

## 数字人文中图像资源的语义化标注研究

陈 涛<sup>1\*</sup>, 单蓉蓉<sup>2</sup>, 李 惠<sup>1,3</sup>

(1. 上海图书馆 (上海科学技术情报研究所), 上海 201100;

2. 上海大学 图书情报档案系, 上海 200444; 3. 南京大学 信息管理学院, 南京 210046)

**摘 要:** [目的 / 意义] 数字人文领域存在的大量图像资源长期以来一直没有得到有效的利用, 图像孤岛现象愈演愈烈。国际图像互操作框架 (IIIF) 为图像资源的深度交互和应用打开了新的窗口, 而图像标注正是打通图像孤岛、提升图像认知的关键。[方法 / 过程] 论文基于 IIIF 框架提出了“图像层面的元数据注入”“对象层面的内容录入”和“语义层面的语义引入”的 3 层标注模型。同时, 文中构建的 本体管理平台为图像资源的语义化标注提供了本体上的支持。[结果 / 结论] 以韩干《照夜白图》为例, 探讨了使用图像标注和 IIIF 技术在图像内容语义增强、对象标签聚类 and 资源精准关联等方面所带来的新的交互体验和研究模式。可以说, IIIF、关联数据 (Linked Data) 和人工智能 (AI) 的结合必将开启数字人文研究的新时代。

**关键词:** 国际图像互操作框架 (IIIF); 语义化标注; 本体; 关联数据; 数字人文

**中图分类号:** G250; TP393

**文献标识码:** A

**文章编号:** 1002-1248 (2020) 09-0006-09

**引用本文:** 陈涛, 单蓉蓉, 李惠. 数字人文中图像资源的语义化标注研究[J]. 农业图书情报学报, 2020, 32(9): 6-14.

## Semantic Annotation of Image Resources in Digital Humanities

CHEN Tao<sup>1\*</sup>, SHAN Rongrong<sup>2</sup>, LI Hui<sup>1,3</sup>

(1. Shanghai Library/Institute of Scientific & Technical Information of Shanghai, Shanghai 201100;

2. Department of Library, Information and Archives, Shanghai University, Shanghai 200444;

3. Department of Information Management, Nanjing University, Nanjing 210046)

**Abstract:** [Purpose/Significance] A large number of image resources in digital humanities have not been in effective use for a long time, and the phenomenon of "isolated image islands" is intensifying. The International Image Interoperability Framework (IIIF) opens a new window for more in-depth interactions with image applications. Image

收稿日期: 2020-01-08

基金项目: 国家社会科学基金项目“数字人文中图像文本资源的语义化建设与开放图谱构建研究”(19BTQ024)

作者简介: 陈涛 (1981-), 男, 工学博士, 高级工程师, 研究方向为知识图谱、关联数据, E-mail: tchen@libnet.sh.cn。单蓉蓉 (1987-), 女, 博士研究生, 研究方向为数字人文。李惠 (1987-), 女, 理学博士, 研究方向为数字人文、社会网络

annotation is the key to breaking image islands and enhancing image recognition. [Method/Process] Based on IIIF framework, this paper proposes three dimensions of image annotation: image-level metadata injection, object-level content transcription, and semantic-level deep annotation connection. At the same time, the ontology management platform built in the paper provides ontology support for semantic annotation of image resources. [Results/Conclusions] This paper takes Night-Shining White, a monochrome ink-on-paper painting by Han Gan, as an example. In this case, the paper discusses the new interactive experience and research mode brought by the use of image annotation and IIIF technology in semantic enhancement of images, object tag clustering and precise associations of resources. It indicates that the combination of IIIF, linked data and artificial intelligence (AI) is bound to open a new era of digital humanities research.

**Keywords:** international image interoperability framework (IIIF); semantic annotation; ontology; linked data; digital humanities

## 1 引言

数字人文, 正从一个极为专门的学术领域成为一种学术时尚; 从强调文化记录的产生、保存、解释和呈现形式, 到对人文学科研究方法的颠覆。大数据、云计算、人工智能迅猛发展, 影响着人类生活的方方面面, 也不可避免地影响着人文学科研究、生产、传播的方式<sup>[1]</sup>。数字信息技术对人文学科研究的资源类型和研究方法产生了巨大影响, 数字人文正是数字技术与人文学科交叉融合形成的跨学科研究领域。数字人文如今已经成为一个活跃的研究领域, 研究显示有越来越多的机构和学者参与到这个领域的研究中来<sup>[2]</sup>。

从古代卷轴到现代报纸, 从欧洲中世纪手稿到中国书法手抄版, 这些图像资源大量存在于图书馆、博物馆、档案馆等各大馆藏机构中。相比格式化资源的普遍利用, 图像资源(尤其是超大超清图像)长期处于信息孤岛状态, 图像资源本身以及图像之间所承载的研究和应用价值没有得到有效体现。2015年6月, 由大英图书馆、牛津大学图书馆、哈佛大学等29个非营利图像资源存储机构共同提出国际图像互操作框架(International Image Interoperability Framework, IIIF), 旨在保证全球图像存储的互操作性和可获取性<sup>[3-5]</sup>。从技术实现角度来看, IIIF框架完全基于关联数据的理念来实现图像资源的共享和交互。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框架将研究对象提升到图像资源的角度。

在国外, IIIF已大量应用于艺术馆、图书馆、档案馆、博物馆等文化遗产机构中<sup>[6,7]</sup>, Getty博物馆、V&A博物馆、大英图书馆、威尔士国家图书馆、Europeana欧洲数字图书馆、美国华盛顿国家艺术画廊等都采用IIIF框架来发布馆藏资源。2019年5月29日发布的“在瑞士文化遗产机构中部署IIIF的建议措施”白皮书<sup>[8]</sup>更是将IIIF提升到了国家基础设施建设的层面, 目前瑞士虚拟手稿图书馆、瑞士图书馆数字化稀有图书平台、数字化报纸收藏、珍贵书本等资源均采用了IIIF标准。日本从2017年开始也逐渐在国内多个机构中推行IIIF框架, 目前已有日本国会图书馆、人文数据开放中心(CODH)、京都大学、关西大学等20多家机构在使用IIIF进行资源的展示, 并且规模仍在不断扩大<sup>[9]</sup>。国内方面, 近两年也陆续出现IIIF相关的研究论文, 耿曼曼<sup>[10]</sup>、张轶<sup>[11]</sup>分别就技术框架和技术特点等方面对IIIF进行了梳理和分析; 付跃安提出了IIIF在数字资源集成中的应用设想<sup>[12]</sup>。应用方面, 上海图书馆、复旦大学、华东师范大学都纷纷尝试使用IIIF来进行馆藏资源的呈现。

目前网络中以IIIF标准发布的图像资源多达数亿之多, 这些资源多用于呈现为主, 然而图像的研究与深度应用少不了图像的标注。标注工具方面, Pundit可以用来对网页内容和图片进行语义标注<sup>[13]</sup>; Recogito可与

IIIF 结合实现地理方面的文本和图像的注释和可视化, 并获得 2018 年数字人文最佳工具奖<sup>[4]</sup>。标注方法方面, LAURA 使用了本体进行图像标注<sup>[5]</sup>; ZHENG 等试用了 CNN 方法进行图像语义标注<sup>[6]</sup>; LISE 提出了适用于构造自然历史档案馆藏中命名实体的语义模型<sup>[7]</sup>。模型研究方面, PIERRE 等从注释的结构复杂度、所用词表的类型和支持的协作类型 3 个角度分析了现有的注释模型的不同特征和功能<sup>[8]</sup>; 王晓光等指出数字图像的语义描述与标注是解决图像检索中语义鸿沟问题的关键<sup>[9]</sup>, 并对敦煌壁画图像中包含的主题进行深度语义标注<sup>[20]</sup>; 同时王晓光等提出了叙事型图像语义标注模型<sup>[21]</sup>; 陈金菊等则指出图像语义标注的基础是图像语义标注模型的构建, 并对主流图像语义标注模型进行了梳理和总结<sup>[22]</sup>。本文借助 IIIF、关联数据、本体和 Web 注释数据模型等标准和规范, 研究了构建图像语义化标注模型和方法, 并探讨其给数字人文所带来的新的交互体验和研究模式。

## 2 图像标注模型研究

根据图像资源的特点, 可从图像、对象及语义 3 个层面对数字人文图像资源进行逐层标注 (图 1)。3 层标注模型之间难度逐层加大, 图像层面从整幅图像进行标注; 对象层面则对应到图像上的具体对象 (文字、事物、人物等); 语义层面强调对象与外部资源的关联和整合。

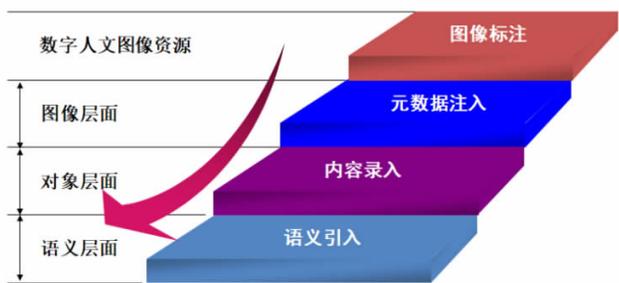


图 1 图像标注模型层级分布

### 2.1 图像层面元数据注入

第一层标注主要用来将一些格式化的元数据信息

添加至整幅图像中, 该层通常通过机器来自动完成。IIIF 中 Presentation API 中制定了一些属性, 如图像题名 (dct:title)、描述信息 (dc:description)、许可声明 (dct:rights)、归属信息 (iiif:attributionLabel) 等。IIIF 中指定的元数据属性有限, 实际使用时可根据需要将图像的更多属性添加至 metadata 部分中, 如作者、出版地、备注等信息, 放入 metadata 中的信息主要为 label-value 对, 其中 value 值可以设置成多语言。

### 2.2 对象层面内容录入

第二层标注用来对图像中具体的内容进行抄录, 数字化的众多图像资源通过 OCR 方式进行文本化以及众包模式进行手稿等资源的抄录都可以看成是对象层面的抄录, 只是针对的对象为文本内容。然而在图像资源中, 除了文本内容, 还有一些图像方面的非文本对象, 此时可以通过众包方式。

对象层面的内容标注模型如图 2 所示, 该模型基于 Web 注释数据模型 (Web Annotation Data Model), 该数据模型提供了一个可扩展的、互操作的框架, 可用于图像内容标注, 以使注释可以在不同的硬件和软件平台之间轻松共享和重用。该模型最为核心的有 anno、body 和 target 三个类, 其中 anno 表示标注的每个对象资源, 每个对象资源都将以独立的 Graph 存在; body 表示具体的标注内容资源; target 则表示标注的

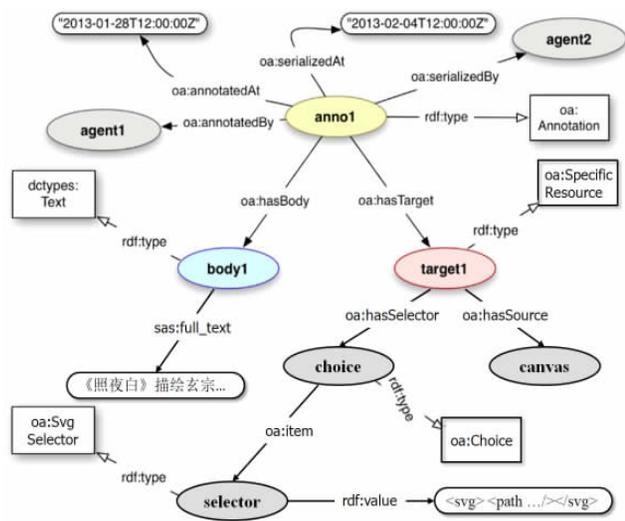


图 2 图像资源内容标注模型

区域资源, 该资源将通过 `oa:hasSource` 属性与具体的 `canvas` (图像页) 进行关联, 而 `canvas` 又存储在相应的 `manifest` (清单) 中。具体的标注区域将以 SVG (可缩放矢量图: Scalable Vector Graphics) 格式存储, 便于在呈现时进行精准的对象定位。

## 2.3 语义层面语义引入

第三层的语义标注则体现了关联数据和 IIF 结合的优势, 主要在第二层标注的基础上使用本体属性将标注的对象与开放的关联数据集进行关联。如何发现不同数据集中相同或者相似的资源, 对于关联数据而言, 可以采用 SPARQL 中的联邦检索来进行不同数据集中资源的发现。使用语义标注的好处是, 当建立了与目标数据集的关联后, 可获取到该数据集中相关资源的所有信息, 尤其是资源的对外关联信息。这里以“韩幹”为例建立不同资源节点之间的关联图谱, 图 3 显示将 DBpedia 知识库中“韩幹”资源关联到 wikidata (维基数据)、freebase (链接数据库)、yago (链接数据库)、dnb (德国国家图书馆) 等外部资源, 同时 wikidata 中又关联到了 worldcat (联机联合目录数据库)、wikiart (视觉艺术百科全书)、viaf (虚拟国际规范文档)、isni (国际标准名称标识符)、loc (美国国会图书馆) 等资源。由此可见, 仅需将资源关联到图谱中的一个节点, 即可获取到其他资源节点的丰富知识。

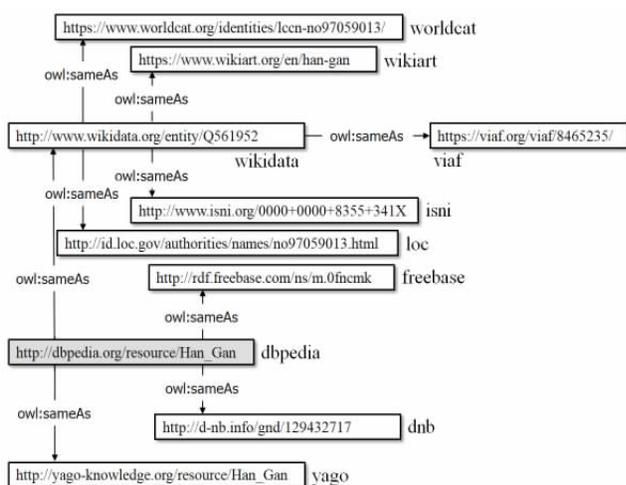


图3 “韩幹”资源关联图谱

## 3 语义知识 (本体) 管理平台构建

从图像语义标注可以看出, 要实现图像对象与外部资源的语义关联, 离不开本体和属性的构建。常用的语义关联属性有 `owl:sameAs`、`rdfs:seeAlso`、`skos:closeMatch` 等, 甚至所有的对象属性都可作为资源之间关联的桥梁, 因此语义知识 (本体) 管理平台的构建显得至关重要。目前, 本体管理平台中含有 foaf、bibframe、schema、skos、dbpedia 等常用的知识本体, 以及上海图书馆、华东师范大学等机构发布的与数字人文相关的知识本体。

### 3.1 本体存储机制设计

本体存储是整个本体管理平台底层数据结构的持久化方案, 整个存储机制设计方案如图 4 所示。设计时主要考虑以下两点。

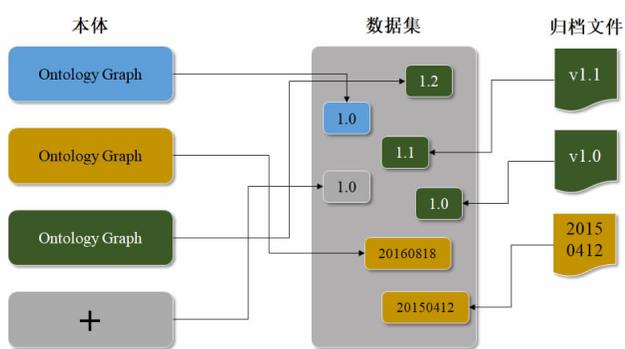


图4 本体存储机制设计

#### 3.1.1 不同本体之间的存储机制

不同本体之间的存储机制遵循既关联又独立的设计原则: 一方面, 关联是为了将来的本体检索考虑, 不同本体很可能属于某同一发布机构或同一研究领域, 这时不同本体之间需要联合操作。另一方面, 独立是为了将来本体的操作考虑, 对于某本体的操作需要完全独立于其他本体。设计时, 采用了 1+N 的存储机制, 其中 1 表示将不同本体的元数据描述信息存储于统一的数据集图 (Graph) 中, 以便管理不同本体以及本体的不同版本信息; N 表示不同的本体图 (Graph), 每个本体需要独立存储于不同的 Graph 中, 以保证本体

操作的便捷性和安全性。

### 3.1.2 本体不同版本之间的存储机制

每个本体 Graph 中都将保存该本体的最新版本，当本体升级到某个新版本后，会将该本体对应的 Graph 中本体数据进行保存，形成归档文件。归档后将本体的新版本重新写入本体对应的 Graph 中，而数据集 Graph 中将会保留历史版本的元数据信息，以便用来跟踪本体的不同版本信息。

## 3.2 本体管理平台框架设计

本体管理平台框架如图 5 所示，主要包括本体查询和本体校验两大中心，以及本体查询、本体导出、本体发布 / 浏览和本体归档四大模块。其中两大中心分别对应于开放数据模型中的第六星和第七星模型<sup>[2]</sup>，本体查询中心托管数据集的本体发布，并针对本体提供多种交互模块；本体校验中心用来校验本体在数据集中的应用状态。

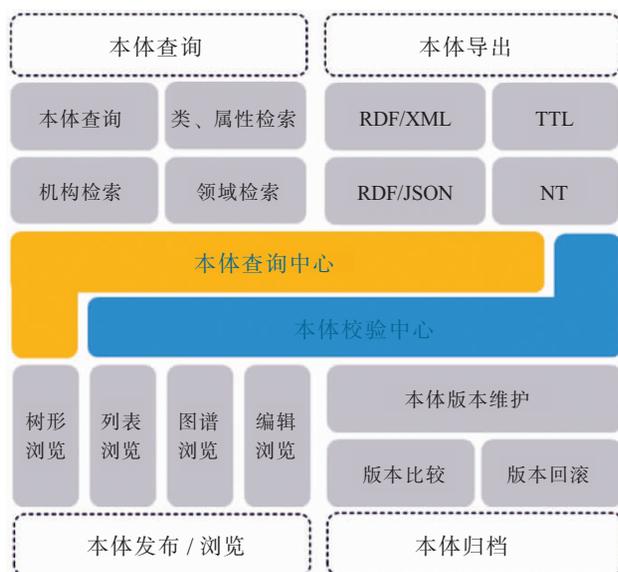


图 5 本体管理中心功能框架

### 3.2.1 本体查询

本体管理平台提供多种维度的本体查询，可以依据本体、本体中类和属性、本体发布机构和本体所属领域进行检索。

### 3.2.2 本体导出

提供 RDF/XML、TTL、RDF/JSON (JSONLD) 和

NT 4 种格式的本体数据导出，4 种格式无需事先生成，而是根据不同格式的实际需求实时生成本体格式。在本体存储机制设计时本体已经独立存储于各自本体图 (Graph) 中，可以方便地对每个本体进行格式生成和导出操作。

### 3.2.3 本体发布 / 浏览

本体发布和浏览是图像语义标注中，选择关联属性时最直接的交互方式。本体管理平台中的本体将会自动生成诸如树形浏览、列表浏览、图谱浏览和编辑浏览等多种本体浏览方式。树形浏览主要提供本体中类之间的上下层级关系，便于了解类之间的继承和复用关系；列表浏览将以表格形式列出所有类以及所有属性，便于快速查找具体的类和属性；图谱浏览提供图谱方式展示类和属性，以及类之间的关联关系，便于关联者理解本体设计理念；编辑浏览是在图谱浏览的基础上，通过图形拖拉操作，实现本体的在线可视化编辑。

### 3.2.4 本体归档

本体归档主要用来记录本体的衍变过程，并追踪本体的历史版本，归档后的本体以静态文件的形式在磁盘中持久化保存。当需要对某个版本进行回滚时，会将该本体重新更新到本体 Graph 中。

本体管理平台为图像资源的语义化标注提供了强有力的本体支持，通过属性的关联，图像与图像之间、图像与资源之间，建立起丰富的语义知识，打破图像孤岛，形成图像智联。

## 4 图像标注案例分析

图像标注采用自主研发的 IIIF-IIP 沉浸式交互平台 (<http://www.usources.cn/sas/list.jsp>) 实现，该平台实现了数字人文图像资源的在线分享、复用、标注、赏析等功能。目前平台中已包含数万幅图像资源，并对其部分资源做了公开发布和内容标注。这里以现藏于美国大都会博物馆唐代韩幹的《照夜白图》为例，“方寸之间，一眼千年”，照夜白图中包含了唐、宋、元、明、清等朝代中多位历史文人和巨匠的题跋和印

章, 对这些信息进行标注后, 可以为数字人文的图像资源带来的新的交互体验和研究模式。这里从图像内容语义增强、全文检索与对象聚类、图像资源精准关联 3 个角度进行探讨。

#### 4.1 图像内容语义增强

图 6 为《照夜白图》中“照夜白”的标注示例, 标注功能采用了文中提出的 3 层标注模型, 从右往左分别对应标注模型的 3 个层面: ①最右侧为第一层标注(元数据注入), 在“展品信息”部分展现了整幅图像元数据注入情况, 包括题名、作者、创作时间、描述、分类、许可声明等字段信息, 利于对图像基本信息有总体的认知。②中间区域为第二层标注(内容录入), 目前通过人工众包方式进行对象标注。除了标注“照夜白”外, 还标注了所有的题跋、印章信息。基于这些标注信息, 可以实现图像内容的检索与发现。③左侧部分为第三层标注(语义引入), 这里显示了第二层标注的具体内容, 同时显示了标注对象与外部资源的关联情况。这里将“照夜白”和 DBPedia、CBDB、SinoPedia 中的“韩干”资源进行了语义关联, 查看关联资源可以获取该关联资源相关信息, 以此达到语义增强的目的。

#### 4.2 全文检索与对象聚类

对于有很多标注内容的图像(或图册), 如果能像 Word 文档或者 PDF 资源一样进行内容检索, 将会在很大程度上提升图像资源的使用率和改善交互体验。图像资源本身并不能直接进行检索, 但是从图像资源内容标注模型中可以看出, 每个标注的内容独立存在, 这给图像内容的全文检索提供了解决方案。进行全文检索时, 将会使用 SPARQL 在该资源所有标注内容中进行检索, 并将检索结果中的标注文本、标注区域、所在图像页等相关信息进行返回。图 7 中显示了在《照夜白图》中进行全文检索的示例, 这里检索了所有题跋中含有“照夜白”的记录, 共有乾隆、危素、曾鲁、沈德潜等人所做题、跋。同时浏览检索结果, 可以在图像资源中进行精准定位。

对图像中的对象资源进行标注时, 除了抄录对象的注释信息, 还可为标注对象添加多个文本标签(Tag)信息。通过标签, 可以实现内容的聚类显示, 如图 8 显示了《照夜白图》中的所有题跋信息, 这里将 22 条题跋记录一并加载和展示。研究者在进行图像资源研究时, 可以对整副图像资源中的标注内容有整体的了解, 迅速地获取和解读图像想传达的深层含义。



图 6 《照夜白图》3 层标注片段

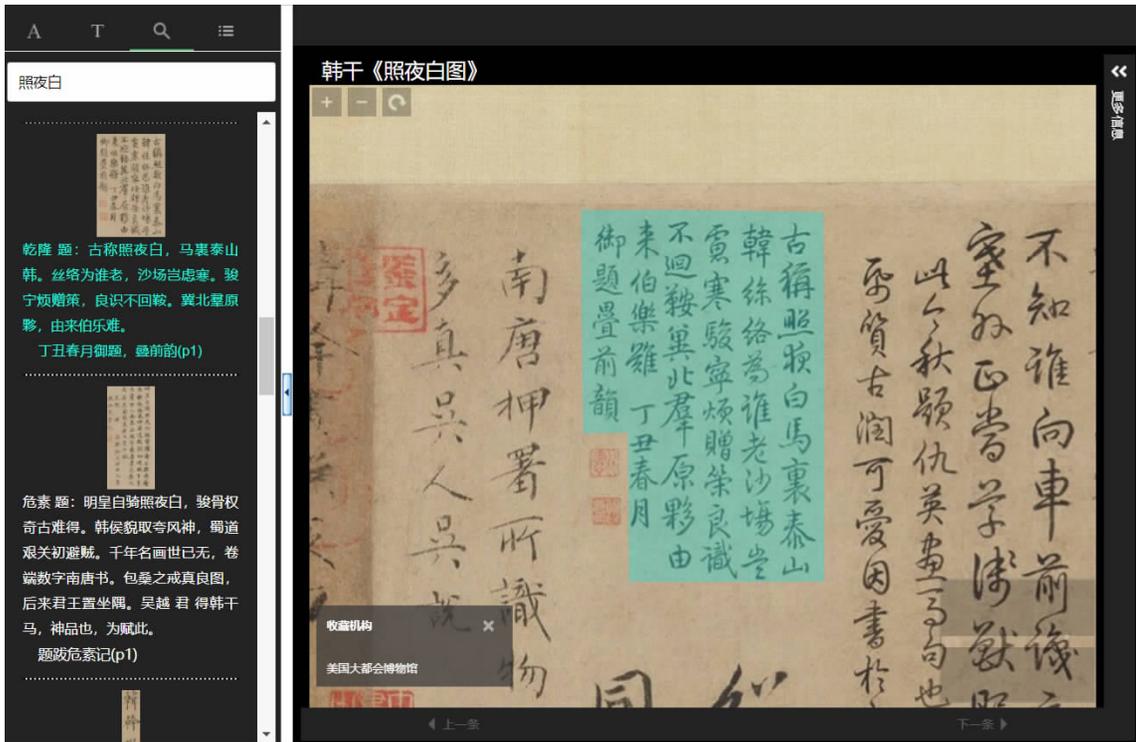


图7 “照夜白”全文检索



图8 《照夜白图》题跋标注

这对数字人文图像资源，尤其是文化遗产方面的图像资源传播起到很好的推动作用。

### 4.3 图像资源精准关联

以往在进行图像关联的时候，都是在两个图像之间建立起超链接，而不能对图像中的具体对象进行关联。对于较小的图像资源，研究者可能会快速找到所需的关联对象，而对于超大尺寸的图像资源，研究者往往很难根据图像超链接而快速查询到所需对象。这里以《照夜白图》中乾隆印章“三希堂精鉴玺”为例，表1为不同资源中“三希堂精鉴玺”印章分布，由于扫描的精度不同，印章尺寸在像素上有所差异。“印

章/图像尺寸比”为该印章在整幅图像资源中的比例大小，《照夜白图》中该印章的尺寸占比为万分之六，而在《清院本清明上河图》中，仅为万分之二左右。

对于如此巨大尺寸的图像，如何快速从一幅图像中的这枚印章迅速找到其在其他图像中的位置，这给图像的应用提出了挑战。采用 IIF 和 Web 数据注释模型，可以方便地建立起不同图像资源相同对象之间的关联，图9显示了《照夜白图》《清明上河图》和《清院本清明上河图》三者之间“三希堂精鉴玺”印章的对象关联。通过《照夜白图》中该方印与其他两幅图像的关联，研究者可以迅速定位到该方印在相关图像资源中的精确位置。这将产生新的应用模式，可根

表1 “三希堂精鉴玺”印章所在图像

图像	图像尺寸/像素	印章	印章尺寸/像素	印章/图像尺寸比
照夜白图	2 000*1 393 (2 786 万)	三希堂精鉴玺	105*160 (1.68 万)	6/10 000
清明上河图	132 778*5 000 (6.639 亿)	三希堂精鉴玺	480*906 (43.488 万)	6.5/10 000
清院本清明上河图	56 531*1 700 (9 610 万)	三希堂精鉴玺	103*189 (1.95 万)	2/10 000



图9 不同图像资源中“三希堂精鉴玺”对象关联

据不同的对象资源建立对象库,如根据书画资源中的印章建立印章库,根据人物头像建立头像库等,访问印章或头像时将会呈现该对象所对应的原始图像及对象所在具体区域。

## 5 结论与展望

数字人文中存在的大量图像资源长期以来一直没有得到有效的利用,IIIF的提出很好地解决了图像资源的共享与互操作,已成为全球文化遗产机构争相追逐和遵循的标准规范。论文基于IIIF技术,提出了图像层面元数据注入、对象层面内容抄录和语义层面语义引入的图像资源3层标注模型。①元数据一般已经存在于各个馆藏机构中,因此可以方便地依据IIIF中呈现API的要求进行元数据的注入,将元数据的信息呈现在整副图像或整套图像中。②对象层面的内容抄录常结合OCR技术和众包模式进行内容的注入和加工,通过抄录可以实现图像中对象资源的细粒度和碎片化,该层的标注常用所见即所录方式,即页面中看到什么标注什么。③语义层面的深度标注则需要借助于本体技术和关联数据技术,将图像中的对象与外部开放资源进行关联,实现对象内容的语义增强和知识扩

充。文中提出的本体服务中心,可以很好地为图像资源的语义标注提供本体支持。

论文的案例部分以韩干《照夜白图》为例,探索了使用IIIF和关联数据技术所能带来的潜在应用场景和深度交互的体验模式。目前的图像标注部分还过多依赖于众包和人工模式,下一步将会思考如何结合云计算、机器学习、图像处理等技术来实现图像对象的自动识别和语义标注。①对象自动识别方面主要针对图像中具有明显结构特征的对象,比如印章、人物头像、虫鸟、花卉等对象,可以使用深度学习进行对象区域的自动化识别和提取(Object Detection),并可自动对提取的区域添加标签<sup>[24,25]</sup>。目前,Google、微软、亚马逊等都提供了图像自动标注接口<sup>[26]</sup>。②语义标注方面很难实现完全的自动化,通过对象识别可以进行对象分类标记,但很难识别出对象的深层语义,如可以识别对象是人或者花卉,但是具体是谁或者是什么花卉,在很多图像资源中很难完成。但是,结合开放的关联数据集,在人工标注时可进行资源的自动化关联,如人、地、事件等资源。可以说,IIIF、关联数据、AI三者的有效结合将成为数字人文研究的新热点,也必将开启数字人文中图像资源研究和应用的新时代。

参考文献:

- [1] 马林青, 韩若画(译). 数字人文: 改变知识创新与分享的游戏规则[M]. 北京: 中国人民大学出版社, 2018.
- [2] 大卫·M 贝里, 安德斯·费格约德. 数字人文: 数字时代的知识与批判[M]. 王晓光, 译. 大连: 东北财经大学出版社, 2019.
- [3] ALBERTO S. International image interoperability framework(IIIF): A panoramic view[J]. JLIIS. it, 2017, 8(1): 50-66.
- [4] STUART S, ROBERT S, TOM C. The international image interoperability framework (IIIF): A community & technology approach for web-based images[J]. Society for imaging science and technology, 2015, 8: 16-21.
- [5] SARAH A L. Review: International image interoperability framework (IIIF)[J]. Journal of the American musicological society, 2018, 71(2): 561-572.
- [6] ANDY C. IIIF: Unshackle your images[C]. MW2016: Museums and the Web 2016, Los Angeles, USA, 2016.
- [7] TRISTAN R, ROB L, JEFFREY C S C, et al. IIIF at Scale[EB/OL]. [2020-04-04]. <https://mw20.museweb.net/paper/iiif-at-scale/>.
- [8] JULIEN A R. Suggested measures for deploying IIIF in swiss cultural heritage institutions[EB/OL]. [2020-04-05]. [https://www.researchgate.net/publication/340966631\\_Swiss\\_institutions\\_climbing\\_up\\_the\\_IIIF\\_ladder](https://www.researchgate.net/publication/340966631_Swiss_institutions_climbing_up_the_IIIF_ladder).
- [9] IIIF collections in Japan[EB/OL]. [2020-04-04]. <https://www.kanzaki.com/works/2016/pub/image-annotator?u=https://nakamura196.github.io/iiif/data/collection/collection.json>.
- [10] 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.
- [11] 张轶. 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.
- [12] 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, 4: 159-166.
- [13] GRASSI M, MORBIDONI C, NUCCI M, et al. Pundit: Creating, exploring and consuming semantic annotations[C]. In proceedings of the 3rd international workshop on semantic digital archives, Valletta, Malta, 2013.
- [14] SIMON R, BARKER E, ISAKSEN L, et al. Linked data annotation without the pointy brackets: Introducing recogito[J]. Journal of map & geography libraries: Advances in geospatial information, collections & archives, 2017, 13(1): 111-132.
- [15] LAURA H, GUUS S, JAN W, et al. Semantic annotation of image collections[EB/OL]. [2020-05-10]. <https://www.cs.vu.nl/~guus/papers/Hollink03b.pdf>.
- [16] ZHENG Y Z, LI Z X, ZHANG C L. A hybrid architecture based on CNN for image semantic annotation[C]. 9th international conference on intelligent information processing (IIP), Melbourne, Australia. 2016,11: 81-90.
- [17] STORK L, WEBER A, GASSZÓM E, et al. Semantic annotation of natural history collections[J]. Web semantics, 2019, 59: 100462.
- [18] PIERRE A, ILYA Z, JUAN P. A Classification of Semantic Annotation Systems[J]. Semantic Web, 2012, 3: 223-248.
- [19] 王晓光, 徐雷, 李纲. 敦煌壁画数字图像语义描述方法研究[J]. 中国图书馆学报, 2014, 40(1): 50-59.
- [20] WANG X G, SONG N Y, ZHANG L, et al. Understanding subjects contained in Dunhuang mural images for deep semantic annotation[J]. Journal of documentation, 2018, 74(2): 333-353.
- [21] 徐雷, 王晓光. 叙事型图像语义标注模型研究[J]. 中国图书馆学报, 2017, 43(5): 70-83.
- [22] 陈金菊, 欧石燕. 数字图像语义标注模型比较与分析[J]. 图书情报工作, 2018, 62(6): 116-124.
- [23] 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.
- [24] ADRIAN R. YOLO object detection with OpenCV[EB/OL]. [2020-04-18]. <https://www.pyimagesearch.com/2018/11/12/yolo-object-detection-with-opencv/>.
- [25] GONZALO V-C, NIELS D L, MUBARAK S. Holistic object detection and image understanding[J]. Computer vision and image understanding, 2019, 181: 1-13.
- [26] FILESTACK. Comparing image tagging services: Google vision, Microsoft cognitive services, amazon rekognition and clarifai[EB/OL]. [2020-05-20]. <https://blog.filestack.com/thoughts-and-knowledge/comparing-google-vision-microsoft-cognitive-amazon-rekognition-clarifai/>.