

—— 上海图书馆开放数据竞赛最佳智能体 ——

基于IIIF标准和智能体技术 的书画研究创新范式



目 录

一、作品概述

- (一) “今秘阁”书画智能体
- (二) IIIF标准与书画数据整合

二、解决的主要问题

三、应用示例

- (一) 使用指南
- (二) 功能演示

四、技术方案

- (一) 数据处理
- (二) 智能体架构

五、创新点

六、潜在价值

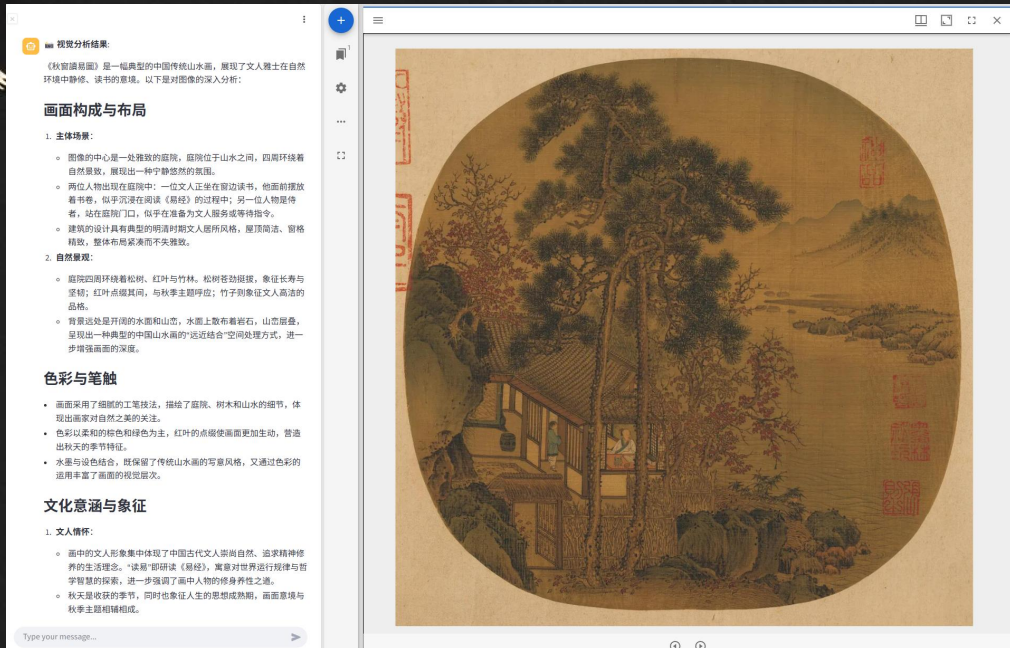
七、参赛感想





作品概述：“今秘阁”书画智能体

▷ 解读《秋窗读易图》

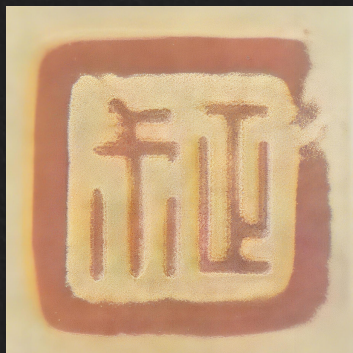


以自然语言交互为核心，基于 IIIF 国际标准的图像数据库，深度整合全球书画古籍资源和相关专业知识，提供问答解析、动态导览、版本对比、OCR识别和图文生成等功能，即时回应关于创作背景、画家生平、艺术风格、题跋印章等问题，并同步调取高清图像，比较不同版本细节，提供图文并茂的动态讲解和指尖上的沉浸式体验。

既为学术研究者提供便捷的书画图像分析工具，也助大众降低鉴赏门槛，感悟中国艺术精神。

访问地址：<https://37b5k38133.vicp.fun/ag/index.html>

宋代秘阁：
内府书画古籍
交流与收藏



“今秘阁”的涵义：

凭借数智时代的海量数据和技术优势，以IIIF标准汇集整理全球书画资源，构建一个契合当今时代潮流、满足大众文艺需求的书画收集、交流与鉴赏空间。



作品概述：IIIF标准与书画数据整合



国际图像互操作框架（International Image Interoperability Framework，IIIF），是一组实现图像资源互操作的技术标准，旨在使用户能更轻松地查看、操纵、比较和注释网络上的数字图像。

IIIF于2015年由欧洲和美国的图书馆等29个非营利图像资源存储机构共同成立，对以图像为载体的书籍、地图、卷轴、手稿、乐谱、档案等在线资源进行统一展示和共享。国际上主要的文化遗产研究机构都采用了IIIF对其图像进行管理和共享。

Funded by global technology leaders

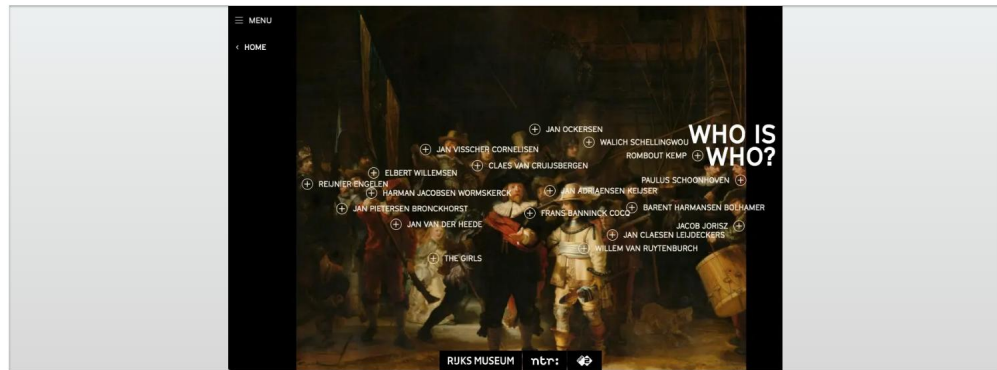
IIIF is funded by a 69-member global consortium, and leveraged by aggregators, research institutions, national libraries, archives, museums, software companies, and digital agencies around the world. Consortium members include:



IIIF解决了文化资源数字化图像难以被发现、再利用、引用、交换、比较分析等问题，为确保全球图像存储的互操作性和可获取性提供了**国际化通用标准**。

使用开放式 API 打破孤岛

许多对研究至关重要的图像和音频/视频资源都存在于孤岛中，访问仅限于本地构建的应用程序。IIIF 让您和您的受众可以自由地跨越障碍工作。



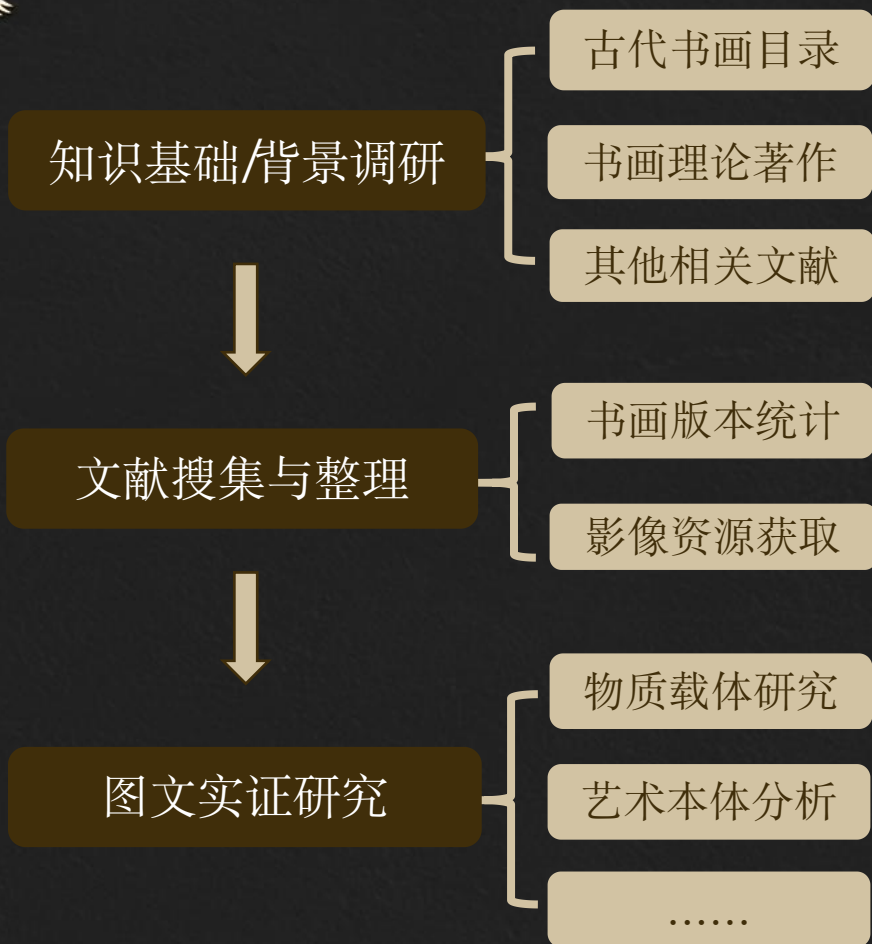
国家博物馆拥有丰富的在线展览，由 IIIF 注释驱动。

IIIF提供了高效可持续管理、分享以及利用数字图像的创新型数字技术标准，可以通过统一的数据框架实现全球书画图像资源的聚合



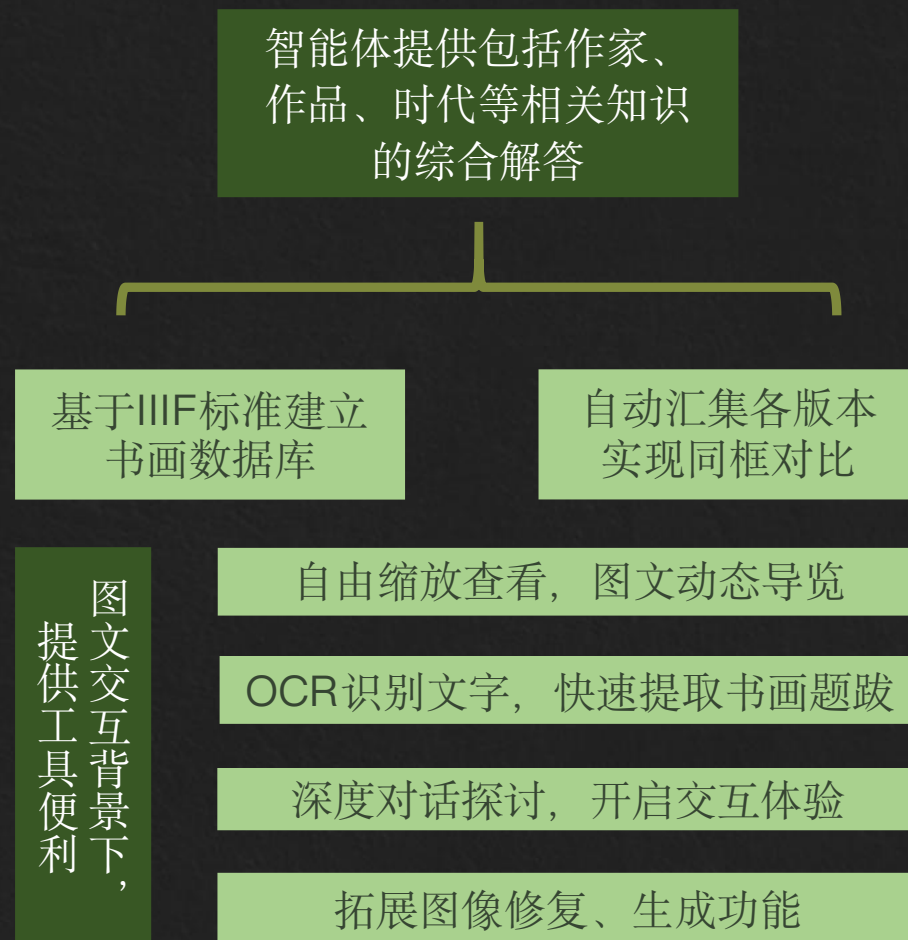
♀解决的主要问题

传统的书画研究路径：实物+文献



依赖于人工搜集、整理和统计
实物与文献分离，导致查阅困难

基于IIIF标准和智能体技术的书画研究创新范式



驱动智能体进行阅读准备和资源聚合
图文交互形式下展开深度挖掘

♀应用示例：使用指南

【问答栏】

【功能栏】

【图像展示窗口】



已经为您打开了东晋顾恺之的《宋摹洛神赋图》(北京故宫博物院藏, 乙本)。此画是中国画史中极负盛名的洛神题材古画之一, 以细腻工笔及诗意情境展现曹植《洛神赋》中的神女形象, 流传有宋代摹本。

如需欣赏画作细节、了解相关注释、绘画布局或画中人物解析, 请随时告知!



放大



> 🔑 Using tool: zoom

左侧输入【图像操作指令】

Type your message...



窗口操作指令:

放大 / 缩小 / 旋转 / 下一页 (针对多页画册而言)

右侧【Mirador窗口】呈现对应效果



♀应用示例： 视频介绍 ——以《洛神赋图》为例

今秘阁 书画智能体



应用示例：功能演示

书画数据索引

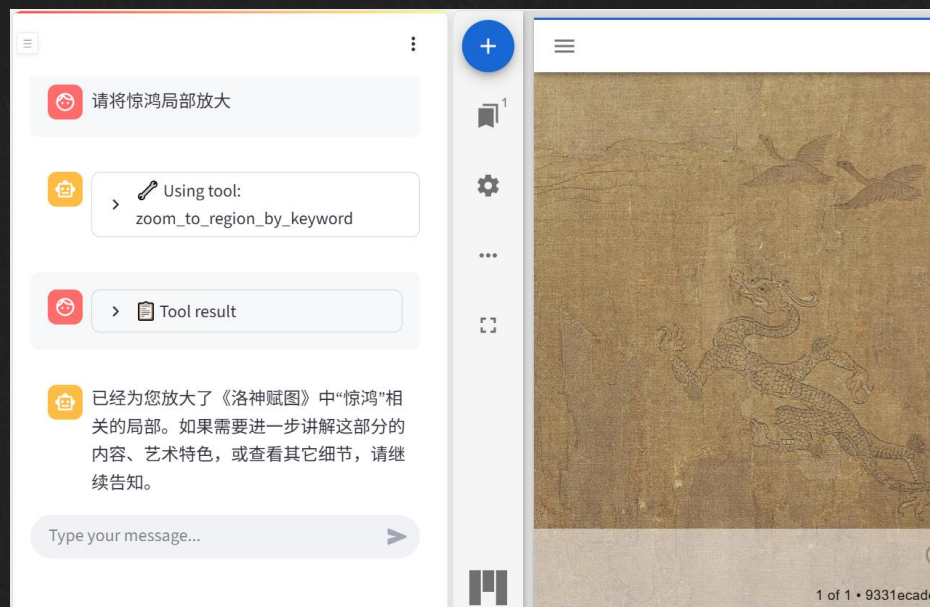
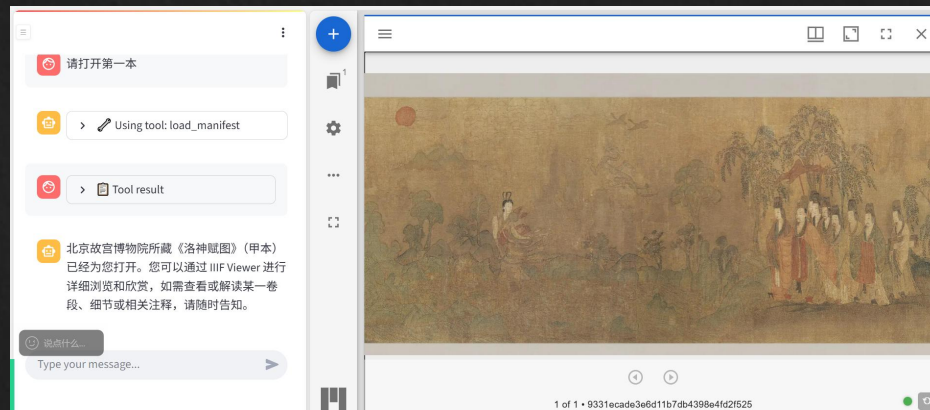
请查找数据库中的洛神赋图

Using tool: find

Tool result

数据库中收录了多件《洛神赋图》，均以顾恺之作为作者，主要藏于以下机构，不同版本如下：

1. 北京故宫博物院 藏“甲本”
 - 时代：东晋
 - 作者：顾恺之（顾恺之）
 - 版本：甲本
2. 英国·大英博物馆 藏
 - 时代：东晋
 - 作者：顾恺之（顾恺之）
 - 版本：未注明
3. 美国·弗利尔美术馆 藏“伪本”
 - 时代：东晋
 - 作者：顾恺之（顾恺之）
 - 版本：伪本
4. 台北故宫博物院 藏“清本”



01
IIIF标准
书画数据库

02
自动统计
版本信息

03
高清图像
缩放自如

☆基于IIIF（国际互操作框架）建立书画数据库

⊙任意检索书画作品，自动统计版本信息

☆图文并茂介绍书画，同步展示全貌或局部

不停留于书画文物的静态保存

→提供自由操作的图像界面

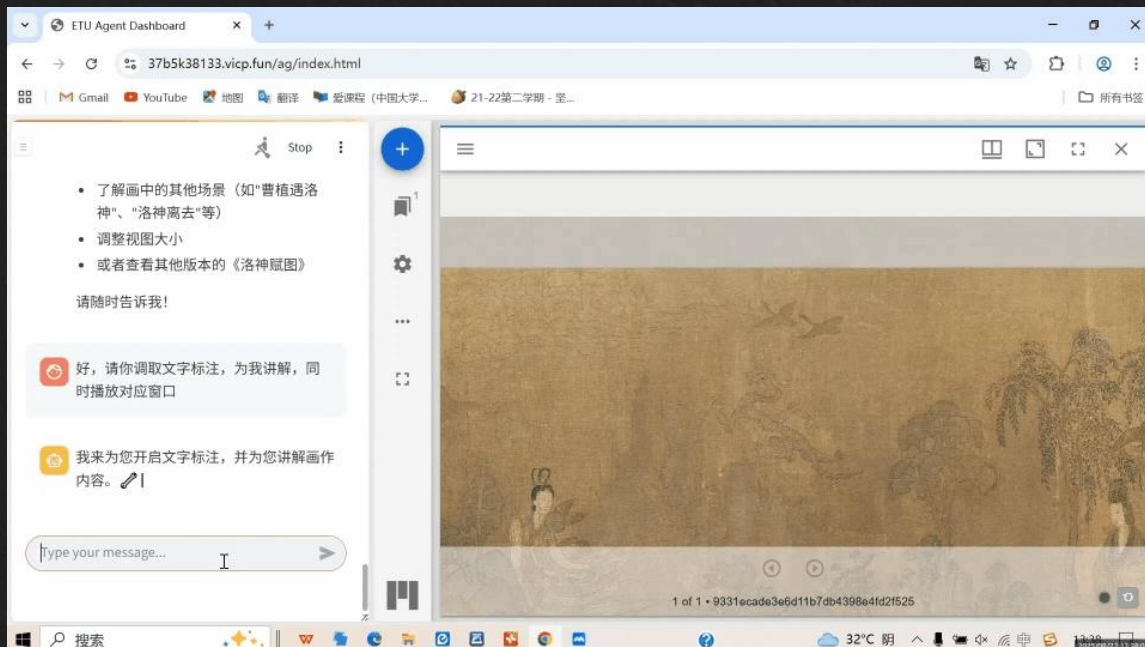
→被机器理解，与世界对话

应用示例：功能演示

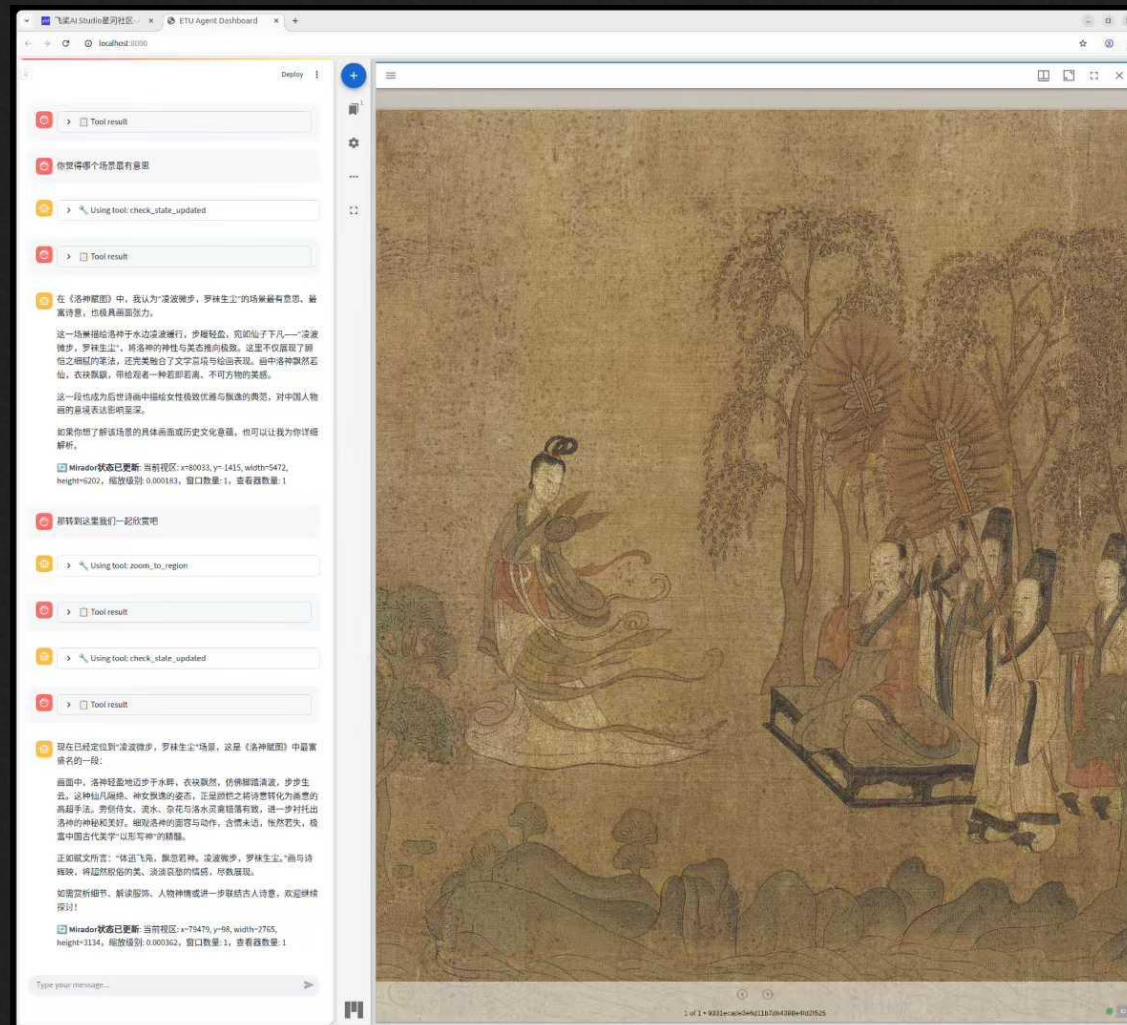
图文动态导览

画卷从右至左展开的故事

让我们先回到画卷的开始部分：### 1. 序章：曹植归途 这是故事的开端。画面展现了曹植从京城返回封地的场景。文字记载：“黄初三年，余朝京师，还济洛川。”曹植的车马行进在洛水之畔，正是在这里，他将遇见传说中的洛神。### 2. 初见洛神 这是曹植初次见到洛神的场景。“远而望之，皎若太阳升朝霞”，描绘了洛神如



根据文字标注，提供动态书画导览 (gif)



图文同步输出，免去页面切换烦恼

应用示例：功能演示

深度问答探索

请你介绍一下洛神赋图

好的，以下是**《洛神赋图》**的详细介绍：

一、作品背景

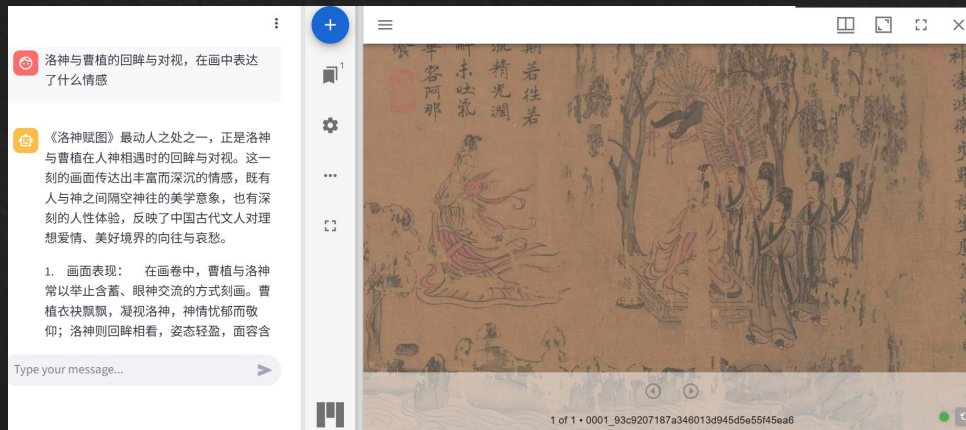
《洛神赋图》是根据东晋文学家曹植的《洛神赋》创作的连环画卷。“洛神”，即洛水之神，曹植在赋中虚构自己与洛神邂逅、倾心但终不得相守的故事，深沉哀婉，极具浪漫主义色彩。这一文学名篇激发了古代画家的创作灵感。

二、创作者及版本

最著名的《洛神赋图》作者是南朝东晋时期著名画家顾恺之。顾恺之有“画圣”之誉，他首创的人物故事画以传神为要，影响中国古代绘画数百年。

现今流传下来的《洛神赋图》多是唐、宋以后的摹本，尤以下三本最负盛名：

1. **故宫博物院藏本**：纸本卷，五代十国时期摹本，画风细腻流畅，人物形象古雅。
2. **辽宁省博物馆藏本**：宋代摹本，画法精致，墨色淡雅，是目前公开展出较多的版本。

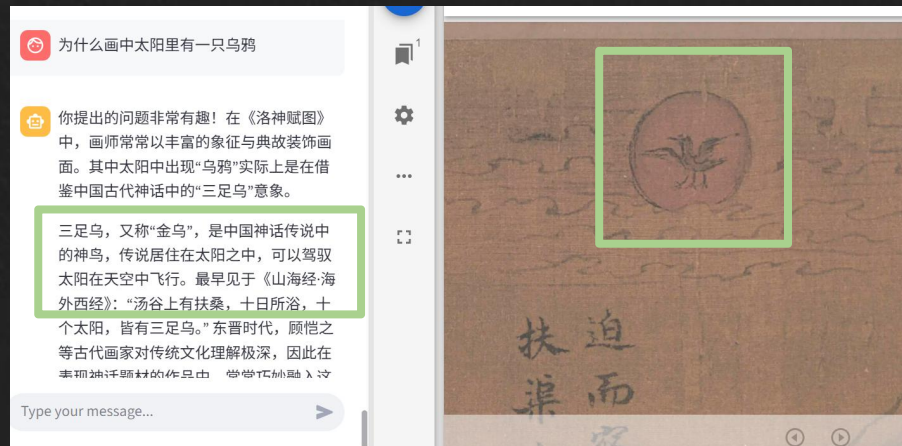


从孤立图像走向可被机器“阅读”的知识体

☆支持用户以自然语言提问，看图说话，自由探索

⊙以图文问答的互动探索，替代单一的读文或读图模式

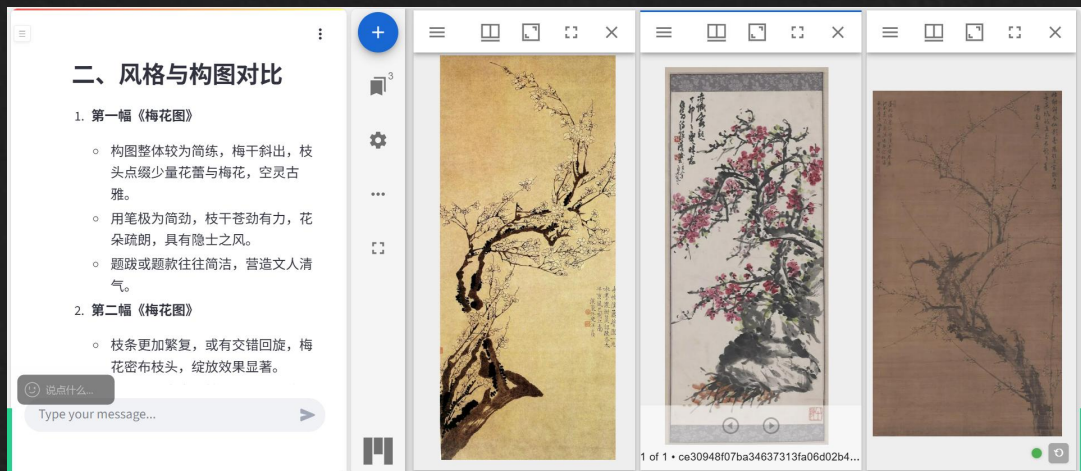
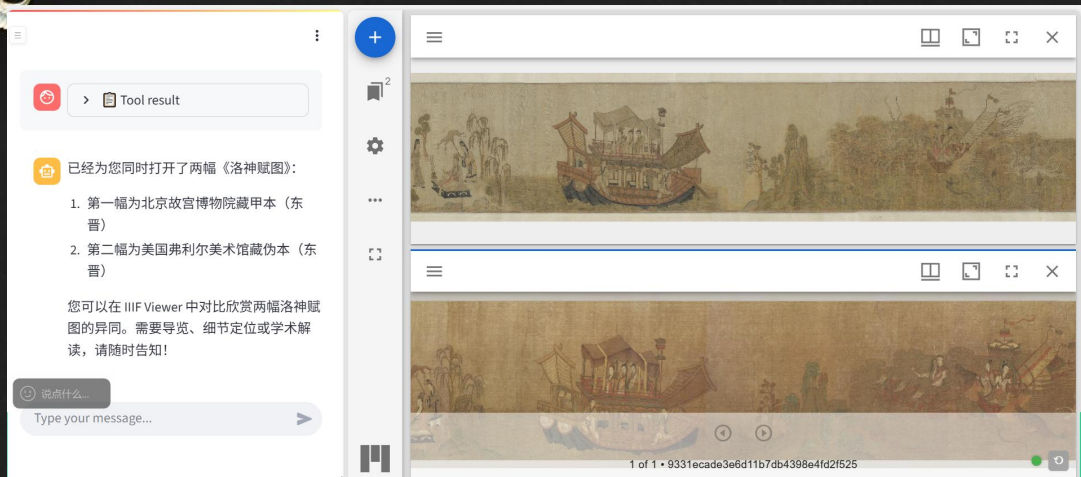
☆关联古籍知识，深度发掘，即时生成专业解答



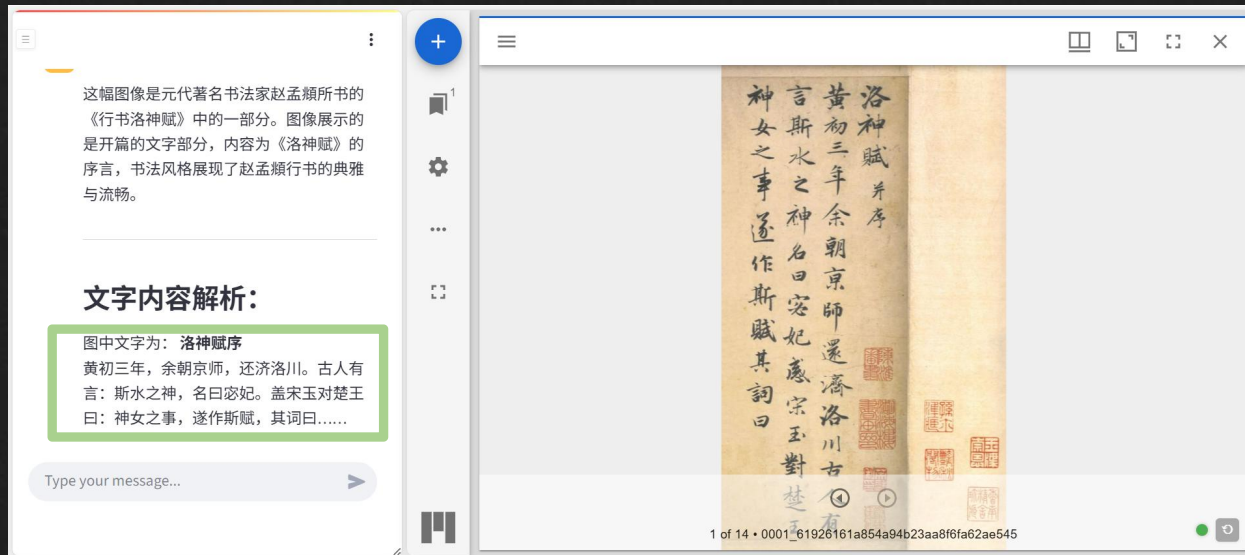


♀应用示例：功能演示

版本比对与识别



同框呈现作品各个版本，彰显细节差异
分析同题画作异同，鉴赏多元艺术风格



1. 书法风格：

- 字体特点：赵孟頫的行书书法以舒展大方、笔画圆润著称。此作品中可以看到笔锋的流畅与转折的精妙，展现出他的书写功底。
- 整体布局：文字排列整齐，字与字之间留白均匀，体现出书法作品的空间美感。
- 章法：开篇“洛神赋序”四字稍微加大，起到了视觉上的强调作用。

2. 印章：

- 图像右侧有多枚印章，其中包括作者的个人印章及收藏者的鉴藏章。这些印章具有重要的历史意义，表明作品的流传与收藏情况。

3. 纸张与墨色：

- 纸张呈现出古旧的黄色，墨迹稍显褪色，说明作品年代久远。
- 墨色浓淡适中，显现出赵孟頫书写时的用墨技巧。

历史与文化背景：

赵孟頫是元代的重要文化人物，被尊为书法“复古之宗”，他的书法深受二王（王羲之、王献之）的影响，讲究古意与优雅。书写《洛神赋》这样的经典内容，展现了他对传统文化的传承与个人艺术的表达。

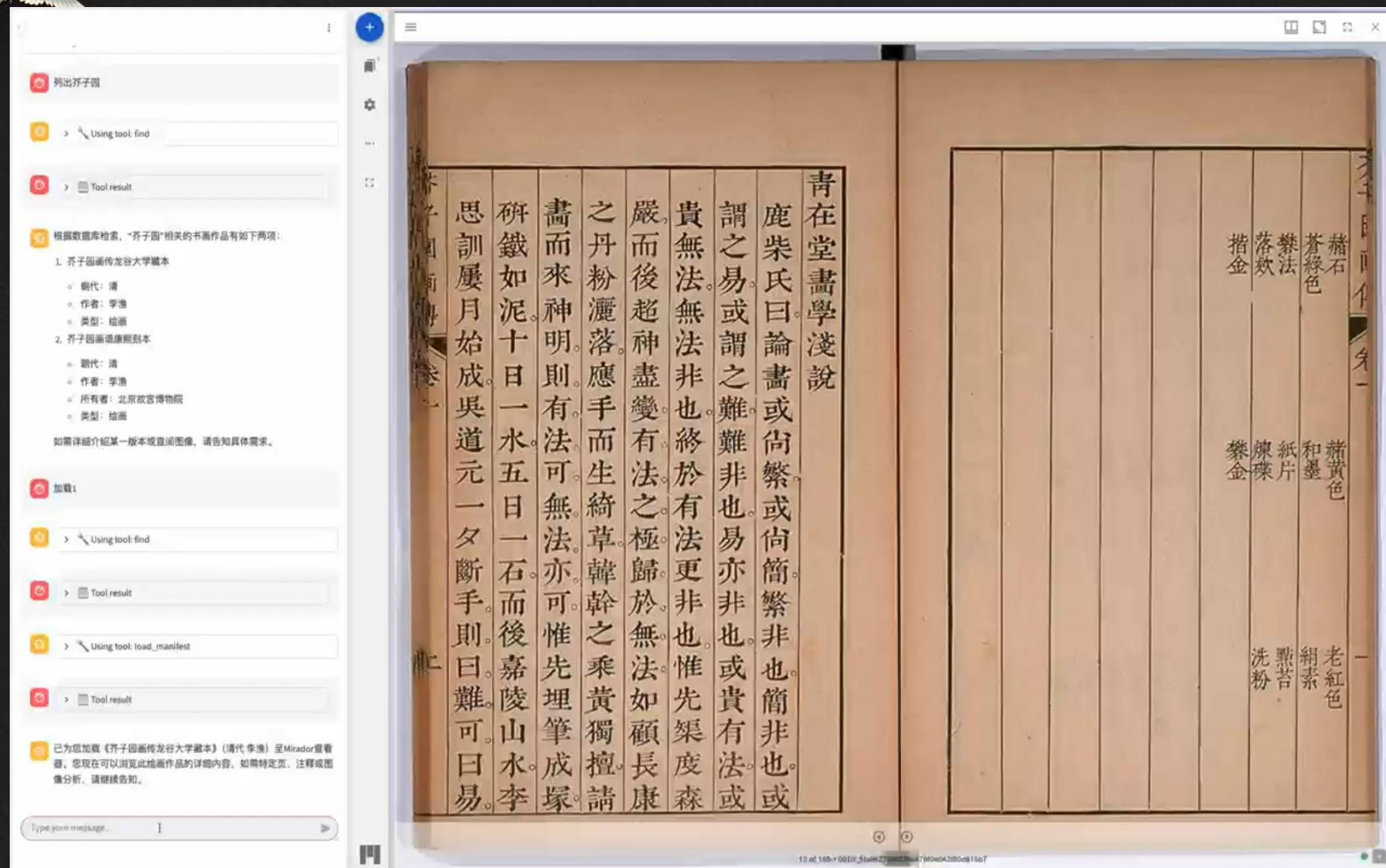
《洛神赋》是曹植的辞赋名篇，讲述了洛水女神宓妃的神话故事，其内容与情感表达成为后世艺术创作的灵感来源，影响深远。

用笔 印章 墨色 文化背景

OCR识别《洛神赋》书法作品并赏析

♀应用示例：功能演示

版本比对与识别



OCR识别《芥子园画传》并翻译（视频演示）

基于IIIF的书画智能体，相比现有的OCR技术实现工具，能带来什么创新？

⊙古籍数字化整理工作对图文复合类文献的回避

e.g. 识典古籍（仅校对文字，不处理图像）；
籍合网-明代集部文献数据库（识别的多为古籍刻本，未见插图本）

- 目前古籍整理过程中，未能实现图像与文本的多模态互动
- 如书画图像上的题跋这类特殊文献，还未进入古籍OCR识别范围内

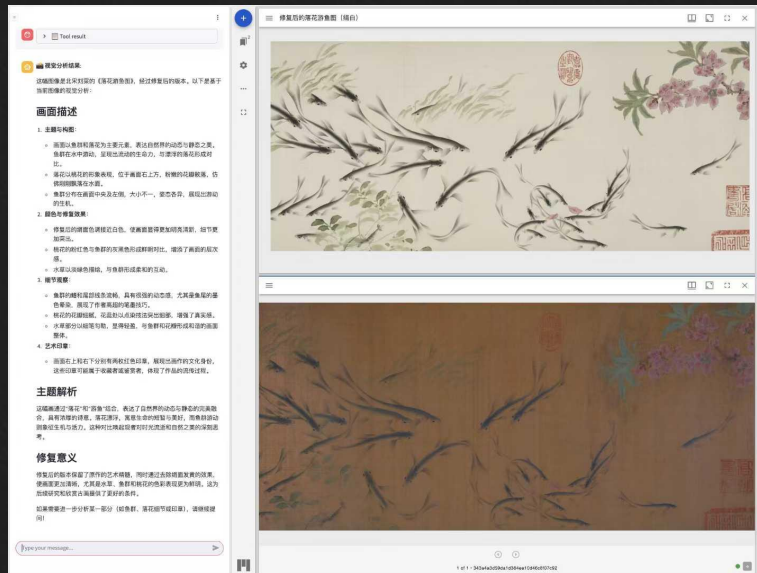
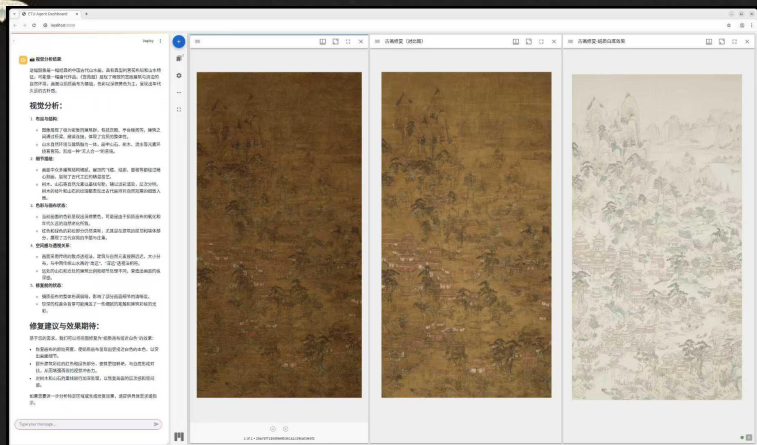
⊙书画文献与作品实物（图像）的离散现象

- 读艺术文献而不见书画图像
- 书画上的题跋文字，除已被辑录的部分外，还未与书画作品建立起较清晰的映射关系



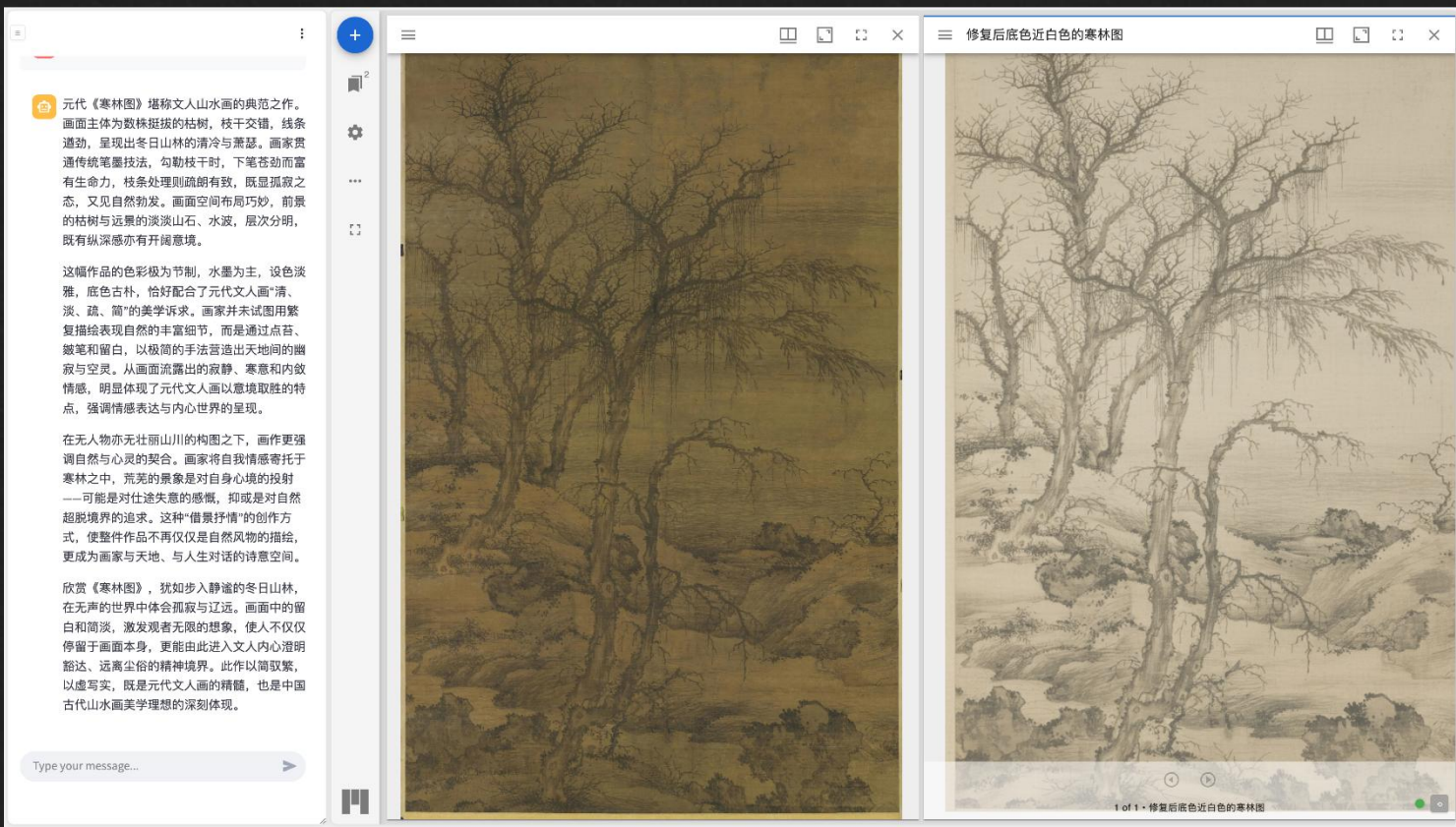
应用示例：功能演示

图像修复与生成



数字化修复（增强清晰度、补全残缺等）

★使用指引：需先在提问栏输入“图文生成：”
(生成时参考的输入图像为当前窗口局部。)



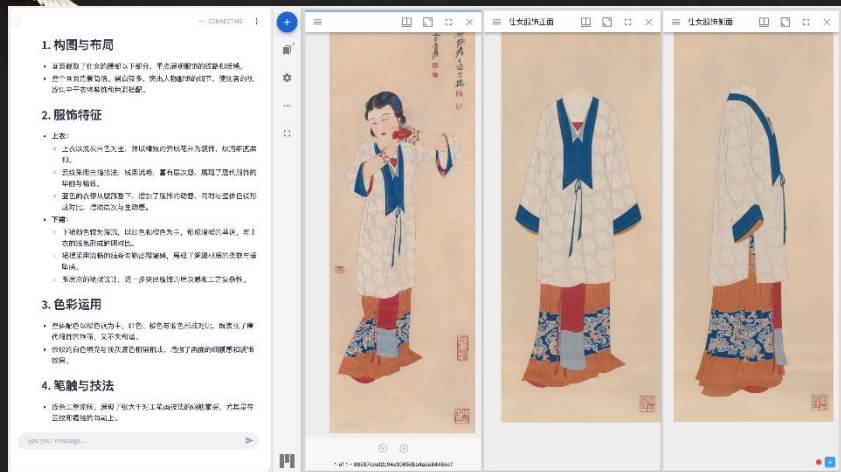
更改画作底色，突出笔墨线条

实现对书画图像的数字化修复，助力文化遗产保护



♀应用示例：功能演示

图像修复与生成



基于《仕女图》生成多角度示意图

立体呈现古代仕女服饰样貌

活化中华经典文化

《洛神赋图》服饰研究学者所绘线稿

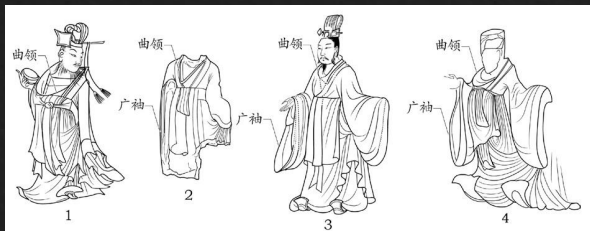


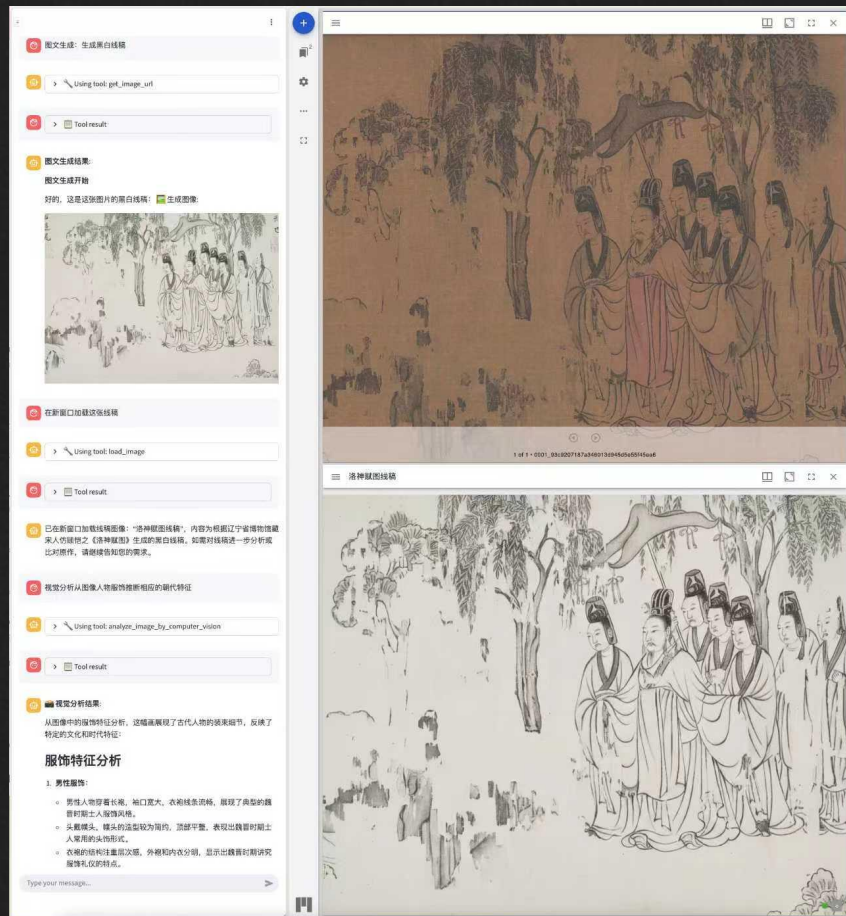
图4 传顾恺之《洛神赋图》“腾文鱼以警乘，鸣玉鸾以偕逝”部分的曹植（从左往右依次为故宫本、辽宁本和弗利尔本）

1. 首服

三卷《洛神赋图》中，曹植的首服应为“远游冠”（图4）。曹植的身份，属“王”级别。《续汉书·舆服志》：“远游冠，制如通天，有展筒横于冠前，无山述，诸王所服也。”^②又《晋书·舆服志》：“远游冠，傅玄云秦冠也。似通天而前无山述，有展筒横于冠前。皇太子及王者后，帝之兄弟，帝之子封郡王者服之。”^④

参考文献：张珊,丁洁雯.传顾恺之《洛神赋图》祖本创作时代再探——从东晋南朝服饰角度谈起[J].南京艺术学院学报(美术与设计版),2018,(06):71-80+210.

“今秘阁”书画智能体 图文生成功能绘制的线稿



为艺术研究者提供学术制图便利，降低美术技能门槛



🌸技术方案：数据处理

著录【朝代】【画家中英文姓名】【作品中英文名称】【注释】【版本】【风格】【藏地】【尺寸】【设色】【形制】【类别】等信息

数据使用情况：

◎上海图书馆 其它基础知识库（藏印知识库等）

◎中国历代人物传记资料库

书画作品来源：

北京故宫博物院、台北故宫博物院、国家图书馆、上海博物馆、辽宁省博物馆、美国弗利尔美术馆、美国哈佛大学、日本早稻田大学等

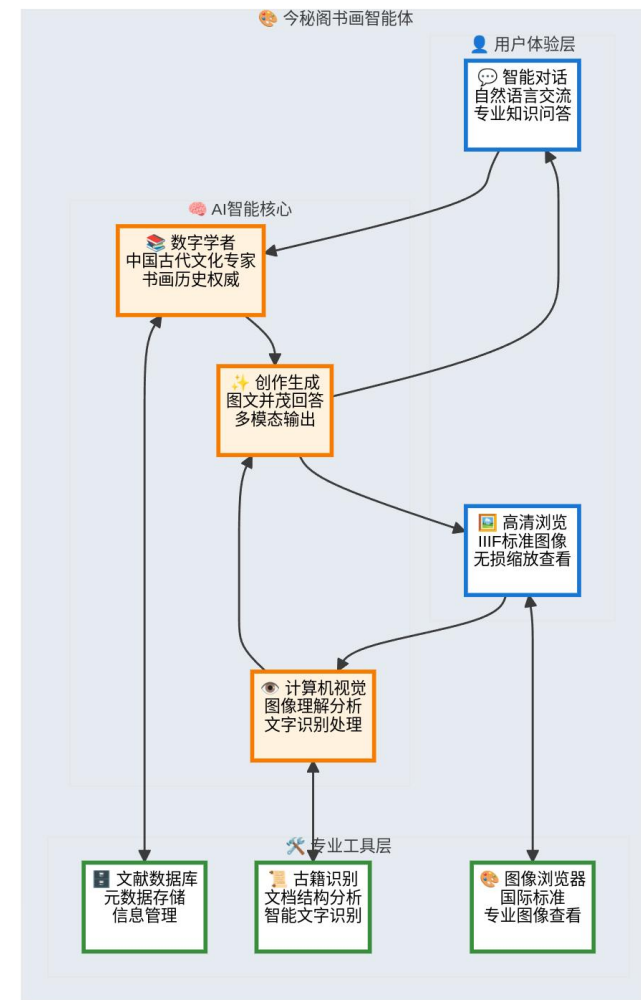
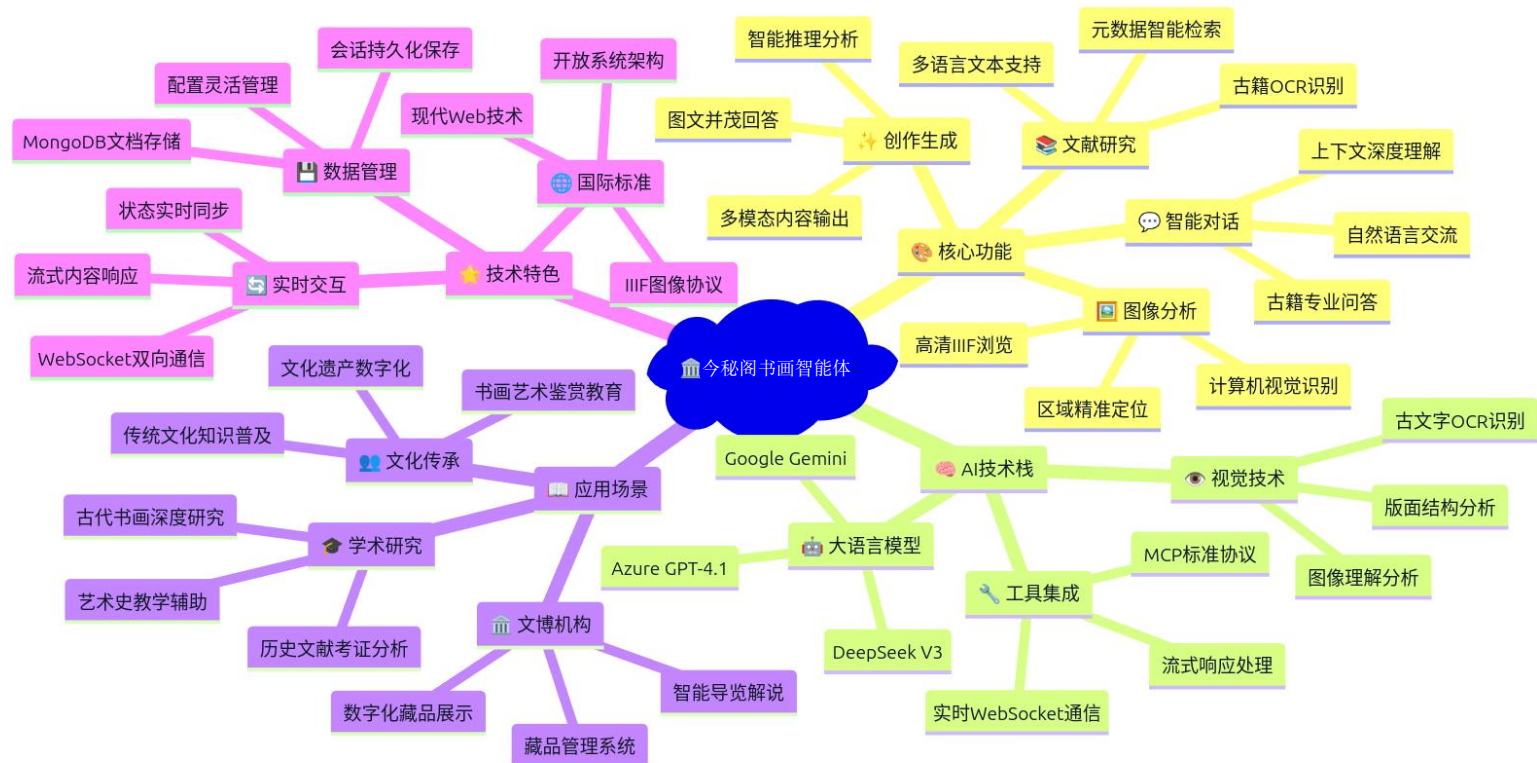
period	crea	nar	nan	nam	nam	title_ht	title_hs	title_en	title_ja	note	conrit	carve_occur	edition	press_c	calligraphy
北朝	北周					田仁訓及妻王氏墓誌銘並蓋	田仁訓及妻王氏墓志銘并蓋	The epitaph of Tian Renxun	田仁訓及妻王氏墓誌銘並蓋					原拓	
北朝	大象元年安道					鐵山刻經	铁山刻经	Iron Mountain Carving Sutra	鉄山刻經						
南朝	宋	2dfbf9c0	谢靈運	谢灵运	Xie Ling	謝靈運	長眉佛像	Long eyebrow Buddha	長眉の仏像	非谢灵运手笔。					
南朝	大明二年					爨龍顏碑	爨龙颜碑	Cuan Longyan Stele	爨竜顔碑			原刻	清末拓本		魏碑
南朝	宋					爨龍顏碑大明二年版	爨龙颜碑大明二年版	Cuan Longyan Stele, second	爨竜顔碑大明二年版						
南朝	宋					爨龍顏碑單頁版	爨龙颜碑单页版	Cuan Longyan Stele single-p	爨竜顔碑単ページ版						
南朝	宋	2dfbf9c1	陸探微	陆探微	Lu Tanwei	陸探微	歸去來辭圖	Return speech	帰りを辞する	非陆探微手笔。					
南朝	宋					刻石泰始七年南陽王泰等殘字	罗漢像	The portrait of the Rohan	羅漢像	南朝劉宋刻石泰始七年南陽王泰等殘字墨拓2			墨拓本		
南朝	宋	2dfbf9c0	謝靈運	谢灵运	Xie Ling	謝靈運	羅漢像			非谢灵运手笔。					
南朝	宋	2dfbf9c1	陸探微	陆探微	Lu Tanwei	陸探微	洛神賦圖	Luo Shen Fu	ロスの賦図	非陆探微手笔。此幅曾误标为《九歌图》卷。谢稚柳：“元人摹本。”徐邦达：“卷					
南朝	大明八年					南朝劉懷民墓誌	南朝刘怀民墓志	Epitaph of Liu Huaimin in the	南朝劉懷民墓誌						
南朝	宋	2dfbf9c1	陸探微	陆探微	Lu Tanwei	陸探微	人物圖	Character Map	人物図						
南朝	宋	2dfbf9c1	陸探微	陆探微	Lu Tanwei	陸探微	五嶽圖	Five mountain chart	五岳図	非陆探微手笔。					
南朝	梁					大般涅槃經	大般涅槃经	Mahaparinirvana Sutra	大般涅槃經						
南朝	梁					供養人畫像石	梁神農伏羲像	Statue of Liang Shennong ar	梁神農伏羲像	南北朝供養人畫像石拓片					
南朝	梁	2dfc20d0	張僧繇	张僧繇	Zhang Se	張僧繇	梁神農伏羲像	Statue of Liang Shennong ar	梁神農伏羲像						
南朝	梁	2dfc20d0	張僧繇	张僧繇	Zhang Se	張僧繇	梁神農伏羲像	Statue of Liang Shennong ar	梁神農伏羲像						
南朝	梁					陵墓梁武帝蕭衍墓石柱東闕西闕	水拓瘞鶴銘二十九字	Twenty nine characters of Sh	水拓瘞鶴銘二十九字	蕭梁陵墓墨拓本梁武帝蕭衍墓石柱東闕西闕			墨拓本		
南朝	梁					水拓瘞鶴銘二十九字	四體千字文	Four Body Thousand Charac	四体千字文	康熙五十二年 明末拓本					
南朝	梁	66cedafe	周興嗣	周兴嗣	Zhou Xir	周興嗣	四體千字文								
南朝	齊梁					宋拓曹娥碑小楷	宋拓曹娥碑小楷	Song Tuo Cao E Stele in Sm	宋拓曹娥碑小楷				宋拓		
南朝	梁					武帝半身像	雪山紅樹圖	Snow mountain mangrove n	雪山の紅樹図	梁武帝半身像軸					
南朝	梁	2dfc20d0	張僧繇	张僧繇	Zhang Se	張僧繇	雪山紅樹圖			左方上有“御府之印”一玺。又“子孙亿世家传之宝”、“合同”二印。下有“墨林					



元数据收集

图像精细化标注

技术方案：智能体架构



通过语言、OCR、图文生成、数据库查询等功能模块，实现大语言模型与IIIF浏览器的双向感知，提供可操作的图像界

♀技术难点：图像任务自治规划

所谓自治规划是指智能体能够分解复杂任务并按步完成，比如古画导览中的任务自动分解：**画面分段-》局部聚焦-》局部解说-》……-》解说完成**

这一能力的基础是语言模型的聊天界面对于传统应用程序的整合，即：

- **连接海量高清图像数据**
- **触发图像局部变化事件**
- **在聊天界面中实现图像互操作**



10月7日Sam Altman在旧金山OpenAI开发者大会上提出的将ChatGPT转换成超级App的Apps SDK

与大语言模型深度整合的III浏览器，彻底释放了模型的多模态能力，使得图像任务的自治规划得到充分发挥



♀ 创新点

建立中国古代书画图像数据整理规范，推进艺术资源聚合

凭借IIIF国际统一标准的优势，实现书画图像资源的自由流动

01

从单一读图或读文模式，走向图文联动的多模态探索

智能查询书画数据，同屏展示图文，OCR识别和图像问答助力书画理解与阅读

02

促进传统文博导览方式革新，提供动态游览、智能交互选择

自由查看图像整体或局部，动态讲解和问答打造沉浸式交互体验

03

满足版本比对、图像标注等需求，创新古代书画研究范式

取代传统人工搜集整理模式，自动统计、对比版本差异，实现图像修复与再生成

04

我们不只是“保存传统”，而是在让传统重新“参与当下”；
我们不只是“展示过去”，而是在用智能化方式，让文化继续生长。



♀潜在价值

假如置身《洛神赋图》前，你会如何“观看”它？

你也会好奇，拂过历史的烟云，曾经的古人如何欣赏一幅书画作品吗？

太阳里面为什么有乌鸦？

《洛神赋图》是怎么来的，《洛神赋》与这幅画有何关联？顾恺之的绘画有何特点？

这些题跋写着什么？书法是好是坏？



洛神是谁，她手里拿着什么？

什么是“翩若惊鸿，婉若游龙”？

这状如柳树的植物在画中可有深意？

这幅画如何装帧？

这是谁的印章？刻的什么字？

也许，是一个晓风和畅的晴日，在西园的荷塘香榭，煮一壶茶，点一炉香，三五好友相聚。小心翼翼地褪下锦套，徐徐展开泛黄的手卷，墨痕在柔和的光影里缓缓苏醒。“气韵”“笔墨”“骨法用笔”“经营位置”……一些并不陌生的词汇在口、笔相传中永恒。



♀潜在价值：学术研究+文化传播

梳理传移模写脉络，
推进书画版本研究

自由对比版本细节，OCR识别
+图文生成 减轻工作负担

从古代文人的高雅趣味到全民参与的“文博热”
从袅袅檀香中的零星残卷到卷帙浩繁的数字资源
从题跋画记的文学互动到口授心传的博物讲解

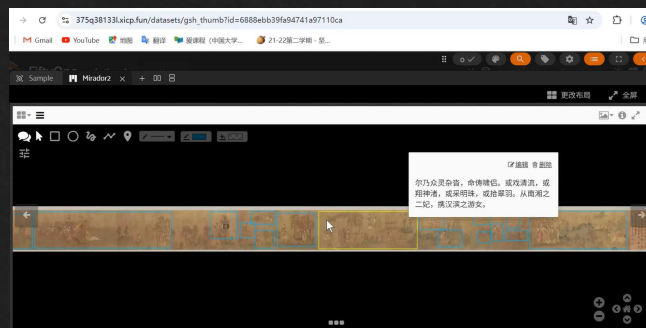


当诗书画印一体的古代语境，迎来数智时代的技术冲击，
面对专业知识壁垒和庞杂的书画资源，
今天的我们如何理解和传承千年笔墨中的中国艺术精神？

- 顾恺之是《洛神赋图》的原创者，现存多个版本：
- 甲本：北京故宫博物院藏
 - 乙本（宋摹本）：北京故宫博物院藏
 - 辽宁省博物馆藏本：宋人仿顾恺之洛神赋图
 - 清本：台北故宫博物院藏
 - 英国大英博物馆藏本
 - 美国弗利尔美术馆藏本（标注为伪本）



小结	阶跃	豆包
人物身份分析	可以区分主要人物和次要人物	可以识别核心人物
空间构图分析	可以根据色彩、服饰和位置区分人物地位（贵族和平民）	一定程度说明和判断人物行为和身份
环境情感分析	可以依据方位分析人物行为活动，阐述构图可能传达的意味（有时也结合一些数学判定方法如黄金分割）	可以依据左右方位描述，不能运用数学方法（如黄金分割）精准说明空间位置
绘画风格分析	可以分析环境的情绪倾向（平静、肃穆）	可以分析环境中的植物，但未关注其暗示的氛围
画作推理判断	对画作风格及其朝代有所判断（准确率存疑）	可以推测画作所处时代（准确率存疑）
视觉识别能力	难以将画作与画家对应，容易出现虚构现象	在充分提示下仍推断错误，难以得出正确的画家和画作信息
	可以对分散的人或物进行较准确的计数	可以对5左右的数量进行较为准确的计数
	有意识地区分动物颜色和植物种类	较少关注动物的数量和植物的种类



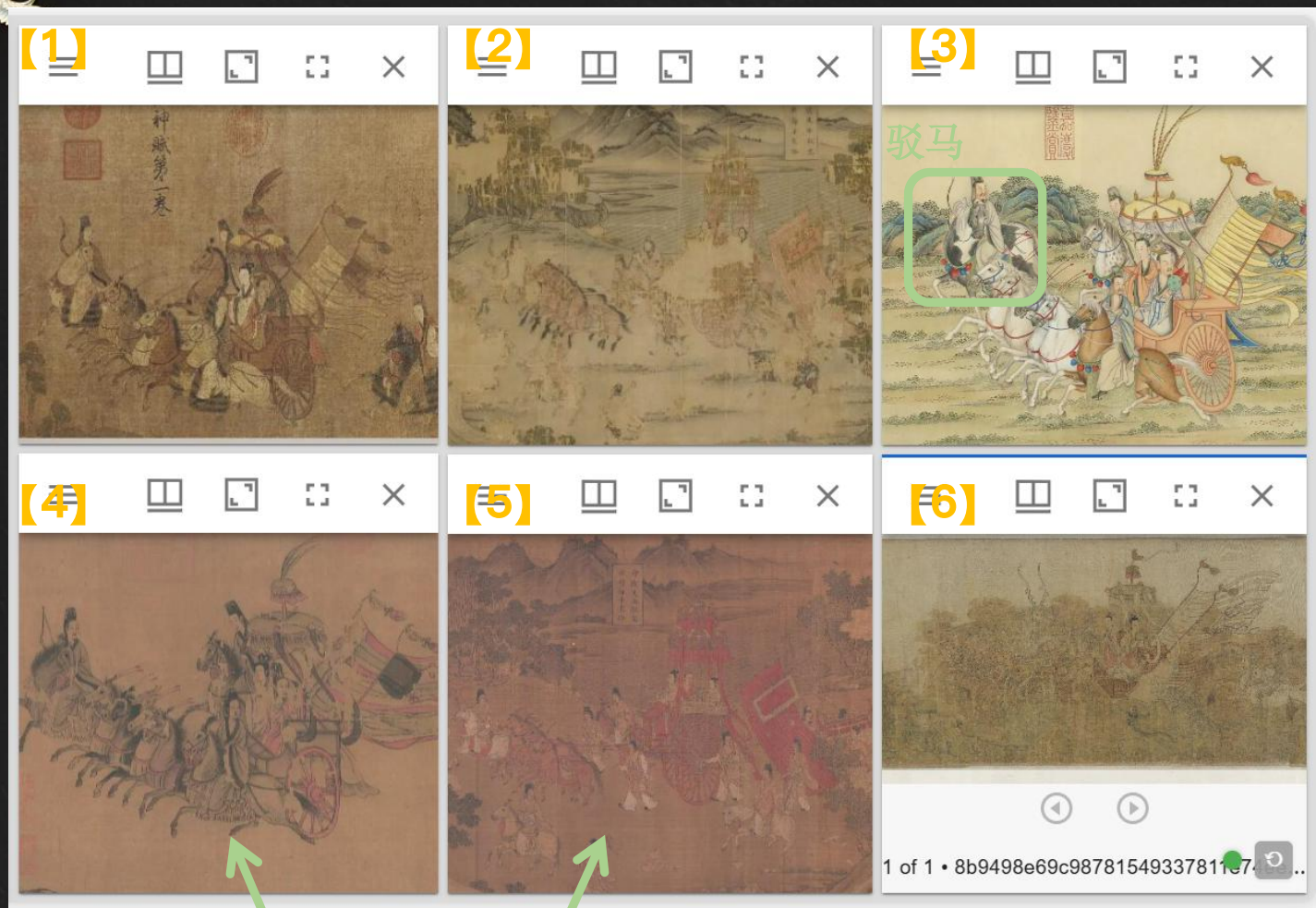
探索AI图像理解能力边界

推进跨学科视野下的题画文学研究

☆帮助大众跨越欣赏古代书画可能面临的知识沟壑（书画理论、名物考证、装裱版式、书法题跋、印章篆刻……）
☆线上提供动态智能导览，实现指尖“卧游”书画天地；
线下作为数字化创新形式，辅助文物展览，打造个性化交互体验。

例如：礼教视野下的《洛神赋图》

《逸礼·王度记》：“天子驾六马，诸侯驾四。”



曹植车架：四马异色 → 四马同色 洛神六龙，曹植四驾：人神之道殊

借助“今秘阁”同框展示《洛神赋图》的六个版本：

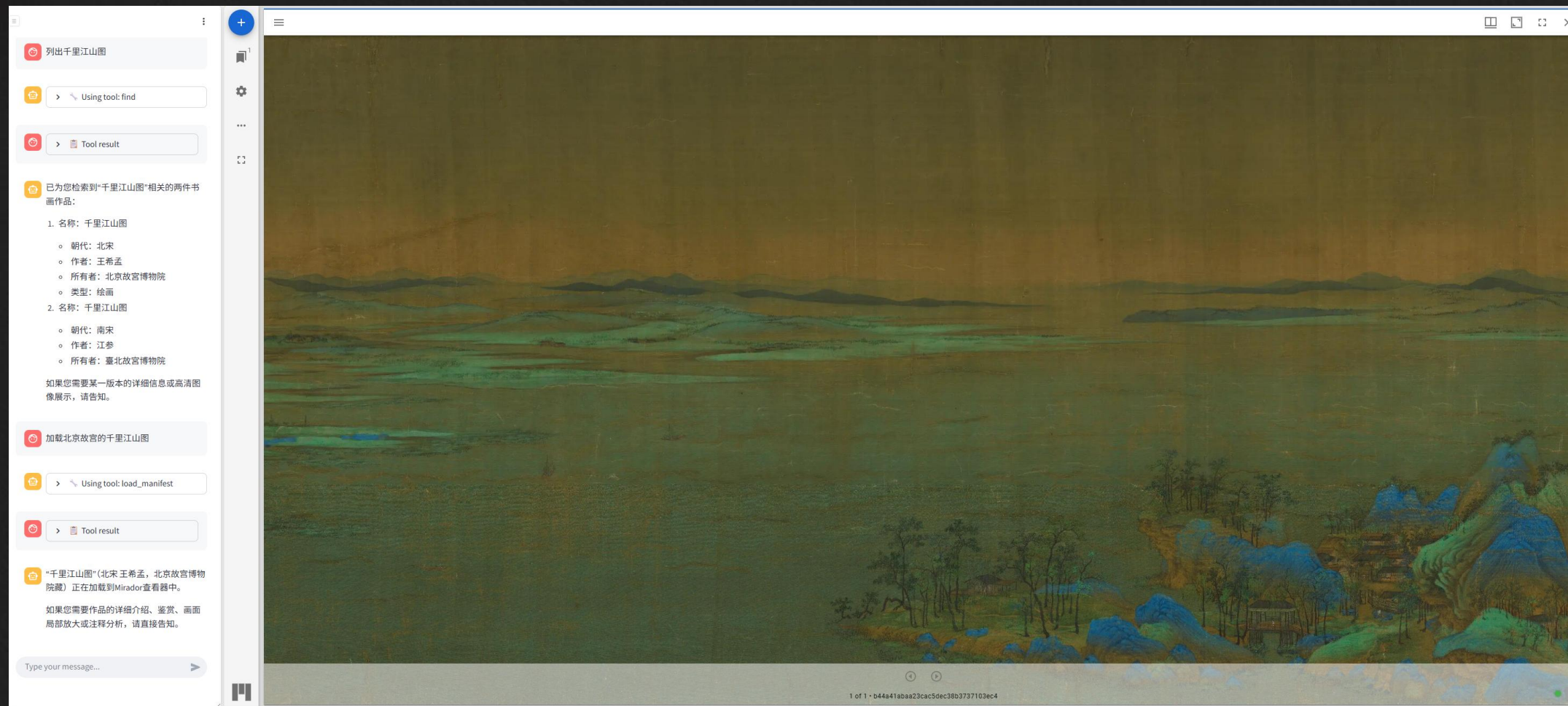
- 【1】 故宫博物院藏甲本《洛神赋图》
- 【2】 英国大英博物馆藏《洛神赋图》（宋摹本）
- 【3】 台北故宫博物院藏丁观鹏摹《洛神赋图》
- 【4】 辽宁省博物馆藏《洛神赋图》（南宋摹本）
- 【5】 故宫博物院藏乙本《宋摹洛神赋图》
- 【6】 美国弗利尔美术馆藏《洛神赋图》

研究基于IIIF（国际图像互操作框架）对《洛神赋图》的六个版本进行文字标注，构建书画智能体以实现对《洛神赋》文图互动关系的动态游览，并提取各版本在人物视线流动、肢体语言和车服规制等细节上的差异，从而阐释从赋到图过程中曹植由诸侯向君王靠近的地位抬升现象，以及男性注视下洛神形象的世俗化和卑微化。

例如，故宫甲本、辽宁本和弗利尔本“天子驾六，诸侯驾四”的规制，经由故宫乙本和大英本以纯色代替杂色马的逾越，最终产生了丁观鹏本曹植乘纯色四马车架、侍从乘驳马的等级秩序。



参赛感想



基于IIIF的今秘阁书画智能体，是千百年来中国书画优秀传统文化的延续与发扬。
愿它在人工智能时代为我们的笔墨丹青谱写出崭新的文化旅程！

T H A N K Y O U

感谢观看

基于IIIF标准和智能体技术的书画研究创新范式

