



面向数字人文的典籍图像深度揭示与利用*

□钱智勇 陈涛 张志美 徐宇红 何书

摘要 典籍图像是馆藏重要资源,图像数字化和内容揭示是数字人文基础建设,可实现“以图证史”“图像叙事”、多媒介传播等应用价值。文章依据图像元数据、资源描述框架(RDF)和国际图像互操作框架(IIIF)等标准规范,研究典籍图像深度揭示与利用的模型架构与实现方法。该模型由图像数字化、内容深度标注和数字人文应用三个相互关联的模块组成。从图像场景、实体对象和语义关联三个层次,揭示图像特征、典籍注释及背景知识,辅助数字人文研究。例证选用经典辞书《尔雅》插图版《尔雅音图》进行数字化,构建《尔雅》多语语义词表并与实体进行关联,深度揭示《尔雅》词汇的古今释义、音韵、分类、典籍注疏、例句典故等知识内容,探究《尔雅》图像的数字人文应用场景。

关键词 典籍图像 数字人文 《尔雅音图》 深度揭示 IIIF 关联数据

分类号 G254.3 TP391.1

DOI 10.16603/j.issn1002-1027.2022.05.005

“典籍”最早是指记载先祖法度或国家法则的重要文献,后被用作各种书籍的统称^[1]。典籍中的图像又称插图,在写本时期,典籍中包含了大量的插图。南宋郑樵《通志·图谱略·索象篇》记载:“古之学者为学有要,置图于左,置书于右,索象于图,索理于书。”可以看出古代学者治学过程中图像与文字不可分割的关系。图书馆、博物馆、档案馆等典藏机构收藏大量典籍图像载体,包括书影、印谱、地图、卷轴、乐谱、手稿、档案图像等类型^[2],这些典籍图像是宝贵的文化遗产,图像数字化建设是数字人文基础设施的重要内容。利用国际图像互操作框架(International Image Interoperability Framework, IIIF)与关联数据描述框架可实现对典籍图像深度揭示和语义关联,使典籍图像在数字人文中发挥重要价值与作用:(1)“以图证史”作用。通过对典籍图像高清图呈现和内容深度揭示,增强了图像还原史实、以图证史的文献考证价值^[3]。(2)“图像叙事”作用。在碎片化阅读时代,“图像正以前所未有的力量从文化的每个层面向我们压来”^[4],典籍中的图像与文字通过互注、互文、互释、互读等方式实现文本叙事功能,

为读者呈现图文同现的叙事效果。(3)增强典籍多媒介传播作用。文学图像化改变了文学传播的场域、传播方式及传播理念,图像与文本的深度关联,可以互相带动彼此在异域文化的深入传播,并可启发不同民族的读者对典籍作品的深层认知^[5]。

1 国内外相关研究

基于元数据规范组织典籍图像,揭示图像的特征、主题、分类、创作者、馆藏、版本等内容,但读者难以快速检索典籍图像内容及其所蕴含的背景知识,典籍与图像分离,成为信息孤岛,难以广泛传播与利用,因此需要对典籍图像内容进行细粒度知识组织。近年,学者围绕数字图像内容组织与利用进行了许多研究。在国内,曾子明等提出面向数字人文的图像语义描述模型^[6]。张永娟等依据 IIIF 整合印谱图像资源,辅助知识发现^[7]。王晓光等构建敦煌壁画主题词表、敦煌石窟本体及相关数据模型^[8]。陈涛等通过 IIIF 与人工智能相结合,构建沉浸式交互平台,实现图像资源标注与发布^[9]。杨佳莹等通过报纸广告本体模型,准确揭示广告图像文本信息^[10]。

* 国家社会科学基金一般项目“基于关联数据的古代辞书知识组织与应用研究”(编号:21TQB094)、江苏省社会科学基金项目“大数据环境下典籍多语语义词典构建与知识发现实证研究”(编号:19TQB003)的研究成果之一。

通讯作者:钱智勇,ORCID:0000-0003-2967-989X,邮箱:qzy@ntu.edu.cn。



在国外,邓斯特(Dunst A)等人研究图形叙事语料库,采用 XML 注释标题和全文示例^[11]。斯托克(Stork L)等人通过全文转录和实体提取,直接标记和注释手写档案中的图像文档内容^[12]。程学芳(Cheng X F)等人构建多层语义描述框架,描述图像的内涵语义信息^[13]。亚勒米苏·阿布加兹(Abgaz Y)等人提出一种利用人工智能技术挖掘文化遗产数字图像中人文信息的方法^[14]。

以上相关研究表明,在标注模型构建、古籍自动识别、实体提取、图像互操作、本体、关联数据与人工智能应用结合等方面,已经取得了许多研究成果和技术工具,可以实现图像元数据聚合、国际图像互操作、图像关联数据与本体词表开放数据服务。在典籍图像中的古籍文本内容深层标注、图像认知计算、语义检索、数字人文深度应用等方面尚有进一步拓展研究的空间。本文参考已有研究方法与技术,依据图像元数据规范、资源描述框架和国际图像互操作标准,研究典籍图像深度揭示与利用的实现路径与方法,构建面向数字人文的典籍图像数字化和深度语义标注的模型架构,并以《尔雅音图》为例进行文本图像的数字扫描、文本识别,通过《尔雅》多语语义词表与《尔雅》图像语义关联,深度揭示《尔雅》词汇和图像中的背景知识,进一步探究《尔雅》图像的以图证史、图像叙事、多媒介传播等数字人文应用场景。

2 典籍图像深度揭示的相关标准规范

典籍图像深度揭示是以图像元数据标准规范、资源描述框架理论、国际图像互操作框架为依据,实现基于图像内容的细粒度语义标注、检索、发现和开

放共享服务。

2.1 图像元数据标准

图像元数据标准是描述和限定图像数据对象所需要的一系列原则的集合^[15]。都柏林元数据标准(以下简称 DC)对图像内容描述的核心元素包括题名、主题、描述、来源、关联和范围。已有的图像元数据标准主要有描述艺术品、建筑物等类目的元数据标准(以下简称 CDWA)、描述视觉及图像资料类目的元数据(以下简称 VRA Core)、数字图书馆图像元数据标准(以下简称 CDL)、描述静态数字图像的元数据标准(以下简称 TMI)等^[16]。现有标准主要关注整体图像的不同属性和图像集合之间的同质性,对图像的外部特征进行详细描述,在对图像内容进行描述时,不同元数据集的元素之间可以建立映射关系,表 1 展示了 DC 核心元素与 CDWA、VRA Core 描述图像内容的元素映射。通过不同元数据标准之间的元素映射,再结合资源描述框架以及国际图像互操作标准,可以对图像数字对象的内容特征进行语义描述,多维度实现典籍图像深度揭示与内容检索。

2.2 资源描述框架

资源描述框架(以下简称 RDF)是国际语义网联盟(以下简称 W3C)推出的用于对结构化元数据进行编码、交换与再利用的基础架构,它提供一致化描述领域资源的机制,并允许不同的使用者根据需要扩展元数据,编制人机可读的领域词汇表,此外还提供结构化的相互兼容机制,为基于 XML 的各种不同元数据提供相互利用与转换的平台^[17]。RDF 提供了图像资源对象描述的开放数据模型,该模型由资源、

表 1 描述图像内容的 DC 核心元素与 CDWA、VRA Core 映射表

DC 核心元素	DC 映射到 CDWA、VRA Core	映射说明
dc: title (题名)	CDWA: title(艺术作品题名) VRAcore: title(图像名称)	DC 题名元素映射到 CDWA 艺术作品题名(CDWA: title)和 VRA Core 作品图像名称(VRA core: title)
dc: subject (主题)	CDWA: subject(艺术作品主题) VRA core: subject(图像主题)	DC 主题元素映射到 CDWA 作品主题(CDWA: subject)和 VRA Core 图像主题(VRA core: subject)
dc: description (描述)	CDWA: context(背景描述) VRAcore: culture(背景文化)	DC 资源描述元素映射到 CDWA 背景描述(CDWA: context)和 VRA Core 背景文化描述(VRA core: culture)
dc: relation (关联)	CDWA: related works(相关作品) VRA core: relation(作品间关系)	DC 关联元素映射到 CDWA 相关作品(CDWA: related works)和 VRA Core 作品间关系(VRA core: relation)
dc: coverage (覆盖范围)	CDWA: style(风格) VRA core: style(风格)	DC 时空特征元素映射到 CDWA 的艺术作品风格(CDWA: style)和图像作品风格(VRA core: style)



属性、声明组成,每个资源都被赋予一个 URI,读者既可获取资源本身,又可获取资源对象的内容描述。属性是指资源对象之间的关系。声明明确了资源对象的属性,RDF 使用图形化方式书写相同声明。如果许多 RDF 文档使用不同的元数据标准标识了相同的图像资源,通过聚合工具自动收集关于该资源的元数据并将所有开放数据融合起来,开放数据词汇表可以通过数据接口被获取并缓存到本地服务器中,以便快速访问,这为数字人文提供了尽可能广泛的图像数据及其属性关联。

2.3 国际图像互操作框架

国际图像互操作框架(IIIF)是由英国国家图书馆、牛津大学图书馆、哈佛大学等 29 个著名馆藏机构协作制定的一组支持馆藏数字图像资源互操作的框架标准,提供操作与访问图像资源的统一标准与方法,对典籍图像资源进行统一的在线组织、展示、检索与应用,以促进全球图像资源的互操作与开放获取^[18]。IIIF 框架通过定义一组通用的应用程序接口(API)规范实现图像资源互操作性与可获取性。IIIF 提供的接口包括:图像 API(Image API)、呈现 API(Presentation API)、检索 API(Search API)与授权 API(Authentication API)^[19]。图像 API 提供了图像处理的方法^[20],呈现 API 提供了用于构造图像和图像相关资源集合的数据模型^[21]。检索 API 通过扩展的标准参数检索图像元数据^[22]。授权 API 通过注册验证,对图像资源进行权限控制^[23]。IIIF 还开发了资源发现应用程序^[24]和具有图像策展功能的应用程序接口^[25]。关联数据和 IIIF 资源之间的双向关系丰富了数据集的语义,这

些规范和标准为典籍图像的深度揭示提供了技术支持。

3 典籍图像深度揭示与利用模型架构

典籍图像深度揭示与利用模型的构建目标是以图像元数据、资源描述框架和国际图像互操作标准为基础,对典籍图像进行数字化和内容深度标注,实现图像与内外部文献实体的语义关联,提高典籍图像的可理解性,促进典籍数字人文应用。图 1 展示的模型框架由典籍图像数字化组织与存储、图像深度标注与语义关联、数字人文应用研究等三个相互关联的模块组成。

3.1 典籍图像数字化

纸媒典籍数字化是采用高清晰度扫描仪等数字设备,对选定的典籍文本、图像进行扫描并转化为能被计算机识别的数字符号,再通过图像 OCR 识别、元数据组织等方式建成典籍全文和图像数据库。

第一步,扫描典籍文本。将文献或图片资料按原貌逐页扫描并存储为图像文件。扫描图像后,可能存在图像倾斜、文字不清晰、版心不正等问题,需要进行图像编辑校正、二值化处理,图像二值化处理可以较为真实模拟还原典籍的原始形态,保证文献的真实性,并可有效提高 OCR 识别率^[26]。在 IIIF 图像 API 中,每个扫描的图像都被设置唯一的图像 URI 标识,图像 API 基于标准 HTTP 请求返回图像的 Web 服务,对图像平移和缩放只需在 HTTP 请求的 URI 中提供区域坐标和大小参数,IIIF 服务器即可转换原始图像。

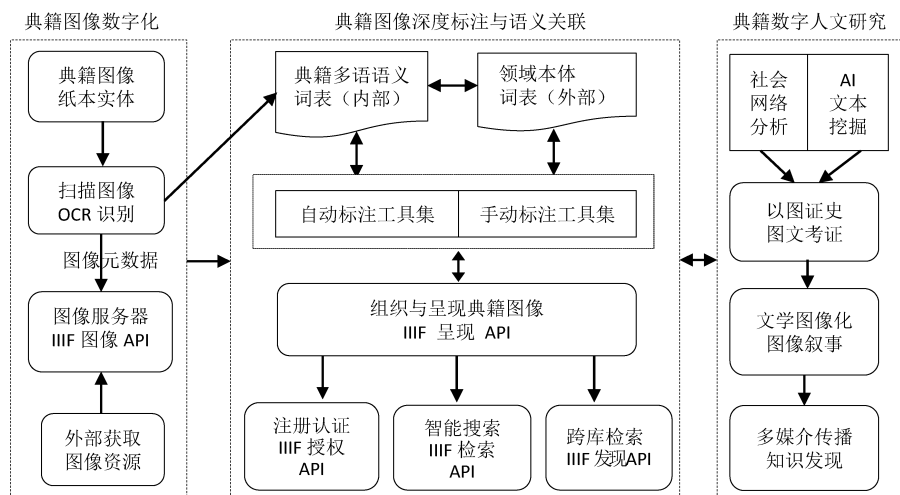


图 1 典籍图像深度揭示与利用模型架构

第二步,OCR识别。对扫描处理后的典籍图像进行文本识别,实现机器可阅读的全文生成,是数字人文的基础工作,利用OCR技术可以进行各种印刷字体图文版面的识别。难点在于如何准确识别典籍中的手写字、异体字等各种变体,为了提高OCR识别的准确率,可将OCR技术与深度学习相结合,通过特征映射让深度卷积神经网络(以下简称Deep CNN)去学习古文字的整体或偏旁等各种变体的一致性特征,然后将训练好的Deep CNN用于识别模型未见过的其他变体^[27]。典籍全文生成可采用OCR机器识别和众包方式的手工输入有效交互的方法进行,添加OCR的众包工具有助于准确识别异体字^[28]。

第三步,元数据组织。参考国外图像元数据标准和国家图书馆元数据规范^[29],结合特定馆藏的描述需求,映射和复用典籍图像核心元素和著录规则,据此进行图像元数据标注。揭示图像内容的核心元素包括:对象主题、对象分类、创作朝代、创作地点、相关典籍记载、风格、文化、语种、责任者相关描述等。IIIF呈现API定义了描述图像的标准模型,提供添加元数据的功能,可将任何图像或区域描述为一个有序的集合,包含清单、序列、画布、内容的基本结构。不同馆藏通过IIIF图像服务器发布出来的图像可利用发布URI接口进行重新组织和图像互操作。

3.2 典籍图像深度标注与语义关联

深度标注是通过元数据集对特定典籍图像进行基于内容的细粒度知识标引,来提高图像的可理解性。语义关联是通过IIIF与资源描述框架对典籍图像进行内部和外部词表的语义关联,以揭示图像的语义知识,并将图像与其他数字资源进行整合。根据典籍图像的存储异构性、格式多样性、内容离散性等特点,我们提出深度揭示典籍图像的实现路径:以构建典籍多语语义词表为基础,嵌入语义标注工具集,对典籍文献进行深度语义标注与多维度关联,并与相关资源进行整合。典籍多语语义词表是对多语种典籍文献中音、形、义相同或相似的字词进行集中揭示的词汇表,整个图像和子图像标注都可以使用词表中的控制词汇术语进行语义索引。通过元数据标注与关联数据转换、存储与发布,完成典籍多语语义词表的构建,再结合外部领域词表对典籍图像文本进行基于内容的深度语义标注。常用语义标注

工具有:CULTURA、CTEXT、TEXTGRID、MARKUS等文本标注系统,提供自动和手动分词标注功能,还包括搜索功能和外部参考功能。我们依据IIIF和本体规范设计了典籍图像深度标注的层级结构,如图2所示,分为场景标注、内容标注、语义标注三个层次。

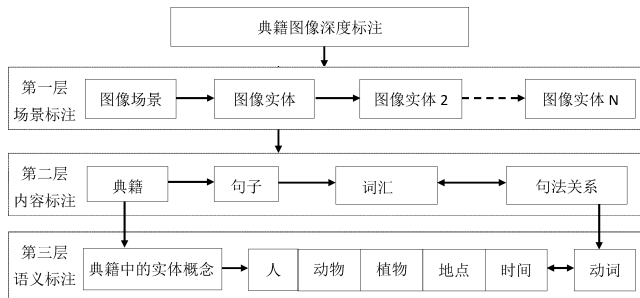


图2 典籍图像深度标注层次

第一层场景标注。根据图像场景,标注不同场景中的图像。采用矩形、圆形、多边形、自由标注等工具进行图像轮廓及其特征标注,并按主题对图像场景进行有序排列。第二层内容标注。对典籍、句子、词汇及其句法关系等内容进行标注。其中,句法关系包括典籍作品篇章之间的关系、句子结构关系以及词汇关系等。第三层语义标注。对典籍实体概念的语义及其关系进行深度标注。典籍中的实体概念包括“人”“动物”“地点”“时间”等概念类。“关系”包括概念层次关系和非层次关系。对典籍中的实体概念进行语义标注和属性关联,可依据典籍多语语义词表和相关本体词表,利用自动或手动标注工具,进行基于自然语言处理的命名实体识别、标注与抽取。实体之间概念层次关系可采用基于模板的方法半自动标注与提取;概念之间非层次关系主要指动作,可通过基于监督学习模型的方法,自动抽取实体关联的动词,实现典籍实体对象之间的语义关联。

典籍图像的深度揭示可以利用IIIF呈现API所定义的数据模型,实现典籍图像资源与资源组织、标注等操作相分离。它提供了用于构造图像和图像相关资源集合的数据模型,以及通过清单文件在IIIF感知应用程序之间进行数据交换的方法。IIIF呈现API中的资源包括集合、清单、序列、画布、注释、注释列表、范围、层和内容等。为了在呈现API中使用关联数据控制的词汇表,可以采用IIIF规范中定义的附加类型Annotation,将关联数据分类中的概念连接到典籍图像资源,可以将受控词汇表中



的主要资源主题 URI 插入 IIIF 画布上的注释资源中。通过 IIIF 检索 API 可实现图像标注层面的检索,还可利用 IIIF 发现 API 定义的基本资源列表、资源变更列表中记录的图像 URI 发布与变更信息,实现跨库与跨条目检索。

3.3 典籍图像数字人文应用

对典籍图像数字化、深度标注及语义关联,再结合深度学习、文本挖掘、时空分析和社会网络分析等技术,可实现典籍图像以图证史、图像叙事、多媒介传播等数字人文应用。

首先,通过图像比对、文献考证等实现“以图证史”。古籍中的图像包括各种手绘插图、印本木版插图等纪实性原创图画,反映了古代社会历史事件、人物形态、服饰、器具、宗教、建筑、音乐、动物、植物以及山川河流等自然地理面貌。经过高清扫描的典籍图像真实还原了各个朝代社会发展形态以及人类衣食住行等物质生活方式,经过文本识别和内容深度标注,与不同馆藏的同类图像建立语义关联,再结合机器学习进行图像特征识别,与出土文物及相关典籍进行图文比对和考证,可以实现以图证史的数字人文研究。借助典籍图像语义标注平台,人文学者在解读古籍文字和图像时可以参考和使用各种典籍数据库资源,进行分词、标注和聚类分析,辅助图文考证的研究。将汉字和图画可视化为知识图谱,可以激发学者探索考证新观点以及公众对典籍学习的兴趣与认知。

其次,通过场景标注、主题索引等实现“图像叙事”功能。在文学图像化语境下,典籍图像高清呈现与浏览改变了传统文字的呈现形式,典籍呈现可以向图文并茂甚至以图为主的多媒体形式转变,图像和文字以互读、互文、互释等方式呈现图像叙事功能。通过元数据映射建立基于故事的图像索引系统,识别图像中每个实体对象,分离并提取目标元素及属性,每一个元素都对应一个语义概念。IIIF 是描述图像故事的基础,提供了不同标记区域之间的坐标、形状、颜色和重叠关系,可描述图像中的主题和内容,图像资源的语义对象通过分类和互连被组织成图结构,每个故事或事件是通过图像中的元素而不是自由文本来表现的,图像中的设置和实体的变化随时间和空间演变,这有助于读者理解图像,标注的数据集使图像通过深度学习可理解,领域词汇和本体可以对标注输出进行标准化。

最后,利用多种媒介实现典籍图像在世界的传播。文学图像化改变了典籍文献传播的场域、方式与观念。利用 IIIF 与 RDF 将分散在世界不同馆藏的典籍图像进行组织、重构和基于主题内容的深度揭示,可开发各种应用 API,如策展 API 可以将画布中的图像进行任意形状的裁剪切割,并添加元数据,还可从 IIIF 图像典藏机构收集相关主题的图像,利用机器识别图像并自动添加标签,将图像数据或文本文件按主题聚合在一起。IIIF 发现 API 可以利用资源变更列表,对不同机构、数据库进行多维检索,分布在世界各地的典籍图像资源在多种平台上以多种格式被获取和分享,再利用流媒体、网站、搜索引擎、微信公众号、微博、头条、推特、抖音等多媒介发布与传播。读者与创作者通过手机等移动设备,利用沉浸式交互网络平台随时进行交流互动、解读书籍典籍图像、通过众包参与图像标注与研究,促进典籍的世界传播与利用。

4 案例研究:《尔雅音图》深度揭示与利用

本文选择《尔雅音图》作为典籍图像深度揭示与利用案例,《尔雅》是我国古代最早的训诂名物的语言专著,汇释了战国秦汉间的语言文字材料,汉代被列入《五经》,是古代儒生诵经的必读工具书^[30],历代《尔雅》注本以郭璞《尔雅注》最具代表性,郭璞在《尔雅注序》中称所作“别为《音》、《图》,用祛未寤。”^[31]据《隋书·经籍志》记载,郭璞注本有《尔雅音》二卷,《尔雅图赞》二卷,当时尚传于世,可惜后来亡佚,仅散见于前代旧籍^[32]。传世《尔雅音图》是清代两淮都转运使曾燠于清嘉庆六年(1801年)据影宋本刊刻,全书有图有注,注后有读音,保留了大量古代语言资料,对音韵、语音及艺术研究均有重要参考价值。本文以浙江人民美术出版社影印出版的《尔雅音图》^[33](该书影印清嘉庆六年曾氏刻本)为底本进行数字化、深度标注与数字人文应用的例证研究。

4.1 《尔雅音图》数字化

《尔雅音图》的数字化过程包括书稿扫描及图像处理、文字 OCR 识别、元数据映射组织。

首先完成书稿扫描与图像处理。将《尔雅音图》中的图像、文字、读音、注释等内容按原貌逐页精确扫描处理并存储为图像文件。选择扫描精度,扫描分辨率一般控制在 300—600dpi,扫描方式包括黑白扫描、灰度扫描和彩色扫描方式,保存格式按存储大



小依次为:tif、gif、jpeg、jpg、png。图像扫描处理过程:导入图片→校正处理→二值化处理→图像抹白处理。其中二值化处理可提高图像的压缩比率,缩小文献的存储空间,提高OCR识别率。图像抹白处理可清除二值化处理后的图片可能会产生的黑色区域。

其次对扫描后的图像文字进行OCR识别。例如识别释诂中的第一句:“初哉首基肇兆祖元胎倣音落權輿始也”。先对其进行矩形框选,点击自动

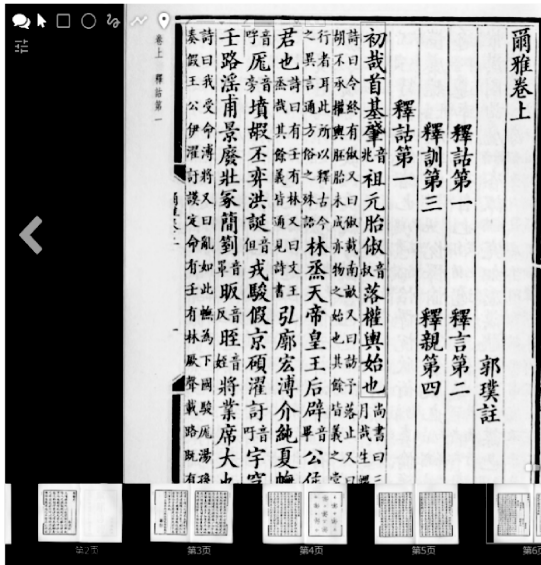


图3 自动OCR识别与JOOCR保存

最后,元数据组织是进行图像核心元素映射与书目数据转化。通过映射DC、CDWA、VRA Core等元数据集:题目(DC:Title)|创建者(DC:Creator)|主题(DC:Subject)|分类(CDWA:Classification)|描述(VRA Core:Classification)|关联(CDWA:Related works)|格式(DC:Format)|分类(CDWA:Classification)日期(DC:Date)|资源类型(VRA Core:Work type)|权限(CDWA:Copyright)|。创建《尔雅》图像元数据核心元素集,依此进行《尔雅音图》实体对象标注。再将标注数据转换成关联数据的RDF格式图数据,存储在三元组数据库中,通过SPARQL进行查询与存取。

4.2 《尔雅音图》深度标注与语义关联

《尔雅音图》的深度标注工具采用自主研发的多维图像智慧系统,该系统是基于IIIF与RDF的沉浸式交互操作平台(<http://www.usources.cn/sas/>),包括藏品检索、藏品中心、集合中心、图像标注、图像管理、SPARQL EDITOR和藏品赏析等7个模块。

OCR,调用外部OCR接口,可将OCR的结果显示在JOOCR部分,再进行人工审核后,保存JOOCR,完成自动文字识别(见图3)。系统可调用百度OCR、书同文OCR、汉王OCR等自动OCR接口,可识别竖版古籍的常用汉字及其符号,对于部分异体字的识别可利用机器深度学习结合众包方式标注完成。将识别结果保存JOOCR,为内容标注和语义关联做好准备。



藏品按照IIIF框架进行组织,每个藏品有一个Manifest清单文件,系统核心功能是图像标注。我们依据典籍图像深度标注层次模型(见图2),利用系统图像标注模块,进行《尔雅音图》的图像标注和语义关联。

第一层,图像轮廓标注。《尔雅音图》中的实体图像包括人、鸟、鱼、兽、畜、草、木、工具、建筑、天、地、山、水等名物实体,这些实体大多带有背景环境,有不同的场景,因此先要根据图像的不同场景进行图像实体的轮廓标注,标注轮廓的工具共有5种:点标、矩形标注、圆形标注、多边形标注、自由标注。《尔雅》插图的实体标注使用较多的是矩形标注和自由标注。一般在标注轮廓较复杂、具有整体性内涵以及都是文字的图像时,使用矩形标注;而在对人、动物等轮廓较为清晰的图像进行标注时,一般使用自由标注,如图4所示。利用图形标注工具,我们完成了《尔雅音图》中630余幅图像场景的轮廓标注。

第二层,实体对象标注。《尔雅音图》中的实体



分类依据《尔雅》19 大类:释诂、释言、释训、释亲、释宫、释器、释乐、释天、释地、释丘、释山、释水、释草、释木、释虫、释鱼、释鸟、释兽、释畜。例如标注一种被称作“𩇛”(别名“乌骃”)的马的实体图像,首先选用多边形标注工具将图像中的马的实体标出来(图 5 左),实体框选出来后,在弹出的标注窗口(图 5

右)上面输入框中标注《尔雅》图像上的文字注解“骊白杂毛𩇛”,以及《尔雅译注》中对这种被称作“𩇛”的马的形象特征描述,在输入框下面的“tag”中填入实体在《尔雅》中所隶属的分类“释畜”“马属”。这样就完成了图像实体的标注。



图 4 图像轮廓标注



图 5 实体对象标注

第三层,深度语义标注。通过构建领域词表可以实现图像深度语义揭示与关联。我们以上海古籍出版社《尔雅译注》^[34](该书是上海古籍出版社邀请名家历经十年完成的简体中文《十三经译注》之一,可帮助读者最大程度读通和理解原著),参考相关词表构建了《尔雅》多语语义词表,内容结构由训释词语、被训释词语、例证 3 部分组成,语种包括中文、英文、日文和韩文。设计了以“词表”“典籍”“句子”“分类”“人”为实体类的《尔雅》词汇知识本体,实体属性关系包括: belongsTo(属于分类)、SubClassOf(子

类关系)、dc:source(句子来源)、name(作者姓名)、address(作者籍贯)、dynasty(作者朝代)等等。通过属性实现了《尔雅》词语释义的语义关联。根据中文释义,进行词汇释义的多语翻译和标注,完成了 3584 个被训释词语和 2219 个训释词语的英、日、韩语的释义翻译。最后通过对《尔雅》词表关联数据转换、存储与发布,实现了语义关联检索和开放链接服务,为《尔雅音图》语义关联做好了准备。

当完成词表构建与发布之后,就可对《尔雅音图》进行内部词表标注与关联,进入标注中心,点击

语义标注按钮,会出现输入关联信息的标注界面,首先输入系统内置的关系属性 URI(<http://www.w3.org/2000/01/rdfschema#seeAlso>);其次,输入关系对象 URI,需要关联到《尔雅》多语语义词表中的训释词,打开词表搜索平台(网址:<http://dh.usources.cn:8080/soopa>)检索训释词语“𩇛”,得到词表中“𩇛”的释义与关联的语义信息网址,在关系对象一栏输入链接关联的网址。进行对象描述时,参照《尔雅译注》中文释义,这样就实现了《尔雅音图》第 286 页图像“骠白杂毛𩇛”与《尔雅》词表内容

的深度语义关联(见图 6)。

利用《尔雅》多语语义词表关联,检索图像“𩇛”时,就可以获得该词在词表中关联的这种马的中、英、日、韩文释义,别称,读音,所属类别以及郭璞《尔雅注》原句“今之乌骊。”邢昺《尔雅疏》原句“毛色黑白而复有杂毛相错者名为𩇛。”《诗经·郑风·大叔于田》例句“叔于田,乘乘𩇛。两服齐首,两骖如手。”《毛诗故训传》注释句“骠白杂毛曰𩇛。”陆德明《经典释文》注句“𩇛,依字作𩇛。”等知识图谱内容(见图 7)。

检索结果 (1):

𩇛 (RDF/XML | JSON-LD | NTriples)
<http://ntu.edu.cn/erya/concept/854c1254694849a995c039220e7ba9ae>
 释义:毛色黑白相杂的马称作“𩇛”。“骠”,黑。“𩇛”或通作“𩇛”。别名“乌骊”。;注释原句:“今之乌骊。”;出处:《尔雅注》(郭璞 晋)

语义标注

【标注内容】骠白杂毛𩇛
 【所在画布】第153页
 【所属藏品】尔雅音图

新增标注

属性/关系	关系对象	对象描述	相关操作
http://www.w3.org/2000/01/rdf-schema#seeAlso 查看定义	http://www.usources.cn:8480/ntu/erya/concept/854c1254694849a995c039220e7ba9ae 查看资源	毛色黑白相杂的马称为𩇛[尔雅词表]	修改 删除

图 6 图像语义关联

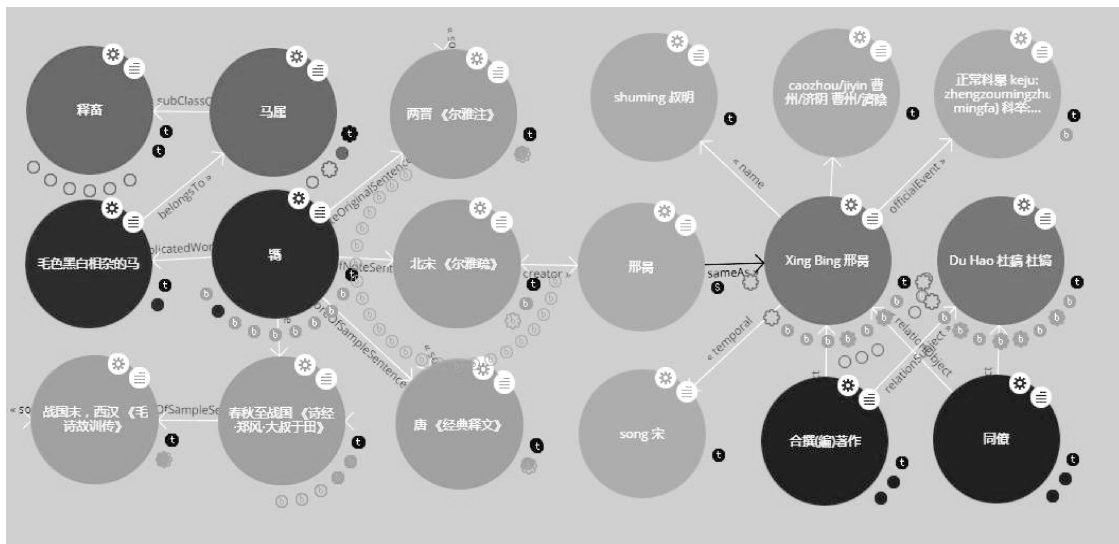


图 7 《尔雅》词表中“𩇛”知识图谱与 CBDB 中《尔雅疏》作者邢昺关联



4.3 《尔雅音图》数字人文应用场景

通过对《尔雅音图》的深度揭示,实现了《尔雅》图文互释、语义标注和关联检索功能,为数字人文应用提供了可以众包参与的交互操作平台,可实现“以图证史”“图像叙事”“多媒介传播”的数字人文应用场景。

首先,“以图证史”应用。利用图像互操作平台,将《尔雅音图》中的插图及文字与相关典籍图像、出土文物图像等相互对比,对考证典籍出处、版本、作者、古代社会生活、经济、军事、典章制度、自然环境等具有重要价值。例如,有学者在对《尔雅》插图研究时,通过高清图的呈现发现《尔雅》“释天—讲武图”中犬的形象和《事林广记》插图中犬的形象非常相似,都有脑袋偏长、身上有条纹、尾巴上卷、颈上佩戴铃铛等特点,在出土元墓壁画中的犬也有同样形态(见图8),据此作为《尔雅》图像有元代渊源的佐证^[35]。《尔雅音图》中的“释草”“释木”还可以和《唐本草》《本草图经》等医药典籍进行对比,再结合文本分析,从时空地多维度考证古代植物特征及其出处。《尔雅音图》中神话人物、动物图像与《山海经》进行对比,考证神话传说的历史地理出处。

语义词表实现了对典籍图像的深度揭示,再通过关联外部词表和其他相关资源,读者可获取更多相关的背景知识,探究作者与典籍知识发现。《尔雅音图》与外部词表进行关联,可通过 SPARQL 联邦检索获取外部数据 API 实现,例如需要了解《尔雅》注疏者的更多背景资料信息时,可以检索并关联中国历代人物传记资料库(CBDB),CBDB API 支持人物 ID 和人名两种查询方式,我们通过人名查询《尔雅疏》作者邢昺,CBDB 可显示与邢昺的所处朝代、籍贯、科举、同僚、合著者、学生、家人及其他相关的人物和作品等背景知识。通过在《尔雅》词表里面加入一条三元组 owl:same As <http://cbdb.library.sh.cn/entity/person/te4pwql98df1pd7s > 关联到 CBDB 的邢昺知识图谱。图7不但显示了古代一种被称作“𩚑”的马的相关注释、注者、作品、分类、注疏原句、例句及注疏等典籍知识图谱,还关联到 CBDB,显示与注疏者邢昺有关的人、时、地、作品等更多背景信息,再进一步关联其他典籍中相关马的高清图像及文本信息,为典籍考证和图像叙事等数字人文应用提供了帮助。



图8 《尔雅音图》与元墓壁画中的犬

其次,图像叙事应用。人类最早的图像符号象形文字就具有叙事属性,国外学者将图像叙事定义为“视觉或绘画表现的文本表现”^[36]。《尔雅音图》中包含了大量反映先秦时期人类生活、自然生态和社会文化等多方面的写实插图,有宫廷建筑、常用器具、工具、乐器、天地四方、物产人物、天象、山川、河流、动物、植物等实体图像 630 余幅,并配有注释文字,对这些珍贵的插图与文字记载,按主题进行深度揭示,建立《尔雅》图像叙事模型,通过图像场景分

层、实体对象标注、情感词语标注和实体关系抽取,建立图像及其背景的文本关联,描述图文混搭的叙事情节,再通过主题关联《诗经》《尚书》《楚辞》《周易》《论语》等相关典籍文献和插图,给读者提供丰富的图像知识及其背景故事情节,在碎片化阅读时代,激发读者阅读典籍的欲望、图像审美和主体参与的探究精神。

最后,典籍多媒介传播。随着语义网络和动态云计算的普及,读者通过多媒介快速访问、传播藏品



图像,为《尔雅音图》创建推特订阅和 RSS 订阅,以吸引更多读者参与,通过语义网聚合引擎和 IIIF 图像互操作标准,将中国典籍与世界各国典藏资源库进行关联和互操作,将翻译、解释、比较、分析等工作流程进行众包,创建作者索引和搜索系统,记录和展示读者对藏品的操作与相关争论,设计移动 APP,使用移动设备标注图像,编辑与呈现《尔雅》图像叙事情节,利用沉浸式多媒介交互平台,创作者、读者、人文学者、留学生等都可以通过网络进行交流互动。还可以利用虚拟现实技术,创设古代自然和人文环境,让人身临其境,学习和体验古代社会生活,突破语言和文化的障碍,促进典籍图像在世界范围的数字人文推广、传播与利用。

5 结论与展望

典籍图像在数字人文中具有以图证史、图像叙事和多媒介传播等作用,国际图像互操作框架(IIIF)和资源描述框架(RDF)促进了典籍图像藏品的共享与利用。本文结合 IIIF 与 RDF,参考图像元数据标准,构建包括图像数字化、深度语义标注和数字人文应用的典籍图像深度揭示与利用的模型架构,实现了图像扫描、OCR 识别与元数据组织的图像数字化过程,提出了典籍图像深度标注的三个层次:(1)利用 IIIF 呈现 API 实现图像的场景标注;(2)通过映射图像元数据核心元素进行典籍实体内容标注;(3)通过构建典籍多语语义词表实现典籍实体概念的深度标注与语义关联。例证采用经典辞书《尔雅》的手绘插图版《尔雅音图》,利用自主研发的基于沉浸式多维图像智慧互操作平台进行《尔雅》图像的数字化,通过构建多语语义词表,实现对《尔雅》图像实体对象的深度语义关联,扩展了对《尔雅》注者、作品等信息的关联。最后探究了《尔雅音图》的数字人文应用场景,包括图文互证、图像叙事和多媒介传播。未来将利用众包工具深度标注更多不同版本的典籍图像,通过领域多语语义词表建立与世界不同馆藏典籍图像、文本的关联;同时开发基于 IIIF 的移动式多媒介典籍传播平台,使分布在世界各地的读者、学者利用移动设备众包参与典籍阅读、解释、标注、分析、发现和研究,典籍文献在公众赏析、解读、研究和讨论的过程中获得新的生命力,真正使书写在古籍里的文字和图像活起来,在典籍图像多媒介传播与利用的过程中,实现中华优秀传统文化

创造性转化与创新性发展^[37]。

参考文献

- 1 夏征农.辞海[M].上海:上海辞书出版社,1989:330.
- 2 颜佳,杨敏,彭梅.面向数字人文的图像数据基础设施建设研究——以我国图博档领域为视角[J].图书馆,2021(5):51-58.
- 3 朱赛虹.清宫殿版图像典籍述略[J].明清论丛,2018(2):393-401.
- 4 W.J.T.米歇尔.图像理论[M].陈永国,胡文征,译.北京:北京大学出版社,2006:7.
- 5 姜智慧.文学图像化语境下中华戏曲典籍对外出版与传播——以昆曲典籍《牡丹亭》为例[J].中国出版,2020(1):63-67.
- 6 曾子明,周知.面向数字人文的图像语义标注模型研究[J].情报理论与实践,2018(1):116-121.
- 7 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].图书情报工作,2020(7):127-135.
- 8 王晓光,谭旭,夏生平.敦煌智慧数据研究与实践[J].数字人文,2020(4):11-23.
- 9 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(2):67-78.
- 10 杨佳莹,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,2020(3):96-102.
- 11 Dunst A, Hartel R, Laubrock J. The Graphic Narrative Corpus (GNC): design, annotation, and analysis for the digital humanities[C]. 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Nov 1, 2017.
- 12 Stork L, Weber A, van den Herik, et al. From handwritten manuscripts to linked data[C]. 22nd International Conference on Theory and Practice of Digital Libraries, Sep 10-13, 2018.
- 13 Cheng X F, Yang J Q, Jiang L, et al. Interpreting and semantically describing Chinese traditional brocade: Xilankapu[J]. Electronic Library, 2019, 37: 474-489.
- 14 Abgaz Y, Souza R R, Japesh M, et al. A methodology for semantic enrichment of cultural heritage images using artificial intelligence technologies[J]. Journal of Imaging, 2021, 7(8):121.
- 15 刘嘉.元数据导论[M].北京:华艺出版社,2002:42-44.
- 16 袁莉,张晓林.数字图像的元数据格式[J].大学图书馆学报,2001(2):27-30,41.
- 17 Powers S. Practical RDF[M]. Published by O'Reilly & Associates, Inc., 2011:1-4.
- 18 Sarah A L. Review: international image interoperability framework(IIIF)[J]. Journal of the American Musicological Society, 2018, 71(2):561-572.
- 19 IIIF Technical details[EB/OL].[2021-07-30].<https://iiif.io/technical-details/#core-api>.
- 20 IIIF Image API3.0.0[EB/OL].[2021-07-30].<https://iiif.io/api/image/3.0/>.
- 21 IIIF Presentation API3.0.0[EB/OL].[2021-07-30].<https://iiif.io/api/presentation/3.0/>.



- 22 IIF Search API API1.0.0[EB/OL].[2021-07-30].https://iiif.io/api/search/1.0/.
- 23 IIF Authentication API1.0.0[EB/OL].[2021-07-30].https://iiif.io/api/auth/1.0/.
- 24 IIF Change Discovery API 1.0.0[EB/OL].[2021-07-30].https://iiif.io/api/discovery/1.0/.
- 25 单蓉蓉,陈涛,刘炜,等.国际图像互操作框架及拓展应用[J].图书馆杂志,2021(5):85-95,123.
- 26 苏东出.一种改进的黑白二值化方法——谈文献扫描图像的数字化处理[J].情报杂志,2003(5),69-70.
- 27 李文英,曹斌,曹春水,等.一种基于深度学习的青铜器铭文识别方法[J].自动化学报,2018(11):10-11.
- 28 Liang J H, Wang H, Li X J. Task design and assignment of full-text generation on mass Chinese historical archives in digital humanities a crowdsourcing approach[J]. Aslib Journal of Information Management, 2020, 72(2):262-286.
- 29 郑巧英,周晨,彭佳.国家图书馆图像资源元数据规范和著录规则[M].北京:国家图书馆出版社,2013.
- 30 顾廷龙,王世伟.尔雅导读[M].成都:巴蜀书社,1990:31.
- 31 郑樵.通志二十略[M].王树民,点校.北京:中华书局,1995.
- 32 周中孚.郑堂读书记补逸[M].北京:中华书局,1993:402-403.
- 33 郭璞.尔雅音图[M].杭州:浙江人民美术出版社,2013.
- 34 胡奇光,方环海.尔雅译注[M].上海:上海古籍出版社,2004.
- 35 吴洪斌.《尔雅》插图研究——兼论中国古代知识传统[J].南京艺术学院学报(美术与设计),2019(2):69-76,209-210.
- 36 James A W H. Ekphrasis and representation[J]. New Literary History, 1991, 22(2):297-316.
- 37 中共中央宣传部.习近平新时代中国特色社会主义思想学习问答[M].北京:学习出版社,2021:318.
- 作者单位:钱智勇、张志美、徐宇红,南通大学图书馆,江苏南通,226019
陈涛,中山大学信息管理学院,广东广州,510006
何书,南通大学文学院,江苏南通,226019
- 收稿日期:2021年10月2日
修回日期:2022年1月26日
- (责任编辑:支娟)

In-depth Disclosure and Utilization of Images in Classics for Digital Humanities

Qian Zhiyong Chen Tao Zhang Zhimei Xu Yuhong He Shu

Abstract: The images in classics are an important resource of the library. The digitization of images and disclosure of content are two essential parts in the basic construction of Digital Humanities, which can realize the applied values such as “proving history with pictures”, “image narration” and multi-media communication. Based on standard specifications such as image metadata, Resource Description Framework (RDF) and International Image Interoperability Framework (IIIF), this article is meant to study the model architecture and implementation methods for the in-depth disclosure and utilization of images in classics. The model consists of three interrelated modules which contain the digitization of images, deep content annotation and Digital Humanities application. From three levels of image scenes, entity objects and semantic associations, it reveals image features, annotations of classics and their background knowledge as well as assists in Digital Humanities research. The illustration uses the classic dictionary *Erya*, the illustrated version of *Erya Yintu* for digitization and in-depth annotation, constructs an *Erya* multilingual semantic dictionary and associates it with the entity objects, also it deeply reveals the knowledge like the ancient and modern paraphrase, phonology, classification, annotation of the classics, examples and allusions of the image entities of *Erya*, and then explores the Digital Humanistic application scenarios of *Erya* images.

Keywords: Images in Classics, Digital Humanities, *Erya Yintu*, In-depth Disclosure, IIIF, Linked Data