_Report.pdf#page=4).
  - 39 All things cataloged. GND-a new authority file[EB/OL].[2023-03-17]. <https://allthingscataloged.wordpress.com/category/authorities/>.
  - 40 Smith M, Stahmer C, Li X, et al. BIBFLOW: roadmap [EB/OL].[2023-02-26]. <https://bibflow.library.ucdavis.edu/roadmap/>.
- 作者单位:陈辰,河北大学管理学院,河北保定,071000  
槐燕,国家图书馆中文采编部,北京,100044  
收稿日期:2023年3月28日  
修回日期:2023年6月7日  
(责任编辑:李晓东)

## Research on Development and Application of Linky MARC

Chen Chen Huai Yan

**Abstract:** In the current hybrid environment where MARC and linked data coexist, it is of great significance to study the deployment and application of identifiers in MARC format to simplify MARC to linked data conversion. Using the identifiers in MARC21 and UNIMARC authority format as research objects, this paper identifies and discusses the problem of alternating application of old and new concepts brought by the hybrid of document network and data network environment. The concept and application of the following objects in MARC authority are discussed in detail: the local authority record control number and persistent record identifier, local and remote identifier, authority record and real-world object identifier, RDF URI and common Web URL. The results show that MARC authority has updated the format and content for the development of linked data. Various identifiers carry different functions in MARC authority, so they should be recorded in different fields and subfields to distinguish. In the future, it still needs further study on the coordination of URI and related objects, the identification and selection of identifiers, the balance between URI link and traditional cataloging work, etc.

**Keywords:** Linked Data; URI; MARC Authority; Unique Identifier; Authority Files