

RDA 继续前行*

克里斯·奥利弗 撰 赵丹丹 译

摘要 简要介绍了 RDA(《资源描述与检索》)这一标准的宗旨、特点以及当前的编制情况。由于客观原因,目前 RDA 的实施仍采用 MARC 格式,但 RDA 通过定义元素及强调关系的记录完全可以被应用于语义网环境中。在 2010 年首次发布后,RDA 一直在自我完善,而且随着技术进步及应用范围的扩大,RDA 前进的脚步将不会停止。

关键词 RDA 规范检索点 RDA 注册 国际化

Moving Forward with RDA

Chris Oliver

Abstract: This article briefly introduces the aim, characteristics and current compilation of the standard, RDA. Due to objective reasons, RDA is now implemented in the format of MARC, but RDA can be completely applied to the context of semantic web through determining elements and emphasizing the record of relations. After its first release in 2010, RDA is always on the road of self-improving, and with technical progress and the enlargement of application scope, nothing can stop RDA from moving forward.

Keywords: RDA; Authorized Access Point; RDA Register; Internationalization

2010 年 6 月,RDA(《资源描述与检索》)首次发布,自此以后,该标准秉承“2005—2009 战略规划”的目标和 RDA 序言的宗旨仍不断进行修订。RDA 第一版达成了最基本的目标:实现了方向性转变,探索了资源描述和检索的新方法,描述内容足以用于发布和早期实施。“首次发布”(first release)的概念在 RDA 工具套件网站和 RDA 开发联合指导委员会(JSC)网站上均没有得到明确说明,但对于 RDA 开发者来说却是心照不宣的。该措辞频繁出现于 JSC 认可的文件中,如常见问题(FAQ)、JSC 成员报告以及文件《议题延期至 RDA 首次发布后》(5/JSC/Sec/6/Rev)。RDA 中也有一些占位章节,这些占位章节的内容只有一句话:待补充。但这并不代表 2010 年 RDA 发布时是不完整的,相反,RDA 是一部可以取代并超

越其前身 AACR2 的完整标准。“首次发布”的版本仅仅是当时开发阶段的写照,并不代表开发工作的终结,而是新阶段的开始。

在 2010 版 RDA 中,我们可以看到,RDA 所代表的方向与 AACR2 是不同的。RDA 旨在成为适用于数字环境的标准,应用范围更为广泛的标准,以及以 FRBR 家族概念模型^①为基础的理论框架更加清晰的标准。在 2010 版 RDA 中,这些方面都已得以实现,且随着 RDA 的持续开发和更新,标准的内容仍在不断完善并逐渐达到更高水准。较之于 2010 版,2015 版 RDA 更加符合其当初设定的基本目标和宗旨。随着技术要求的进步和标准逐渐广泛地为国际编目领域采用,RDA 会继续得到完善。

* 本文基于 2015 年 10 月 26 日在中国国家图书馆的报告,经作者授权翻译并在本刊发表,英文原文详见《国家图书馆学刊》网站(<http://gtxk.nlc.gov.cn>)2016 年第 1 期。

① FRBR 家族概念模型是指 IFLA(国际图书馆协会联合会)开发的三个概念模型:FRBR(书目记录的功能需求)、FRAD(规范记录的功能需求)和 FRASD(主题规范数据的功能需求)。

自开发之初,RDA 就定位为用于记录数据的内容标准,而不是编码标准或记录显示标准。RDA 定义了用于著录和检索的元素,规定了元素中数据的表达格式。数据不再是长长的文本字符串,而是记录在定义清晰、标记明确的各个元素中。RDA 不再与任何编码格式或显示方式相关联。无论是 2010 版还是 2015 版,RDA 的本质为用以记录实体书目数据以及实体间关系的数据元素集。

RDA 开发时已着眼于未来。数据的粒度和精度可以由大规模的机器操作来实现,也可以应用在新的技术环境中,尤其是在关联数据的语义网中使用。RDA 特别强调准确记录关系的重要性,在后 MARC 时代,关系将成为使数据有关键要素。同时,RDA 具有内在的灵活性,预计未来将会出现多种不同的整合数据的方式。然而,我们目前不得不在 MARC 环境中实施 RDA,从而掩盖了 RDA 的前瞻性。尤其是在实施的第一阶段,编目员对于规则的转换无所适从,对 RDA 的诠释比较僵硬死板,所有人遵从相同的选项,采用相同的数据编码方案。

从 RDA 的第 6、9、10 和 11 章内容中可见其灵活性。在这些章节中多次出现下列规则:“识别属性或作为独立元素,或作为检索点的一部分,或兼作两者予以记录”。这些章节论述了有关实体作品、内容表达、个人、家族或团体的数据元素,每一章节的最后均为规范检索点的构建规则。RDA 着重于收集有关各个实体的数据,而后,编目机构可以使用其中一些数据元素来构建规范检索点,但 RDA 的主旨仍是收集关于实体的数据。

在 MARC 格式中,尤其在共享规范文档中,规范检索点是预先构建的,且必须遵循特定的顺序。但是,由于网络环境和新技术的流动性,这一固定模式终将成为过去时。RDA 的建构师汤姆·德尔塞有自己的观点,在 2010 年发给笔者的电子邮件中,重申了其观点:

独立元素的定义意味着,传统上融入带有预设结构的检索点的数据元素被当作 RDA 元素记

录时,不需要在那样的结构中进行存储和显示……过去被锁定在预设的“统一模式”检索点结构中的元素,通过未来个性化的显示方式,能更好地服务于特定用户群体的需求,或快速构建检索点,使用自动程序生成能更直接响应用户检索要求的结果显示集。

例如,一部作品的作者是“Geoffrey Brown”,这是一个常见的名字,可能有许多个叫“Geoffrey Brown”的人。现行要求是首先通过日期进行区分:

Brown, Geoffrey, 1926-

Brown, Geoffrey, 1935-

Brown, Geoffrey, 1941-

Brown, Geoffrey, 1954-

Brown, Geoffrey, 1960-

Brown, Geoffrey, 1963-

Brown, Geoffrey, 1964-

用户一定知道他们检索的作者的出生日期吗?如果一名学生想要查询一本他的教授的著作,他会知道该选哪个“Geoffrey Brown”吗?也许在未来,用户可以用更加有意义的数据来区别个人,如职业或工作:

Brown, Geoffrey (Sailor)

Brown, Geoffrey (Psychologist)

Brown, Geoffrey (Banker)

Brown, Geoffrey (Mechanical engineer)

Brown, Geoffrey (Computer scientist)

Brown, Geoffrey (Biologist)

Brown, Geoffrey (Author)

生物学专业的学生可以很快地识别和选择出他想要查询的“Geoffrey Brown”。RDA 中构建规范检索点的小节就是针对当前的环境而撰写的,但其他章节则是为未来更加灵活的环境而设计的。

编目界正在逐渐摆脱通用性,向着灵活性的方向迈步。虚拟国际规范文档(VIAF)就是一个实例。个人名称等的不同记录方式由一个标识符控制。通过将世界各国的规范文档集合在一起,VIAF 展示了一种新的模型,在该模型中没有统一

的首选形式,而是人名在不同地理区域、不同语言环境下的不同首选形式。

Confucius, 0551?-0479? av. J.-C.

Confucius

Kongzi, 551-479 v.Chr.

Confuci, ca. 551-479 a.C.

Confucius, 551-479 a.C.

Kong, Qiu v551-v479

孔子, 552-479 B.C.

Kongzi

Konfucius, 551-479 f.Kr

(来自于 VIAF)

在 RDA 开发过程中,重要举措之一是不再像 AACR2 那样使用字符串和段落,而是定义数据元素。我们使用 RDA 规则时通常是通过工具套件中的标签阅读线性文本,线性文本是依据资源描述和提供检索的要求按必要步骤组织的。在工具套件的工具体标签中,有另一种阅读模式可供选择,即将 RDA 作为数据字典。在这种模式下,RDA 规则作为元素集呈现,其中的元素和子元素按照 FRBR 模型组织。

每个 RDA 元素都是准确和清晰的,只包含一类数据,并尽可能地使用受控词汇。因此,未来任何数据元素都有可能用于为数据作索引、查找和检索以及创建有意义的记录显示方式。同时,由于数据经过精确解析,任何元素均可由人或计算机使用。数据越精确粒度就越细,因而也就越能可靠地被计算机所处理。2010 年这种方式已初见端倪,而且随着 RDA 开发工作的推进,RDA 数据的机器可操作性的提高将面临新的机遇。目前,有讨论稿提议将数量和尺寸数据拆解为更加精细和准确的数据块,以便未来更好地实现机器处理。

这一讨论稿为《机器可操作测量数据元素,载体数量、页码和叶码、尺寸、内容数量和持续时间——讨论稿》(6JSC/ALA/Discussion/5)。讨论稿提出只允许使用文本串或文本串与粒度更细的测量子元素集相结合的方式,但未来可能只采用粒度更细的测量子元素集。如果想要用不同

的编号或语言术语来呈现数据,粒度越细操作越方便。

例如:

字符串	测量子元素	
1 volume	测量类型	载体数量
	测量单位	volume
	测量值	1
approximately 600 slides	测量类型	载体数量
	测量单位	slides
	测量值	600
	测量限定词	approximately
25×35cm	测量类型	高度
	测量单位	cm
	测量值	25
	测量类型	宽度
	测量单位	cm
	测量值	35

目前针对这一讨论稿尚无决议,但已经过多次讨论,且是 JSC 会议正在进行中的议题。

JSC 创建了 RDA 命名空间,在 RDA 注册网站发布了元素集和受控词汇。RDA 注册网站为 RDA 元素集和 RDA 受控词汇通过网络获取提供了基础,适用于关联数据环境。就像关联数据环境中的使用者预期的那样,数据以计算机可读格式使用,而且数据是公开获取的。这样数据不仅可以为图书馆界所用,也可为其他文化传承机构使用。不论 RDA 元素集是被全部应用,还是被部分应用,都有助于数据的互操作。

RDA 是一套实用的规则,目前作为 AACR2 或其他编目标准的替代品可以在许多环境中应用,而图书馆仍然采用传统的 MARC 环境。但 RDA 在将图书馆已经创建和正在创建的数据迁移到新技术环境中发挥着重要作用,确保数据在未来可以被有效地利用。

RDA 旨在成为应用范围更广的标准,有两层含义:一是与除图书馆外的其他文化传承机构相

关联,如档案馆和博物馆;二是打破前身的英美限制,使得 RDA 在不同语言和文化背景中更易使用。在 2010 年首次发布中,RDA 的 0.11 小节声明如下:RDA 是为应用于国际化环境而设计的。这一声明彰显和表明了 RDA 的一个重要意图。RDA 的前身 AACR2 虽被翻译成多种语言,并在除澳大利亚、加拿大、英国和美国四个编制国之外的许多国家应用。但是,AACR2 是从英美角度编写的,偏向于一系列特别的参数集——语言、文字、编号、日历、文化等。在 2010 年以前,RDA 的大部分内容也是由上述四个拥有相同语言和相似文化的编制国开发完成的。作为一个双语国家,加拿大在四个编制国中具有唯一的发言权,时常提醒编制组要意识到规则只适用于英语国家的局限性。国际化的重要性是毫无疑问的,但真正付诸实施是需要时间的,也需要来自其他国家的反馈。随着 RDA 实施者队伍的不断壮大,将有更多的机会识别英美偏见的残留并将之消除。

变更作品名称中首冠词的著录规则就是一个成功的案例。2012 年 4 月,RDA 对一条原有规则进行了变更,而这源于德国国家图书馆的一项提案。德国国家图书馆在翻译 RDA 的过程中发现,其中 6.2.1.7 的主要规则沿袭了英美的编目实践。主要规则规定省略首冠词,而交替规则规定记录作品题名的首冠词。鉴于首冠词在有词尾变化的语言如德语中的重要作用,这一规定会为其实施 RDA 带来许多问题。德国国家图书馆提议将这两个规则互换位置,消除 RDA 中的英美偏见,最终获得通过并付诸实施。

JSC 采取了两项举措来支持和扩大国际化:成立翻译工作组,对负责 RDA 开发的 JSC 和负责人委员会实行新的管理结构。

随着 RDA 翻译语种数量的增长,JSC 发布了一份关于翻译政策的文件(Translation Policy for RDA and RDA Toolkit 6JSC/Policy/6)。在这份文件中,JSC 决定设立翻译工作组,这是因为过去几年间来自翻译团队的反馈一直是富有启发性的,而且也为标准的修订做出了积极贡献。下面截

取了关于设立新工作组的部分文字:

JSC 希望能够从翻译者处获得关于如何完善标准的信息,尤其是对非英语母语者来说存在问题的短语或词汇的选择,以及可能体现文化偏见的规则方面的信息。为了方便这一信息的获得,JSC 成立了翻译工作组,负责提供语言方面的报告,并帮助 JSC 确定解决问题的方法。希望各语言翻译小组可以向 JSC 提供此类信息,各语言翻译小组的成员可以作为代表加入翻译工作组。

(RDA 与 RDA 工具套件翻译政策 6JSC/Policy/6)

工具套件的结构中比较值得一提的是各翻译语言已经整合于其中,而且可以实现整个文本或文本中某个规则在不同语言间的切换,可以比较不同翻译语言对某一特定规则的表达方式,也可以通过分屏对不同语言间的文本进行比较。不同翻译语言的文本通过段落标识符相互关联,这些标识符也包含了语言版本和内容版本(如,与哪一年更新相符)的信息。各翻译版作为同一作品的多样化表现形式,在同一地址、用同一订阅号呈现,肯定了翻译版作为主要文件的重要性,而不是将其作为分离的、辅助的文件。

“2015—2020 战略规划”的首要目标是使 RDA 成为国际认可的标准,其次是促进 RDA 在世界范围内的实施。为了实现这些目标,该规划建议进行根本性的变革,实行新的管理结构,扩大 RDA 规则制定的参与范围,保证决策过程中的多角度考虑。JSC 和负责人委员会重新被命名为 RDA 指导委员会(RSC)和 RDA 董事会,将由专家和地区代表共同组成。地区代表将不再仅限于原有的编制国,而是基于联合国地理区域。RDA 董事会将包括来自国家机构的代表,来自非洲、拉丁美洲和加勒比地区、北美洲、亚洲、欧洲和大洋洲六大地区的代表各一位。RSC 也将包括以上六大地区代表各一位,此外,还包括来自翻译工作组的代表。RSC 和 RDA 董事会为实现国际化正在努力构建与之相匹配的管理结构,以便在范围更广泛的国际代表中进行决策制定。

RDA 的本质特征之一为它是一部基于完善理论框架的实用标准。在开发的过程中,RDA 就

明确地与 FRBR 和 FRAD 相协调。在 RDA 的导言中,即声明了这一标准与 IFLA 概念模型之间的协调性。RDA 的结构,用词,用户任务的重要地位,依据实体、属性和关系理解书目数据以及对关系的强调,这些都是其与 IFLA 概念模型相协调的佐证。在 2015 年 4 月的更新中,RDA 完成了与 FRAD 之间的协调,使得其与 FRBR 家族概念模型完全相符。实际上,RDA 并没有增加实体和属性用于定义主题,或增加记录这一数据的相应规则,而是保持在高水平的通用性层面。主题可以依据语言、文化、学科、资源的特殊性等不同方式分类。RDA 没有定义主题元素,主要为第 23 章增加了有关主题关系的文本内容。

在 2010 年版中,RDA 简要地确认了主题关系的重要性:至少应包括一个主题关系元素,且包含在 RDA 0.6(核心元素)中。当时,第 23 章还没有文本。2015 年,与 FRBR 家族模型的结合得到扩展并囊括了 FRAD。关于与模型相协调的部分从 RDA 0.3 变换到了 0.2,处于 RDA 导言章节中更重要的位置。第 23 章新增了文本,定义了术语、功能目标和原则、核心元素、记录主题关系和关系标识符的规则(遵循其他章节中的模式),还增加了附录 M,罗列出用于主题关系标识符的受控词汇。

第 23 章的文本中引入了一个新概念——可识别的主题系统,指的是“创建数据的机构使用的主题检索点和/或分类号的标准”。这可以实现与外部主题方案或受控词表的关联,而无需预先决定主题检索如何组织或使用何种术语。这也证明了 RDA 并不旨在成为统一模式,而是满足多种需求,促进标准化框架的灵活性。

与 FRBR 家族三个概念模型相协调标志着 RDA 之前未完成工作的完结。然而这并非结束,RDA 仍需要不断改进以保持与 FRBR 家族概念模型之间的协调。IFLA 的 FRBR 评估组正在将

三个模型合并为一个模型,合并中需要重新建模以解决模型间的冲突或不一致。建模过程中也为统一模型引入了流线型,取消了一些实体并引入了新的实体,统一模型没有完全颠覆前三个模型,而是沿袭了其实质内容。统一模型可能会对 RDA 的元素和结构产生影响,可能会改变我们对数据某些部分的理解,但不会从根本上影响数据的记录。这一模型正在开发中,期望于 2016 年接受世界范围内的评估。在模型没有发布前,其对 RDA 的影响还不确定。JSC(现在称为 RSC)掌握了统一模型的研发情况,并与 FRBR 评估组保持着密切的沟通。JSC 认为,统一模型会是对 RDA 的内容和结构产生影响的因素之一。2015 年初,JSC 决定采用务实的工作原则:如果对 RDA 某一领域的修订建议有可能会被外部变化影响,那么该领域的修订则推迟或交由工作组考虑特定外部变化的总影响。

与首次发布的版本相比,2015 年的标准已大不相同。为了推进开发并使之与愿景和目标更好地匹配,RDA 已经做了许多改变。这些改变并没有使 RDA 偏离原来的定义特征:数字世界的编目标准;与更大范围的文化遗产机构和国际编目界联合;通过与 IFLA 概念模型相协调形成完善的理论框架。自从首次发布开始,RDA 始终致力于成为灵活性和适应性强的标准,并面向其他机构开放。为了实现国际化的目标,RDA 将会不断地完善以适应技术进步以及不同语言和文化地区的需求。

(克里斯·奥利弗 渥太华大学图书馆资源描述与元数据服务中心主任、加拿大编目委员会主席、国际图联 FRBR 评审组主席,赵丹丹 副研究馆员 国家图书馆外文采编部)

收稿日期:2015-12-19