

由中国人民大学信息资源管理学院冯惠玲教授、刘越男教授、严承希博士和哲学院王小伟副教授主讲的《数字人文导论》是中国人民大学本科生通识核心课程之一，面向全校本科生开设。课程内容包括数字人文的基本概念、历史源流、发展脉络、基本方法和技术、前沿议题等。本系列推文是该课程学生的学习成果展示。

本期作者

薛涵 明德书院2021级

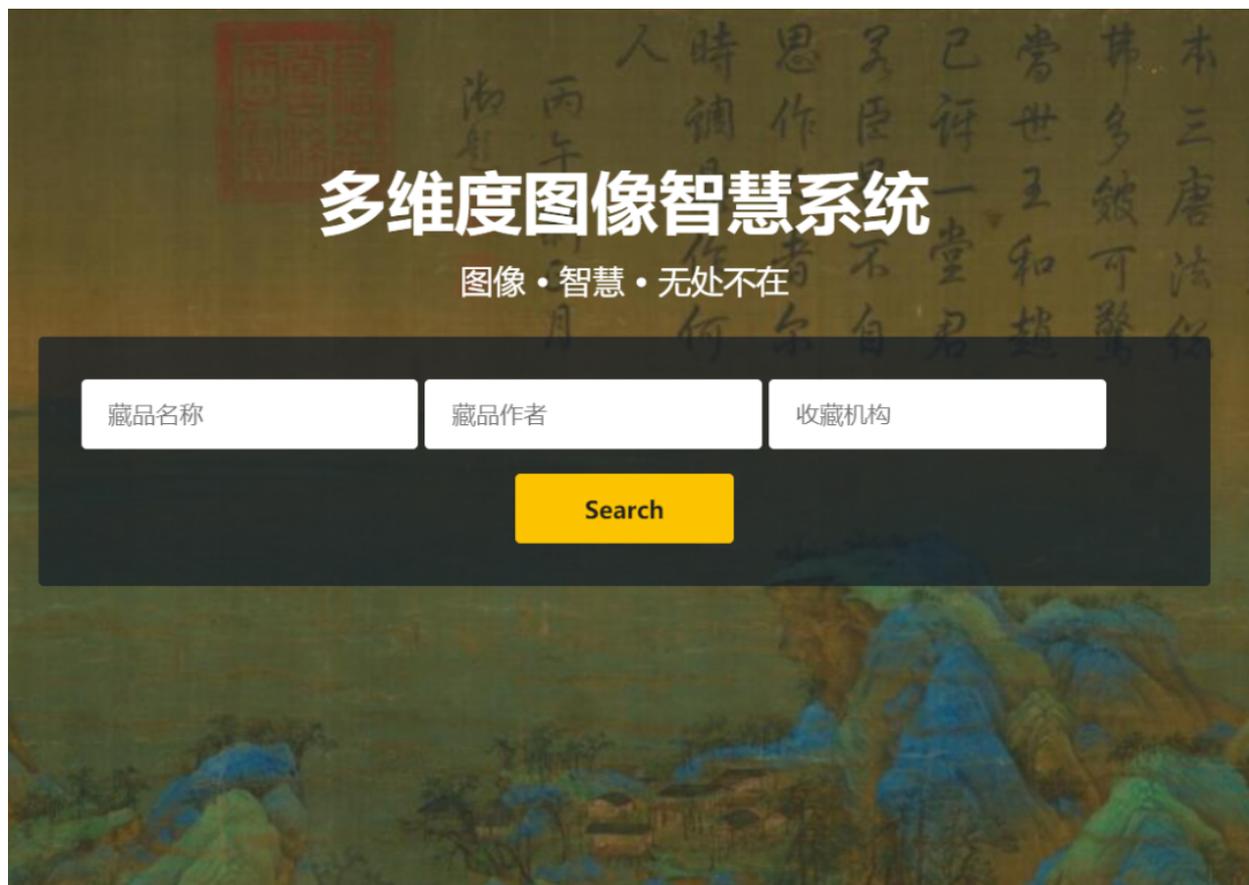
赵振真 新闻学院2021级

魏梦凡 国际关系学院2020级

翟宇晴 艺术学院2019级

“多维度图像智慧系统”项目案例评介

案例引入



“多维度图像智慧系统项目”（项目网址：<http://miss.newwenke.com/sas/>，曾获2020年数字人文奖“最佳数字人文工具奖”提名奖），“多维图像智能系统”是一个展示、分享、浏览、检索藏品的开放式平台。用户可以通过“藏品名称”“藏品作者”“藏品机构”和“藏品分类”实现多维度检索。该平台基于LIBRA技术理论，可实现古籍、文物、藏品等图像资源的在线复用、文字识别、语义标注、特征提取和光谱分析等功

能，将成为数字人文和新文科建设中图像资源的深度组织和研究平台，有助于文化遗产知识的传播传承和研究利用。

项目背景

01研究背景

长期以来，图像资源由于物理对象衰变、信息封闭、语义化程度低、资源重复建设且利用率低及资源交互和深度应用实现困难等原因，形成了图像信息孤岛和价值洼地现象。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）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。而图像文本资源的语义化组织与建设将丰富数字图书馆元数据建设体系的研究。

国内高校和研究机构纷纷使用国际图像互操作框架来呈现古代书画，但基于题跋标注的知识发现仍为数不多。国内书画题跋类文字的内在联系和书画的传承轨迹研究较少。通过题跋分析书画的传承轨迹，既可纵览中国书画史，也可以深入到个案，对于推进传统研究有相当的意义。

02团队介绍

【项目发布机构】

华东师范大学数据与调查中心

上海市多维度信息处理重点实验室

【项目重点负责人】

陈涛：上海图书馆、南京大学联合培养博士后，现任中山大学信息管理学院副教授

关键技术

01两大相关技术



●国际图像互操作框架

International Image Interoperability Framework

IIIF是2015年成立的，目的是为了确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、档案资料等在线资源进行统一展示和使用。

IIIF 定义了一系列应用程序接口（Application Programming Interfaces，简称 API），图像，呈现，搜索，验证四类API。

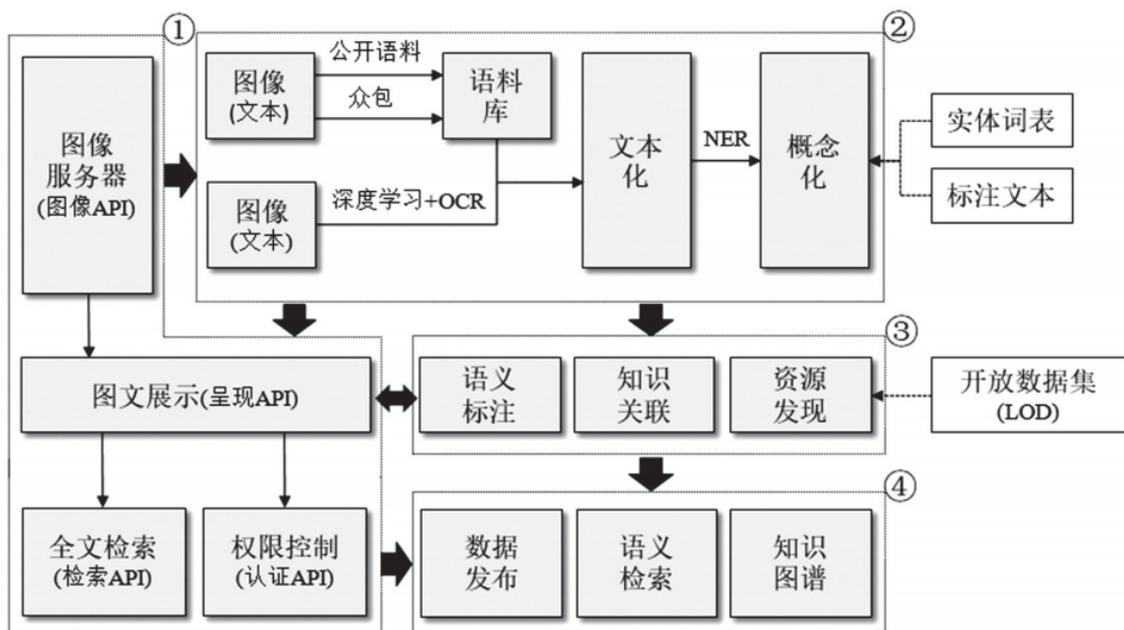


图1 图像资源语义化框架

从这幅图来看IIIF提供的四类API可以贯穿图像资源语义化建设也就是本项目研究的始终。

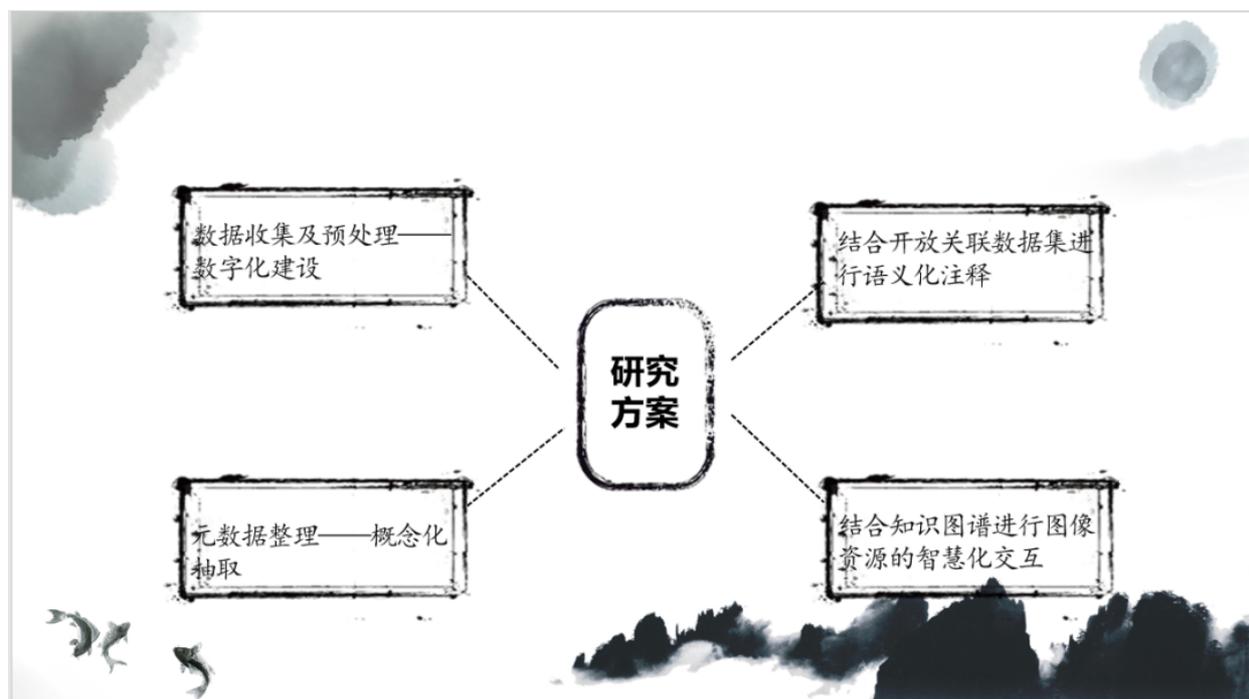
●关联数据

关联数据（Linked Data）是一种简单的语义网实现技术，价值在于通过资源描述框架数据模型，将网络上多数据节点的非结构化数据和采用不同标准的结构化数据转换成遵循统一标准的结构化数据，以便机器理解。数据之间的关联越是丰富，数据的价值就越能得到体现。

关联数据在整合孤立数据、提供开放的元数据服务、实现语义互操作等方面具有广阔的应用前景。

02研究方案

该项目的研究方案大致可以分为四步



第一步：数据收集和预处理

●采用众包模式实现语料库收集

图像中的文字识别是语义化建设的基础，开放的和自建的文字语料库，可直接作为图像文本识别的训练集。对于识别困难的古籍，可借助众包思路：将图像文本资源中的文字从图像中切分出来，混杂到较易识别的文字语料中，开放给大众进行校验，如作为某些平台登录时的验证码。达到一定的设置阈值后，即可利用这些图像文本的文字语料，并辅以一定的人工审核来提高文字识别的精度。

●基于深度学习的OCR字符识别

使用收集的语料构建机器学习模型，可以采用卷积神经网络（Convolutional Neural Networks, 以下简称CNN）和长短期记忆网络（Long Short-Term Memory, 以下简称LSTM）的联合算法，其中CNN用来做图像文本的特征提取，LSTM则用来生成描述。训练时，为了避免多拟合，可以采用多次迭代的方式。训练好

模型后，对于图像文本的OCR就转为一个个文字图像的多标签学习问题，可以用来自动识别大量的图像文本资源。

第二步：元数据整理

●实体标注与抽取

实体标注与抽取。需要借助实体语料库，可以用很多公开的规范词表（人名、地名、时间等）。此外，对于一些无法总结出规范词表的实体，如古地名、建筑名等，可以采用BIO标注，并使用RNN（循环神经网络）-CRF（条件随机场）模型对文本中的实体概念进行抽取，这也是目前深度学习的NER方法中的最主流模型。

●NER

命名实体识别(Named Entity Recognition, 简称NER), 又称作"专名识别", 是指识别文本中具有特定意义的实体, 主要包括人名、地名、机构名、专有名词等。一般来说, 命名实体识别的任务就是识别出待处理文本中三大类(实体类、时间类和数字类)、七小类(人名、机构名、地名、时间、日期、货币和百分比)命名实体。

主要过程组成(1)实体边界识别;(2) 确定实体类别(人名、地名、机构名或其他)。

第三步：结合开放的关联数据集进行图像资源的语义化注释

这里涉及到我们一开始介绍到的关联数据。

●IIIF框架并不包含语义标注标准和模型，因此要实现图像资源的语义标注，需要结合另一个通用的国际标准“开放标注数据模型（Open Annotation Data Model, 以下简称OADM）”，通过注释可以传达有关资源或资源之间关联的信息。OADM规范描述了一种结构化的模型和格式，以使注释能够在不同的硬件和软件平台上共享和重用。

●语义化注释将实现不同数据集之间资源的发现与知识关联机制，并和IIIF框架进行结合，以知识链接形式关联到外部开放的关联数据集（Linked Open Data, 以下简称LOD），从而对之前抽取的实体进行自动化语义标注。语义标注后，抽取出来的每个实体不再以字符串形式存在，而是直接对应到相应的资源URI，而且可以链接到大量的开放数据集。这些开放的数据集组成了实体资源丰富的数据池，对实体内容进行语义层面的增强。

(2) 语义检索实现。图像有了语义注释，就可以通过信息检索、关联数据、语义分析、自然语言处理等技术实现基于语义（概念）的复杂的图像文本信息检索。

(3) 知识图谱构建。馆藏机构可以结合自身馆藏资源和外部的开放关联数据集来绘制实体知识图谱（以实体为中心）、集合知识图谱（以藏本为中心）、领域知识图谱（以领域为中心），创新服务模式，提高科研效率与质量。

项目成果

● 网站

<http://miss.newwenke.com/sas/>

● 应用

多维度图像智慧系统

● 论文

[1] 陈涛,张永娟,单蓉蓉,刘炜.数字人文图像资源语义化建设框架研究[J].数字人文, 2020(02): 106-115.

[2] 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(09): 6-14.DOI:10.13998/j.cnki.issn1002-1248.2020.09.20-041.

[3] 李惠,陈涛,单蓉蓉.跨越时空的对话——构建基于IIIF-IIP语义标注平台的书画记忆链[J].农业图书情报学报, 2020, 32(09): 15-21.DOI:10.13998/j.cnki.issn1002-1248.2020.09.20-0504.

[4] 陈涛. 目标检测在数字人文图像中的应用尝试[J]. 数字人文研究, 2021, 1(3): 39-50.

● 网站内容

网站将上千项古代的国画、油画和字帖等通过数字技术，提供书画数据的本体结构、关联数据发布、多视角的可视化展示、统一检索和阅览。

基于 IIIF 技术，结合MISS平台与VoTT工具，提出了图像层面元数据注入、对象层面内容抄录、分类标签和语义层面语义引入的图像资源三个层次四个维度的标注模型。

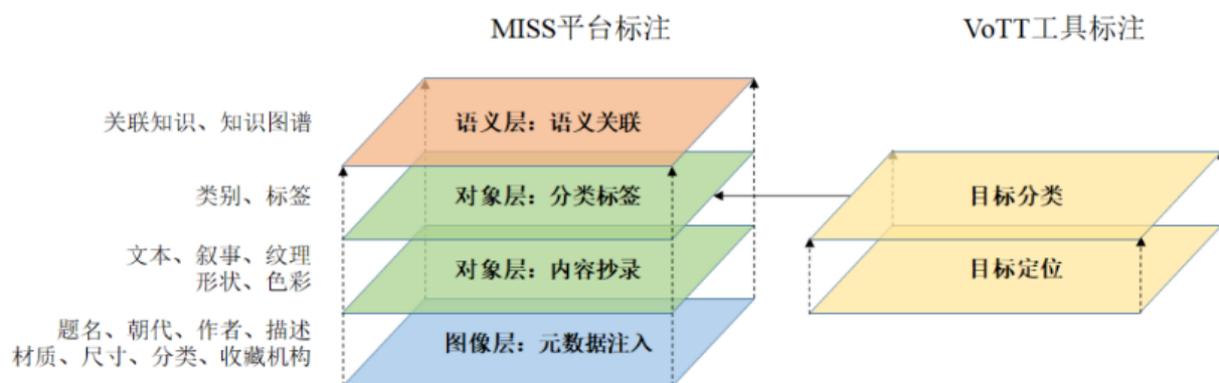


图3 MISS与VoTT结合的图像标注框架

项目特色与评价

01项目特色

- 包容性强，支持多种格式的图象资源

IIIF-IIP 平台从“元数据注入”“对象内容抄入”和“语义增强引入”3个层面对图像资源进行深度标注。依据 IIIF 和关联数据的标准和要求建立，支持多种格式图像资源的在线分享、复用、发布、语义标注等功能。

对于不同格式的私有资源，可以通过IIIF-IIP生成符合IIIF交互标准的图像资源；对于已经按照MF标准公开的藏品资源，可以在IIIF-IIP中建立各自的索引数据。

- 数据收录系统性强，挖掘记忆链网络隐含人物关系

以题跋时间、人物年份、朝代最后一年作为三个时间节点参数设置，研究者建立书画记忆链网络，并基于此将发现的人物关系与其他已经构建好的古代人物关系网络进行整合，旨在构建中国古代社会的学者网络。如图1所示，本文结合IIIF-IIP沉浸式交互平台，呈现可视化的书画记忆链《照夜白图》，根据时间信息，以朝代为主序，由古至今，展示由创作者、题跋者和印章者共同构成的时间链，依据此时间链作为线索收集作品，网站的收录更加体系化。



分享您的藏品

01.

上传/选择图片

上传您收藏的图片，或者浏览已经发布的 IIF 资源并收藏资源中的图片。

02.

创建/标注藏品

根据选择的图片创建自属的藏品 (Manifest)，并可对创建的藏品进行语义标注 (Semantic Annotation)。

03.

分享/赏析藏品

通过二维码分享您的藏品，并可分享标注后的藏品，用于赏析之用。

● 作品质量，整合了高清电子古画资源

平台目前收录的书画均提供高清的图片模型，并且在每个图象下面搭配四个按钮，可以一键进行全览以及放大观赏细节，同时在 Resource Discovery 中，点击可以看到收录的图片来自的数据库网址，用户也可以以此作为索引进行新的探索。

该网站目前可以免费使用，为鉴赏者提供更为友好、开放的环境。



同等资源:

http://dbpedia.org/resource/Wang_Ximeng
<http://rdf.freebase.com/ns/m.065zj1>
http://yago-knowledge.org/resource/Wang_Ximeng
<http://viaf.org/viaf/73816153>
<http://d-nb.info/gnd/1065416520>
<http://www.wikidata.org/entity/Q626360>
http://ca.dbpedia.org/resource/Wang_Ximeng
http://es.dbpedia.org/resource/Wang_Ximeng
http://et.dbpedia.org/resource/Wang_Ximeng
http://fr.dbpedia.org/resource/Wang_Ximeng
http://nl.dbpedia.org/resource/Wang_Ximeng
http://ru.dbpedia.org/resource/Ван_Симэн
http://sv.dbpedia.org/resource/Wang_Ximeng
http://uk.dbpedia.org/resource/Ван_Симен

02项目评价

●项目优点

从用户的角度讲，该项目有将传统书画作品电子资源化，这利于为书画研究者、鉴赏者提供丰富、全面的书画信息库，同时，多维度的搜索系统，也大大提升了查找的便利性。

从研究者的角度讲，在基于书画记忆网络链的查找、统计方式的基础上，研究者可以利用此工具更加清晰地计算出古代书画学者人物关系图谱，这会使得一些曾经书画研究中未被传统学者注意到的人物线索逐渐发掘，从而有利于推动文化遗产知识的新研究发展。

总体而言，该项目是计算技术与人文领域深度交互而达成数字人文项目范例，其中的古画高清图片、多个数据库网站联络、人物关系图表....让我们看到了基于计算机的算法技术可以为传统古代字画提供的新的活力与可能。

●项目改进与前景

在创办者的论文中看出，目前网站还在进一步建设中，目前就特定历史时期的特定领域的人物网络编织大多为独立研究，而网站下一步希望可以将碎片化古代人物关系进一步整合，达到更加系统的统计效果。同时在技术上也在不断完善，新加入了视觉目标标注工具VoTT，尝试构架更加完整的图像应用生态。

网站目前只能提供图片赏析、传送，依靠藏品名称、作者、机构三部分进行索引。而依据上述提到的人物图表等、画中其他元素等更多维度的索引还未加入体验。另外，对目标的检测的精确性还有待提升，期待网站下一步的建设，更加落实“多维度图象智慧系统”的名称。



数字人文研究

左 数 字 ， 右 人 文

