



AI 赋能的数字人文新生态

上海图书馆开放数据：资源、技术及应用示例

仇开域

上海图书馆（上海科学技术情报研究所）· 2026

一、数字人
文开放数据
建设历程

二、2026
年开放数据
资源全景

三、核心技
术及「AI+数
据」双向赋能
体系构建

四、数据获
取与调用方
式

五、应用案
例

一、开放数据建设历程

2014

启动建设

2016-2022

数字筑基 知识挖掘

2023-2026

智能跃迁

数据
+AI

历史纪
年表
881

上海地
名志
1w+

舆图
目录
1w+

避讳字
697

姓氏
600+

机构
2k+

地理名
词表
4k

上海年
华机构
2w+

历史文
化事件
1.3w+

印章
2w+

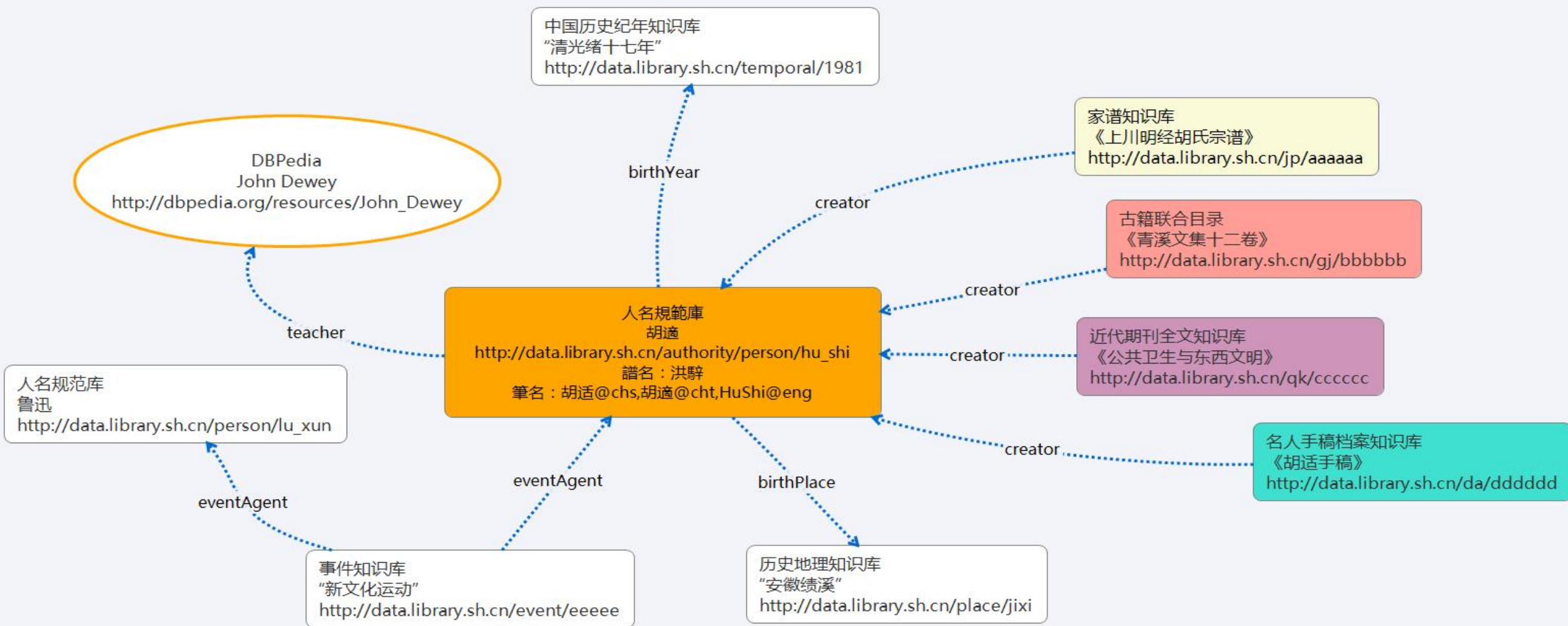
人名
规范库
135万

古籍
刻工
5.8k+



将不同种类、来源的数据通过知识建模等方式组织关联，建设成为面向人文研究的数据基础设施的一部分

开放数据的内容组织框架（例）



二、2026年开放数据资源全景

1 上海图书馆、2 合作机构单位

基础知识库

1. 人名规范库：135万余条
2. 地理名词表：3800余个
3. 华人姓氏：600余个
4. 上海优秀历史建筑：1086处
5. 上海市不可移动文物：3452处
6. 上海地名志：10000余个
7. 上海历史文化事件：1.5万余条
8. 上海年华机构名录：2.1万余个
9. 中国历史纪年表：879条
10. 收藏机构：2000余个
11. 藏印知识库：印章（22000余个）、藏书楼（54处）、藏书家（455位）
12. 避讳字知识库：697个
13. 古籍刻工名录：5800余人

文献知识库

1. 家谱：10.5万余种 原文14625种
2. 盛宣怀档案：15万余种
3. 古籍循证数据：130万余种 原文10802种
4. 上海联编中心书目数据：150万余种
5. 红色文献书目数据：8500余种
6. 红色旅游景点：1199处
7. 电影库数据：4000余种
8. 碑帖 45件碑帖目录数据及相关背景知识、收藏历史
9. 近代图书 原文620种
10. 戏单 897条

《近代上海城市文化》专题片、中文期刊论文数据、用户与借阅数据、东馆预约数据、活动预约数据

二、2026年开放数据资源全景

上海图书馆、合作机构单位

公共图书馆

南京图书馆、浙江图书馆、安徽省图书馆、苏州图书馆、南通市图书馆、扬州市图书馆、嘉兴市图书馆、深圳图书馆、上海市静安区图书馆、上海市闵行区图书馆、上海市嘉定区图书馆、上海市金山区图书馆、上海市奉贤区图书馆、上海市崇明区图书馆

平台/机构/单位

上海市公共数据开放平台、《全国报刊索引》、CADAL项目管理中心、东方社区信息苑、复旦大学图书馆、广州搜韵文化发展有限公司、杭州弘雅科技有限公司、华东师范大学图书馆、上海宋庆龄研究会、上海韬奋纪念馆、苏州市公共文化中心、中国人物传记资料库项目（CBDB）、中国人民大学图书馆

三、核心技术概念介绍

1

Ontology (知识本体)

- 对领域知识进行抽象的、形式化的知识建模过程
- 可共享的概念模型 (古籍、家谱...一体化本体模型)

2

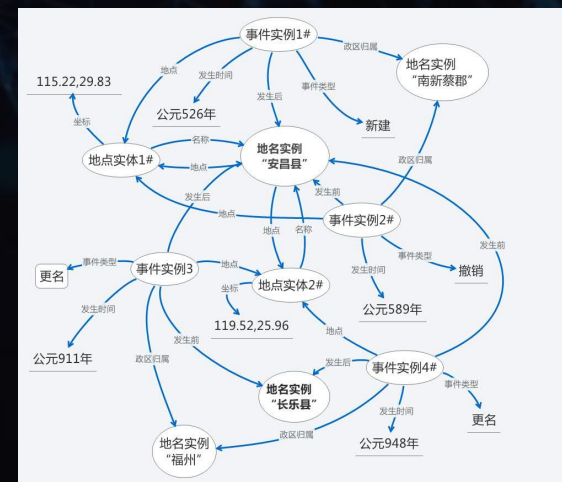
RDF (资源描述框架 Resource Description Framework)

- W3C提倡的一个数据模型
- 三元组形式, 用来描述万维网上的资源及其相互间的关系

3

Graph Data Store (图数据仓储)

- 以图的结构形式来存储数据信息, 节点、边
- 很好地表现数据的联系



三、核心技术概念介绍

4

RAG (Retrieval-Augmented Generation)

- 检索-增强-生成：基于知识库，抑制AI幻觉
- RAG on RAG：组合多种检索策略，根据问题类型动态调度

5

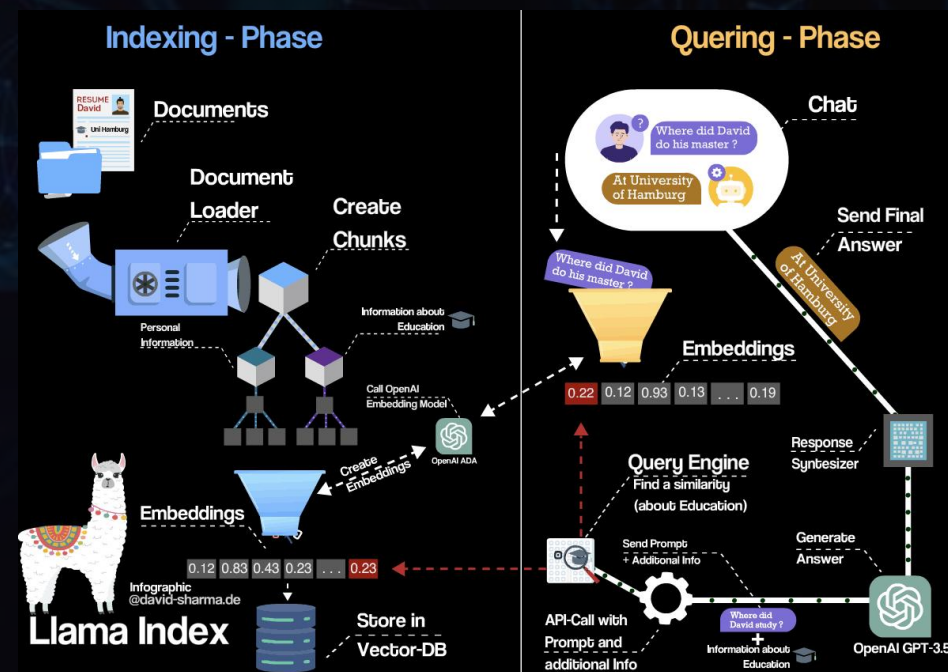
Graph RAG

- 用知识图谱替代向量检索，解决多跳推理问题

6

Agentic RAG

- 将智能体 (Agent) 技术与RAG结合的高级AI框架，能动态规划多步骤操作，并根据结果质量自动调整策略



AI 与数据的双向赋能体系



四、数据获取与调用方式

1.进入竞赛官网
opendata.library.sh.cn



2. 进入【开放数据】，根据
需要获取数据

接口
调用

3.1. Restful API

特点：Web调用，跨语言，门槛低。

方法：http://data.library.sh.cn/sg/person/
data?[参数1]&[参数2]&[参数
3]?key=YourAPIKey

3.2. HTTP URI内容获取

特点：Web调用，跨语言，访问实体的HTTP
URI即可获得其结构化数据。

方法：需先通过调用接口获得实体的HTTP
URI

返回数据格式：
JSON-LD

2026开放数据竞赛
开放所有全文数据：

- 每类文献资源每天调用上限：2000页
- 单次调用最大上限为500条。
- 每位用户调用API时间间隔：2秒或以上



五、开放数据应用示例

- 自建文献知识库与数据库
- 展陈项目
- 开放数据平台

馆藏方志智慧服务平台

上海图书馆藏方志智慧平台

方志资源

责任者

专家

上海图书馆藏方志智慧平台

方志资源

责任者

请输入方志题名、责任者、版本等字段检索或直接问我

Q

高级检索

📍

⋮

上海图书馆藏方志智慧平台是一个专门收录和展示地方志资源的数字化平台，目前包含 500 余种上海图书馆收藏的珍贵古籍方志，作为综合应用数字人文技术和人工智能技术的服务平台。该平台除了目录、影像、全文信息外，还结合采用知识图谱/关联数据技术开发的各类知识库应用，以及历史地图，提供更加准确的语义服务。用户可以便捷地访问和检索这些珍贵文献，以及利用丰富的工具进行句读标点、文白互译、实体提取、图像理解、数据分析和可视化展示等，也可以利用其内容生成能力帮助从事具体的研究工作。

馆藏方志智慧服务平台

AI 增强检索

有庙宇记载的上海方志有哪些

展开分面信息

宋

元

明

[正德]松江府志三十二卷

责任者：陈威(修) 顾清(纂)

版本：明正德7年

全文浏览

智慧平台

方志资源 责任者

上海简称“沪”的来源和螃蟹有关吗

关于上海简称“沪”的来源，根据上下文中的信息，并未直接提到与螃蟹有关。上海古称“滬澹”，“滬”字意指“独出其水而入海”，反映了古代黄浦江入海时水流较急，吴淞江为主要水道东流入海的情况^[1]。因此，可以认为上海简称“沪”来源于这一古代名称，而非与螃蟹有关。不过，文中确实提到了上海地区的一些水生生物，包括螃蟹等，但这些描述与“沪”字的由来无直接关联^[3]。

来源：

1. [正德]松江府志三十二卷
应略而不言也新唐史書望九域志書繁元既即縣為府因為屬邑國朝因之上海縣在府東北九十里本華亭縣地舊曰華亭海後以人煙浩穰海船輻輳遂成大市宋於其地立市舶提舉司及權貨場曰上海鎮其立以為縣則知府僕散翰文之請也以地居海之上洋故名曰上海亦稱為上洋云分野歷代星書皆以斗中為吳越之分野雖互有同異而大要不出乎此松江本吳地故具列占驗以俟知星者茲為費帝分星次斗十一度至婺女七度日須女又日星紀於辰在丑於律為黃鐘周禮保章氏註星紀吳越也爾雅星紀斗牽牛吳分分野西漢天文志斗江潮牽牛婺女揚州地理志吳地斗分野越地牽牛婺女之分野

2. [正德]松江府志三十二卷
應略而不言也新唐史書望九域志書繁元既即縣為府因為屬邑國朝因之上海縣在府東北九十里本華亭縣地舊曰華亭海後以人煙浩穰海船輻輳遂成大市宋於其地立市舶提舉司及權貨場曰上海鎮其立以為縣則知府僕散翰文之請也以地居海之上洋故名曰上海亦稱為上洋云分野亡★代星書皆以斗中為吳越之分野雖互有大要不出乎此松江本吳地故具列占驗以俟知星者茲為費帝分星次斗十一

明崇禎4年

單種阿婆 書志 全文浏览

清康熙22年

输入问题询问大模型...

馆藏方志智慧服务平台

册数 《 1册



全屏

上一页

下一页

第3页/共64页

原文

翻译

句读

繁体

简体



实体标注



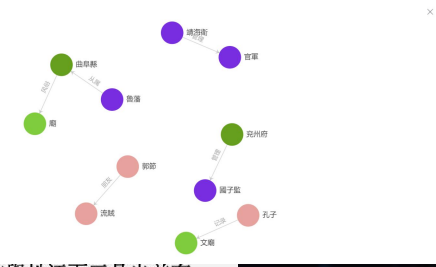
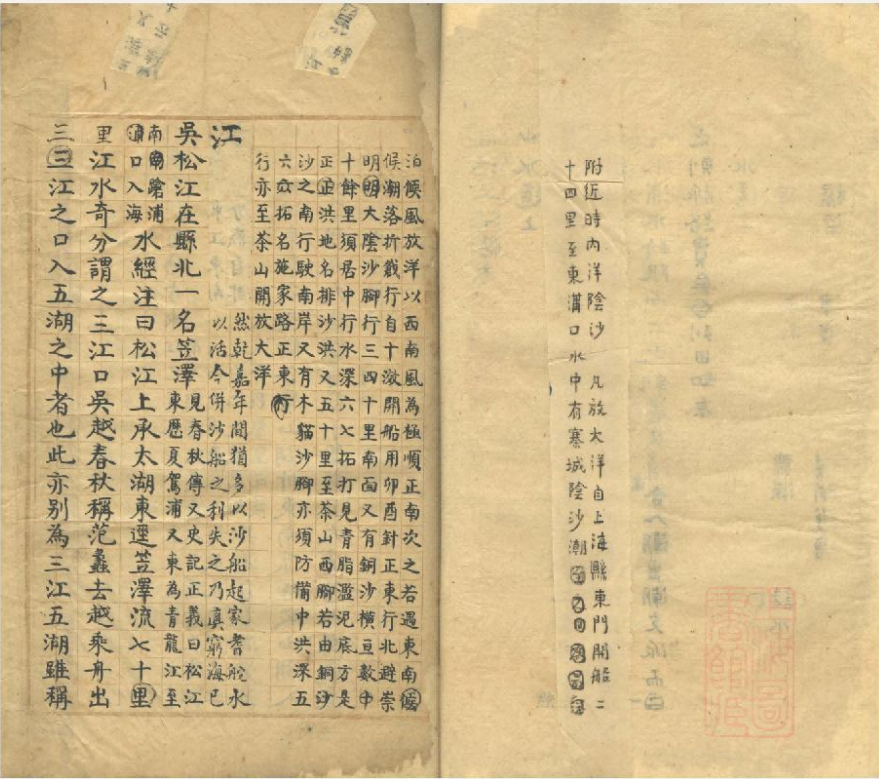
同治上海縣志 行政区划 志 俞樾 人 稿本 文獻 行政区划

同治上海縣志、水道上然邑瀕、工一環海朝朝夕、實江浙之尾閭東南已澤國地德緯、水道者查以幹河通舟楫者為經支河通潮汕者為委分、江水浦水喻限而二之則眉目青合入潮出潮支派而一日之則脈絡貫每合列日如左水道上每、吳松...工黃甫...諸之水橋梁聿、渡堰閘護唐土。

附近時內洋陰沙凡放大洋自上海縣東門開船二十四里至東溝口水中有寨城陰沙潮由○日圖每風候順正南極風放洋以為南西次泊之若遇東南候開十自用卯候潮落折行船西針正東行北避崇滋四十三有行又里南面明明大陰沙銅沙橫旦數中深七六水行中十餘里須居見青拓打脂泥底方是名正又地西沙洪五十里至茶山排銅沙並洪腳若山沙之南南有駛又行岸木貓沙腳亦須防備中洪深五六众名家正路拓施東何行亦大洋至茶放山開然乾嘉年間猶多以沙船起家者舵水江以活今併沙船之利失之乃真窮海已見春秋傳又史記正義白鶴山松文松江吳文松江在文縣兆一名笠澤東歷夏時東浦又時東為龍江至南陰浦流口等澤十里七水經注曰松江上承太湖東遷入海南浦里江水奇分謂之江口吳越春秋稱范蠡去越乘舟入三江五湖雖稱者也此亦別為三江五湖雖稱

、今名龍江蟠附近時內洋陰沙凡放大洋自上海縣東門開船二四里至東溝口水中有寨城陰沙潮留○日恩過察十風西南候放洋以極順正風泊南次南候之若遇東十為開自候落潮用卯船折行西針正東行北避崇滋四十三明里行明沙大腳南面又陰有銅沙拓打脂泥方泥底是正並洪名沙洪五又里至茶山山山西十駛又中沙腳亦須防備洪深五六众名拓路東正家施府行亦至起家者舵水江以活今併沙船之利失之乃真窮海已見春秋傳又笠澤又東歷夏駕浦東為青龍江至南南陰浦局行志云大入海十里至海凡五日水奇分謂之三江口吳越春秋稱范蠡去越乘舟入三江五湖雖稱

相相亂不與職方同庚仲初揚都賦注曰今太湖東注為松江上海為東前志引史記吳濞通江淮通渠三江五湖以證鄭氏別為三江五之湖泣與松江而三是也並有吳松江非禹貢三江之一辨今見拾遺志又引說曹允儒海塘考東江已湮塞以今地言之在浙界而選之引諸志亦及蘇東金山入海者近且有以蘇東之老浦東入海而昔浦東入海故也



馆藏方志智慧服务平台

AI书志及实体图谱

上海图书馆藏方志智慧平台

方志资源 责任者 | **城(专家)

[道光]普洱府志二十卷

责任者: (清)郑绍谦纂修 版本: 清道光20年

书志模板

格式:

预览

 引用

重新生成

复制书志

书名: 普洱府志

责任者: 鄭紹謙

書籍介紹

根據所提供的圖書資訊，現對《普洱府志》一書的背景與內容進行全面、嚴謹的綜合性總結，嚴格依據所給資料，不作主觀推斷或補充未載內容：

一、書籍基本屬性與文獻定位

《普洱府志》屬中國傳統四部分類法中「史部」之重要子類——「方誌類」，具體歸入「地誌之屬」（見《中國古籍總目》分類）及「地理類—方誌類」（見上海圖書館古籍資料庫分類）。此雙重分類明確標示其性質為清代雲南省普洱府一級行政區劃的官修或私纂地方誌，是研究清代西南邊疆政區建置、社會經濟、民族風俗與地理環境的核心原始文獻。

二、編纂背景與時代語境

全部7部存世版本中，6部明確標註為「清」代版本（佔85.7%），僅1部屬「其他」朝代（未具體說明，可能為民國影抄或後世整理本，但無證據支援其為明代或更早）。結合「普洱府」建置史實：清雍正七年（1729年）正式設立普洱府，隸屬雲南承宣布政使司，至民國二年（1913年）廢府存縣。故該志必成書於雍正七年之後、清代終結之前，極可能為乾隆、嘉慶或道光年間纂修，反映清廷強化滇南邊疆治理、推行改土歸流、發展茶馬貿易（尤以普洱茶為核心）的歷史背景下所產生的地方知識體系。

三、責任者與編纂團隊考辨

主要責任者為鄭紹謙，在全部7部版本中均列名，且6部明確標為「鄭紹謙」，顯示其為全志主纂或總纂官，地位穩固；

李熙齡（3部）與李熙齡（1部）並存，其中「熙」「熙」為異體字關係（「熙」為「熙」的舊字形），可確認為同一人，即李熙齡參與了至少3-4部版本的編纂或校訂工作，應為協纂、分纂或刊刻主持者之一；

無「其他責任者」記錄，表明編纂群體高度集中於鄭、李二人，符合清代府志常由知府主導、本地士紳或幕友協理的慣例。

四、版本形態與文獻流傳特徵

版本型別：以「刻本」為絕對主體（6部，佔85.7%），僅1部為「抄本」。刻本優勢在於權威性、流通性與官方認可度，佐證該志具有較高行政地位與區域影響力；抄本存在則反映其在藏書家或學者間的手工傳播需求（如用於批校、輯佚或地方教學）。

館藏分佈：7部存世本分藏於6家省級以上公共圖書館——貴州省圖書館（2部）、上海圖書館（2部）、國家圖書館（1部）、福建省圖書館（1部）、天津圖書館（1部）。此跨地域收藏格局（覆蓋華東、華北、西南、東南）說明該志自清代以來即受全國性學術機構重視，

館藏研究

郑成功的历史背景与影响

郑成功（1624年－1662年），是中国历史上著名的军事和政治人物，以其在台湾的统治和对荷兰殖民者的抵抗而闻名。他的成功攻克台湾岛，标志着台湾历史上的一个重要转折点，深刻影响了岛上的政治格局和社会结构。

攻克台湾的历史事件

在辛丑年（1662年）夏天的五月，郑成功成功攻克台湾岛，迫使荷兰殖民者逃离。这一事件不仅结束了荷兰在台湾的统治，也为郑成功及其后代在台湾的统治奠定了基础。郑成功的军事才能和领导能力使他成为了台湾历史上的重要人物，影响了当地居民的生活和社会结构 [Data: Entities (984); Sources (68)]。

对台湾的统治与影响

郑成功及其子相继统治台湾，对岛上的居民结构产生了深远的影响。这一时期不仅改变了台湾的政治格局，也对当地的社会和文化发展起到了重要作用。郑成功的统治为台湾带来了相对的稳定，促进了经济的发展和文化的交流 [Data: Entities (984); Sources (68)]。

结论

郑成功的历史地位不仅体现在他对台湾的军事征服上，更在于他对台湾社会的深远影响。他的成功攻克台湾的事件，标志着一个新的历史时期的开始，影响了后来的政治、经济和文化发展。郑成功的事迹至今仍被后人铭记，成为台湾历史上不可或缺的一部分。

1. 郑经与台湾政治动态

- 观察习俗的变化不如在乡间 ...
- ◆，是前任宰相文恭公世孙 ...
- 的姜出现，偶然被樵夫发现 ...
- 隐于天际。 **怪兽之 ...
- 哆囉满、买猪、抹里和沙晃 ...

AI年谱服务系统

  盛宣怀年谱

时间: 1844年 (道光二十四年甲辰) 一岁 ▾ 模型: qwen2.7b ▾  


盛宣怀


翁方纲


顾炎武


钱谦益


徐禎卿


唐伯虎


翁同龢


王铎


傅山


万斯同


黄道周


赵之谦


朱彝尊


袁枚


宣继佐


查慎行


祝允明


钱大昕


文徵明

盛宣怀出生

事件内容: 出生于江苏武进, 字杏荪、幼勖等, 晚号止叟。他的出生正值中国社会由封建向半殖民地半封建社会的转型时期。

事件类型: 家庭与个人生活
相关地点: 江苏武进

1842年9月13日 (九月十三日)

事件类型: 思想与著述
相关人物: 林则徐、魏源

<https://names.library.sh.cn/ainianpu?uri=http://data.library.sh.cn/entity/person/zoxuby93fxpx6ji&name=袁枚>

AI年谱系统



演出单实体关系编目审校平台



WELCOME

演出单实体关系智能编目平台

演出单是中国式现代化艺术变迁的独特样本，演出单实体关系包含演员、剧种等丰富信息，可用于公众艺术普及和专业学术研究。上海图书馆演出单实体关系智能编目平台是在演出单大模型抽取数据基础上进行实体关系揭示的工具平台，旨在通过数字化手段与大模型技术，优化演出单编目流程，提升演出单编目效率，同时基于实体关系模型构建演出单文献知识体系，规范文本信息组织，深化多维度元数据内容挖掘，实现演出单的数据关联与知识发现。

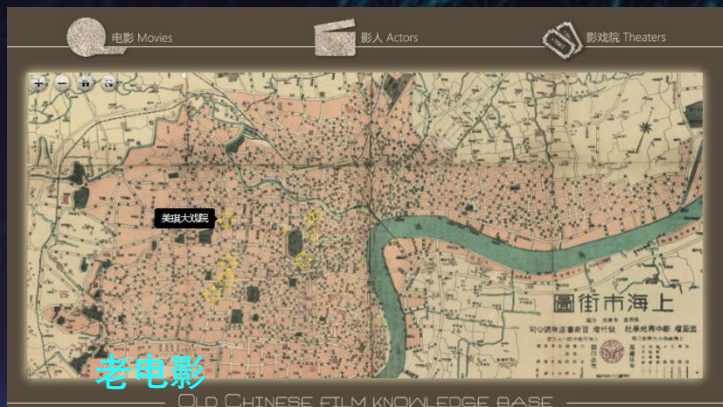
登录 →

注册 →

01/03

← →

上海年华主题数据——技术及交互创新



上海年华机构路

上海历史文化事件

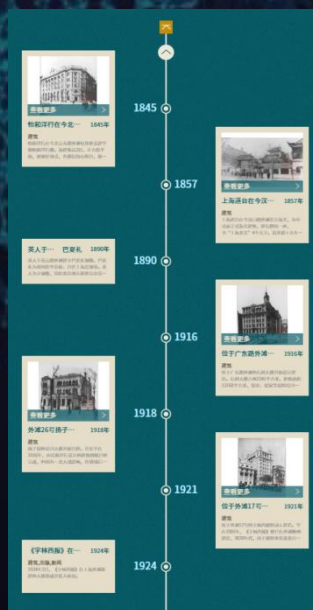
电影库数据

上海优秀历史建筑

⋮



从武康路出发



上海历史文化年谱



系列互动展项

家谱数据——动态迁徙图大屏展演

君从何处来

南北朝 420-589年 共175起迁徙事件

南汉 三国 南北朝 南 唐 五代 北宋 南宋 元 明 清 民国 当代

当前时间点各项统计数据



175

迁徙事件



58

收录姓氏



87

始迁祖



23

涉及地名

迁出、迁入统计

上海图书馆开放数据平台

<http://data.library.sh.cn>





2026 开放数据竞赛
期待您的参与

开放数据竞赛官方网站