

传统民族服饰数字化采集元数据构建

周 伟¹, 赵海英^{1,2}

(1. 北京邮电大学数字媒体与设计艺术学院, 北京 100876;
2. 北京邮电大学世纪学院移动媒体与文化计算北京市重点实验室, 北京 102101)

摘 要: 针对当前传统民族服饰元数据数字化采集没有统一的标准和规范以及采集后的元数据没有进行文化内涵深入挖掘的问题, 以国际通用的 DC 元数据为基础, 提炼出传统民族服饰数字化采集基础元素, 深入挖掘传统民族服饰文化内涵进行基础元素扩展, 同时与服装、刺绣以及纺织图片元数据作对比映射, 构建出一套基于文化本体、内涵挖掘、数字化色彩描述的多层可扩展的民族服饰的元数据, 以蝙蝠纹蜡染背扇为例, 说明传统民族服饰数字化采集元数据的标注过程, 并通过 RDF/XML 的元数据标注和 SparQL 的元数据查询仿真, 实现了一套更具可操作、可扩展的传统民族服饰数字化采集规范。

关 键 词: 民族服饰; 数字化; 元数据; RDF/XML; SparQL

中图分类号: G 250.73

DOI: 10.11996/JGJ.2095-302X.2018061183

文献标识码: A

文 章 编 号: 2095-302X(2018)06-1183-09

Metadata Construction of Digital Collection of Traditional Costumes

ZHOU Wei¹, ZHAO Haiying^{1,2}

(1. School of Digital Media and Design Arts, Beijing University of Posts and Telecommunications, Beijing 100876, China;
2. Beijing Key Laboratory of Mobile Media and Cultural Computing, Beijing University of Posts and Telecommunications, Beijing 102101, China)

Abstract: In the process of current traditional national costume metadata digital collection, there is no uniform standard and specification, and the collected metadata does not carry on the deep excavation of cultural connotation. Firstly, based on the international common DC metadata, the basic elements of digital collection of traditional national costume are extracted. Then the basic elements of the traditional national costume culture are deeply excavated and expanded. At the same time, a contrastive mapping with clothing, embroidery and textile picture metadata is made, and a set of cultural ontology, connotation mining, and the excavation of the connotation are constructed. The metadata of multi-layer and extensible national costumes is described in digital color. Finally, with the bat pattern batik back fan as an example, the annotation process of digital collection metadata of traditional national costume is illustrated. Through the RDF/XML metadata annotation and SparQL query metadata simulation, the paper implements a set of more operational specifications on the digital collection of traditional national costumes.

Keywords: national costume; digitalization; metadata; RDF/XML; SparQL

传统民族服饰是人类社会发展和进步的一种 载体, 是人类不断在实践活动中探索出来的艺术之

收稿日期: 2017-11-27; 定稿日期: 2018-04-23

基金项目: 北京市科委课题(D171100003717003); 甘肃省人才引进项目(2015-RC-47)

第一作者: 周 伟(1994-), 男, 安徽安庆人, 硕士研究生。主要研究方向为文化计算与媒体数据挖掘。E-mail: 1556520619@qq.com

通信作者: 赵海英(1972-), 女, 山东烟台人, 副教授, 博士。主要研究方向为文化计算与媒体数据挖掘。E-mail: zhaohaiying@bupt.edu.cn

美,是人类劳动文明的智慧结晶。民族服饰的文化凝聚着历史的记忆,寄托着我国人民的民族情怀,重视传统文化的继承与发展,是一个国家文明程度高低的标志^[1]。然而,随着现代化进程不断加快,对于传统民族服饰的保护工作相对滞后,当前的传统民族服饰元数据采集没有统一的采集规范和标准,以及采集后的元数据没有进行内涵深入挖掘,不利于传统民族服饰的传承和发展。正是在此背景下,本文提出一种具有科学性和适应性的面向传统民族服饰数字化采集的元数据构建方案。

1 相关工作

元数据(meta data)被定义为“关于数据的数据”,或是描述和限定其他数据的数据。元数据作为描述信息资源的特征和属性的结构化的数据,具有定位、发现、证明、评估和选择信息资源等功能^[2]。元数据是描述任何 INTERNET 信息资源组织和发现的数据,借鉴图书馆处理图书资料的方法,采用在现有网页上添加一组 meta 标记的方法来描述该网页的作者、关键词等信息,从某种意义上说,meta data 是一种扩大化、普及化的编目方法^[3]。

1.1 国内外常见元数据标准

目前国际上比较有影响且受到广泛应用的标准有:CDWA (categories for the description of work of art)、FGDC (federal geographic data committee)、VRA (visual resources association)、DC (dublin core)、GILS (government information locator service)。CDWA 由艺术信息专业组织(art information task force, AITF)颁布实施,适用于艺术品及数字图像资源描述;FGDC 是美国国家层面的一个标准化机构委员会,其宗旨是协调发展国家基础地理空间数据的使用、共享和传播^[4];VRA 由美国视觉资料协会制定,适用于艺术、建筑、史前古器物以及民间文化等艺术类三维实体的可视化资源描述;DC 由美国 OCLC 公司发起,国际性合作项目 Dublin Core Metadata Initiative 设计,适用于网络资源描述;GILS 由美国政府提出并实施,即政府信息定位服务,是一种支持公众搜寻、获取和使用政府公开信息资源(包括电子信息资源)的开放环境下的分布式信息资源及利用体系^[5]。以上 5 个国际化常见的元数据标准,CDWA、VRA 和 DC 元数据标准从其适用类型看对构建传统民族服饰数字化采集有重要的参考价值,本文将元数据的元素进行提取,对

比元数据间的映射关系,总结出本文传统民族服饰数字化采集的元素为:题名、日期、创建者、主题、类型、描述、格式、来源、权限、标识符、语种、分类、地点、文化环境、时代样式。这些元素对传统民族服饰数字化采集元数据的构建具有重要的参考价值,但对民族服饰数字化采集特性描述的针对性、准确性还远远不够,无法深入挖掘民族服饰文化的内涵,也无法具体有效的描述民族服饰的数字化色彩,因此需要对元数据进行扩展。

与国外元数据相比,我国对元数据的研究相对较晚,但在国外元数据标准推动下发展迅速。由中国国家图书馆牵头制定的《中文元数据方案》,适应于海内外中文数字资源的建设、保存和共建共享服务交流;由中国高等教育数字图书馆(China academic digital library & information system, CADLIS)发布一系列技术标准与规范,项目研究成果“专门数字对象描述元数据规范”,适用于古文献系列、论文系列、图书序列、视音频资料系列和网络资源系列^[6];由北京大学数字图书馆研究所主持的“中文 metadata 标准研究”重点研究项目建立了拓片和舆图元数据研究、拓片元数据著录实验系统和古籍元数据著录实验系统 3 个数据库,开展中文元数据的研究^[7];由中国 21 实际议程管理中心、国家科委、国家经贸委和中国科学院等共同开发的“中国可持续发展信息示范系统”,推出适用于地理、海洋、植物、自然灾害等元数据的数据库^[8]。尽管国内各领域元数据的研究取得不少成就,但对各个领域的分支研究明显不足,仍有很大的发展空间。

1.2 国内服饰元数据标准研究

我国对于服饰元数据的研究工作较为深入,取得不少研究成果。文献[9]对服装图片元数据方案进行长期研究,提出了 23 个元素的服装图片元素集,其中有 4 个元素是服装图片的特有属性;文献[10]对纺织品图片元数据进行标准规范,提出了 22 个元素的纺织品图片元素集,其中有 4 个专门核心元素是根据纺织品的特点制定,有 6 个元素是以特定的纺织品为基础制定的;文献[11]对刺绣图片元数据做了著录规范研究,提出了 21 个元素的刺绣图片元素集,其中有 3 个元素是系统自动生成;袁莉和张晓林^[12]选取 MOA2、CDL、VRA Core 等对数字图像元数据进行了系统介绍;杨西宁等^[13]在开展敦煌学数字图书馆的建设项目中,

参考 CDWA、DC 等元数据标准,提出了敦煌彩塑元数据和敦煌壁画元数据标准。此外,还有很多学者和研究机构对服饰、图片、刺绣以及文化遗产等元数据做了相关研究,推动了相关领域元数据著录规范和标准的研究,为我国文化遗产的保护做出了贡献^[14]。但具体针对传统民族服饰数字化采集的元数据标准研究十分欠缺,无法深入挖掘民族服饰文化内涵以及有效描述民族服饰的数字化色彩,不利于促进传统民族服饰的传承和发展。如果完全复用 DC 元数据,则无法准确的、深入的描述和挖掘民族服饰的内涵,因此本文针对传统民族服饰提出一种数字化采集元数据的构建方案。

1.3 形式化描述语言

LEE 等^[15]基于 XML 的描述语言开发出叙述标记语言(TML),构建了叙述描述 Ontology 的框架;任磊和谭跃生^[16]通过研究 RDF 资源描述框架和网格应用中的元数据,提出了一个基于 RDF 元数据描述的网格资源统一描述方法;曾新红^[17]依据我国情报检索语言的权威理论和构建标准,基于 SKOS 标准对中文叙词表及分类法中各项语义元素的 SKOS 描述进行了具体的规定或建议;FISCHER^[18]基于 OWL 制作出美国国家癌症研究所(NCI)叙词表的 NCI-OWL 文档。虽然 XML 提供明显结构和数据类型约束的支持,但是只提供很少的语义支持,而这种语义支持对于实现元数据域的可扩展性和动态交互是十分必要的;RDF 虽然提供丰富的语义描述支持,但对本地使用约束规格只提供有限的支持^[19];SKOS 虽然丰富和扩展了 RDF 的描述能力,但缺乏对元素的限制和约束;OWL 在 RDF 的基础上添加了更多的用于描述属性和类的词汇,但描述过程的人工干预增强,需要人员的参与和确认,增加了实施难度和成本^[20]。因此,本文利用 RDF 模式和 XML 模式的互补性:RDF 以一种标准化、互操作的方式来规范 XML 的语义,XML 文档可以通过简单的方式实现对 RDF 的引用,从而便于 Web 数据的检索和相关知识的发现。

2 传统民族服饰数字化采集元数据构建方法

元数据格式的设计应从4个方面出发:①要对用户的需求及资源的关系进行全面分析。②对资源

组织、管理与保存,资源发现、评价与检索,提高资源利用的便利性、价值性与共享性方面要有一定的目的性。③对资源的著录要有客观性、专指性。④要采用自然语言与叙词语言相结合,集中著录与分散著录相结合,手工标引与计算机辅助标引相结合,完善标引规范的建立^[21]。参考这4个方面元数据的设计格式,本文首先提炼出传统民族服饰数字化采集的核心元数据集,然后从文化内涵挖掘出专门元数据集,最后针对传统民族服饰色彩的特点提出数字化色彩描述元数据集。从3个层次划分采集元数据集,目的是为了保证在元数据标准一致性和整体性的基础上,最大限度的体现不同资源的特征^[22],更大限度的挖掘其背后的文化内涵,对传统民族服饰的传承和发扬起到推动作用,同时有利于标准的推广和应用。

2.1 传统民族服饰数字化采集元数据定义与对比分析

考虑到元数据标准规范的通用性、可扩展性、互操作性和数据库可查询性等需求以及目前应用最广的网络资源描述型 DC 元数据,本文的核心元数据将基于 DC 核心元数据集设计。

首先对 DC 核心元数据集的部分核心元素进行复用,然后对服装图片元数据、刺绣图片元数据和纺织品图片元数据进行元数据映射,同时结合传统民族服饰数字化采集元数据的特点,以深入挖掘其背后的文化内涵,对元素进行扩展和新增,使得构建的传统民族服饰数字化采集元数据方案更具有科学性和互操作性。本文的元数据构建与以前研究者设计的方案有所不同,具体包括:

(1) 在面向文化本体元数据层级上,复用了 DC 元数据的7个元素,即题名、日期、创建者、类型、权限、标识符和语种。另外,将 DC 元数据中的“主题”拆分为“检索主题词”和“主题分类号”,将“格式”拆分为“载体格式描述”和“数字格式描述”,将“出版者”修改为“其他责任者”,删去“覆盖范围”字段,新增“民族”字段,从而更好记录所采集服饰的具体民族,以至于不会混淆。同时,由于是数字化采集,需要记录描述服饰实体的物理存储位置和描述服饰数据源的虚拟存储位置,新增“实物存储位置”和“数字存储位置”。因此,本文基于文化本体的元数据有14个元素,同时与 DC 元数据、服装图片、刺绣图片以及纺织品图片元数据对比映射(表1)。

表 1 传统民族服饰数字化采集面向文化本体元数据及对比表

核心元素	元素定义	与 DC 元数据 映射	与服装图片元数 据映射	与刺绣图片元数 据映射	与纺织图片元数 据映射
标识号	便于管理赋予的资源标识号	标识符	标识	标识	唯一标识符
标题	服饰名称	题名	题名	题名	标题
类型	服饰所属类型	类型	分类	分类	分类
创建者	传统服饰采集过程的责任人和创作者	创建者	创建者	创建者	创建者
创建日期	服饰数据创建发表并登记的特定日期	日期	日期	日期	日期
其他责任者	采集该传统服饰的其他责任者	出版者	-	-	-
描述语种	描述该服饰数据源的语言类型	语种	-	-	-
民族	服饰所属的民族	-	-	-	-
检索主题词	描述服饰的主题内容	主题	主题	主题	主题
主题分类号	依据服饰主题内容进行分类赋予编号				
实物存储位置	描述服饰实体的物理存储位置	-	-	-	-
数字存储位置	描述服饰数据源的虚拟存储位置	-	-	-	-
载体格式描述	描述服饰数据源的传统载体信息	-	-	-	-
数字格式描述	通过数字格式描述服饰 数据源的逻辑和物理形态	格式	-	-	-

(2) 在面向文化内涵挖掘元数据层级上，将 DC 元数据中的“来源”拆分为“来源时间”和“来源出处”，将“关联”拆分为“隶属关系描述”和“包含关系描述”，将“描述”细化为“领域专家描述内容”，使得对采集物件的描述更具有专业性和权威性。为了突出描述民族服饰的特点，本文参考李晓玲等^[9-11]对服装图片、刺绣图片以及纺织品图片元数据的研究，增加“穿着方式”和“使用场合”，同时为记录各个民族服饰的着装人群，新增“使用人群”；将“色彩”、“款式”、“材质”、“工艺”和“度量”合并成“风格特征”；将“文化”

修改为“价值特征”，可描述民族服饰同时具有的“文化”和“艺术价值”。为了更大限度的挖掘传统民族服饰的文化内涵，需要标识出该服饰数据源代表的时间范围和有效地域范围，因此本文新增“时间跨度”和“地域跨度”，同时为了体现在时间、空间下的文化价值变迁和发展脉络，新增“时空演变”，为了表达服饰文化形成与发展的历史过程，新增“历史演变”。因此，本文基于文化内涵挖掘的元数据有 15 个元素，同时与 DC 元数据、服装图片、刺绣图片以及纺织品图片元数据对比映射(表 2)。

表 2 传统民族服饰数字化采集面向文化内涵挖掘元数据及对比表

核心元素	元素定义	与 DC 元数 据映射	与服装图片元 数据映射	与刺绣图片元 数据映射	与纺织图片元 数据映射
权限	描述服饰文化数据源的权限信息	权限	权利	权利	权限
领域专家描述	描述服饰内涵并反映其文化主题内容	描述	描述	描述	描述
隶属关系描述	标识和定义服饰文化数据源在文化不同层次之间的关系	关联	-	关联	关联
包含关系描述	标识和定义服饰文化数据源在文化同一层次之间的关系		-		
来源出处	描述该服饰数据源的物理出处信息	来源	来源	来源	来源
来源时间	描述该服饰数据源的时间出处信息	-	-	-	-
穿着方式	衣着方式	-	穿着方式	-	-
使用场合	服饰的用途，或是适用不同场合	-	使用范围	使用范围	-
适用人群	着装人群，用户群体	-	-	-	-
价值特征	描述服饰具有的文化和艺术价值	-	文化	文化	文化
风格特征	描述服饰色彩、款式、材料、工艺和度量	-	色彩/款式/材料 /工艺/量度	色彩/款式/材料 /工艺/量度	色彩/材料 /量度
时间跨度	标识该服饰数据源代表的时间范围		-	-	-
地域跨度	标识该服饰数据源代表的有效地域范围		-	-	-
时空演变	在时间、空间下文化价值变迁和发展脉络		-	-	-
历史演变	表示服饰文化形成与发展的历史过程		-	-	-

(3) 在面向数字化色彩描述元数据层级上，由于在数字化采集过程中，运用相机拍摄或者扫描仪扫描，从而导致拍摄中曝光对色彩的影响，集中体现在对色相、纯度和明度的影响(色相、纯度和明度也称色彩三要素)，只有正确运用这些知识，正确的控制

曝光，才能准确的再现被拍摄物的色彩，更加真实、准确记录传统民族服饰的色彩^[23]，因此本文新增“色相”、“纯度”和“明度”，本文基于数字化色彩描述的元数据有 3 个元素，同时与 DC 元数据、服装图片、刺绣图片以及纺织品图片元数据对比映射(表 3)。

表 3 传统民族服饰数字化采集面向数字化色彩描述元数据及对比表

核心元素	元素定义	与 DC 元数据映射	与服装图片元数据映射	与刺绣图片元数据映射	与纺织图片元数据映射
色相	描述服饰颜色的类别	-	-	-	-
纯度	描述服饰颜色的类别	-	-	-	-
明度	描述服饰颜色的亮度	-	-	-	-

2.2 传统民族服饰数字化采集元数据规范构建

元素限定词能够缩小元素的含义和描述对象的范围，使其更具有专指性。限定词分为描述对象限定词、复合限定词和基本限定词。其中，基本限定词是元素限定词中最基本的信息单元，不能再进行元数据的扩充；复合限定词是由若干基本限定词、或者基本限定词与其他复合限定词、或者若干其他限定词共同组成的；描述对象限定词是对描述对象差异性较大的几个元素加以描述的元素限定词^[24]。

本文提出的模型采用 32 个核心元数据和基本限定词，同时结合不同传统民族服饰元数据数字化采集属性的必备性，限制为 3 个选择项，即“必填，由系统生成”、“必填”和“有则必填”。同时选取邹加勉^[25]的著作《中国少数民族图案与配色》中蝙蝠纹蜡染背扇为例，如图 1 所示，说明传统民族服饰数字化采集元数据的标注过程，见表 4。



图 1 蝙蝠纹蜡染背扇

3 传统民族服饰数字化采集元数据应用实例

3.1 基于 RDF/XML 的数字资源元数据标注

本文所构建的元数据主要是面向传统民族服

饰进行数据采集而提出的一套数据描述结构；需要 RDF/XML 对元数据属性进行标记，便于计算机理解。共有 32 个核心元数据和基本限定词，利用 RDF 的三元组模型^[26]对元数据的 3 个层次分别进行描述，并采用自定义命名空间，对逐个核心元素进行命名。

三元组模型为

{N, P, N or V} (1)

其中，N(node)为节点；P(Properties)为属性；V(values)为取值。

本文以面向文化内涵挖掘元数据描述为例，采用 CD 命名空间，对核心元素“权限”、“来源出处”分别命名为“Power”、“Source”，以白彝族刺绣为例，如图 2 所示。



图 2 白彝族刺绣

对图 2 进行元数据文化内涵挖掘层标注，部分代码如下所示：

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:tnc="http://traditionalnationalcostume.edu.cn/charm/"
  xmlns:cd="http://www.dress.org/cd/elements/">
  <rdf:Description
    rdf:about="http://traditionalnationalcostume/">
    <cd:Power>公开</cd:Power>
```

表 4 传统民族服饰数字化采集核心元数据元素限定词

层级	核心元素	元素限定词	必备性	案例	参考项
面向文化本体元数据	标识号	系统生成	必填, 系统自动生成	BFWLRBS-1	-
	标题	正式名称/英文名称 /缩略名称	必填	蝙蝠纹蜡染背扇	-
	类型	实物/手稿/资料/图书/ 图片/视频/其它	必填, 系统自动生成	图片	-
	创建者	创建者标识号/所在单 位/联系方式	必填	项目组	-
	创建日期	创建日期/发表日期 /登记日期	必填	2017-11-03	-
	其他责任者	-	有则必填	-	-
	描述语种	简体中文/繁体中文/地 区方言/少数民族语言	必填	简体中文	-
	民族	56 个民族	必填	苗族	-
	检索主题词	核心词/标注字段	必填	苗族服饰、背扇、襁褓	-
	主题分类号	主管部门编制的 分类体系	必填	-	-
	实物存储位置	具体位置/地理坐标	必填	-	-
	数字存储位置	具体位置	必填	C:\传统民族服饰\苗族	-
	载体格式描述	数字/实物	必填	数字	-
	数字格式描述	电子形态/数据格式 /文件大小	必填	<数据格式> .jpg <文件大小> 1.02 M	-
	权限	公开/限制/秘密 /机密/绝密	必填	公开	参见《文献保密等 级代码与标识》
	领域专家描述	由各行业各领域的领 域专家进行标注描述	有则必填	母爱、聪慧、健康、坚定执着、忠贞不渝、 自然和谐、对太阳的精神崇拜	-
	隶属关系描述	组成、部分、时期	有则必填	-	-
	包含关系描述	被...组成/被...部分 /被...时期	有则必填	-	-
	来源出处	-	有则必填	-	参见《馆藏文物登 录规范》
	来源时间	-	有则必填	-	
面向文化内涵挖掘元数据	穿着方式	上衣下裙/上衣下裙 /裙装式/其他	必填	背扇/襁褓	-
	使用场合	日常/盛装/祭祀/其他	必填	日常	-
	适用人群	老人/妇女/儿童 /成年男子/其他	必填	贵州少数民族劳动妇女 /未婚的少女	-
	价值特征	文化/艺术	有则必填	少数民族工艺、 非物质文化遗产 背扇、襁褓 冰纹、动物纹、太阳纹 蓝色、白色 棉麻 点蜡、缸染、脱蜡	参见《文物馆藏定 级标准》
	风格特征	结构/纹样/色彩 /材质/工艺/度量	必填		-
	时间跨度	公元 XXXX 年...	有则必填		参见《博物馆藏品 信息指标著录规 范》
	地域跨度	时期版图: 秦、汉等	有则必填	-	
	时空演变	公元 XXXX 年— 公元 XXX 年	有则必填	-	-
	历史演变	按照历史路线轨迹, 分析各个时期的路线 变化与迁移	有则必填	-	参考《丝绸之路》
	色相	色相距离[0, 180°]	必填	-	-
面向数字化色彩描述元数据	纯度	纯度等级[0, 10]	必填	-	-
	明度	亮度等级[1, 9]	必填	-	-

```
<cd:Source>中国少数民族图案与配色</cd:Source>
<cd:Wearingstyle>马甲、坎肩</cd:Wearingstyle>
<cd:RetrievalofKeywords>大白彝、彝族，挑花刺绣、传统、服饰、刺绣</cd:RetrievalofKeywords>
<cd:Expertdescription>圆领外镶花背面图案十分精彩，
绣工精致，图案中有纵横十字轴和斜线轴，各种图案中心对称，
对称中心形似亭台，斜线轴上为树，纵轴上为花瓶，中心
为八角，第二环为花开石榴
</cd:Expertdescription>
<cd:Stylecharacteristics1>
    结构：对称、线形、纵横十字轴、斜线轴、菱形
</cd:Stylecharacteristics1>
<cd:Stylecharacteristics2>
    纹样：菱形、蝴蝶、花纹、马缨花、犬齿纹、石榴
    花、石榴、花开石榴、几何纹、蝶恋花、八角纹
```

```
</cd:Stylecharacteristics2>
<cd:Stylecharacteristics3>
    色彩：黑色、蓝色、蓝色、粉色、玫红色、黄色，
    绿色、红色，白色、褐色、紫色
</cd:Stylecharacteristics3>
<cd:Stylecharacteristics4>
    工艺：纺线、捻线、织布、染色、缝制、刺绣、镶
    嵌、挑花、散花、按花、扑花、穿花、堆花
</cd:Stylecharacteristics4>
</rdf:Description>
</rdf:RDF>
```

将上述的代码放在 W3C 的 RDF 在线验证器上进行验证，如图 3 所示。在验证器的显示结果选项中选择有向图模式，并做编码中文化处理，在线生成的有向图如图 4 所示。

Check and Visualize your RDF documents



图 3 RDF 在线验证器

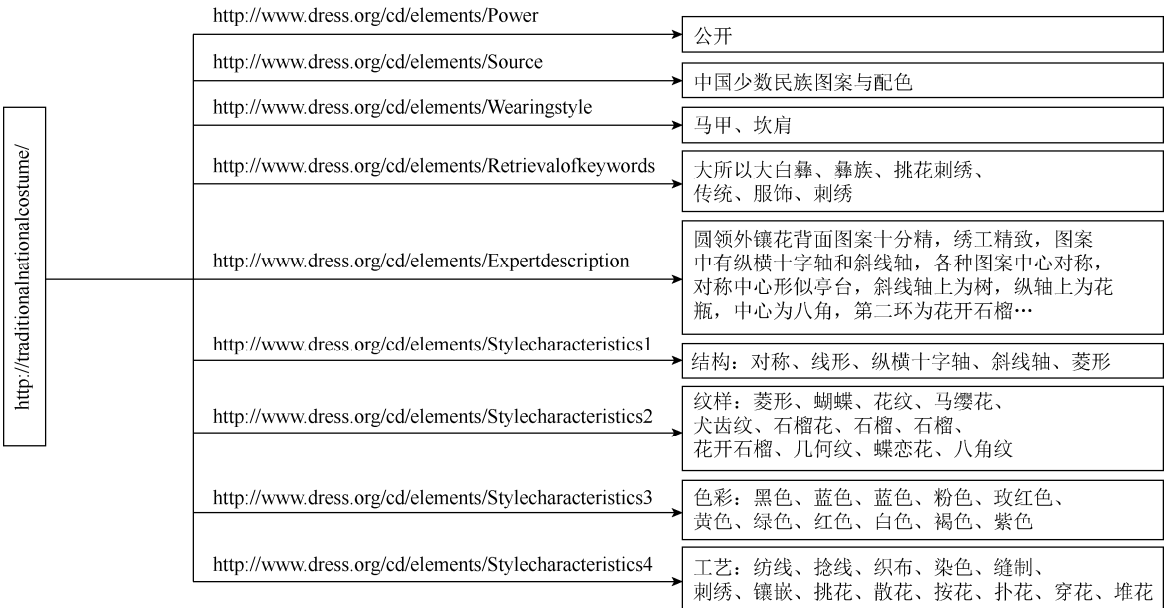


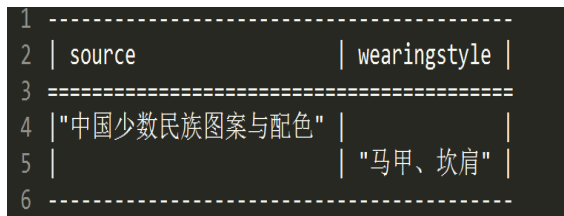
图 4 RDF 在线生成有向图

3.2 基于 SparQL 的数字资源元数据查询

SparQL (simple protocol and RDF query language) 查询语言通过图形模式(Graph Pattern)匹配实现查询功能^[27], 是一种面向于 RDF 模型的查询语言和数据访问的协议^[28], 是为 W3C 所开发的 RDF 数据模型所定义, 但是可以用于任何可以用 RDF 来表示的信息资源, 其查询结果以变量的绑定集合或 RDF 图的形式返回。SparQL 类似于 SQL 查询语言的 select-from, 区别在于其以三元组模式作为查询基础, 典型的 SparQL 查询语句由 SELCET、FROM、WHERE 三部分组成。在 oracle 官网中下载 JDK 然后安装, 并配置 Jena 开发环境。例如, 查询白彝族刺绣中的来源和穿着方式, 相关查询语句如下所示:

```
PREFIX cd: <"http://www.dress.org/cd/elements/">
SELECT ?source ?wearingstyle
WHERE
{
    [] cd: Source ?source .
    [] cd:Wearingstyle ?wearingstyle
}
```

其对应的查询结果如图 5 所示。



source	wearingstyle
"中国少数民族图案与配色"	"马甲、坎肩"

图 5 查询结果

3.3 选用 RDF/XML 标注和 SparQL 查询的优点小结

元数据标注和查询功能的实现必须依靠相应的软件平台进行信息的组织和管理。作为一种结构化的数据, 元数据可应用的系统和软件平台有很多, 但要在各应用中有效的使用元数据, 必须对元数据规范进行描述和定义。而 XML 是一种元标注语言, 简化了网络中数据交换和表示, 作为数据交换的标准格式, 使得代码、数据和表示分离^[29]。同时, RDF 定义了一个独立于任何特定的语法结构的数据模型, 可用来描述 Web 上的各种资源, 可以将各种资源进行统一识别和规范^[30]。将 RDF 与 XML 的优势互补封装成 RDF/XML, 就是以 RDF 作为约束规则, 同时用 XML 对异构资源进行形式化描述, 从而真正实现重用和共享资

源库, 解决资源库的建设中存在标准不一, 数据格式多样, 结构复杂等问题^[30]。由图 4 可发现 RDF 有向图能很好地展示被描述的资源 and 用来描述资源的属性值及其之间的相互关系。与此同时, SparQL 查询是针对 RDF 数据的语义查询技术, 能以 URIS、空白结点、无格式和类型文字的形式提取信息、可以提取 RDF 子图、能在查询图中构造基于信息的新 RDF 图^[31], 如图 5 所示, SparQL 能很好查询到 RDF 中的内容。

4 结束语

本文以国际通用的 DC 元数据为基础, 并针对传统民族服饰的特点以及深入挖掘传统民族服饰文化的内涵进行元素扩展, 提出了传统民族服饰数字化采集的元数据构建方案。该构建方案不仅满足科学性和适用性, 并且能实现互操作性, 同时采用 RDF/XML 和 SparQL 进行元数据统一的标注和快速的查询, 实现了对元数据的语义化描述、挖掘、存储和查询。今后将通过数字化采集充分利用元数据建立传统民族服饰资源库, 推动传统民族服饰的传承和发展。同时在后续的研究中, 将进一步优化标注和查询方式, 使其更加便捷、高效。

参考文献

- [1] 王晶. 基于民族服饰文化传承与创新的分析[J]. 西部皮革, 2016, 38(16): 37.
- [2] 刘嘉. 元数据导论[M]. 北京: 华艺出版社, 2002: 67.
- [3] 张学福, 肖荣荣, 王晓谦. 元数据及其在网络信息资源组织开发中的应用[J]. 现代情报, 2002, 122(5): 48-49.
- [4] 周倜. FGDC 空间数据标准概况[J]. 航天标准化, 2010(4): 19-22.
- [5] 赵志荣, 张晓林. GILS 元数据的结构 and 应用[J]. 情报理论与实践, 2000, 23(6): 445-448.
- [6] 赵悦. 数字图书馆元数据应用研究[D]. 武汉: 武汉大学, 2005.
- [7] 吴显义. 我国元数据研究现状分析[J]. 情报科学, 2004, 124(1): 55-58, 62.
- [8] 赵庆峰, 鞠英杰. 国内元数据研究综述[J]. 现代情报, 2003 (11): 42-45.
- [9] 李晓玲, 黎焰. 服装图片元数据标准初探[J]. 图书馆杂志, 2008, 127(6): 63-65, 34.
- [10] 李晓玲. 刺绣图片元数据著录规范研究[J]. 情报探索, 2012(2): 111-113.
- [11] 黎焰, 李晓玲. 纺织品图片元数据研究[J]. 纺织科技

- 进展, 2013(1): 23-26, 43.
- [12] 袁莉, 张晓林. 数字图像的元数据格式[J]. 大学图书馆学报, 200119, (2): 27-30, 41.
- [13] 杨西宁, 杜义涛, 赵书城. 敦煌石窟艺术元数据标准的设计及实现[J]. 上海交通大学学报, 2003, 37(S1): 221-225.
- [14] 吴晓波. 民族服饰元素在现代服装设计中的应用研究[J]. 天津纺织科技, 2012(4): 53-56.
- [15] LEE M, BAILLIE S, DELL'ORO J. TML: a thesaural markup language [EP/PL]. [2017-10-11]. <http://pdfs.semanticscholar.org/050d/44e723ff10279e15a8d7d15e720cda35e9bf.pdf>.
- [16] 任磊, 谭跃生. 基于 RDF 元数据的网格资源统一描述方法[J]. 内蒙古科技大学学报, 2009, 28(2): 148-151.
- [17] 曾新红. 传统受控词表的 SKOS 描述规范及应用指南[J]. 数字图书馆论坛, 2014(4): 9-19.
- [18] FISCHER D H. Converting a thesaurus to OWL: notes on the paper "the national cancer institute's thesaurus and ontology" introduction: from thesaurus towards ontology [EP/OL]. [2017-11-01]. https://www.researchgate.net/publication/267402951_Converting_a_Thesaurus_to_OWL_Notes_on_the_Paper_The_National_Cancer_Institute%27s_Thesaurus_and_Ontology_1_Introduction_From_Thesaurus_towards_Ontology.
- [19] 岳鹏飞. 基于 RDF/XML 的教育资源元数据解析和查询方案研究[D]. 呼和浩特: 内蒙古大学, 2010.
- [20] 贾君枝, 卫荣娟. 叙词表形式化描述语言的比较研究[J]. 图书馆杂志, 2010, 29(1): 27-30, 26.
- [21] 王梦菊. DC 元数据的应用及思考[J]. 现代情报, 2005, 128(8): 5-7.
- [22] 张勇, 蔡璐, 李月明. 非物质文化遗产数字资源元数据标准应用的研究和思考[J]. 图书馆, 2016(2): 38-42.
- [23] 张纪州. 曝光对色彩三要素的影响[J]. 中国校外教育: 理论, 2007(8): 158.
- [24] 易军凯, 周育彬, 陈刚. 可扩展的数字博物馆元数据规范研究与实践[J]. 数字图书馆论坛, 2014 (2): 43-53.
- [25] 邹加勉. 中国少数民族图案与配色[M]. 大连: 大连理工大学出版社, 2010: 23.
- [26] 罗彬, 张健. RDF 数据可视化系统设计与实现[J]. 图书情报工作, 2013, 57(S1): 281-285.
- [27] ELBASSUONI S, RAMANATH M, SCHENKEL R, et al. Searching RDF graphs with SPARQL and keywords [EP/OL]. [2017-09-21]. https://www.researchgate.net/publication/220282494_Searching_RDF_Graphs_with_SPARQL_and_Keywords.
- [28] 肖竹军. 基于 SPARQL 的 RDF 数据节点间关系路径检索[J]. 微型机与应用, 2011, 30(9): 50-53.
- [29] 曹洁. 基于 XML/RDF 的数字图书馆互操作模型研究[J]. 图书馆学研究, 2016(11): 39-44.
- [30] 李满玲, 吴坚. 基于 RDF/XML 公共元数据的数据交换模型[J]. 湖南科技学院学报, 2014, 35(5): 115-118.
- [31] 谢桂芳. SPARQL——一种新型的 RDF 查询语言[J]. 湘南学院学报, 2009, 30(2): 80-84.