

地方诗词资源的关联组织研究

——以苏轼镇江诗词为例

李永丹^{1,2} 周树斌¹ 黄柏芝¹

(1. 江苏大学科技信息研究所, 镇江 212013; 2. 江苏大学图书馆, 镇江 212013)

摘要: 本文通过对地方诗词资源关联聚合模式进行设计, 包括实体属性确定和本体构建, 以苏轼镇江诗词为例进行实证研究, 对相关人物和诗词数据进行采集, 引入关联数据进行关联组织, 通过Drupal结合XAMPP集成包实现苏轼镇江诗词资源平台的搭建。本研究实现对地方诗词资源本体模型的构建, 规范并揭示了地方诗词研究领域内知识的概念、属性和相互关系。设计苏轼镇江诗词资源实例化应用平台, 对于地方诗词文化的开发和知识服务具有一定的推动作用。本研究所采用的方法和视角可以为地方诗词文化推广提供参考, 为地域文化和城市史研究提供思路, 为关联数据等技术拓展到诗词研究领域提供借鉴。

关键词: 地方诗词; 诗词资源; 知识组织; 关联数据

中图分类号: G122; G250.7

DOI: 10.3772/j.issn.1673-2286.2020.02.004

诗词作为我国文化资源的重要组成部分, 包含丰富的地理信息和人文情怀, 而其地域属性将历史文化名人与区域历史文化紧密相连, 因此对地方诗词资源的开发与研究, 对于促进地方文化建设具有一定的作用。在当前文化与科技融合趋势下, 借助大数据、数据可视化、GIS空间可视化^[1]、用户贡献内容(UGC)、语义网等现代信息科技为地方诗词在数字时代的重构提供新的可能, 使地方诗词资源的知识组织向语义化方向发展。

目前, 对于诗词的研究和整理多是停留在传统理论研究或简单的数字化层面, 诗词的数字化实践已取得了一定成果, 如古诗文网等诗词网站对大量中国古典诗词进行了分门别类的收录; Chinese-poetry数据库较为全面地收录了中文诗歌古典文集; 此外, 北京师范大学的唐诗别苑、武汉大学的KnowPoetry^[2], 以及搜韵平台的唐宋文学编年地图^[3]等引入了知识图谱及可视化的技术手段, 对传统诗词资源有了更直观的展示。近年来, 国内学者对关联数据^[4]的创建、发布、维护、检索、可视化^[5]及标准规范^[6]等均进行了理论与技术研究。随

着以关联数据为代表的语义网技术在国内外众多领域的运用和逐渐成熟, 使其在资源聚合与知识服务上呈现出明显的技术优势, 并在历史文化研究领域表现出形成可观的研究成果。如刘美杏等^[7]构建古道线性文化遗产信息资源的关联数据模型, 并以潇贺古道为例进行实证分析; 周溢青等^[8]运用关联数据技术对“江海文化”资源进行语义聚合研究; 张乐等^[9]使用Drupal实现对民国建筑知识库关联数据的组织与发布; 侯西龙等^[10]通过关联数据技术对非物质文化遗产进行知识管理研究; 祝帆帆等^[11]以中国十大绘画为例进行了馆藏文物资源关联数据的创建与发布。此外, 芬兰的MUSEFINLAND项目^[12]、欧盟的数字图书馆项目Europeana^[13], 以及国内华东师范大学图书馆的数字方志项目^[14]、上海图书馆的开放数据平台项目^[15]、北京大学信息管理系的宋代学术语义网络平台项目^[16]等也均验证了将关联数据应用于文化资源组织研究上的实践意义。已有研究表明, 将关联数据技术应用于历史文化领域实现资源语义融合具有优势和可能性, 因此, 引入关联数据来进行地方诗词资源的关联组织研究具

有一定的创新性和科学性。

本文通过对地方诗词关联组织模式进行设计,构建地方诗词资源本体模型;基于关联数据技术,以苏轼镇江诗词为例展开实证研究,完成关联数据模型的构建以及实例化展示,并对研究结果进行分析和总结。本研究通过对地方诗词资源进行直观、形象地组织和可视化呈现,以期对地方诗词的相关研究提供新的参考,为地方诗词资源的分析与组织提供新的思路,为城市旅游、文化、经济开发与发展提供新的方案。

1 地方诗词资源关联组织模式设计

1.1 地方诗词资源实体和属性

本文提出并强调诗词资源这一概念,所谓诗词资源不等同于诗词,其指代包含诗词相关要素的集合,而本文所研究的地方诗词资源是诗词资源的延伸,其在诗词资源的基础上进一步强调并凸显其地域属性特

征。本文所研究的地方诗词资源主要包括3个组成部分,即诗词内容、人物和地点。诗词内容是地方诗词资源的基础部分和核心,是对诗词资源的核心表征;人物是诗词的创作者、诗词内容的相关者,是诗词创作这一事件的发起者,是资源中的主体;地点则体现出诗词的地域属性及地方特色,之所以强调地点内容而不是地方或地域,是因为地点内容突出了对于地方或地域内容的把控,若把地方或地域看作一个区域的整体,地点则是其区域内的标志性组成单元,是诗词地域属性的精确表征。通过上述内容,可以对地方诗词资源主题对象进行一定的诠释,更深度地关注地方文化,为地方文化和旅游提供参考与服务。提取诗词资源中内容、人物、地点信息,设定内容、人物、地点为基本类目。3个基本类目包含各自的构成要素和属性信息,如诗词内容包含题目、正文、作者、创作时间(历史纪年/公元年)、创作背景等内容;人物包括姓名、字、号、介绍等内容信息;地点包含地名、坐标、地理信息展示、简介等信息。具体的地方诗词资源实体属性如图1所示。

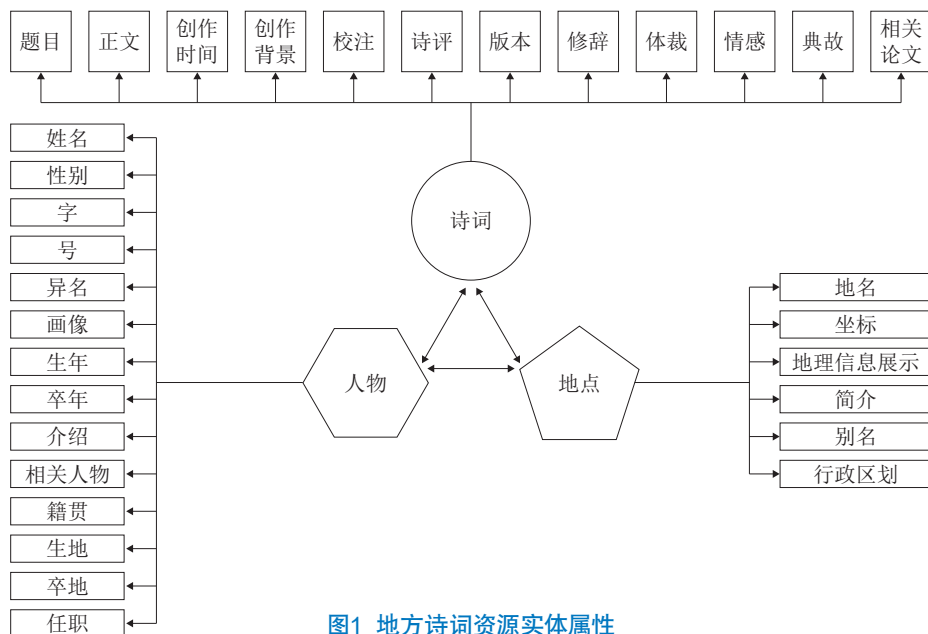


图1 地方诗词资源实体属性

1.2 地方诗词资源本体模型设计

本体是对概念体系的明确、形式化、可共享的规范说明^[17]。本体的设定关系到领域内知识的开放共享,因此本体的构建对于地方诗词资源的组织十分重要。根据本文所研究地方诗词资源的特点,现有的本体无法对其进行良好的描述和诠释,所以需要地方诗词

资源本体模型进行设计从而实现对多种类型数据的集成。现阶段有两种类型的主流本体构建方式,即多本体型和混合型^[18]。根据上文所述,本文将地方诗词资源量化分为人物、诗词内容、地点3个组成部分,多本体型灵活便捷的特性更适用于本文的研究,因此本研究选择多本体型作为本体设计的基础。

构建本体的一个重要参考原则是要尽可能对现

有本体词表进行复用,因此针对地方诗词资源本体词表的制定和设计,需要考虑对国内外已有本体和相关词表进行复用,并根据实际情况和自身需要进行扩展并额外添加相应的属性元素。根据本文的研究特点,考虑到本体的易用性和通用性要求,主要选择对DC、FOAF、GeoNames、上海图书馆人名规范库本体(SHLNames)等本体进行复用^[19],其中,DC是描述数字文献的通用元数据标准框架,包含15个核心元素集;FOAF是一种遵循W3C体系标准的资源描述框架词表,主要用于对人物相关数据内容属性的描述;GeoNames是最广泛被使用的地理关联数据集之一,其中相关术语广泛适用于地理区域本体的构建,主要用于对地点

相关属性的描述;SHLNames对历史人物相关属性进行了较为详尽的描述和规范,其在FOAF的基础上对本研究人物相关内容属性的描述进行了很大程度上的补充及完善。然而,根据地方诗词资源的结构特点,单纯地对现有本体词表进行复用并不能满足本研究的需求,还需要根据研究的实际需要额外构建词表来进行补充,本文根据自身需要构建了RPR(Regional Poetry Resources)词表来进行地方诗词资源实体属性的描述。完成对本体属性的确定后,通过本体间相互映射,可实现不同本体间的关联。最终确定的本体类及其属性如表1所示。

表1 地方诗词资源本体类及其属性

类别(class)	属性标识符(property)	标签(note)	域(domain)	取值范围(range)	出 处
人物 (rpr:Person)	foaf:name	姓名	rpr:Person	rdfs:Literal	FOAF
	foaf:gender	性别	rpr:Person	bf:Category	FOAF
	shl:briefBiography	介绍	rpr:Person	rdfs:Literal	SHLNames
	shl:speciality	社会身份	rpr:Person	rdfs:Literal	SHLNames
	shl:deathDay	卒年	rpr:Person	rdfs:Literal	SHLNames
	foaf:birthday	生年	rpr:Person	rdfs:Literal	FOAF
	foaf:img	画像	rpr:Person	foaf:Image	FOAF
	shl:courtesyName	字	rpr:Person	rdfs:Literal	SHLNames
	shl:pseudonym	号	rpr:Person	rdfs:Literal	SHLNames
	foaf:familyName	姓	rpr:Person	shl:FamilyName	FOAF
	shl:temporal	朝代	rpr:Person	shl:Temporal	SHLNames
	shl:name	异名	rpr:Person	rdfs:Literal	SHLNames
	shl:nativePlace	籍贯	rpr:Person	shl:Place	RPR
	shl:officialEvent	任职事件	rpr:Person	shl:OfficialEvent	SHLNames
	foaf:knows	相关人物	rpr:Person	rpr:Person	FOAF
	shl:creatorOf	创作诗词	rpr:Person	rpr:Poetry	RPR
	shl:birthPlace	出生地	rpr:Person	shl:Place	SHLNames
	shl:deathPlace	死亡地	rpr:Person	shl:Place	SHLNames
	rpr:visitPlace	到访地点	rpr:Person	rpr:Place	RPR
诗词内容 (rpr:Poetry)	dc:title	题目	rpr:Poetry	rdfs:Literal	DC
	rpr:background	创作背景	rpr:Poetry	rdfs:Literal	RPR
	rpr:note	校注	rpr:Poetry	rdfs:Literal	RPR
	rpr:content	正文	rpr:Poetry	rdfs:Literal	RPR
	rpr:emotion	情感	rpr:Poetry	rdfs:Literal	RPR
	rpr:rhetoric	修辞	rpr:Poetry	rdfs:Literal	RPR
	dc:date	创作时间	rpr:Poetry	rdfs:Literal	DC
	rpr:version	版本	rpr:Poetry	rdfs:Literal	RPR
	dc:type	体裁	rpr:Poetry	rdfs:Literal	DC
	rpr:allusion	典故	rpr:Poetry	rdfs:Literal	RPR

续表

类别 (class)	属性标识符 (property)	标签 (note)	域 (domain)	取值范围 (range)	出 处
诗词内容 (<i>rpr:Poetry</i>)	<i>rpr:review</i>	诗评	<i>rpr:Poetry</i>	<i>rdfs:Literal</i>	RPR
	<i>dc:creator</i>	作者	<i>rpr:Poetry</i>	<i>rpr:Person</i>	DC
	<i>rpr:releventPerson</i>	相关人物	<i>rpr:Poetry</i>	<i>rpr:Person</i>	RPR
	<i>rpr:releventPoetry</i>	相关诗词	<i>rpr:Poetry</i>	<i>rpr:Poetry</i>	RPR
	<i>rpr:createPlace</i>	创作地点	<i>rpr:Poetry</i>	<i>rpr:Place</i>	RPR
	<i>rpr:describePlace</i>	描述地点	<i>rpr:Poetry</i>	<i>rpr:Place</i>	RPR
地点 (<i>rpr:Place</i>)	<i>gn:name</i>	地名	<i>rpr:Place</i>	<i>rdfs:Literal</i>	GeoNames
	<i>gn:locationmap</i>	地理展示	<i>rpr:Place</i>	<i>gn:locationMap</i>	GeoNames
	<i>rpr:othername</i>	别名	<i>rpr:Place</i>	<i>rpr:Place</i>	RPR
	<i>dc:description</i>	介绍	<i>rpr:Place</i>	<i>rdfs:Literal</i>	DC
	<i>rpr:releventPerson</i>	相关人物	<i>rpr:Place</i>	<i>rpr:Person</i>	RPR
	<i>rpr:releventPoetry</i>	相关诗词	<i>rpr:Place</i>	<i>rpr:Poetry</i>	RPR

地方诗词资源的3类基本实体均根据自身特点复用并补充了相应的属性。人物 (*rpr:Person*) 类, 在对诗词资源进行描述和管理中, 人物是本体模型中的主体, 在进行词表设定时, 继承和复用了FOAF、SHLNames使其可以对历史人物属性进行充分描述, 将人物相关信息进行描述标识, 与其相关的诗词内容及地点实体相关联; 诗词内容 (*rpr:Poetry*) 类, 诗词内容是地方诗词资源的基础部分, 对DC进行了复用, 其他属性均使用自建词表对其进行规范, 这是本文对于地方诗词资源研究的创新点和特色, 通过上述属性可以对诗词内容

较为完整地进行规范性描述; 地点 (*rpr:Place*) 类, 对GeoNames、DC进行了继承复用, 同时通过自建的方式对地点这一实体进行补充性描述, 通过对诗词资源地理信息的描述凸显了诗词资源的地域属性。

在本体词表设定的基础上, 对地方诗词资源关联数据模型进行设计, 充分考虑可扩展性、可复用性以及可共享性等因素, 对涉及的概念、属性、关系进行定义说明, 最终构建人物-诗词内容-地点本体模型如图2所示, 基本可以描述地方诗词资源的共同属性特征, 具有一定的普适性特点。

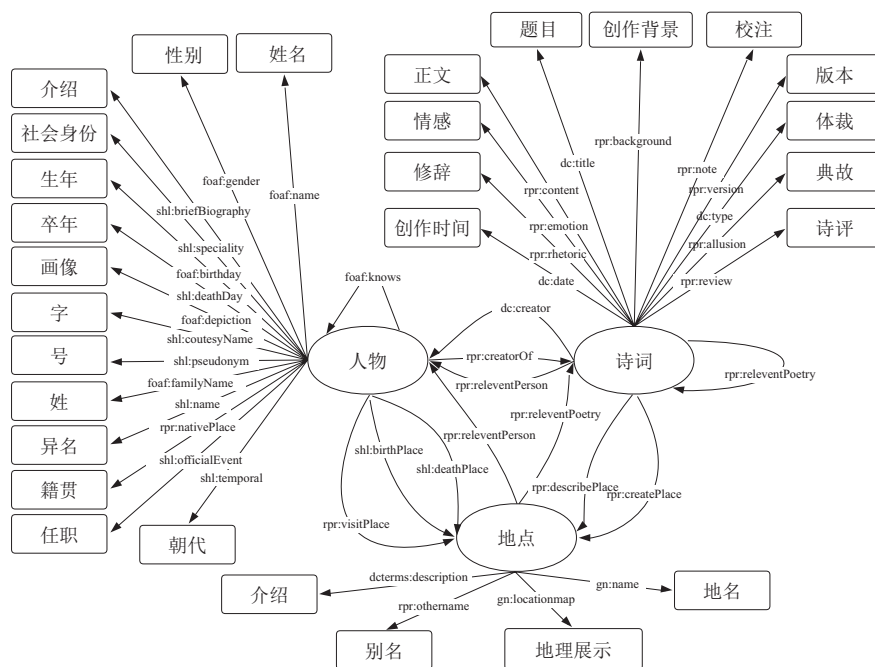


图2 人物-诗词内容-地点本体模型

本文以Berners-Lee制定的关联数据内容描述的四项基本原则^[20]为理论基础,作为构建诗词资源关联数据模型的依据,结合一定的技术手段实现关联数据模型的构建和实例化展示。下文以苏轼镇江诗词为例,对相关资源进行了关联组织的实证研究以及实例化的应用展示。

2 实证研究

2.1 苏轼镇江诗词资源及其数据采集

镇江作为一座国家级历史文化名城,自古便是知名的游览胜地,有很多历史典故和传说故事,具有丰富的地方文化资源。自大运河开凿后,镇江便成为江河交汇处江南运河段的起点^[21],是当之无愧的“黄金十字路口”,有着沟通南北的重要地位,大批文人墨客的驻足为镇江累积了丰富的诗词资源。

北宋著名诗人苏轼虽然并未在镇江担任过官职,但是他的一生与镇江有着深厚的渊源。作为“唐宋八大家”之一,苏轼一生途经镇江十余次,留下了百余篇诗词,在一定程度上反映了镇江的历史、经济、文化等方面的发展,折射出士人心态、人生理想、思想状态的变迁。这些诗词有着鲜明的地域与文化特色,构成了镇江地方文化资源的重要部分,同时也提升了镇江地域文化的品味和格调。进一步延伸和开发苏轼诗词资源,可深化镇江地方历史文化研究,推动镇江旅游资源开发。

本研究所需数据主要通过人工采集和网络收集两种途径来共同完成和保证,优先通过网络途径获取并通过纸质书籍对数据内容进行校对以保证数据内容的可靠性。诗词的题目、正文数据主要以中华书局1984年版《苏轼诗集》、中华书局1986年版《苏轼文集》、北京大学出版社1998年版《全宋诗》等权威书籍作为版本内容依据;诗词的其他内容数据则通过相关网站、书籍来采集确定,如百度百科、北京大学数据分析研究中心研发的全宋诗分析系统、古诗文网、《中国历代人名大辞典》、《中国古今地名大辞典》等。人物相关数据来自《中国历史人物生卒年表》、国图规范档、上图古籍数据库、《全宋诗》中诗人简介等。采集的数据均有版本标注以保证数据来源的规范可靠,对各部分数据逐条采集后,将采集到的数据根据诗词资源的类和属性特点结构化保存并录为CSV格式文件存储,以供后续导入平台和发布。本文以喻世华^[22]针对苏轼与润州(镇江)

相关诗文篇目的考证研究为依据,收集了苏轼与润州的相关诗词114首,其中诗97首、词17首。

2.2 基于Drupal平台的技术实现

地方诗词资源关联数据的发布与应用需要借助软件平台来实现,已有研究表明,D2RQ^[23]、Virtuoso^[24]、Drupal^[25]等均适合作为关联数据发布的平台,其中Drupal相较于其他软件平台更适用于轻量级数据的实现与发布,且更适合非专业人员学习掌握。结合本案例采集到的诗词资源的数据特点,本文选取Drupal 7作为数据发布的平台,其可以在Windows、Linux等多种操作系统下运行,本文所使用的操作系统为Windows 10,平台主要依靠XAMPP建站集成软件包与Drupal 7搭建完成^[26],XAMPP为Drupal 7提供了Apache+MySQL的服务器环境。配置完成后的Drupal 7主要依托其模块化的设计思路来完成对诗词资源的组织管理,主要用到的模块包括Drupal 7Core(RDF)、RDFx、CCK、FIELDS、evoc、RDFUI、ARC2、SPARQL Endpoint等。

2.2.1 资源标识

关联数据在发布前,首先需使用统一资源标识符(URI)对资源进行标识,使得各地方诗词资源的各实体具有唯一性。根据关联数据内容描述的四项基本原则,通过URI的“http://”形式为资源数据提供唯一标识。本文的系统平台是在本地搭建的,其URI基地址为“http://localhost:8000/drupal7”,各实体资源均有唯一标识,均以“http://localhost:8000/drupal7/node/”开头,如本文中人物实体“苏轼”的URI为“http://localhost:8000/drupal7/node/36”,诗词内容实体“《大风留金山两日》”的URI为“http://localhost:8000/drupal7/node/38”,地点实体“金山”的URI为“http://localhost:8000/drupal7/node/39”。

2.2.2 内容类型和属性构建

在对资源进行标识后,需要对其相关的内容类型和属性进行构建。Drupal的基本存储单元是Node(节点),内容全被存储为节点形式,在内容节点下包含多种类型字段,如文本、字符、图像、Node Reference等。通过使用Drupal中的CCK模块可创建各种节点类型和

字段。在内容类型的设置上,根据本研究本体模型的设计,建立人物、诗词内容、地点3个内容类型。属性的构建是在内容类型创建的基础上来展开,以诗词内容为例,根据诗词内容的属性添加了多种字段,如“field_creator”代表作者,“field_background”代表创作背景,“field_date”代表创作时间,“field_createplace”表示创作地点等。内部数据的关联主要通过Node Reference来实现,通过FIELDS模块,设置人物、诗词内容、地点等字段类型为Node Reference,从而使其与其相关的数据关联起来。以诗词为例,其部分内容节点设置如表2所示。

表2 诗词内容节点设置(部分)

标 签	机读名称	字段类型
题目	title	节点模块元素
作者	field_creator	Node Reference
正文	body	长文本和摘要
创作背景	field_background	文本
创作时间	field_date	文本
创作地点	field_createplace	Node Reference
描述地点	field_describeplace	Node Reference
版本	field_version	文本
体裁	field_type	文本
校注	field_note	长文本
典故	field_allusion	长文本
诗评	field_review	长文本
修辞	field_rhetoric	文本

2.2.3 RDF映射

通过Drupal对地方诗词资源的内容类型及其属性字段进行设置后,需要对关联资源内容类型及属性与相关本体之间进行RDF映射。通过evoc模块实现对外部本体库的导入,并利用RDFUI模块提供的界面设置使平台内部数据与外部本体库建立映射关系。以诗词内容为例,其RDF映射如表3所示。主要对DC本体库title、date、type等属性进行了复用,其他属性来源于本文自建本体RPR。相关属性使用RDF模块中的MAPPINGS功能与相对应的本体库中的属性建立关联,如DC、FOAF等。

根据已搭建的实体关系模型来添加实体间的RDF链接。将数据转化为RDF类型后,利用ARC2模块完成对转化后数据的存储。人物-诗词内容-地点关联实

表3 诗词内容RDF映射(部分)

标 签	RDF Type
类型	rpr:poetry
题目	dc:title property
正文	rpr:content property
创作背景	rpr:background property
校注	rpr:note property
情感倾向	rpr:emotion property
修辞	rpr:rhetoric property
创作时间	dc:date property
版本	rpr:version property
体裁	dc:type property

例如图3所示,可以看出,3类实体间已实现了一定程度上的互联互通。图3展示了人物实体“苏轼”、诗词内容实体“《大风留金山两日》”、地点实体“金山”相互间的关联。通过对诗词内容与人物和地点的整合,可使用户能够凭借诗词内容的资源信息全面了解到相关的人物与地点信息,反之亦可从人物或地点关联到其诗词内容与地点或人物信息。

2.2.4 可视化展示

以图3所示关联为样本,本部分以苏轼在金山作《大风留金山两日》为例,进行相关内容发布效果的可视化展示,如图4、图5分别是对人物、地点属性信息的可视化效果展示。图4展示了苏轼关于姓名、个人介绍、性别、社会身份、生年、卒年、画像、字、号、姓氏、朝代、籍贯、异名、主要官职、相关诗词、到访地点等方面的属性信息。图5展示了金山关于地名、地理展示、简介、相关人物、相关诗歌以及别名等方面的属性信息。此外,对于《大风留金山两日》关于题目、正文、创作背景、情感倾向、创作时间、版本、体裁、诗评、作者、创作地点、描述地点等方面的属性信息也可进行诗词内容的实例化展示,实现了各实体的属性信息及其相互间的关联。通过SPARQL Endpoint模块所提供的SPAQL访问端口可实现对所发布关联数据的查询获取。

需要说明的是,图5中的地理展示是依托于百度地图API来实现的^[27],通过调用百度地图API可实现对基础地理信息数据的获取和应用,以满足个性化的研究需求。本文通过调用百度地图API完成了对金山的地理位置在地图上的标注和展示,基于数据来源的可靠性

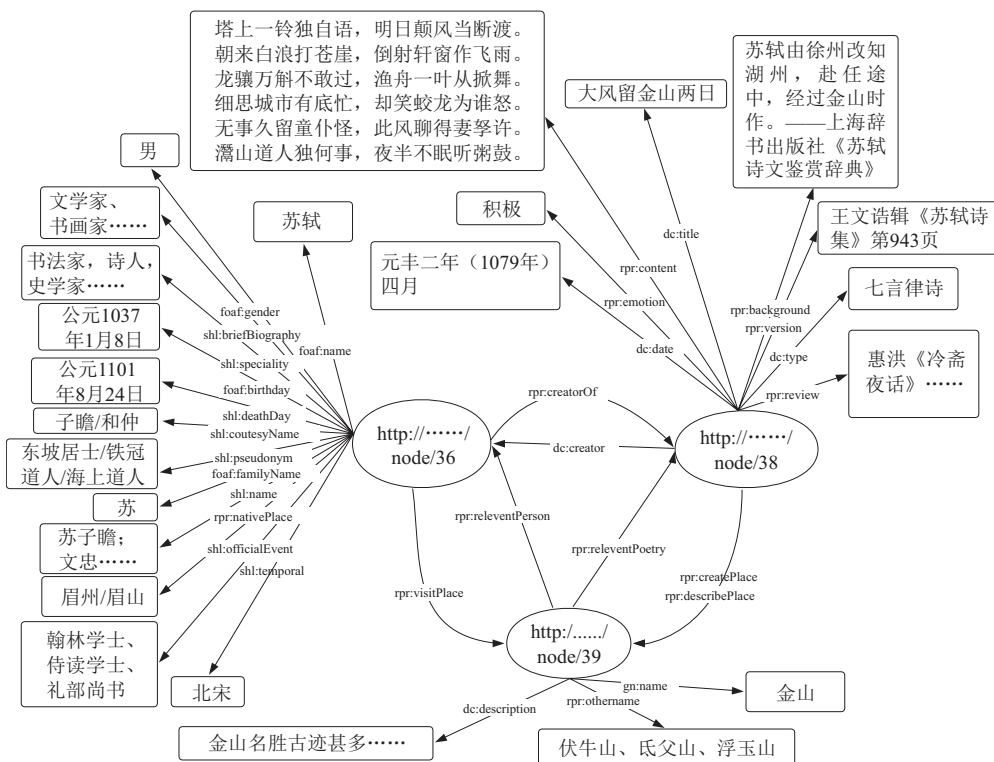


图3 人物-诗词内容-地点关联实例

与资源搜集的广泛性，本研究使得最终用户面向的资源内容相比常规检索获取的内容更为丰富和真实。通过关联数据的方式使得面向用户的数据呈现语义化、结构化特点，而非传统方式的碎片化、非结构化，这为用户充分利用地方诗词资源提供了可能，也为后续的知识发现服务提供了基础和条件。本研究的技术方案主要依托于Drupal来实现，Drupal凭借其可扩展性、可伸缩性、简便性等优势，使得非计算机专业人员也可以进行关联数据的发布，方便用户发布和管理站点上的内容，其引入后控词表的管理机制也是其优势所在。作为开源软件工具，模块化的设计思路在增加了Drupal便捷性与可扩展性的同时，也有不足之处，其在功能实现上主要依托于第三方模块来实现，而第三方模块更新较慢难以保障，且目前在处理大数据方面的模块有所欠缺。今后进一步考虑的研究方向是引入第三方工具来对其功能进行拓展，克服其在模块化方面的不足，并通过语义推理技术实现对诗词资源的知识发现，从而更好地服务用户并提高资源的利用率。

3 结论与展望

随着文化与科技融合趋势的不断推进，全新的科



图4 人物“苏轼”实例展示

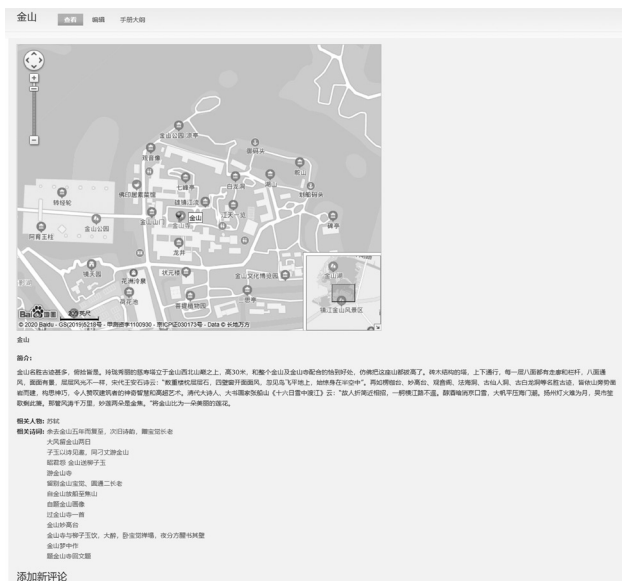


图5 地点“金山”实例展示

学技术越来越多地被用于文化资源研究领域,传统文化资源组织方式也受到更为广泛的关注,然而将诗词上升为一种资源进行组织研究并不多见,本文将诗词视为一种地方文化资源,对其知识组织方案进行探讨和设计尝试。为对地方诗词资源进行有效的知识组织,结合关联数据的理论方法对地方诗词资源本体模型进行了设计和构建,通过对其实体属性的定义,建立实体间的语义关联,使得人物、诗词内容、地点实体相互关联。用本体构建人物、诗词内容、地点间的关系,使研究者可以定向地获得特定诗词在特定地域背景下的关系脉络。本文以苏轼镇江诗词为例进行了实证研究,在XAMPP的系统环境下进行Drupal内容管理系统的搭建,在技术层面实现了对苏轼镇江诗词资源的知识发布和可视化展示。综上,本文的研究从理论到技术方法都具有高度的可复用性,一方面,地方诗词资源领域本体的构建为领域内知识在不同环境下的共享和重用奠定了基础,为不同系统间的高层互操作提供了工具方法;另一方面,Drupal对关联数据的支持及其易操作和可扩展性的特点也为地方诗词资源的关联组织提供了有力的技术保障,有利于领域内的研究人员在此基础上进一步研究。因此,本研究拓宽了诗词资源组织研究的视野与领域,对地方诗词资源由信息资源向知识化资源的转换提供了方法上的创新。实例的应用既可提供规范的数据用作学术研究用途,面向学者和研究人员,满足其研究地方诗词文化的需要,同时也可用作地方文化普及用途,保证普适性的特点;面向

社会大众,满足社会大众普及文化的需要,使地方诗词资源成为大众了解地方历史的一面镜子,其可应用于地方图书馆、文化旅游网站等多种场景,为地方文化研究和地方文化品牌的宣传和推广提供了可以使用的平台体系架构。

地方诗词资源中潜藏着海量的知识,本文仅梳理了人物、诗词内容、地点等基本知识要素,知识组织的粒度还有待进一步细化。尽管本研究在一定程度上实现了对地方诗词资源的知识组织体系构建,但是尚未实现知识发现服务,且收集的地方诗词资源数据还不够全面,部分属性字段内容还未采集和填充,因此后续还会进行更为深入的研究。今后,一方面通过对地方诗词资源数据进行更全面的收集,增加更多的资源内容和对象,引入更全面的技术手段和方法对数据资源进行加工处理,从而提高平台资源内容的全面性以及数据整合的自动化;另一方面引入语义推理技术使服务拓展到知识发现的层面,从已有资源数据中推导和挖掘出其中隐含的知识,借助数据关联和共享实现对相关数据的复用和共享。通过这些改进可以对地方诗词资源进行更好的保护和开发,也为地方文化资源的开发与利用提供一定的方案和思路。

参考文献

- [1] 张建立,李仁杰,傅学庆,等. 古诗词文本的空间信息解析与可视化分析[J]. 地球信息科学学报, 2014, 16(6): 890-897.
- [2] 周莉娜,洪亮,高子阳. 唐诗知识图谱的构建及其智能知识服务设计[J]. 图书情报工作, 2019, 63(2): 24-33.
- [3] 王兆鹏,郑永晓,刘京臣. 借器之势,出道之新——“数字人文”浪潮下的古典文学研究三人谈[J]. 文艺研究, 2019(9): 79-88.
- [4] 刘炜. 关联数据: 概念、技术及应用展望[J]. 大学图书馆学报, 2011, 29(2): 5-12.
- [5] 曲佳彬,欧石燕. 关联数据可视化研究进展分析[J]. 图书与情报, 2018(4): 51-61.
- [6] 陈涛,张永娟,刘炜,等. 关联数据发布的若干规范及建议[J]. 中国图书馆学报, 2019, 45(1): 34-46.
- [7] 刘美杏,徐芳. 古道线性文化遗产信息资源关联数据模型构建及其实证研究[J]. 图书馆学研究, 2019(14): 40-50.
- [8] 周溢青,徐晨飞. 基于关联数据的“江海文化”资源多维语义聚合研究[J]. 情报科学, 2019, 37(9): 72-78, 89.
- [9] 张乐,常娥. 基于Drupal的民国建筑知识库关联数据的组织与发布研究[J]. 图书馆学研究, 2018(19): 64-70.

- [10] 侯西龙, 谈国新, 庄文杰, 等. 基于关联数据的非物质文化遗产知识管理研究 [J]. 中国图书馆学报, 2019, 45 (2): 88-108.
- [11] 祝帆帆, 高劲松, 梁艳琪. 馆藏文物资源关联数据的创建与发布——以中国十大绘画为例 [J]. 图书馆理论与实践, 2018 (4): 96-101.
- [12] 梁艳琪. 基于关联数据的文物数字资源语义融合与服务研究 [D]. 武汉: 华中师范大学, 2017.
- [13] 王萍, 黄新平. 基于关联开放数据的数字文化资源语义融合方法研究——欧洲数字图书馆案例分析 [J]. 图书情报工作, 2016, 60 (12): 29-37.
- [14] 张毅, 李欣. 面向数字人文的特藏资源揭示研究——以方志数据库建设为例 [J]. 图书馆, 2019 (6): 100-105.
- [15] 夏翠娟, 刘炜, 陈涛, 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, 42 (3): 27-38.
- [16] 杨海慈, 王军. 宋代学术师承知识图谱的构建与可视化 [J]. 数据分析与知识发现, 2019, 3 (6): 109-116.
- [17] 金家琴, 夏翠娟. 数字人文数据基础设施建设中机构本体的构建: 研究和应用 [J/OL]. 图书馆论坛: 1-12 [2020-02-18]. <http://kns.cnki.net/kcms/detail/44.1306.G2.20191021.0748.002.html>.
- [18] 刘佳音. 国内高校数字资源利用效能综合评价研究 [D]. 武汉: 武汉大学, 2013.
- [19] 姚天泓, 陈艳梅, 刘革, 等. 基于CIDOC-CRM的数字人文史料资源语义化知识组织研究——以张学良史料资源为例 [J]. 图书馆学报, 2019, 41 (7): 35-43.
- [20] BERNERS-LEE T. Design issues: Linked data [EB/OL]. [2020-01-11]. <https://www.w3.org/DesignIssues/LinkedData.html>.
- [21] 严其林, 程建. 京口文化 [M]. 南京: 南京大学出版社, 2001.
- [22] 喻世华. 苏轼与润州有关的诗文篇目考 [J]. 江苏科技大学学报 (社会科学版), 2012, 12 (2): 20-27.
- [23] 陈婵. 基于关联数据的数字图书馆Web服务语义关联与发布研究 [D]. 武汉: 华中师范大学, 2018.
- [24] 石泽顺, 肖明. 基于引文关联数据的引文路径识别和引文网络可视化研究 [J]. 情报科学, 2019, 37 (2): 89-94.
- [25] 夏翠娟, 刘炜, 赵亮, 等. 关联数据发布技术及其实现——以Drupal为例 [J]. 中国图书馆学报, 2012, 38 (1): 49-57.
- [26] 伍革新. 基于关联数据的数字图书馆资源聚合与服务研究 [D]. 武汉: 华中师范大学, 2013.
- [27] 胡昌平, 石宇, 查梦娟. 少数民族特色文化资源的聚合与可视化展现 [J]. 图书馆论坛, 2017, 37 (4): 32-39.

作者简介

李永丹, 女, 1979年生, 博士, 副研究馆员, 通信作者, 研究方向: 数字人文、宋史研究, E-mail: 15862973682@163.com。

周树斌, 男, 1996年生, 硕士研究生, 研究方向: 数字人文、历史文献学研究。

黄柏芝, 女, 1995年生, 硕士研究生, 研究方向: 数字人文、历史文献学研究。

Research on the Related Organization of Regional Poetry Resources: Taking Su Shi Zhenjiang Poetry as An Example

LI YongHui^{1,2} ZHOU ShuBin¹ HUANG BoZhi¹

(1. Institute of Science and Technology Information, Jiangsu University, Zhenjiang 212013, China; 2. Library of Jiangsu University, Zhenjiang 212013, China)

Abstract: This paper designs the regional poetry resource association aggregation mode, including entity attribute determination and ontology construction. Take Su Shi Zhenjiang poetry as an example to conduct empirical research, collect relevant person and poetry data, and introduce linked data for related organization. Through Drupal and XAMPP integration package to complete the construction of Su Shi Zhenjiang poetry resource platform. This research has realized the construction of the regional poetry resource ontology model, standardized and revealed the concepts, attributes and interrelationships of knowledge in the field of regional poetry research. By designing an example application platform for Su Shi Zhenjiang poetry resources, it will promote the development of regional poetry culture and knowledge services. Some methods and perspectives adopted by this research can provide references for the promotion of regional poetry culture, provide ideas for the study of regional culture and urban history, and provide reference for the expansion of linked data and other technologies into the field of poetry research.

Keywords: Regional Poetry; Poetry Resources; Knowledge Organization; Linked Data

(收稿日期: 2020-01-18)