

# 基于 Wikidata 开展身份管理的模式研究\*

陈 辰 钟雨轩

**摘要** Wikidata 作为开放数据知识库和标识符链接中枢,在改善规范控制工作方面具有较大的潜力,逐渐受到图书馆及相关机构的关注。文章系统调研国内外基于 Wikidata 开展身份管理的研究与实践进展,在文献调研和案例分析的基础上将其模式划分为链接模式、融合模式和集中模式三种类型,并对三种模式的含义、实践案例、应用价值及模式关系等问题进行了系统探讨。基于对 Wikidata 身份管理模式最新实践的认识,尝试提出对于我国名称规范控制带来的几点启示,包括在 MARC 记录中融入外部标识符、在发现系统中进行元数据增强、积极参与集中聚合平台建设。参考文献 61。

**关键词** 身份管理 规范控制 Wikidata 转型模式

## Research on Identity Management Model Based on Wikidata

Chen Chen Zhong Yuxuan

**Abstract:** As an open data knowledge repository and identifier link hub, Wikidata has great potential in improving authority control work, and has gradually attracted the attention of libraries and related institutions. This article systematically investigates the research and practice progress of identity management based on Wikidata at home and abroad. On the basis of literature research and case analysis, this paper refines and divides the types of modes. Finally, three modes are identified, including linked mode, integrated mode and centralized mode, and the meaning, practice cases, application value and the relationship of three modes are systematically discussed. Based on the understanding of the latest practice of Wikidata's identity management mode, the article tries to bring some enlightenment to the authority control in our country, including integrating external identifiers into MARC records, enhancing metadata in the discovery system, and actively participating in the construction of centralized aggregation platforms. 61 refs.

**Keywords:** Identity Management; Authority Control; Wikidata; Transformation Mode

基于文本字符串的传统名称规范控制向依靠唯一标识符的身份(或实体)管理转型是当前编目的发展趋势<sup>[1]</sup>。规范控制向身份管理转型尚没有成熟的模式或路径可供参考,在规范控制研究和实践上一直走在世界前列的欧美等国家和地区,也在积极探索改进传统规范控制模式的方法和路径。Wikidata 作为维基媒体基金会托管的开放数据知识库和标识符链接中枢,具有参与门槛低、便于协作编辑、语义互联等特点,在改善规范控制工作方面具有较大的潜力,逐渐受到图书馆及相关机构的关注。

2019 年,美国研究图书馆协会(Association of Research Libraries, ARL)发布“Wikidata 白皮书”,

指出应鼓励更改书目记录标准,允许包括 Wikidata 在内的其他外部数据源发挥名称标目作用;应以协作性更好和开放性更强的方式创建结构化数据,使图书馆将精力“集中于特藏”和“专注于创建和连接名称网络”<sup>[2]</sup>。联机计算机图书馆中心(OCLC)开展了名为“项目通道”(Project Passage)的研究,探索使用 Wikibase(Wikidata 的基础软件平台)协调、创建和管理基于关联数据的实体和关系的书目和规范数据<sup>[3]</sup>。LD4(Linked Data For)的维基数据关联组(Wikidata Affinity Group)则研究了将 Wikidata 作为发布、链接和丰富图书馆关联数据的平台的方法<sup>[4]</sup>。国际图联(IFLA)Wikidata 工作组进一步探讨如何将 Wiki-

\* 本文系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)研究成果。

data 和 Wikibase 与图书馆系统进行集成,并将 Wikidata 本体与图书馆元数据格式进行对齐,如 BIBFRAME、RDA 和 MARC<sup>[5]</sup>。图书馆如果选用 Wikidata 作为描述、链接和丰富图书馆资源的方法,则可不必要仅关注本地图书馆目录和数据存储建设,还可以选择 Wikidata 集中数据仓储的实体管理,有利于实体在更大网络范围内的可见性和可重用性<sup>[6]</sup>。图书馆也可以将 Wikidata 作为开放关联数据的存储库,并与维基社区进行协作,互相丰富共同关注的书目数据,创建互利共赢的相关应用。

近年来学术共同体开始关注基于 Wikidata 开展协作编目和身份管理的问题,在研究与实践上均有新的探索,但尚缺少专门针对 Wikidata 身份管理模式的系统研究,缺少关于 Wikidata 身份管理应用案例的理论化模式识别和分类,从而难以以为 Wikidata 的进一步应用提供理论逻辑支撑。鉴于此,本文对基于 Wikidata 开展身份管理的优势进行分析,系统调研相关研究与实践进展,通过文献调研和案例分析进行基于 Wikidata 身份管理模式的分类识别,以期将 Wikidata 身份管理的理论研究和实践应用引向深入,并为我国规范控制向身份管理的转型提供有益参考。

## 1 基于 Wikidata 开展身份管理的适用性和优势分析

作为不同于传统规范控制模式的新身份管理方式,Wikidata 可改善传统规范控制面临的不足与缺陷,适应网络环境对图书馆规范控制的发展需求。

### 1.1 基于 Wikidata 开展身份管理的适用性研究

图书馆及相关机构利用 Wikidata 开展了多样化应用,其中使用 Wikidata 进行规范数据的集成和互操作是最受关注的领域<sup>[7]</sup>。Wikidata 平台是一种可轻松管理规范数据的存储库<sup>[8]</sup>,不仅可链

接同一实体的不同规范名称,还可生成尚未在图书馆规范文档中表示的名称实体项目<sup>[9]</sup>。作为通用的开放关联数据资源和“链接中枢”,Wikidata 已链接了越来越多的权威资源,这使其不仅成为权威数据的聚合平台,而且会演变为一个关联的开放数据生态系统<sup>[10]</sup>。各种权威数据库和服务可利用 Wikidata 建立实体并形成互连,实体范围包括人、机构、地理名称、作品、主题和流派等<sup>[11]</sup>,这超出 VIAF(Virtual International Authority File)书目名称的范围而面向更广的社区,使得 Wikidata 成为网络环境下图书馆实施身份管理的理想选择<sup>[12]</sup>。

为了加速实现网络层面无处不在的标识符创建和身份管理,联合编目项目(Program for Cooperative Cataloging, PCC)的名称规范合作项目(Name Authority Cooperative Program, NACO)身份管理任务组在 2020 年启动 Wikidata 试点项目,研究将图书馆数据贡献给 Wikidata 并在 Wikibase 平台中创建标识符(身份)的可行性<sup>[13]</sup>。图书馆和相关机构可以直接使用 Wikidata 标识符,而无需对本地数据库的基础结构进行更改。德国国家图书馆和德国维基基金会共同评估了 Wikibase 作为描述集成规范文档(Gemeinsame Normdatei, GND)软件的适用性,实验表明 GND 的关系和属性都可以在 Wikidata 中进行映射和表示<sup>[14]</sup>。美国国会图书馆也开始探索权威数据和维基数据的连接整合问题,并构建了一个名为“PnP Wikidata”的概念验证工具<sup>①</sup>,展示了将图书馆权威数据与维基数据关联起来的优势。

### 1.2 Wikidata 在改善传统规范控制方面的优势

(1)开放、自由的协作特点,拓展规范数据的创建主体

图书馆传统规范控制的参与门槛较高,编目员需要经过严格的培训与审核才能编制规范数据,这导致规范控制的主体力量有限。且由于传

① <https://labs.loc.gov/work/experiments/pnp-wikidata-tool>.

统规范控制较为强调初次创建的完整性,使得规范控制的规模和效率都不能满足不断扩大的规范控制需求。而 Wikidata 具有开放、自由的协作特点,能够吸引更多人和机器参与,任何人只要拥有注册账号就能对实体进行编辑和管理。这种众包形式的数据编辑机制,可以实现图书馆社区与其他社区协作创建和维护规范数据,跨领域增加规范数据的创建主体,可依靠社区力量不断完善、更新和丰富规范数据<sup>[15]</sup>。

(2) 基于唯一标识符识别模式,提高规范控制的实施效率

传统规范控制中,图书馆需要根据特定的编目规则选择规范标目,对名称标题的过度重视影响了规范控制的效率,且在不同语言文化背景下,不同国家、机构选择统一名称形式并不现实。PCC 的 NACO 任务组重新审查了规范控制工作,提出与精心制定的规范文本字符串相比,消除实体歧义的更好方法是标识符<sup>[16]</sup>。通过 Wikidata 可立即创建基于稳定持久标识符的实体描述,供其他人重复使用。QID 是 Wikidata 身份管理服务的唯一标识符,图书馆等机构可以通过链接 QID 进行实体规范控制<sup>[17]</sup>。此外,Wikidata 提供了 350 多种语言的实体标签、别名和描述,支持用户以喜欢的语言和文字显示名称或标题,这与统一标目相比具有较大的灵活性。

(3) 结构化描述和关联特点,促进规范数据的语义化应用

图书馆的 MARC 规范数据缺少 URI,不支持语义网环境下与其他数据集的关联和互操作,且以平面化记录为单位难以展现和描述立体化语义关系。图书馆利用相对独立和封闭的 Z39.50 协议,使得规范数据难以供图书馆以外的社区重用和共享。而 Wikidata 采用结构化的开放关联数据格式进行实体信息描述,形成由主语、谓语、宾语组成的 RDF (Resource Description Framework) 三元组,并允许对三元组信息进行描述和限定,其数据描述粒度和精度超过图书馆的 MARC 数据<sup>[18]</sup>。Wikidata 数据语句相互链接,支持发现和推理,且数据重用不需要繁琐的归属许可要求,

这些有助于扩展实体身份数据在网络环境下的语义应用,使其成为规范关联数据生态系统的关键组成部分。

### 1.3 Wikidata 相对于其他标识符系统的优势

Durocher 在分析图书馆对唯一标识符需求的基础上,认为未来的身份管理平台可能是 ORCID (Open Researcher and Contributor ID)、ISNI (International Standard Name Identifier)、LC/NAF (Library of Congress/Name Authority File)、VIAF 和 Wikidata 等<sup>[19]</sup>。而 Wikidata 与其他标识符系统相比,具有如下优势:(1) 作为标识符中枢,Wikidata 汇集指向同一实体对象的不同标识符方案,链接了不同领域使用的规范文档和本体等受控词汇。截至 2023 年 8 月,Wikidata 已链接到 1500 多个被称为规范控制属性的权威信息来源<sup>[20]</sup>。一旦链接的权威数据源到达临界点,Wikidata 将成为真正的全球数据网络中心。(2) 实体规模和范围大。Wikidata 是国际性、多语言、以社区为基础的项目,截至 2023 年 8 月已包含近 1 亿个“项目”实体的结构化信息(LC/NAF 约包含 1100 万,ISNI 约包含 1000 万)<sup>[21]</sup>,且范围涵盖超出图书馆传统规范控制对象的多种实体类型<sup>[22]</sup>。(3) 平台支持开放协作。Wikidata 支持来自不同背景的创建主体在协作环境中创建关联数据,并且数据可以自由使用和开放共享。与 ISNI 的会员制以及 VIAF 自上而下的规范数据创建模式相比,这种开放协作平台降低了数据创建和共享的障碍,能满足更大范围的实体管理需求<sup>[23]</sup>。Wikidata 参与的低门槛、便于协作编辑、语义互联、多语言支持以及作为标识符聚合中枢等优势,使其逐渐成为一种有潜力的身份管理平台。

## 2 基于 Wikidata 开展身份管理的模式

基于 Wikidata 开展身份管理的模式,是指相关机构或个人探索利用 Wikidata 开展规范控制的方法或路径,以改善或替代传统规范控制模式。本文在分析已有研究成果和实践案例的基础上,

从如下三个方面考虑对身份管理模式类型的识别和确定:第一,Wallis 提出了书目数据向关联数据转换的三种模式,即链接型 MARC、基于 BIBFRAME 书目本体和基于 Schema.org 本体的书目记录转换<sup>[24]</sup>,本文借鉴该模式划分的方式。同时,鉴于 Wikidata 作为图书馆潜在关联数据平台的研究和实践探索越来越多,本文将基于 Wikidata 本体模型进行规范数据的关联数据转换也作为一种类型,主要体现在本文提出的融合模式中的“规范文档的维基化”和集中模式类型。第二,根据图书馆及相关机构利用 Wikidata 标识符或元数据的方式和程度,以及所达到的应用效果进行模式划分。第三,根据已有规范记录和未有规范记录的实体在利用 Wikidata 开展身份管理方面的不同进行模式划分。基于上述考虑,本文最终将 Wikidata 身份管理的模式划分为链接模式、融合模式和集中模式三种类型。

## 2.1 链接模式

### 2.1.1 含义

链接模式是指将文本记录(如 MARC 书目/规范记录)的命名实体与 Wikidata 中有关同一实体的信息进行匹配,通过建立链接关系实现数据的交互共享。机器人通过链接跟踪和查询实体,获取标识符下的身份信息,辅助实现命名实体的消歧。MARC 记录和 Wikidata 的链接有两种方式:(1)书目/规范记录中链接 Wikidata 标识符,即在 MARC 相应字段或子字段中插入 Wikidata 唯一标识符,通过外部标识符辅助书目实体进行规范控制;(2)Wikidata 中链接规范文档或书目记录标识符,即在 Wikidata 页面通过数据属性建立到 MARC 书目/规范记录的反向链接,利用 Wikidata 的链接中枢地位,增加图书馆规范数据和书目数据的曝光率和使用率。两种方式方向相反,可称为链接和反向链接关系,当前这种互惠链接关系的同时建立尚不普遍<sup>[25]</sup>。

### 2.1.2 实践案例

当前多数书目/规范数据以 MARC 格式存在于图书馆系统中,如要增加用于揭示外部标识符

的字段,需对 MARC 标准进行调整研究。MARC 咨询委员会定义了 MARC 21 规范格式 024 字段(其他标准标识符)中的子字段 \$0 和 \$1,用以实现对外部实体标识符的关联和链接,UNIMARC 则通过 017 字段容纳外部标识符。MARC 2017-01 提案规定,\$4(关系)可用在书目和规范格式相应字段中,用于标识 MARC 记录与外部 Wikidata 等特定实体标识符资源的语义关系。MARC 2019-02 提案则用 \$2 来标识外部标识符的来源。

在进行 MARC 标准调整研究的同时,图书馆和相关机构开始尝试在书目/规范记录添加 Wikidata 等外部标识符。纽约大学<sup>[26]</sup>和哥伦比亚大学图书馆通过 024 字段中的 \$1 来容纳 Wikidata 标识符,哥伦比亚大学还使用特定属性(P5008)将 Wikidata 页面链接至原有项目页面,而不是复制粘贴 Wikidata 的详细信息<sup>[27]</sup>。史密森图书馆探索在 MARC 记录中添加关联数据元素,将 \$0 和/或 \$1 中的 URI 添加到标目字段以容纳 Wikidata 标识符,将 \$4 添加到 758 资源标识符字段,来说明 Wikidata 标识符和书目实体的关系<sup>[28]</sup>。

美国国会图书馆批量加载 Wikidata 标识符到关联规范数据页面(id.loc.gov),通过属性“紧密映射”将规范记录链接至 Wikidata 标识符;同时,Wikidata 通过属性“国会图书馆规范 ID(P244)”建立了指向规范记录的反向链接,一旦出现数据更新,Wikidata 标识符将会自动出现在 id.loc.gov 页面,实现标识符的动态增加和更新<sup>[29]</sup>。瑞典国家图书馆在 Wikidata 上创建了新属性“Libris-URI(P5587)”,通过该属性将图书馆目录项目标识符链接至 Wikidata,增加书目资源的访问和曝光率。莱布尼茨经济信息中心开展实验性项目,通过将其规范文档链接到 Wikidata,获得维基百科页面、图像和其他权威资源的链接<sup>[30]</sup>。由美国犹他大学主导的西方区域名称规范文档(WNAF)也链接至 Wikidata 平台,显示了将本地区域性规范文档链接至国际词汇表的可能性,同时将 WNAF 中已存在的 LC/NAF ID 上载到 Wikidata 项(item)中,进一步增强了现有 Wikidata 元数据的内容<sup>[31]</sup>。国内薛秋红等利用 PARIS 算法将中文名称规范

数据与 Wikidata 进行语义匹配,利用 Wikidata 中心化、开放、共享的特点扩大中文名称规范数据的影响力<sup>[32]</sup>。

### 2.1.3 应用价值

(1) 链接模式可在 MARC 记录中建立到 Wikidata 标识符的链接。一方面, Wikidata 作为规范控制源可辅助书目实体进行命名实体消歧,使编目员不用再将精力重点放在名称标题构建和选择规范检索点上,从而提高规范控制的效率。另一方面, Wikidata 标识符中枢聚合了指向同一实体的多个标识符,图书馆用户可进一步访问更广泛的资源。比如“鲁迅”的 Wikidata 页面有指向全世界各个组织的近 140 个权威资源链接,通过这些链接可间接访问和获取更多有关实体的上下文信息。

(2) 链接模式可在 Wikidata 中链接书目/规范记录,借助 Wikidata 中枢可增加图书馆数据的曝光率和使用率。正如 Smith-Yoshimura 所指出的,链接到“标识符中枢”的次数越多,本地数据被暴露和重复使用的次数就越多<sup>[33]</sup>。因此,建立图书馆元数据记录与外部 Wikidata 标识符的链接,将极大促进 MARC 数据在网络环境中的重用和共享,提高 MARC 数据在网络环境中的应用价值。

## 2.2 融合模式

### 2.2.1 含义

融合模式是指在实现命名实体匹配的基础上,进一步将图书馆及相关机构的数据与 Wikidata 系统进行元数据层面的协调和融合,从而实现各自系统在内容上的丰富和功能上的互补。从图书馆及相关机构融合数据的视角看,有两种融合方式:(1)由外而内,即图书馆及相关机构利用外部 Wikidata 元数据强化记录,为目录或发现系统用户提供更多的来自 Wikidata 的实体上下文信息,提高目录或发现系统的实体描述能力,提升用户体验并增加发现更多信息的机会;(2)由内而外,即内部规范记录转化融合为开放的 Wikida-

ta 项目描述,即 MARC 或本地数据的维基化,从而向网络社区开放书目领域规范数据,借助 Wikidata 平台分享和发现书目规范数据。

### 2.2.2 实践案例

图书馆及相关机构与 Wikidata 在元数据层面融合所展开的各类应用,主要包括如下几个方面:

(1) 增强规范记录。利用 Wikidata 数据丰富已有的规范记录信息,有助于消除实体歧义。如欧洲大屠杀研究基础设施提取了 Wikidata 元数据来增强大屠杀时代营地和贫民窟的名称权威记录,增强的元数据包括更多语言形式的名称标签、地理坐标、位置地图和层次关系,以及到外部百科全书的链接<sup>[34]</sup>。贾君枝等利用 Wikidata 中的职业等信息丰富人物名称规范中的属性元素,并采用关联技术实现名称规范档中人物信息的语义化表示,一定程度上解决其属性信息匮乏、同名人物识别困难等问题<sup>[35]</sup>。此外,贾君枝等还将 Wikidata 作为构建机构规范文档的数据来源,基于机构实体条目及类型属性构建了通用机构范畴树,以应对机构名称混乱、机构间关系模糊等问题<sup>[36]</sup>。在进行机构名称实例的自动聚类时, Wikidata 中机构实体的相关属性被作为上下文环境,用以消除机构实体名称中特征词共现对聚类算法的影响<sup>[37]</sup>。

(2) 知识卡片服务。在用户搜索目录时,图书馆以知识卡片的形式为用户提供来自 Wikidata 和其他开放数据源的信息。加拿大劳伦森大学图书馆为目录 Evergreen 构建了知识卡片服务,该卡片集成了来自 Wikidata 的音乐家和乐队的详细数据,包括许可的照片、简介和标识符链接等<sup>[38]</sup>。威斯康星大学麦迪逊图书馆开发了 BibCard 项目,将包括 Wikidata 在内的外部来源引入到图书馆目录的关联数据实例中,向用户显示关于作者的详细信息,如简介、教育和著作信息等,而不是传统目录仅有的作者名称信息<sup>[39]</sup>。类似项目还包括斯坦福大学 SearchWorks 目录环境中的“集

成知识面板”<sup>①</sup>,香港科技大学图书馆的“知识卡片”目录<sup>②</sup>,密歇根大学数字存储库中的“主题知识卡片”<sup>③</sup>,纽约现代艺术博物馆在线目录中的艺术家“知识卡片”<sup>④</sup>。此外,图书馆技术提供商 Zepheira 也在为美国公共图书馆馆藏中的实体开发创建“关于作者”的工具等<sup>⑤</sup>。这些服务类似维基百科或谷歌知识图谱的信息卡片,是以实体为中心的数据建模和信息访问的体现。

(3) 语义增强服务。将记录中的文本字符串链接到 Wikidata 等关联开放数据集或词汇表的受控术语,这种用附加术语扩充元数据的过程称为语义增强<sup>[40]</sup>。芬兰广播公司使用 Wikidata 概念来标记其数字档案(包括在线新闻和专题文章等)中的内容,通过机器可读的格式将新闻内容和维基项目相互链接,从而有助于标记并丰富语义知识内容<sup>[41]</sup>。在构建本体驱动的标注系统中,Colla 等将 Wikidata 语义模型与 PRiSMHA 本体进行映射,进而从 Wikidata 中自动提取信息,为历史档案文档中的相关实体进行语义标注<sup>[42]</sup>。

(4) 语义查询和扩展服务。通过将 Wikidata 身份数据源进行融合,构建以实体发现为中心的搜索结果显示,提高搜索的效率和精确性。斯坦福大学图书馆的 SearchWorks 目录直接在 solr 中(一个独立的企业搜索服务器)建立 Wikidata 元数据索引,然后使用 solr 连接来增强书目记录和过滤搜索结果<sup>[22]</sup>。荷兰国家图书馆将报纸文本中的命名实体与其外部链接一起构建索引,以便支持基于 Wikidata 属性的语义搜索,比如支持类似“不在荷兰出生的议员的的文章”的语义搜索<sup>[43]</sup>。此外,利用来自 Wikidata 的元数据可进行浏览器的查询扩展,Entity Explosion 即是针对 Chrome 和 Firefox 开发的扩展工具。当用户浏览某一实体的信息时,该工具将通过调用 Wikidata

查询服务 API,将 Wikidata 项(item)与正在浏览的 URL 事物或实体进行匹配,如匹配成功则在浏览器端增强显示来自 Wikidata 的数据和链接扩展<sup>[44]</sup>。

(5) 规范文档的维基化。将 MARC 规范记录的相关字段与 Wikidata 数据模型进行映射,进而将规范数据集成到 Wikidata 平台。Fernsebner 等利用 MarcEdit 和 OpenRefine 工具进行哈佛大学图书馆 MARC 规范记录和 Wikidata 项目的初步映射,从而将 MARC 规范记录加载到 Wikidata 中<sup>[45]</sup>。德国国家图书馆和德国维基媒体基金会研究 Wikibase 作为承载 GND 规范数据的可行性,创建适用于规范数据描述所需的规则集和模型,开发基于 Wikibase 的 GND 并投入使用,满足图书馆和其他社区的协作式规范文档构建和维护需求<sup>[46]</sup>。

### 2.2.3 应用价值

(1) 融合模式可在书目系统或其他发现系统中融合 Wikidata 元数据。一方面,增强的实体信息有助于改善稀疏的规范数据,辅助进行书目实体的名称识别和消除歧义,更好地实现目录的汇集、揭示和检索功能;另一方面,丰富书目实体信息,可为用户提供更多的访问点,进而构建以实体为中心的多源聚合发现服务,如知识卡片、语义搜索和语义标注等增强用户体验、提高搜索精确性的服务。

(2) 融合模式可实现规范数据的维基化,将图书馆规范数据融合到 Wikidata 平台。一方面,为开放关联数据社区贡献书目规范数据,扩大名称规范数据的共享范围,履行世界书目控制(UBC)原则;另一方面,将基于 MARC 的规范记录转换为基于唯一标识符、RDF 三元存储和 SPARQL 查询的关联数据形式,Wikidata 和 Wiki-

① <https://searchworks-ld.stanford.edu/>.

② [https://julac-hkust.primo.exlibrisgroup.com/discovery/fulldisplay?context=L&vid=852JULAC\\_HKUST;HKUST&search\\_scope=HKUST\\_catalog\\_primo&tab=Everything&docid=alma991005090699703412](https://julac-hkust.primo.exlibrisgroup.com/discovery/fulldisplay?context=L&vid=852JULAC_HKUST;HKUST&search_scope=HKUST_catalog_primo&tab=Everything&docid=alma991005090699703412).

③ <https://doi.org/10.1080/10691316.2021.1901634>.

④ <https://medium.com/digital-moma/bringing-art-knowledge-to-everyone-who-seeks-it-899ec257a55c>.

⑤ <https://twitter.com/InformaticMonad/status/1068310528710213633>.

base 提供了便利的基础设施和平台,一定程度降低了图书馆将规范记录转换为关联数据的技术门槛和障碍。

## 2.3 集中模式

### 2.3.1 含义

集中模式是指图书馆等相关机构将 Wikidata 或 Wikibase 作为替代传统规范控制的模式,直接在该平台协作创建基于唯一标识符的实体管理数据。鉴于 Wikidata 从图书馆吸引和整合权威数据的能力不断增强,从“链接中心”演变为全球“权威中心”的可能性增大,当平台达到一定的数量和质量临界值时,可直接将 Wikidata 作为全球独特、统一、通用的标识符<sup>[47]</sup>。根据软件平台的不同,这种集中模式可分为两种方式:(1)基于 Wikidata 开展身份管理,即在 Wikidata 平台上创建实体项目。(2)基于 Wikibase 开展身份管理,即本地安装 Wikibase 软件来开发用例。

### 2.3.2 实践案例

#### (1) 基于 Wikidata 开展身份管理

图书馆等机构在 Wikidata 平台中创建和增强关于人员、组织及其作品等实体的结构化数据,并将该数据链接到其他标识符系统,主要应用领域如下:

##### 1) 科研人员身份管理

科研人员数量庞大且重名现象较为普遍,影响了数字学术环境下成果汇集的全面性和准确性。Wikidata 作为结构化开放知识库可管理学者身份信息、构建学术简历档案,解决科研人员的名称规范问题<sup>[48]</sup>。如印第安纳大学与普渡大学印第安纳波利斯分校,利用 Wikidata 平台补充了教员和合著者的出版物信息,这些信息为结构化的、可消费的 RDF 数据,随后利用 Scholia 进行数据显示,并生成了学术简历(身份)<sup>[49]</sup>。田纳西图书馆测试了 Wikidata 对研究人员档案的适用性,探讨了如何应用 Wikidata 数据在 Scholia 中建立本地作者的个人资料<sup>[50]</sup>。Robare 在 Wikidata 中创建了俄勒冈州重要人员和组织及其作品的结构化数据,以提高其知名度<sup>[51]</sup>。PCC 的 NACO 身份管理

任务组的参与机构为相关的个人和法人团体创建了 Wikidata 身份,如斯坦福大学图书馆在 Wikidata 中创建了斯坦福大学附属人员的身份描述档案<sup>[52]</sup>。

#### 2) 特藏资源的实体管理

作为一个国际性、多语言和基于社区的项目, Wikidata 平台可为图档博机构展示特藏资源。哥伦比亚大学图书馆基于 Wikidata 创建了口述历史档案资源中的实体<sup>[27]</sup>。威尔士国家图书馆在 Wikidata 中描述了 4800 张威尔士肖像,用户可利用该平台查询肖像资源,并创建可视化展示,通过 Wikidata 连接到 Wikipedia 文章或《威尔士传记词典》<sup>[53]</sup>。约克大学与 ARL 开展合作项目,创建 200 多名土著表演者的数据集,并评估和扩展 Wikidata 属性、限定词和其他术语,对档案和特殊馆藏描述进行数据建模<sup>[38]</sup>。史密森学会 2019 年启动开放数据试点项目,重点向 Wikidata 提供机构收藏的开放数据,其中跨不同学科的身份和数据模型协调是项目实施的难点<sup>[54]</sup>。

#### (2) 基于 Wikibase 开展身份管理

Wikibase 作为 Wikidata 的基础软件平台,是用于存储和管理结构化数据的开源包。它既可以作为社区公共基础设施,也可以进行独立安装部署并开发针对图书馆的数据模型,并在权限和质量方面给予相应的本地控制。

2019 年法国国家图书馆和法国高等教育书目机构(Abes)启动实验,重点测试使用 Wikibase 基础设施支持法国国家实体文档(French National Entities, FNE)管理的可能性<sup>[55]</sup>。FNE 是根据 IFLA LRM 模型构建的规范数据平台,其目的是建立一个共享生产实体,但经过测试,发现 Wikibase 在大规模数据加载、修改操作方面未能满足 FNE 的预期要求<sup>[56]</sup>。英国国家图书馆与英国媒体基金会合作开展 Wikibase 项目,对土耳其手稿和库尔德人印刷的藏书中涉及的实体进行管理,该项目显示出 Wikibase 在多语言和文字标签方面的优势。OCLC 主导的 Project Passage、CONTENTdm 关联数据试点以及梅隆基金资助的共享实体管理基础架构(Shared Entity Management Infrastructure),都使用安

装在本地的 Wikibase 进行相关实例探索。随着本地实例数量的增加, Wikibase 推动了开放关联数据网络的发展, 但创建本地 Wikibase 实例的联合生态系统也非常关键, 否则本地 Wikibase 实例可能被“淹没”淘汰<sup>[57]</sup>。

### 2.3.3 应用价值

(1) 集中模式可实现直接利用已有的开源平台进行身份管理探索。一方面, 不需要对图书馆技术基础架构进行修改, 这符合图书馆现实发展需要, 因为单个图书馆的技术力量薄弱, 难以构建大规模身份管理设施。另一方面, 为适应未来的 IFLA-LRM 实体管理模式, 需要一个持久的、共享的、集中的实体管理基础架构, Wikidata 和 Wikibase 提供了满足实体管理需求的平台, 为改善图书馆 MARC 记录质量、支持协作的元数据 workflow 等方面提供坚实基础。

(2) 集中模式可利用 Wikidata 扩展规范控制的规模和范围。一方面, 进行本地实体及特藏资源实体的管理, 提高这些实体在网络环境下的显示度和可发现性。传统规范控制主要关注目录中涉及的专著中的人名(包括责任者和主题)或机构实体, 对于本机构的附属人员和机构实体的关注度不高, 而 Wikidata 实体创建并不受“文献保证规则”的限制, 从而可提高本地实体身份管理的覆盖度。另一方面, Wikidata 支持对代理、作品、概念、时间跨度、地点等更广泛范围实体的描述和识别, 契合未来 IFLA-LRM 涵盖的实体范围和类型。

## 2.4 模式总结

基于 Wikidata 开展身份管理的链接模式、融合模式和集中模式具有不同的应用场景, 相关机构可根据实际需求进行决策选择。

### 2.4.1 模式特点

链接模式需要在 MARC 记录的相关字段(比如 024 字段)或 Wikidata 的“外部标识符”属性中著录关联彼此的标识符, 通过标识符链接实现规范文档和 Wikidata 的关联和互操作, 且需要用户点击 HTTP 形式的链接实现到各系统的数据访

问。融合模式是在系统实体数据关联匹配的基础上, 进一步利用彼此的元数据实现系统的服务增强, 其主要特点是各系统在元数据层面实现协调和互操作, 用户在一个系统界面即可访问来自外部系统的元数据。集中模式主要针对图书馆未有规范记录的实体, 如特藏资源实体或者本地附属人员及机构实体等, 图书馆直接利用 Wikidata 集中聚合平台开展实体描述和标识符链接工作, 从而替代传统的规范控制模式。

这三种模式的区别主要在于: 链接模式和融合模式都需要解决已有规范记录和 Wikidata 实体描述之间的协调对齐问题, 而集中模式主要针对未有规范记录的实体而言, 图书馆在 Wikidata 或 Wikibase 平台开展实体描述与身份管理工作。链接模式仅在标识符层面进行关联和互操作, 不涉及元数据层面的协调或复用; 而融合模式则是在实体匹配的基础上, 进行元数据层面的协调和相关服务的增强。如链接模式中“Wikidata 中链接规范文档或书目记录标识符”与融合模式中的“规范文档的维基化”, 虽然两者都涉及实体的协调匹配问题, 但前者仅在 Wikidata 的“外部标识符”属性中链接规范文档或书目记录的标识符, 不涉及元数据层面的利用; 后者则是在 Wikidata 属性和规范文档字段映射匹配的基础上, 将规范文档字段融合进 Wikidata 实体描述属性中, 实现的是各系统元数据层面的协调和融合。

### 2.4.2 模式关系

链接模式、融合模式和集中模式的实施具有关联关系, 书目系统记录和 Wikidata 实体的协调匹配是所有模式的前提基础, 只有在实体匹配链接的基础上, 才能进一步考虑各系统在元数据层面的协调和增强, 确定是创建新的 Wikidata 实体还是增强已有的 Wikidata 实体。相关机构开发了协调外部标识符的自动化或半自动化工具, 如 MarcEdit 的 Linked Identifiers 模块、图书馆关联数据(LD4L)项目的“查询规范”(Questioning Authorities, QA)、美国西北大学图书馆的“权威工具包”(Authority Toolkit)以及开源社区维护的工具 OpenRefine, 都支持对 Wikidata 等外部标识符的

查找和匹配。

三种模式并不互斥,可根据应用场景进行选择、组合。图书馆在规范记录或书目记录系统、机构知识库等系统中链接外部唯一标识符,辅助进行名称规范控制的同时,也可参与 Wikidata 等集中聚合身份管理平台建设,将名称规范控制力量加入到协作化身份管理工作中。当聚合平台的数量和质量达到一定的要求后,可以利用网络链接方式将 Wikidata 标识符直接添加到规范记录或书目记录中。此外,也可直接将 Wikidata 标识符及其元数据导入到本地身份管理系统,并需要根据添加额外的元数据,构建本地应用概要文件,进而开展多源聚合的数据搜索和发现服务。

哥伦比亚大学图书馆参与 PCC Wikidata 试点项目,同时应用三种模式建立身份管理流程,其探索实践如下:(1)编目员首先搜索 Wikidata,如规范记录与 Wikidata 项目匹配成功,则以 024 字段中的 \$1 容纳 Wikidata 标识符,并在 Wikidata 中使用特定属性(P5008)链接回规范记录。(2)对于未创建规范记录的责任者,编目员则创建一条新 Wikidata 项目。(3)编目员使用新创建的 Wikidata 元数据创建新的规范记录并制定正确的访问点和交叉参考,以便在图书馆书目系统中运行。然后重复步骤(1),在规范记录和 Wikidata 项之间建立链接关系<sup>[27]</sup>。同样,美国国会图书馆开展的 BIBFRAME-Wikidata 试点项目也是类似的身份管理模式和流程。再有,德国莱布尼茨经济信息中心(ZBW)向 Wikidata 捐赠其 PM20 (20th century press archives) 数据,一方面,在 Wikidata 中利用外部标识符属性“PM20 folder ID (P4293)”,建立 Wikidata 项目到 PM20 人员和机构信息描述文件夹的链接;另一方面,利用 PM20 人员和机构的元数据来补充现有的 Wikidata 项目<sup>[58]</sup>。

未来的身份管理将会利用全局 ID 取代本地规范标目或其他各类标识符,全局 ID 自动聚合不同社区的规范数据和身份描述数据<sup>[59]</sup>。该全局 ID 可能是 Wikidata、ISNI 或其他未出现的标识符,在确定全局 ID 之前,无论是各类系统的分布式链

接,还是融合外部元数据构建服务系统,还是积极参与集中聚合平台的身份管理,都是在为未来加入全局 ID 系统做准备,最终可实现三种模式的融合发展。

#### 2.4.3 面临的问题

图书馆等机构可根据需要选择相应的模式开展身份管理工作,但在应用过程中也面临不同的问题:(1)在链接 Wikidata 标识符时,Wikidata 项目也存在条目内容稀疏、信息不丰富的问题,影响实体的匹配效果,但这也是图书馆为 Wikidata 作出贡献或吸引个人用户编辑 Wikidata 条目的潜在机会<sup>[60]</sup>。(2)在基于 Wikidata 开展实体管理的应用场景下,未经授权的用户可对实体项目进行编辑修改,元数据不受控制可能会导致数据质量不一致,影响数据的权威性<sup>[61]</sup>。Wikidata 平台的内嵌限定符和约束类型规则集可一定程度改善该问题,通过智能、动态的控制保证数据输入的完整性、一致性和准确性。如在输入属性值时平台会显示一些警告性提示框,提示属性值约束和规则约束内容,以保证输入数据的质量和完整性。Wikidata 还对大多数陈述语句的来源进行跟踪,以此提高数据的可靠性。(3)Wikidata 新项目存在重复的风险,不同的组织或国家对特定项目有不同的看法,Wikidata 项目的长期可持续性发展等问题需要社区进一步研究协调。

### 3 对我国名称规范控制的启示

本文通过对 Wikidata 身份管理模式的系统调研,发现 Wikidata 作为一个强大的身份管理平台或工具,对图书馆及相关机构具有较大的应用潜力。虽然当前的实践案例和理论研究多集中在国外,且处于初步研究测试阶段,还面临诸多问题需要进一步探讨,但其新型的身份管理思想和模式给我国规范控制工作带来如下启示。

#### 3.1 在 MARC 记录中融入外部标识符

目前书目/规范数据是以 MARC 格式存在于图书馆系统中,在 MARC 记录中融入外部数据源的 URI,标志着传统规范控制和身份管理方法的

融合,可为未来的关联数据化转换做准备。虽然 Wikidata 项目中已经添加了来自我国的 CALIS 和国家图书馆等机构的规范记录标识符,但国内的 MARC 记录并未建立到 Wikidata 标识符的链接,因此应积极推动 MARC 记录到外部标识符链接,利用外部权威资源辅助进行书目实体规范控制。在 MARC 记录融入外部标识符时,编目员面临复杂的唯一标识符生态系统,需要结合中文名称实体的分布特点和规律,确定可链接的外部标识符源。各类实体在各类标识符系统中的分布情况、哪些系统拥有中文实体的数量较多、是否具有学科分布或人口分布上的统计特点、各个系统存在哪些功能差异、其元数据质量如何,对于这些问题的把握和研究,可为编目员链接更加有意义的标识符源提供参考,帮助编目员明确标识符选择的优先级,有计划、分步骤地在 MARC 记录中融入外部标识符。

### 3.2 在发现系统中进行元数据增强

图书馆发现系统应充分集成外部标识符系统数据,从而开展基于实体的元数据增强服务。当前国内多数图书馆目录并未融合外部元数据并开展增强式目录发现服务,应积极从 Wikidata 等外部数据源获取馆藏实体相关的额外信息,并将其融合到馆藏目录中,从而开展知识卡片、语义搜索等增强用户体验的服务。由于存在复杂多样的标识符系统,不同系统在功能侧重、服务目的、类型、范围和学科主题等方面存在差异,图书馆需要研究各系统元素集的特点及功能差异,为系统之间元数据增强和重用提供参考。任何一种独立的系统都不可能满足所有的身份管理需求,各系统需要利用彼此的元数据构建完整的元数据生命周期,才能实现内容的丰富和功能的互补。

### 3.3 积极参与集中聚合身份管理平台

我国图书馆应努力加入 VIAF、ISNI 和 Wikidata 等大规模集中聚合中心,将实施传统规范控制的力量加入到集中聚合身份管理平台。一方面,利用这些平台进行中文规范数据的增强;另

一方面,提供集中的 Web 站点供用户查询,扩大中文名称规范数据的共享范围。正如 Scott 所指出的,图书馆不应只加强自己的本地数据仓库,应该把有限的资源投入到集中数据仓库 Wikidata 中,以最大化增强实体的可见性和可重用性,然后将数据拉回到图书馆本地存储库,以丰富目录显示和集成外部信息<sup>[6]</sup>。那么在众多身份管理平台中,如何选择适合我国身份管理需求的集中聚合平台,需要进行探索和决策分析。同时,基于集中聚合平台或系统进行协作化身份管理,面临与传统规范控制不同的工作流、数据模型和规则,如何将传统规范控制工作流融入集中聚合平台也有待研究和设计科学合理的方法体系。

### 参考文献

- 1 胡小菁. 规范控制:从名称选择到实体管理[J]. 数字图书馆论坛,2018(1):2-7.
- 2 Association of Research Libraries Task Force. ARL white paper on Wikidata: Opportunities and recommendations[R/OL]. [2023-02-02]. <https://www.oclc.org/research/publications/2019/wikidata-opportunities-and-recommendations.html>.
- 3 OCLC Research. Linked Data Wikibase Prototype[EB/OL]. [2023-07-13]. <https://www.oclc.org/research/themes/data-science/linkedata/linkedata-prototype.html>.
- 4 Thorsen H. LD4-Wikidata Affinity Group[EB/OL]. [2023-07-13]. <https://wiki.duraspace.org/display/LD4P2/LD4-Wikidata+Affinity+Group>.
- 5 IFLA. IFLA Wikidata Working Group[EB/OL]. [2023-08-01]. <https://www.ifla.org/node/92837>.
- 6 Scott D. Wikidata, Canada 150, and music festival data[EB/OL]. [2023-08-06]. <https://coffeecode.net/wikidata-canada150-and-music-festival-data.html>.
- 7 Tharani K. Much more than a mere technology: A systematic review of Wikidata in libraries[J]. The Journal of Academic Librarianship, 2021, 47(2):1-8.
- 8 Vrandečić D, Krötzsch M. Wikidata: A free col-

- laborative knowledgebase[J]. *Communications of the ACM*, 2014, 57(10):78-85.
- 9 Smith-Yoshimura K. The coverage of identity management work[EB/OL]. [2023-08-06]. <http://Hangingtogether.org/?p=6805>.
  - 10 Fischer B, Ohlig J. Could you wikify an authority file? Wikibase has been evaluated for the Integrated Authority File (GND)[EB/OL]. [2023-03-29]. <https://blog.wikimedia.de/2020/03/04/wikibase-and-gnd>.
  - 11 IFLA. Opportunities for academic and research libraries and Wikipedia[EB/OL]. [2023-08-06]. <https://www.ifla.org/files/assets/hq/topics/info-society/iflawikipediaopportunitiesforacademicresearchlibraries.pdf>.
  - 12 Chew C N. Expanding the use of linked data value vocabularies in PCC cataloging[J]. *Cataloging & Classification Quarterly*, 2020, 58(3/4):1-9.
  - 13 Beckett M, Durocher M. PCC Identity Management; Wikidata Pilot[EB/OL]. [2023-08-06]. <https://wiki.lyrasis.org/display/pccidmgt/Wikidata+Pilot>.
  - 14 Fischer B, Ohlig J. Report “GND meets Wikibase”[EB/OL]. [2023-07-13]. <https://wiki.dnb.de/pages/viewpage.action?pageId=167019461>.
  - 15 Riemer J. New prospects for library authority data[J]. *Technicalities*, 2012, 32(4):1,7-9.
  - 16 PCC task group on identity management in NACO; Wikidata Affinity Group. The program for cooperative cataloging & a Wikidata pilot[EB/OL]. [2023-08-08]. [https://docs.google.com/presentation/d/1NpkAQdGGft1Wi2vX0zgMtIwXWjPq96NtXx4MmyXFFL&edit#slide=id.g884e18190a\\_0\\_36](https://docs.google.com/presentation/d/1NpkAQdGGft1Wi2vX0zgMtIwXWjPq96NtXx4MmyXFFL&edit#slide=id.g884e18190a_0_36).
  - 17 Wikidata. Wikidata for authority control[EB/OL]. [2023-08-06]. [https://www.wikidata.org/wiki/Wikidata:Wikidata\\_for\\_authority\\_control](https://www.wikidata.org/wiki/Wikidata:Wikidata_for_authority_control).
  - 18 OCLC Research. Works in progress webinar; Case studies from Project Passage focusing on Wikidata’s multilingual support[EB/OL]. [2023-08-06]. <https://www.oclc.org/research/events/2019/061819-case-studies-project-passage-wikidata-multilingual-support.html?>
  - 19 Association for Library Collections & Technical Services. NACO, authority control, and identity management: Evolving strategies for a changing name authorities landscape[EB/OL]. [2024-03-11]. <https://www.ala.org/alcts/conferences/upcoming/webinar/120617>.
  - 20 Wikidata. Help:Data type[EB/OL]. [2023-08-06]. [https://www.wikidata.org/wiki/Help:Data\\_type#External\\_identifier](https://www.wikidata.org/wiki/Help:Data_type#External_identifier).
  - 21 维基数据:统计[EB/OL]. [2023-08-06]. <https://www.wikidata.org/wiki/Wikidata:Statistics/zh>.
  - 22 Thorsen H. LD4 Wikidata Affinity Group[EB/OL]. [2023-08-06]. <https://docs.google.com/document/d/1z1SSAp4c4tftOGW3BbJ6FxfD8oRIhfzveh0zjebDhkk/edit#>.
  - 23 Stevenson J. Thoughts on the Heritage PIDs Project[EB/OL]. [2023-03-25]. <https://blog.archiveshub.jisc.ac.uk/2022/01/18/thoughts-on-the-heritage-pids-project>.
  - 24 Wallis R. The three linked data choices for libraries[EB/OL]. [2023-09-13]. <https://www.dataliberate.com/2018/05/22/the-three-linked-data-choices-for-libraries>.
  - 25 Bianchini C, et al. Beyond VIAF; Wikidata as a complementary tool for authority control in libraries[J]. *Information Technology and Libraries*, 2021, 40(2):1-31.
  - 26 Urban R. Identity management beyond the LC/NACO Authority File[EB/OL]. [2023-08-07]. <https://hangingtogether.org/identity-management-beyond-the-lc-naco-authority-file>.
  - 27 Ryan Mendenhall T. PCC Wikidata Pilot; Expanding Authority Control to Identity Manage-

- ment[EB/OL]. [2023-6-19]. <https://blogs.cul.columbia.edu/tsl/2021/06/08/pcc-wikidata-pilot-expanding-authority-control-to-identity-management>.
- 28 Berthoud H, Hartley J. Practical approaches to linked data[J]. *The Serials Librarian*, 2021, 80(1-4):19-26.
- 29 Ferriter M. Integrating Wikidata at the Library of Congress[EB/OL]. [2023-07-23]. <https://blogs.loc.gov/thesignal/2019/05/integrating-wikidata-at-the-library-of-congress>.
- 30 ZBW labs. Wikidata for authorities[EB/OL]. [2023-07-23]. <https://zbw.eu/labs/en/project/wikidata-for-authorities>.
- 31 Myntti J, et al. Regional connections to national authority files[J]. *Cataloging & Classification Quarterly*, 2020, 58(1):76-89.
- 32 薛秋红,等. 中文名称规范数据与 Wikidata 语义关联实现[J]. *情报理论与实践*, 2019(10):146-150.
- 33 Smith-Yoshimura K. From authorities to identifiers——Bridging the silos[EB/OL]. [2024-03-11]. <https://www.oclc.org/research/events/2019/120319-oclc-symposium-authorities-and-identifiers.html>.
- 34 Cooley N. Leveraging Wikidata to enhance authority records in the EHRI portal[J]. *Journal of Library Metadata*, 2019, 19(112):83-98.
- 35 贾君枝,赵宇飞. Wikidata 与名称规范档数据聚合实现[J]. *情报科学*, 2018(11):72-77,82.
- 36 贾君枝,叶壮壮. 基于 Wikidata 的机构类目范畴树构建与优化[J]. *国家图书馆学刊*, 2018(1):56-64.
- 37 贾君枝,叶壮壮. 基于潜在语义索引的 Wikidata 机构实体聚类研究[J]. *数据分析与知识发现*, 2019(10):56-65.
- 38 Allison-Cassin S, Scott D. Wikidata: a platform for your library's linked open data[J/OL]. *Code4lib Journal*, 2018(40)[2024-03-11]. <https://journal.code4lib.org/articles/13424>.
- 39 Meyer S. Enhancing library discovery with linked open data[EB/OL]. [2023-06-23]. <https://osf.io/h2cav>.
- 40 Europeana Pro. Europeana semantic enrichment[EB/OL]. [2023-08-07]. <https://pro.europeana.eu/page/europeana-semantic-enrichment>.
- 41 Wikimedia Suomi. Yle <3 Wikidata[EB/OL]. [2023-06-23]. <http://wikimedia.fi/2016/04/15/yle-3-wikidata>.
- 42 Colla D, et al. Wikidata support in the creation of rich semantic metadata for historical archives[J/OL]. *Applied Sciences*, 2021, 11(10), 4378 [2024-03-11]. <https://www.mdpi.com/2076-3417/11/10/4378>.
- 43 Van Veen T, et al. Linking named entities in Dutch historical newspapers[M]//*Metadata and Semantics Research*. Bolzano:Springer, Cham. 2016:205-210.
- 44 Chrome. Entity explosion[EB/OL]. [2023-08-07]. <https://chrome.google.com/webstore/detail/entity-explosion/bbfeffclligkmfioanodamdjclgejcn>.
- 45 Fernsebner Eslao C, Moody H. Local authority file conversion from MARC to Wikidata[EB/OL]. [2023-08-05]. <https://wiki.lyrasis.org/display/LD4P2/2019+LD4+Conference+on+Linked+Data+in+Libraries>.
- 46 Barbara F. Authority control meets Wikibase[EB/OL]. [2023-04-15]. <https://wiki.dnb.de/display/GND/Authority+Control+meets+Wikibase>.
- 47 Veen V. Wikidata: From “an” identifier to “the” identifier[J]. *Information technology and libraries*, 2019, 38(2):72-81.
- 48 Neubert J. Wikidata as a linking hub for knowledge organization systems? Integrating an authority mapping into Wikidata and learning lessons

- for KOS mappings [C]. Proceedings of the 17th European Networked Knowledge Organization Systems Workshop, 2017, 14-25.
- 49 Lemus-Rojas M, Odell J. Creating structured linked data to generate scholarly profiles; A pilot project using Wikidata and Scholia [J/OL]. Journal of Librarianship and Scholarly Communication, 2018 (6) : eP2272 [2023-12-21]. <https://doi.org/10.7710/2162-3309.2272>.
- 50 Panigabutra-Roberts A. An experiment with name entities in Wikidata @ University of Tennessee Libraries [EB/OL]. [2023-07-21]. <https://wikiduraspace.org/display/LD4P2/2019+LD4+Conference+on+Linked+Data+in+Libraries>.
- 51 Robare L. Exploring the use of Wikidata; Highlights of a sabbatical project [EB/OL]. [2023-07-13]. <https://connect.ala.org/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=5b1c7cb6-62fb-d490-f598-eb8f1080330e&forceDialog=0>.
- 52 Wikidata; WikiProject Stanford Libraries [EB/OL]. [2023-07-13]. [https://www.wikidata.org/wiki/Wikidata:WikiProject\\_Stanford\\_Libraries](https://www.wikidata.org/wiki/Wikidata:WikiProject_Stanford_Libraries).
- 53 Welsh Portrait Collection. 4800 Welsh portraits added to Wikimedia Commons and Wikidata [EB/OL]. [2023-06-07]. <https://blog.library.wales/welsh-portraits-wikimedia-wikidata>.
- 54 Snyder R, Nguyen D. Smithsonian Open Data Pilot [EB/OL]. [2023-01-20]. <https://confluence.si.edu/display/LODPP/Smithsonian+Open+Data+Pilot>.
- 55 Transition Bibliographique. French national entities file (FNE) [EB/OL]. [2023-08-04]. <https://www.transition-bibliographique.fr/fne/fichier-national-entites>.
- 56 Transition Bibliographique. Questions and answers about the FNE; #4 [EB/OL]. [2023-08-04]. <https://www.transition-bibliographique.fr/2023-03-31-questions-reponses-fne-4>.
- 57 Smith-Yoshimura K. Experimentations with Wikidata/Wikibase [EB/OL]. [2023-06-27]. <https://hangingtogether.org?p=8002>.
- 58 Neubert J. 20th Century Press Archives; Data-donation to Wikidata [EB/OL]. [2024-04-11]. <https://zlw.eu/labs/en/blog/20th-century-press-archives-data-donation-to-wikidata.html>.
- 59 Niu J. Evolving Landscape in Name Authority Control [J]. Cataloging & Classification Quarterly, 2013, 51 (4) : 404-419.
- 60 Thorsen H. LD4P; Wikidata and the Linked Data for Production Project [EB/OL]. [2023-08-07]. <https://docs.google.com/presentation/d/1-KKKDy2EZ1kK1uBIqTns8ydcq1DGIQun/edit#slide=id.p6>.
- 61 Van Veen, T. Wikidata as a universal (library) thesaurus [EB/OL]. [2023-08-08]. [https://commons.wikimedia.org/wiki/File:Wikidata\\_as\\_universal\\_library\\_thesaurus\\_-\\_tvv.pdf](https://commons.wikimedia.org/wiki/File:Wikidata_as_universal_library_thesaurus_-_tvv.pdf).

(陈辰 讲师 河北大学管理学院, 钟雨轩 河北大学管理学院图书馆学专业2021级本科生)

收稿日期: 2023-08-15