

碑刻资源语义化组织研究*

李永卉^{1,2} 刘沁芃¹ 周树斌¹ 屠纪军³

(1. 江苏大学科技信息研究所, 镇江 212013; 2. 江苏大学法学院, 镇江 212013; 3. 镇江焦山碑刻博物馆, 镇江 212013)

摘要: 碑刻作为一种不可再生的宝贵历史文化遗产, 承载着文明和文化的记忆, 对其进行语义化组织研究具有重要意义。本文结合关联数据与国际图像互操作框架 (IIIF) 进行碑刻资源语义化组织的探索。首先, 对碑刻资源语义化组织框架进行设计, 为碑刻资源从采集描述到组织发布提供一套完整可行的方案。其次, 以镇江焦山碑林为例, 基于Drupal内容管理平台进行关联数据应用与IIIF服务集成, 实现碑刻资源概念间关系的关联发布及可视化展示, 验证碑刻资源语义化组织框架的可行性, 为碑刻资源的深度开发和服务提供借鉴。

关键词: 碑刻资源; 关联数据; 国际图像互操作框架

中图分类号: G250 **DOI:** 10.3772/j.issn.1673-2286.2021.12.007

引文格式: 李永卉, 刘沁芃, 周树斌, 等. 碑刻资源语义化组织研究[J]. 数字图书馆论坛, 2021 (12): 41-49.

碑刻作为我国古代一种常见的文化载体, 具有丰富的文化内涵, 承载着中华文明的宝贵信息, 兼具历史、文化和艺术价值, 在书法史、绘画史和绘画理论等领域均具有很高的价值, 是一种特殊的实物资源, 也是中国传统文化重要组成部分^[1]。然而, 由于其载体的特殊性, 随着时间推移与环境变化, 很多碑刻已经风化并不断损坏, 使得碑刻资源的“藏”与“用”面临巨大挑战。在大数据时代, 数字化技术与新媒体发展为历史文化遗产的保存与利用带来了全新的方法和手段, 也为碑刻资源的数字化保护和开发提供了支持。数字技术不仅可以突破时间与空间因素的限制, 还可对实体资源进行深度开发。因此, 数字化保存与传播是新时代碑刻资源保护的重要途径之一, 符合当前文化与科技融合发展的趋势。

本文基于关联数据技术, 将国际图像互操作框架 (International Image Interoperability Framework, IIIF) 与关联数据相结合, 探讨碑刻资源在数字化时代的语义化组织模式和构建方案, 并以焦山碑林为例进行实践研究。通过实例化的研究, 一方面验证了碑刻资源语义化组织框架的合理性与可行性, 为碑刻资源语义化组织提供实践依据; 另一方面, 为碑刻资源的深度开

发提供了现实的参考路径, 也为以焦山碑林为代表的地方历史文化和旅游资源的推广与传播提供助力。

1 相关研究现状

1.1 碑刻资源数字化的实践与理论探索

21世纪以来, 学术界在碑刻资源数字化理论与实践研究方面有了一定探索。实践方面, 北京师范大学民俗典籍文字研究中心的碑刻数字化典藏与应用平台, 对典藏碑拓进行了多重属性的界定^[2]; 国家图书馆使用CNMARC格式对石刻拓片进行编目, 形成“碑帖精华”《中文拓片资源库》, 使石刻拓片书目数据更为标准规范^[3]; 孔庙和国子监博物馆利用三维数字化扫描技术探索了石刻文物的数字化保存方案^[4]。除此之外, 一些学者也进行了理论研究: 戴萍^[5]对大理州碑刻文献在数字时代的整理策略进行了探讨; 潘笑^[6]以宝顶山大佛湾摩崖石刻为例, 通过建立可描述各方面语义信息的模型, 完成了石质文物信息管理系统的设计; 邱孝凤等^[7]在涪溪摩崖石刻保护过程中运用数字化技术进行分析探讨; 徐晨希^[8]、杨祥民等^[9]、唐兰林^[10]均针对南朝石刻

* 本研究得到国家社会科学基金青年项目“标准化推动乡村公共数字文化服务可及性的模式及实施路径研究” (编号: 20CTQ003) 资助。

进行了数字化保护方案的设计与研究;方兴林^[11]在结合数字化处理技术提出徽州碑刻资源特色数据库的构建实施方案。尽管目前已经取得一定的成绩,但随着新技术的发布,碑刻数字资源的组织模式可进一步创新,资源的组织关联、深度融合等问题有望进一步深化。

1.2 关联数据在文化遗产数字化保护中的应用

关联数据作为语义网技术的代表,它采用RDF数据模型,利用URI(统一资源标识符)命名数据实体,来发布和部署实例数据,通过HTTP协议可以揭示并获取这些数据,同时强调了数据的相互关联以及有益于人机理解的语境信息^[12]。近年来,关联数据在资源聚合与知识组织上呈现出明显的技术优势,并在文化遗产领域表现突出并形成可观的研究成果。例如:在非物质文化遗产(以下简称“非遗”)领域,韩洪帅^[13]结合关联数据技术进行传统美术类非遗视频资源的语义化组织方案的研究;李世新^[14]在探索印人和印章实体概念的基础上构建印人和印章本体模型,利用关联数据技术完成相关实体发布,实现了篆刻艺术的可视化知识发现;赵雅倩^[15]以河北大运河非遗为例,进行了非遗资源知识组织及语义发现研究;魏清华等^[16]以甘肃省国家级非遗为例,结合关联数据等数字人文技术进行了非遗知识库构建研究。在物质文化遗产领域,高劲松等^[17]通过关联数据进行了文物信息资源融合模型构建方法的研究探索,并以九大博物馆镇馆瓷器文物信息资源为例对模型进行了验证;王晓光等^[18]构建了敦煌壁画叙词表,就其关联数据发布过程进行研究分析。通过上述研究可以发现,关联数据在文化遗产领域已经有了较为广泛的应用,可以实现资源的深度融合。其中,关联数据在物质文化遗产领域应用的研究实践为将关联数据引入碑刻资源语义化组织研究提供可行性参考,同时面向碑刻资源这一研究对象又是对关联数据应用场景的拓展延伸。因此,基于关联数据技术进行碑刻资源语义化组织研究具有一定的创新性。

1.3 国际图像互操作框架应用现状

IIIF是一组用于定义数字图书馆互操作框架的标准,通过标准的应用程序编程接口(API),提供一种在网络上描述、发布和访问图像的统一方法。IIIF最初是由29个国际非营利图像资源存储机构于2015年6月共同

提出,旨在确保全球图像存储的互操作性和可获取性,对以图像为载体的书籍、地图、卷轴、手稿、乐谱、档案资料等数字在线资源进行统一展示和使用。通过IIIF可以突破图像资源自身的限制,实现与其他馆藏机构图像资源的互操作,极大地提高了在网络数据环境下图像资源的利用率。IIIF被提出后迅速成为GLAM(艺术馆、图书馆、档案馆和博物馆)领域研究的热点,目前国际上主要的文化遗产机构均已加入IIIF,如世界数字图书馆^[19]、哈佛艺术博物馆、大英图书馆、盖蒂(Getty)博物馆、欧洲虚拟博物馆(Europeana)、美国艺术档案馆等,均开始使用IIIF来进行馆藏图像资源的展示与分享交互^[20]。

2019年以后,国内部分图书馆尝试将IIIF应用于馆藏资源展示,如上海图书馆在家谱、古籍、印谱数据库建设项目中使用IIIF框架进行网络展示^[21]。台北“中央研究院”数位人文研究平台^[22]、厦门大学图书馆的特藏数字资源建设^[23]、华东师范大学ECNU-DHRS平台^[24]、复旦大学印谱文献虚拟图书馆^[25]等,均将IIIF引入文献展示。此外,杨佳颖等^[26]选取民国报刊《新闻报》上的越剧广告为研究对象,结合语义模型以及IIIF技术支持平台,揭示图像所涵盖的文本信息;张永娟等^[27]基于IIIF和语义知识图谱,进行了印章资源整合与知识发现的研究;付跃安^[28]探讨了IIIF及其在数字资源集成中的应用;耿曼曼^[29]将IIIF引入图书馆图像资源开发与利用;张轶^[30]对IIIF及其应用前景进行了分析。

尽管IIIF出现时间不长,相关的理论探讨和实践应用还处于探索阶段,不是十分成熟,但是其在文化遗产领域发展潜力巨大,可以帮助图像资源存储机构突破自身资源的限制,实现和其他馆藏机构之间的图像资源互操作,有效提高相关机构在网络数据环境下的研究能力^[21]。碑刻是一种特殊的文化资源,更注重图像资源的保存、利用和传播,IIIF恰好可以满足其对数字图像资源的利用需求。

2 碑刻资源语义化组织的框架设计

为保证资源的互联互通,充分实现碑刻资源的保护、共享、传播,满足不同人群的利用需求,本文通过确定实体对象及其属性关系,在本体模型构建基础上引入关联数据技术与IIIF技术,进行碑刻资源语义化组织框架的设计。

2.1 碑刻资源相关概念及其实体属性分析

“碑”的概念最早出现于周代,原意是指没有文字的竖石,作为一种标志,后经发展演变逐渐成为刻有文字的碑石。清代知名金石学家叶昌炽在《语石》中说:“凡刻石之文皆谓之碑,当是汉以后始”^[31]。关于碑刻,本文主要采用黄永年的论释:“所谓碑刻,均以石上刻有文字,供阅读识别者为限”^[32],即刻有文字的历代碑、碣等皆可称作碑刻或石刻^[33]。碑刻分类对于碑刻资源组织与描述很有必要。本文采用毛远明^[34]《碑刻文献学通论》中的分类方法,将碑刻从形制与内容两方面进行分类:按形制,分为碑、碣、石阙、摩崖、墓志、经幢石柱铭刻、造像题记、石刻画像题字等类别;按内容,分为记事赞颂碑刻、哀诔纪念碑刻、祠庙寺观碑、诗歌散文、图文碑刻、应用文碑刻、石经、题名题记、特殊碑刻等9类。碑刻资源这一概念,目前未见明确定义。根据上文对碑刻概念的探讨,可以将碑刻资源这一概念的内涵从狭义与广义两种角度进行归纳说明。从狭义来看,碑刻资源即指碑刻本身;从广义来看,碑刻资源是表示碑刻相关要素的集合,即以碑刻本身为主体,同时涵盖与其相关的人物、时间、地点、关联资源等诸多要素。本文主要从广义的角度对碑刻资源进行研究。

根据上述对碑刻资源概念的定义,本文结合相关调研结果对碑刻资源的实体要素及其相关属性信息进行分析,将碑刻资源实体划分为碑刻、人物、时间、地点与相关资源5类。其中,碑刻实体是对碑刻本身的阐释,是碑刻资源的核心要素,属性元素包括名称、作者、书体、尺寸、碑文、碑刻图像、出土地、藏地、形制、镌刻人、材质等,这些属性信息对碑刻的外部特征及内部特点进行了充分描述。人物、时间、地点与相关资源作为相关实体要素,为碑刻资源内涵和外延的扩展提供了可能。根据上述分析并进一步细化,可以构建碑刻资源实体属性表(见表1)。

2.2 碑刻资源本体模型设计与构建

在确定碑刻资源实体与属性的基础上,便可通过本体模型的设计与构建对领域内知识进行模型化的梳理描述。所谓本体是对领域知识进行抽象的、形式化的概念模型。碑刻资源本体模型设计与构建的步骤主要包括在明确碑刻资源领域内概念内涵的基础上分析本体的类和属性关系、复用现有本体模型、设计词表、在

表1 碑刻资源实体属性表

实 体	属 性
碑刻	名称、作者/书写人、书体、尺寸、碑文(包括碑额、碑阳、碑阴文字)、碑刻图像(包括原照或拓片)、出土/藏地、碑刻类型/形制、镌刻人、材质等
人物	姓名、性别、籍贯、生年、卒年、朝代、字、号、社会身份、画像、任职事件、作品等
时间	朝代、年号、公元纪年等
地点	地名、别名、地理坐标等
相关资源	如书画、诗词、古籍、方志(包括题名、创作背景、正文、校注、创作时间、版本、体裁、典故等)

Protégé工具上完成本体模型的创建。

首先,对已有本体类和属性关系进行确定。上文确定了碑刻资源的实体可与本体中的类相对应,实体属性与本体属性相对应,类包括碑刻、人物、时间、地点、相关资源5个部分,其核心类为碑刻。其余类均与碑刻类具有直接联系,从而构成碑刻类的对象属性,成为与碑刻实体直接关联的实体。除对象属性外,各实体本身所特有的属性亦构成其数据属性,是以数值或文本形式存在,如碑刻实体中的尺寸、碑文、碑刻图像等内容。其次,对现有本体进行复用。通过对现有本体模型进行调研,在底层受控词表的基础上决定对都柏林核心元素(Dublin Core)、Friend of A Friend词表(FOAF)、视觉资源核心类目(VRA Core)、艺术品描述类目(CDWA)、地名本体(GeoNames),以及上海图书馆开放数据平台相关本体,包括数字方志本体(Fangzhi)、人名规范库(Shlnames)、古籍本体(Shlancient)、中国历史纪年表、地理名词表等国内外通用本体进行复用。最后,根据自身需要进行词表的自建补充。以碑刻资源本体模型中核心词表(见表2)为例,自建词表Regional Culture Resource(缩写为rcr)定义域为rcr:Inscription,关联目标属性包括人物rcr:Person、时间rcr:Time、地点rcr:Place及相关资源rcr:Resource,相关资源可以包括古籍sh:Work、方志sh:Fangzhi、诗词rpr:Poetry等。

在本体复用的基础上实现对实体、属性与关系的建模,通过Protégé工具实现本体模型的构建,如图1所示。后续通过制定规则、导入数据生成RDF数据,通过数据属性对其内在属性充分描述的同时,通过对象属性将与碑刻相关的实体要素关联,为后续实现内部与外部关联提供依据,也为语义推理提供基本的逻辑结构。

表2 碑刻资源碑刻类属性词表的设定

核心属性 (Property)	标签 (Labels)	定义域 (Domain)	值域 (Range)	属性类别 (PropertyType)
dc:title	名称	rcr:Inscription	rdfs:Literal	DataProperty
shl:temporal	朝代	rcr:Inscription	rcr:Time	DataProperty
dc:date	公元纪年	rcr:Inscription	rcr:Time	DataProperty
time:temporalEntity	年号纪年	rcr:Inscription	rcr:Time	DataProperty
vra:location	所处地点	rcr:Inscription	rcr:Place	ObjectProperty
dc:creator	作者/书写人	rcr:Inscription	rcr:Person	ObjectProperty
rcr:typeface	书体	rcr:Inscription	rdfs:Literal	DataProperty
vra:measurements	尺寸	rcr:Inscription	rdfs:Literal	DataProperty
rcr:content	碑文	rcr:Inscription	rdfs:Literal	DataProperty
vra:image	碑刻图像	rcr:Inscription	xsd:anyURI	DataProperty
rcr:site	出土/藏地	rcr:Inscription	rcr:Place	ObjectProperty
vra:type	碑刻类型	rcr:Inscription	rdfs:Literal	DataProperty
pmb:engravedBy	镌刻人	rcr:Inscription	rcr:Person	ObjectProperty
rcr:releventPerson	相关人	rcr:Inscription	rcr:Person	ObjectProperty
vra:relation	相关资源	rcr:Inscription	rcr:Resource	ObjectProperty
dc:description	内容描述	rcr:Inscription	rdfs:Literal	DataProperty
vra:material	材质	rcr:Inscription	rdfs:Literal	DataProperty
dc:type	碑文类别	rcr:Inscription	rdfs:Literal	DataProperty

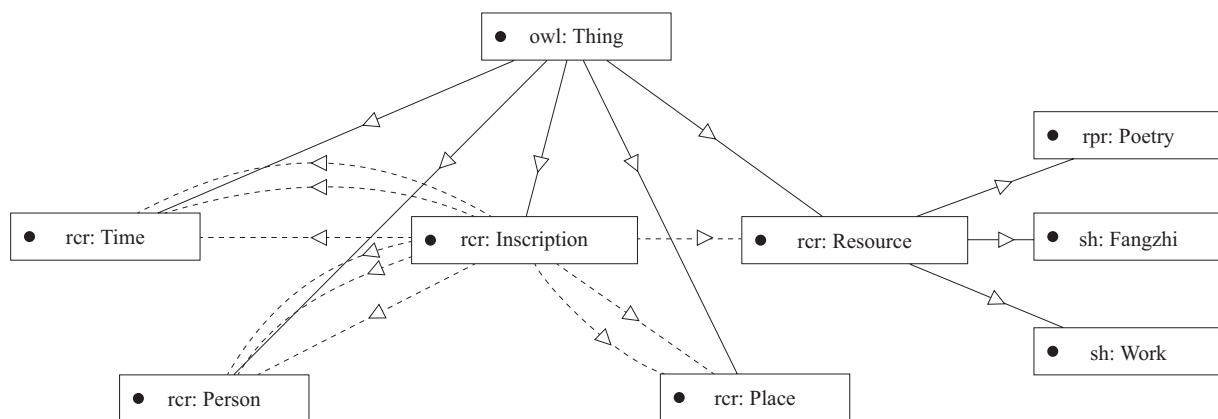


图1 碑刻资源本体模型

2.3 IIIF服务集成

IIIF统一了传递图像的标准,该标准允许使用共享软件对图像的共同操作,解决了图像资源存储分散、机构重复建设的问题,从系统层实现图像资源的数字化重组、数据化识别、语义化标注、智慧化关联。IIIF通过定义一组通用应用程序编程接口规范,来支持不同图像存储库间的互操作,并通过访问用于图像的标准URL结构,使得图像一经发布便可多次复用,解决了创建、

重新创建访问映像以及为不同平台复制文件的繁重工序等问题。IIIF目前有4个标准API,即图像API(Image API)、呈现API(Presentation API)、内容搜索API(Content Search API)以及认证API(Authentication API),均可用于图像元数据的规范、图像的呈现以及语义注释、共享与检索,并且可根据现实需求进行新API扩展。

图像API通过常规图像服务器即可完成,能够对区域、尺寸、质量、格式进行调整,如局部放大、剪切、色

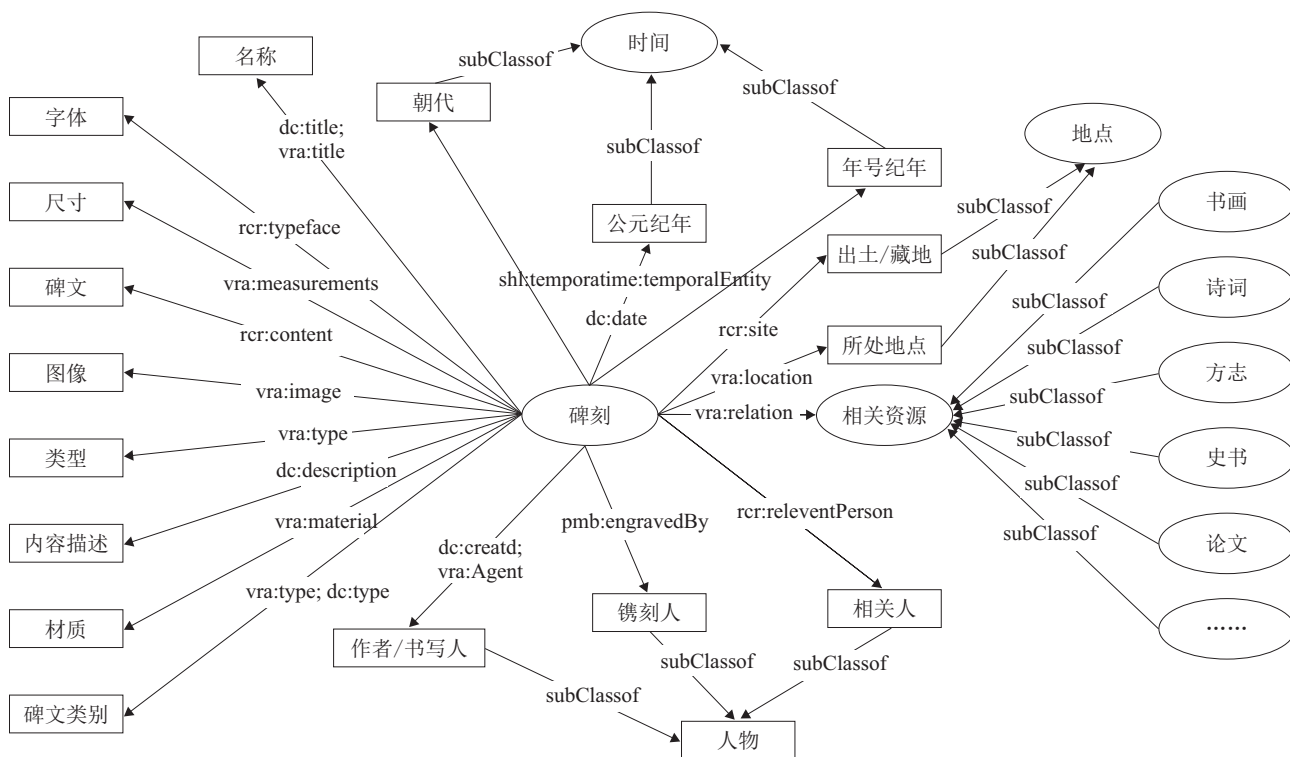
度调节等, 允许获取源图像, 应用URL参数可以提取该图像区域, 然后进行基本图像操作; 呈现API支持将图像以及相关资源和元数据呈现给用户, 描述了如何通过标准的方法提供图像资源的组织布局; 内容搜索API旨在帮助在资源中搜索注释内容, 即一旦找到支持搜索的资源, API将在资源(或资源的注释)中查找文本, 搜索的内容通常是资源文本, 作为可以绘制到画布上的注释, 常见的用例是返回OCR文本或转录; 认证API有助于协调工作流, 以便使不受信任的应用程序可以执行必须的身份验证步骤。

要完成IIIF服务的构建, 首先需要对图像服务器进行配置, 以便能够对图像请求和info.json请求进行响应, 即实现对图像API的应用, 获得图像和info.json后, 便可以轻松使用平移/缩放查看器。目前大多数开源IIIF图像服务器或浏览器可以在Awesome-IIIF^[35]列表

中找到, 本研究选择Cantaloupe作为图像服务器来实现IIIF服务的构建。

2.4 碑刻资源语义关联模型构建

关联数据通过建立实体内外部关联来完成数据的关联组织, 资源实体定义了唯一URI标识, 实体资源之间的访问通过HTTP URI规范来实现^[36]。根据本体词表构建形成以碑刻为核心的语义关联模型(见图2)。通过实体间的语义关联以及相关属性的设定, 实现碑刻资源较为全面而完善的描述, 形成较为完整的逻辑体系架构, 由此可以完成对碑刻资源知识内涵的深度挖掘与揭示。此外, 由唯一的URI将不同数据密切联系在一起, 通过引用其他平台的URI, 也为其他平台引用提供相互连接的接口技术, 实现不同平台间资源的开放共享。



关联数据为分类法管理提供了一个自然的平台, 当与IIIF配合使用访问图像资源时, 关联数据可以提供适合图像注释的应用程序。从技术实现角度来看, IIIF框架主要基于关联数据的理念来实现图像资源的共享和交互。例如, 关联数据较多应用于格式化文本资源之间的信息关联和交互, IIIF框架则将研究对象提升为图像资源。通过利用Drupal或Omeka等内容管理平台(或软

件), 能够在实现IIIF服务集成的基础上完成关联数据发布。

通过以上碑刻资源语义化组织框架搭建的平台, 可使资源更加便捷地传播与利用、更好地服务用户, 实现面向用户的可视化展示、知识查询、知识推理、知识获取, 最终在碑刻资源数字化及语义发布的基础上, 为用户提供关联数据检索和访问服务, 包括关联数据浏

览、SPARQL查询、语义检索、个性化服务等多种形式的服务。既可以为专业人员提供查询检索,也可以为普通用户提供一站式导航服务。该平台支持导航栏访问和关键词检索等功能,方便普通用户获取资源信息与知识,同时IIIF还可以使用户随时随地高效利用大图像文件。在关联数据基础上引入IIIF的设计思路,一方面可以优化纷繁复杂的碑刻资源数据,另一方面可以提供个性化的图像服务。下文以焦山碑林为例展开具体的实证研究。

3 碑刻资源语义化组织的实证研究

3.1 焦山碑林碑刻资源概况及其采集处理

焦山碑林位于江苏省镇江市焦山公园,是全国重点文物保护单位。碑林汇聚了大量江南碑刻精粹,由摩崖石刻与碑林两部分组成,展示了历代书法风格的变迁,具有一定的地域特色和较高的艺术与文化价值。

摩崖石刻经统计共有97方,主要环集在焦山西侧峭壁,包含六朝以来真、草、隶、篆不同种类刻石。其中,浮玉崖段有61方,巨公崖段有25方,雷轰石段11方。形式上多以题名题刻类为主,内容包含抒发忧国之愤、怀古颂今、寄托抱负、阐述佛经教义、摘录道家微言、张扬个人情怀等多种类型作品,其中“大字之祖”瘞鹤铭原址便在此处。此外,还有唐刻《金刚经偈句》,宋刻《壮观亭址瘞鹤铭》《米芾摩崖题名题刻》《陆游踏雪观瘞鹤铭》《吴踞春游焦山诗》,以及方豪、洪亮吉、陶澎、康有为题记等明清作品。碑林现藏碑约500方,包含记事、赞颂、哀诔、纪念、诗歌、散文等多种体裁,有苏东坡、米芾、赵孟頫、董其昌、文征明等名家墨痕,以及唐《魏法师碑》、宋《章岷墓志》、明刻宋米芾临《兰亭契贴》、清《澄鉴堂法帖》等^[37]知名碑刻。笔者实地调研发现,许多碑刻损毁严重,亟需进行科学保存与保护,而数字化保护是重要的一种方式。

我们将收集到的碑刻数据,按照分类对相关概念属性进行逐条筛选摘取,汇总后保存为结构化数据,作为实例化操作的基础数据。其中,摩崖石刻关键词词频统计的高频词包括正德、道光、万历、光绪、嘉庆、同治这类年号以及民国、庚寅、仲春、十年、九月、三月、四月等时间类型词语,体现游览地点、游览者籍贯的焦山、镇江、浙江、宜兴、江宁、丹徒、长沙等地点类型词语,还有如监察、卫指挥、布政使、御史、大学士等

体现官职的词语,展现了历代文人、官员等游览焦山之盛况。

3.2 IIIF服务构建

IIIF服务的构建主要通过Cantaloupe图像服务器来实现,Cantaloupe可以根据研究需要生成高分辨率图像和动态缩略图,可进行自由裁切、缩放和旋转等操作,通过Image API提供规范的URI语法,实现对图像的互操作与共享。

IIIF图像的URI由定义部分组成,从而实现人类可读性和机器可处理性。使用不同的参数组合对相同的图像进行个性化调用,如从全尺寸图像中提取区域、使用尺寸参数缩放图像、图像旋转、颜色质量设置、格式选择等,本研究IIIF参数模板及瘞鹤铭实例(见图3)的下部显示了在不同URL下瘞鹤铭的显示状态,左侧为目前瘞鹤铭残石实景图,右侧则是通过区域选择对其一块残石进行展示,上方为残石实景灰度图,下方为清拓本灰度图。通过拓本与实景图的对比,可较为清晰地辨认出瘞鹤铭刻字的细节、字体等,对研究书法史、艺术审美等提供了绝佳的对比,亦为欣赏“大字之祖”的气韵与灵动提供了较为便捷的方式。

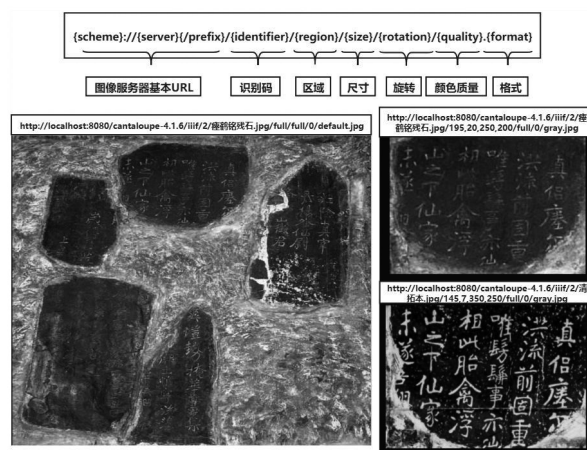


图3 IIIF参数模板及瘞鹤铭实例

3.3 焦山碑林碑刻资源语义化组织的实现与发布

本研究主要立足于开放共享的角度,故鉴于Drupal内容管理平台对关联数据、IIIF服务的兼容性而选择Drupal内容管理平台对IIIF服务进行支撑,在完成图像服务器的搭建并对图像资源进行指向后,利用Drupal平

台对焦山碑林碑刻资源进行个性化集成。

3.3.1 基于Drupal的设计思路

基于Drupal内容管理平台对关联数据、IIIF服务的兼容性,在2.4节的基础上,本文选用Drupal作为焦山碑林碑刻资源的发布实施平台。由于Drupal平台是基于模块化的设计思路,主要功能是借助模块来实现,故需要对模块进行合适的选取以保证功能的实现。本文主要选用ARC2、RDF、SPARQL、IIIF等作为核心模块来实现关联数据的发布功能,其中ARC2为RDF数据提供存储空间,RDF构建关联数据最基本的组织形态结构,SPARQL提供RDF数据的查询检索功能,IIIF模块则通过对图像服务器内图像API的调用,实现对碑刻图像的高效利用与互操作。

Drupal的核心要素主要包括内容类型、字段、节点三部分,分别与本体模型中的类、属性、对象一一对应。通过内容类型、字段、节点与本体中类、属性、实例的对齐,从而实现2.4节碑刻资源语义关联模型到Drupal站点内容的映射。节点创建后,每个实例化数据均会产生唯一的URI与之对应,从而实现将Drupal内容转化为语义化数据,实现碑刻资源的关联数据发布。内容从多维度进行呈现,一方面可以从朝代、人物、书法风格、字体、地点等多种角度对碑刻进行分类展示;另一方面通过数据实体间的关联,可使碑刻与更多的资源(如相关的书画、诗文等)相关联,使得面向用户的资源全面、形象和具体。

3.3.2 实例化发布

按照碑刻资源语义化组织框架的设计思路,在明确碑刻资源实体概念和概念间关系的基础上,根据设计的碑刻资源实体属性词表以及碑刻资源本体模型,结合Drupal内容管理系统中内容类型、字段、节点的构建方法以及Drupal对IIIF服务的集成,以碑刻资源语义关联模型为基础完成焦山碑刻关联数据发布。

首先,根据碑刻实体概念在Drupal站点内部构建碑刻资源内容类型,主要包括碑刻、人物、时间、地点、相关资源5种内容类型,用以表示碑刻资源的实体概念,并按照各实体的属性添加内容字段,如碑刻的内容字段包括碑刻的名称、朝代、时间(公元纪年/年号)、所处地点/位置、作者/书写人、书体、尺寸、碑文(碑额、

碑阳、碑阴文字)、碑刻图像(原照或拓片)、出土/藏地、碑刻类型/形制、镌刻人、相关人、关联资源(诗词、书画、拓本、方志、古籍、研究论著等)、内容描述、材质等字段信息。碑刻的名称、书体、尺寸、碑文、图像、碑刻类型/形制、内容描述、材质等为数据属性,展示了碑刻自身所具有的独特属性。而其成碑朝代、时间、所处地点/位置、作者/书写人、出土/藏地、镌刻人、相关人、关联资源等为对象属性,这样碑刻实体通过对象属性与人物、时间、地点、相关资源等实体相关联,碑刻在充分展示其自身属性特点的同时,完成对不同实体要素的关联。需要说明的是,其中碑刻图像资源主要通过添加IIIF模块来完成对Cantaloupe上发布图像资源的调用。对象属性通过Node reference字段来完成不同实体间的关联。

其次,在碑刻资源内容类型以及字段创建后,通过为内容类型和字段填充具体实例内容,从而完成碑刻资源在Drupal平台的集成和关联数据的可视化发布。每个内容实例的创建均具有唯一性且与唯一的URI相对应。本文以瘞鹤铭为实例进行关联数据在本地平台发布。发布内容一方面聚焦瘞鹤铭的核心属性,包括对碑名、内容描述、碑刻图像、尺寸、碑文、碑刻类型/形制、内容描述、材质、书法字体等数据属性的描述,充分体现瘞鹤铭的自身特性。对象属性显示瘞鹤铭成碑朝代为“南梁”(存争议),具体成碑时间无从考证,所处地点或位置为“焦山”,藏地为“焦山碑刻博物馆”,作者或书写人为“陶弘景”(存争议),这些对象属性均有具体的实体节点相对应。另一方面,发布内容展示了与其相关联的其他类型资源,包括方志、诗词、典籍、拓本、相关碑刻以及相关论文等。从方志来看,现存镇江相关志书基本都对瘞鹤铭这一历史知名碑刻有所记载。从诗词来看,有一系列瘞鹤铭相关诗词流传于世,其中不乏名人佳篇,如吴琚(宋代)所作的《春日焦山观瘞鹤铭》,值得一提的是这篇诗作至今可见于焦山的摩崖石刻。从典籍来看,历史上较为知名的金石刻录典籍上均有对瘞鹤铭的记载,如宋赵明诚《金石录》、清顾炎武《金石文字记》等。从拓本来看,目前现存较知名的瘞鹤铭拓本有宋拓本和清拓本,这些传世拓本再现了瘞鹤铭碑刻在历史上的样貌,对于瘞鹤铭的研究具有重要意义。从相关碑刻来看,既有碑刻是对瘞鹤铭的重摹,又有碑刻是对瘞鹤铭相关事件的记载。还有一系列的学术论文对瘞鹤铭进行了研究。通过瘞鹤铭碑刻的发布实例可以看出,一定程度实现了瘞鹤铭碑刻内容

全面、具体的呈现,同时也对其知识内涵进行了深度揭示,对相关资源进行了较为全面的关联,验证了碑刻资源语义化组织框架的可行性,完成了对碑刻资源语义化组织模式的实验性探索。

4 总结与展望

中国碑刻文化源远流长,与书法、文学、绘画等学科关系密切,有着丰富的历史文化价值。在开放共享的大背景下,关联数据、IIIF等技术方法的日趋成熟为碑刻资源的数字化实践和研究奠定了理论与技术支撑。本文通过关联数据与IIIF的结合,对碑刻资源语义化组织进行了探索性研究,有助于实现碑刻资源的语义聚合,深度揭示碑刻资源的语义内涵和文化价值。平台化的场景实现,一方面可以为研究人员提供精准、全面的碑刻资源内容,满足其研究需求;另一方面可以通过提供丰富的碑刻资源内容,满足大众文化普及和地方文化推广的需要,为地方精品文化的推广、传播和我国历史文化遗产的保护提供参考。但是,由于本文在构建碑刻资源语义关联模型过程中没有可复用的最优本体模型,仅对碑刻核心要素进行分类探索,模型构建还需要进一步扩展。同时,IIIF的应用仅进行了图像API的应用实践,呈现API、内容搜索API、认证API以及图像标注的深度应用将是后续研究的重点。

参考文献

- [1] 李岩,顾涛. 碑刻文化与历史记忆 [N]. 中国社会科学报, 2020-08-21 (4).
- [2] 王宁. 数字化时代的碑刻与碑刻学研究 [J]. 陕西师范大学学报 (哲学社会科学版), 2017, 46 (2): 119-121.
- [3] 袁玉红. 国家图书馆石刻拓片的数字化 [J]. 图书馆理论与实践, 2014 (5): 17-18.
- [4] 马琛. 试谈新时期石质文物保护的新方式——以孔庙和国子监博物馆石刻文物三维数字化扫描为例 [J]. 文物鉴定与鉴赏, 2020 (1): 58-61.
- [5] 戴萍. 数字化时代大理州碑刻文献的整理策略 [J]. 山西档案, 2017 (6): 148-150.
- [6] 潘笑. 石质文物数字化信息参考模型研究及应用 [D]. 北京: 北京建筑大学, 2020.
- [7] 邱孝凤, 唐金娟. 3Dmax的三维建模技术在涪溪摩崖石刻数字化重建中的应用 [J]. 数字技术与应用, 2019, 37 (12): 67, 69.
- [8] 徐晨希. 南朝陵墓石刻的整体性保护与多样化展示 [J]. 东南文化, 2020 (4): 20-25.
- [9] 杨祥民, 张新. 南朝石刻艺术的数字化保护与设计 [J]. 装饰, 2020 (2): 130-131.
- [10] 唐兰林. 南京六朝陵墓石刻的数字化保护技术应用研究 [D]. 南京: 南京艺术学院, 2019.
- [11] 方兴林. 徽州文化遗存资源特色数据库建设研究——以徽州碑刻为例 [J]. 西昌学院学报 (自然科学版), 2017, 31 (4): 71-74.
- [12] BERNERS-LEE T. Linked Data [EB/OL]. [2021-11-09]. <https://www.w3.org/DesignIssues/LinkedData.html>.
- [13] 韩洪帅. 基于关联数据的传统美术类非遗视频资源语义化组织研究 [D]. 武汉: 华中师范大学, 2019.
- [14] 李世新. 基于关联数据的近代篆刻艺术印人和印章知识发现研究 [D]. 保定: 河北大学, 2020.
- [15] 赵雅倩. 非遗资源知识组织及语义发现研究 [D]. 保定: 河北大学, 2020.
- [16] 魏清华, 刘勐. 非物质文化遗产知识库构建——以甘肃省国家级非遗为例 [J]. 图书馆学研究, 2020 (6): 33-38.
- [17] 高劲松, 彭博. 基于主题识别的文物信息资源知识发现方法研究 [J]. 情报科学, 2021, 39 (4): 9-14.
- [18] 王晓光, 侯西龙, 程航航, 等. 敦煌壁画叙词表构建与关联数据发布 [J]. 中国图书馆学报, 2020, 46 (4): 69-84.
- [19] World Digital Library [EB/OL]. [2021-11-09]. <https://www.wdl.org/zh/>.
- [20] 陈涛, 张永娟, 单蓉蓉, 等. 数字人文图像资源语义化建设框架研究 [J]. 数字人文, 2020 (2): 106-115.
- [21] 陈涛, 刘炜, 孙逊, 等. IIIF与AI作用下的文化遗产应用研究新模式 [J]. 中国图书馆学报, 2021, 47 (2): 67-78.
- [22] 数位人文研究平台 [EB/OL]. [2021-11-09]. <http://dh.ascdc.sinica.edu.tw/member/index.html>.
- [23] 陈晓亮, 苏海潮, 刘心舜. 图书馆特藏数据结构化的探索 [J]. 图书馆杂志, 2019, 38 (6): 44-48, 91.
- [24] 陈涛, 单蓉蓉, 张永娟, 等. 数字人文研究的语义支撑平台构建研究——以ECNU-DHRS平台为例 [J]. 图书馆杂志, 2021, 40 (3): 69-77.
- [25] 印谱文献虚拟图书馆 [EB/OL]. [2021-11-09]. <http://yin.fudan.edu.cn/>.
- [26] 杨佳颖, 许鑫. 民国报纸广告图像资源的语义标注——以《新闻报》所刊的越剧广告为例 [J]. 图书馆杂志, 2021, 40 (3): 96-102.
- [27] 张永娟, 刘炜, 于建荣, 等. 基于IIIF和语义知识图谱的印章资源整合与知识发现研究 [J]. 图书馆情报工作, 2020, 64 (7): 127-135.

- [28] 付跃安. 国际图像互操作框架 (IIIF) 及在数字资源集成中的应用 [J]. 图书馆论坛, 2020, 40 (4): 159-166.
- [29] 耿曼曼. 图书馆图像资源开发利用: 国际图像互操作框架 [J]. 图书馆学研究, 2019 (18): 37-45.
- [30] 张铁. 国际图像互操作框架及其应用分析 [J]. 数字图书馆论坛, 2019 (5): 42-49.
- [31] 叶昌炽. 语石 [M]. 杭州: 浙江大学出版社, 2018: 76.
- [32] 黄永年. 碑刻学 [J]. 新美术, 1999 (3): 3-5.
- [33] 张希. 山西碑刻书法艺术 [D]. 南京: 南京师范大学, 2015.
- [34] 毛远明. 碑刻文献学通论 [M]. 北京: 中华书局, 2009: 7.
- [35] Awesome International Image Interoperability Framework (IIIF) [EB/OL]. [2021-11-09]. <https://github.com/IIIF/awesome-iiif>.
- [36] 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.
- [37] 王同顺. 焦山碑林及摩崖石刻述略 [J]. 文教资料, 1994 (3): 58-64.

作者简介

李永卉, 女, 1979年生, 博士, 副研究馆员, 研究方向: 数字人文、地方文献。

刘沁芃, 男, 1997年生, 硕士研究生, 研究方向: 数字人文、关联数据、知识组织。

周树斌, 男, 1996年生, 硕士研究生, 通信作者, 研究方向: 数字人文、关联数据、知识组织, E-mail: zshubin001@163.com。

屠纪军, 男, 1970年生, 副研究馆员, 研究方向: 碑刻文化、地方文化。

Research on Semantic Organization of Inscription Resources

LI YongHui^{1,2} LIU QinPeng¹ ZHOU ShuBin¹ TU JiJun³

(1. Institute of Science and Technology Information, Jiangsu University, Zhenjiang 212013, P. R. China;

2. School of Law, Jiangsu University, Zhenjiang 212013, P. R. China;

3. Zhenjiang Jiaoshan Stele Museum, Zhenjiang 212013, P. R. China)

Abstract: As a precious non-renewable historical and cultural heritage, stele inscriptions carry the memory of civilization and culture. Therefore, it is of great significance to explore the semantic organization of it. This research combines IIIF to explore the semantic organization model of inscription resources. First of all, the semantic organization framework of inscription resources is designed to provide a complete and feasible plan for the collection and description of inscription resources to the organization and release of inscription resources. Secondly, taking the Jiaoshan Stele Forest in Zhenjiang as an example, based on the Drupal content management platform to integrate linked data applications and IIIF services, to realize the related publishing and visual display of the relationship between the concept of inscription resources, and verify the feasibility of the semantic organization framework of inscription resources, which provides reference for the in-depth development and service of inscription resources.

Keywords: Inscription Resources; Linked Data; IIIF

(收稿日期: 2021-11-10)