

Sally T. WongAvery Digital Collections for Chinese Materials

Yan He (Cambridge U)
Shi Deng (UCSD)
Xi Chen (UCSD)



Presentation Outline

1. A three-year joint project between UCSD and Cambridge libraries;
2. Preparation and learning from each other through staff exchange
3. Metadata creation
4. Metadata batch and single record upload
5. Digital collections - showcase

A Three-Year Joint Project Between UCSD and Cambridge libraries

- A three-year collaborative project launched in June 2023 between the University of California San Diego Library and University of Cambridge Library
- To broaden the awareness, access, and use of the extensive East Asian collections held by the two institutions
- The sponsorship of the Avery-Tsui Foundation (Sally WongAvery and Natasha Wong)
- Digitization and metadata of Cambridge and UC San Diego's unique collections
- Librarian exchange and scholar exchange

Preparation and Learning from Each Other through Staff Exchange



Cambridge and UCSD teams with members of the Foundation during the Library exchange trip to San Diego in January 2024



Cambridge and UCSD teams during the library exchange trip to Cambridge in September 2025



Preparation and Learning from Each Other through Staff Exchange

UCSD team's library exchange trip to Cambridge in September 2025



Cambridge Visit and Cataloging Deep Dive Session

In Jan. 2024: UCSD Shared Metadata Workflow for Digital Collections

3:00-3:05 - Settle in and get to business

3:05-3:35 - Cataloging - individuals and their roles, workflows and processes

- Pre-cataloging workflows: Getting Started (Xi Chen) (3 minutes)
- DAMS Overview & Setting up (Arwen Hutt) (3 minutes)
- Cataloging: Manuscripts & Archives (Laurel McPhee) (6 minutes)
- Cataloging: Special Collections (Michelle Mascaro) (6 minutes)
- Cataloging: Chinese Digital Collections (Michelle Woo) (6 minutes)
- Post-cataloging workflows (Arwen Hutt) (6 minutes)

3:35-3:55 - Questions & discussion

CEAL Cataloging Workshop on Digital Objects, March 2024

In March 2024, The UCSD metadata workflows shared with Cambridge were expanded into a [training session](#) at the CEAL Cataloging Workshop

March 8, 2024

Cataloging Digital Objects

Arwen Hutt, Digital Object Metadata Librarian

Michelle Woo, Chinese Language Metadata Specialist

Shi Deng, Head, CJK Metadata Unit

Metadata Services, UC San Diego Library

UC San Diego

LIBRARY



Bear: model: front view, from the Stuart Collection Photographs Collection
Photographer: Gregoire, Mathieu (American conceptual artist, born 1953)

UCSD visit Cambridge and attending EASL Conference

In September 2025, UCSD presented [UCSD-Cambridge collaboration project](#) at EASL conference in London and learned about Cambridge workflows in oracle bone 3D digitization and metadata workflows.

UC San Diego

The Sally T. WongAvery Digital Collections of Chinese Materials at the UC San Diego Library

Digitization, Cataloging, Metadata Creation, and Content Overview

Xi Chen, Sally T. WongAvery Librarian of Chinese Studies and East Asia Collection
Strategist, UC San Diego Library

Shi Deng, Head, CJK Metadata Unit, UC San Diego Library

September 12, 2025 at 45th EASL Conference, the British Library

UCSD Digital Collection Metadata Workflows

Developing a digital collection is like a ball game, with digital objects as the ball passed between players. They are the game's focus, and players work together to move them toward the goal.

This model obtains descriptions directly from subject experts or creators, especially for non-text materials like photos. Then, a cataloging specialist or metadata analyst reviews and refines descriptive and subject metadata.

From Digitization to Discovery: Roles, Responsibilities, and Workflows

UC San Diego

Primary Roles

- Subject Specialist
- Project Manager
- Cataloger
- Metadata Analyst

Also...

- Donor / Data Provider
- Rights determination
- Reformatting Specialist
- Quality Assurance Reviewer

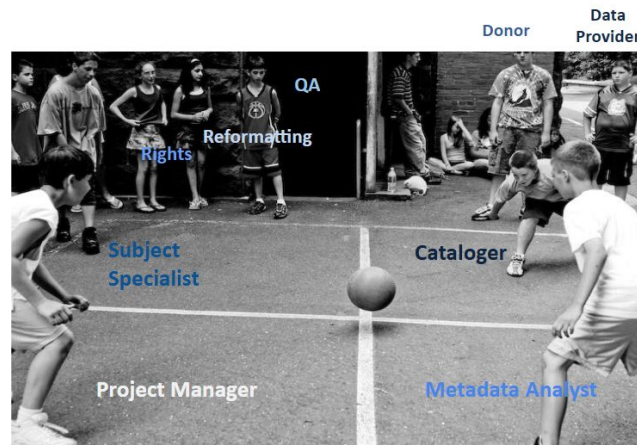


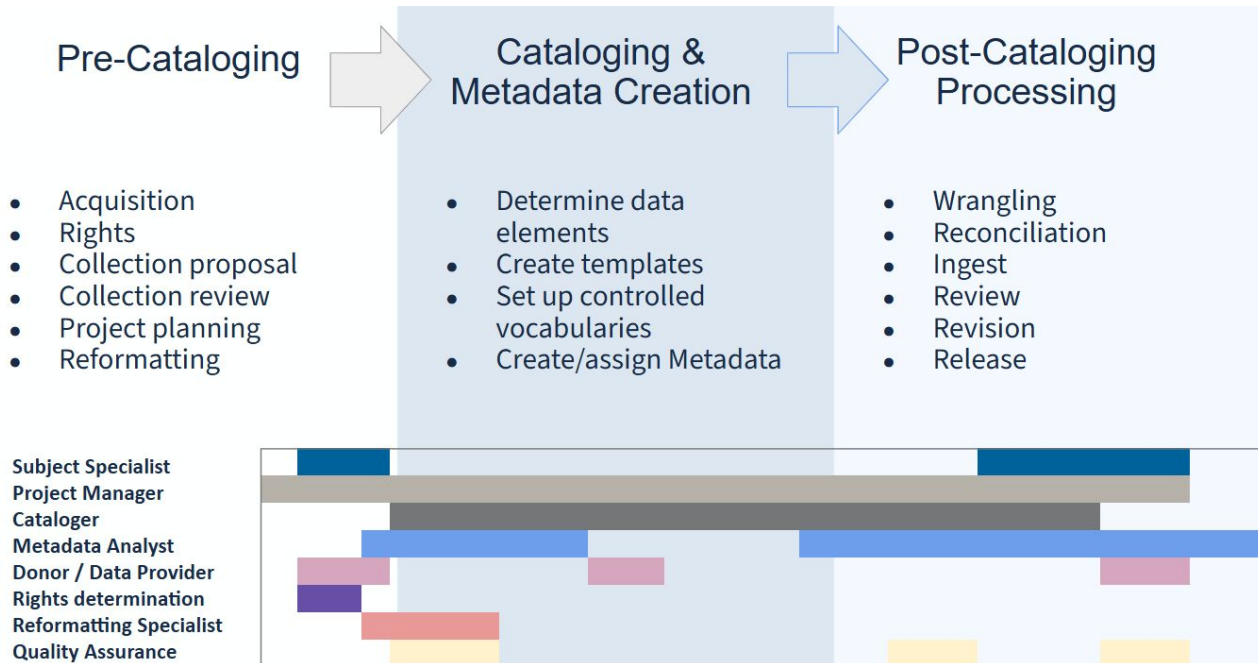
Image credit [Alex Calic](#)

UC San Diego
LIBRARY

UCSD Digital Collection Metadata Workflows

Here is an illustration of the cataloging workflow at UCSD. Different roles engage at various points, and some tasks return to certain roles at multiple stages.

It requires teamwork and coordination to succeed.



UCSD Digital Collection Metadata Workflows Models

Model 1. MARC Cataloging to Prepare Metadata

Collection: [The Anti-Confucius Book Collection](#) (Library owns physical collection, MARC 21–MODS mapping, Data Linking using Persistent ID such as library system control number)

Model 2. Spreadsheet for Metadata Creation

Collection: [We Chinese in American Magazine](#) (pre-designed template, structured for easy reformatting metadata into MODS)

Model 3. Extraction and Reuse of Existing Metadata

Collection: [Carol Pletcher Collection on the U.S. Women Scientists China Study Tour](#) (Finding Aids of this manuscripts & archives collection already developed; efficient mapping of metadata: EAD (traditional description) XML to CSV sheet; data linking back to Finding Aids)

Model 4. Utilize Commercial Tools for Metadata Creation

Collections: [1973 Guardian Tour of China](#), [Committee of Concerned Asian Scholars Friendship Delegations](#) (JSTOR Forum Easy for donor and catalogers to add descriptions, efficient to build controlled list for adding metadata for each digital object; metadata and controlled terms can be extracted and reformatted for batch ingesting into the Library Digital Collection)

Model 5. Utilize AI API for Batch Metadata Creation

Collection: Thea Breggren Photo Collection (Very limited description, rely on AI to generate descriptive and subject metadata)

Utilize AI API for Batch Metadata Creation

- Collaborative project within UC and within UC San Diego Library
- Learned from Haiqing Lin of UCB (October 2025-February 2026)
 - His presentation on AI Tool he developed at the UC Forum
 - His demo to UC San Diego Library staff
 - His AI prompts for generating description and LCSHs he generously shared
- Developed by UC San Diego Library IT staff and collaborated in a team effort (February 11-March 6, 2026)
 - Identified our needs for batch metadata creation, preferable using spreadsheet and local controlled lists, determined metadata elements needed for the end results
 - Developed and tested prompts with limited information (Dec. 1980 and six locations) as well as Haiqing Lin's prompt,
 - Team effort in the iterative fine-tuning and testing prompts with different models, and batch output the results
- Resulted in rich metadata for titles, descriptions, and subject headings, as well as inscriptions from the images, rating of confidence, reasoning, etc.
- Keep learning from the process: building sustainable practices, make it accessible for all, etc.

Utilize AI API for Batch Metadata Creation (example)

Test and fine-tune the metadata output:

-title and description in both Chinese and English

-subject headings with reasoning and uncertainty

Title	Title (Chinese)	Date	Description	Description (Chinese)
Children performing group calisthenics exercises in a schoolyard	儿童在校园中做集体广播体操	December 1980 – January 1981	A group of children performs synchronized stretching exercises in an outdoor paved courtyard, bending sideways with arms extended overhead. The children wear heavy padded winter jackets, mostly in navy blue, along with caps and layered clothing. Some girls wear floral-patterned coats and red hair bows. A brick wall is visible in the background. The exercise appears to be radio calisthenics (广播体操), a standard daily practice in Chinese schools.	一群儿童在室外铺砌的院子里进行集体伸展运动，身体侧弯，双臂举过头顶。孩子们穿着厚实的冬季棉衣，多为深蓝色，戴着帽子，衣着多层。部分女孩穿着碎花棉衣，扎红色头花。背景可见一面砖墙。该活动似为广播体操，是中国学校的日常集体锻炼项目。



Location	Location Reasoning	Subject Headings	FAST Topic Headings	Local Keywords	Uncertainty
Beijing, Luoyang, or Xi'an (low confidence)	The heavy padded winter clothing and caps indicate a cold climate, consistent with northern cities such as Beijing, Luoyang, or Xi'an during December–January. No signage or distinctive architectural features allow further narrowing.	Physical education and training--China; Children--China; Schools--China; Stretching exercises--China; Clothing and dress--China--History--20th century	Radio gymnastic exercises Children Stretching exercises School children Physical education and training Courtyards	radio calisthenics, guangbo ticao, schoolyard exercises, padded jackets, children's clothing	Exact city cannot be determined from the image. Whether this is a school courtyard or another communal space is assumed from context but not confirmed by signage. The specific exercise routine (radio calisthenics vs. another form) is inferred from the synchronized posture but cannot be confirmed.

Sally T. WongAvery Digital Collection of Chinese Materials

“要知天下事，须读古人书。”

*To know all that happens under the sun, be sure
to read books by our ancestors.”*

— Feng Menglong, *Stories to Awaken the World*



The Sally T. WongAvery Digital Collection of Chinese Materials grows out of the WongAvery East Asian Collections Collaboration between Cambridge University Libraries and UC San Diego Library (June 2023 to June 2026).

This digitisation project has been made possible through the generous sponsorship of the Avery-Tsui Foundation in honour of Sally T. WongAvery. With their support, we have been able to preserve and share digitally an array of items from the University Library's Chinese collections, which are now accessible for the first time to both scholars and the public worldwide. We believe that this project will have a profound impact on research in fields such as history, archaeology, palaeography, linguistics, art, religion and geography, while simultaneously expanding opportunities for scholarship, education, and engagement with collections.

The selection of materials for digitisation has been carefully curated to represent pre-modern and late imperial history of China and highlight significant moments and developments in Chinese culture and intellectual life. The collection includes oracle bones, the earliest known form of Chinese writing, offering a window into ancient Chinese divination practices and valuable insights into language development. The rare books and manuscripts encompass some important literary, artistic, and historical texts from different periods. We have also chosen to digitise a selection of single-sheet prints and documents, maps as well as an array of hanging scrolls and examples of Chinese calligraphy. This selection of diverse materials is but a glimpse into the breadth and the significance of the Chinese collections at Cambridge University Library, which reflect the intellectual, artistic, religious and cultural history of China over millennia.

1 2 3 4 5 ... 79

剑桥大学金璋甲骨收藏 (CUL1)

Over three thousand years old,
the Chinese Inscribed Oracle
Bones in the Hopkins Collection
are by far the oldest written ...
[more](#)



剑桥大学金璋甲骨收藏 (CUL2)

Over three thousand years old,
the Chinese Inscribed Oracle
Bones in the Hopkins Collection
are by far the oldest written ...
[more](#)



剑桥大学金璋甲骨收藏 (CUL3)

Over three thousand years old,
the Chinese Inscribed Oracle
Bones in the Hopkins Collection
are by far the oldest written ...
[more](#)



剑桥大学金璋甲骨收藏 (CUL4)

Over three thousand years old,
the Chinese Inscribed Oracle
Bones in the Hopkins Collection
are by far the oldest written ...
[more](#)



Cambridge Metadata batch and single record upload

1. The Hopkins Oracle Bone Collection and the Taiping Heavenly Kingdom Collection were processed using a batch upload workflow. Metadata were first created in Excel spreadsheets and then transformed into TEI/XML files in batches. These files were subsequently reviewed and manually edited to incorporate special features or additional information where necessary.
2. For rare books and rubbings, the XML files were edited directly in Oxygen using the TEI standard.
3. For minor corrections or small updates, edits were made directly through the content loader in the backend management system of the Cambridge Digital Library.

Highlights of Functions and Features:

1. **Cross-collection search** across 80+ digital collections at the Cambridge Digital Library.
2. **Browsable categories and internal subject headings** to facilitate thematic exploration.
3. **IIIF compatibility** enabling image reuse for digital exhibitions, annotation, and research.
4. **User-curated collections**, allowing scholars and educators to assemble digital items for teaching and exhibitions.
5. **Cross-linking of related materials** across collections to reveal historical connections.
6. **Rich item-level metadata**, including descriptions, transcriptions, and translations where available.
7. **Multimedia support**, including 3D images, embedded videos, and other digital formats.

Categories

[Oracle Bone Collection](#)

[Rubblings](#)

[Rare Books](#)

[Taiping Heavenly Kingdom Collection](#)

[Manuscripts](#)

[Painting and Calligraphy](#)

[Painting Manual Collection](#)



Cambridge UL Oracle Bone CUL1 on Sketchfab

In 1982 Mme. Qi Wenxin 齐文心, an oracle bone specialist from the Chinese Academy of Sciences in Peking, came to Britain to make rubbings of all the oracle bones in public and private collections in the UK, which are the best in Europe. These have been published in five large volumes and include the whole of the Hopkins Collection. During her stay in Cambridge a video of Mme Qi made in which she demonstrates the process of making rubbings of the oracle bones and talks to them.



How to make rubbings of Chinese Oracle Bones (YouTube)

Hopkins Oracle Bone, CUL.1

Chinese Collection, Cambridge University Library

- **Physical Location:** Cambridge University Library
- **Classmark:** CUL1
- **Alternative Title(s):** Hopkins Oracle Bone Collection CUL1
- **Subject(s):** [Inscriptions, Chinese](#); [China -- Civilization](#); [Astronomy and Calendar](#); [Oracle Bone Collection](#)
- **Origin Place:** [Yinxu 殷墟, Anyang 安阳 \(China\)](#)
- **Date of Creation:** 1400-1200 BCE
- **Language(s):** Chinese
- **Former Owner(s):** [Hopkins, Lionel Charles, 1854-1952](#)
- **Note(s):** [Full list of contents](#); The inscription on the verso has been identified as the lunar eclipse of [27-28 December 1192 BC](#) (Dubs 1951, Chou 1965). Other possibilities include the eclipses of [24 November 1311 BC](#) (董作宾 1950, Pang 1989) and [16 November 1218 BC](#). See especially [张培瑜 1999:3 p.60.](#); This bone bears an inscription which has been identified as the lunar eclipse of 27-28 December 1192 BCE.;
- **Theme:** [Astronomy and Calendar](#)
- **Physical description:** Inscriptions on obverse partially damaged.

- **Extent:** height: 115 mm, width: 23 mm, depth: 9.5 mm. Weight: 14.7 g
- **Material:** Ox scapula
- **Script:**
 - 甲骨文 Jiaguwen
- **Provenance:** Bequeathed to the University Library by Lionel Charles Hopkins (1854-1952) in 1952. Purchased by Hopkins from Frank Herring Chalfant (1862-1914).
- **Funding:** Digitized with the generous support of the Avery-Tsui Foundation
- **Data Source(s):** F.H. Chalfant: *The Hopkins Collection of Inscribed Oracle Bone*, New York, 1939
- **Bibliography:**
 - 董作宾 殷代月食考 中央研究院历史语言研究所集刊 22 (1950:7) 139-160
 - 董作宾 殷代月食考 中央研究院历史语言研究所集刊 22 (1950:7) 139-160
 - H. Dubs 'The date of the Shang period' *Toung Pao* 40 (1951) 322-335
 - Chou Fa-kao 'On the dating of a lunar eclipse in the Shang period' *Harvard Journal of Asiatic Studies* 25 (1965) 243-247
 - R. Newton *A canon of lunar eclipses for the years -1500 to -1000* Maryland 1977
 - K.D. Pang et al. 'Shang Dynasty oracle bone eclipse records and the earth's rotation rate in 1302 BC' *Bulletin of the American Astronomical Society* 21 (1989) 753
 - Xu Zhentao et al. 'Astronomical records on the Shang dynasty oracle bones' *Archaeoastronomy* 14 (1989) S61-71
 - Xu Zhentao et al. 'Astronomy on oracle bone inscriptions' *Quarterly Journal of the Royal Astronomical Society* 36 (1995) 397-406
 - 刘学顺 殷墟卜辞所记月食的年代 殷都学刊 67 (1998:1) 21-23,72
 - 张培瑜 甲骨文日月食与商王武丁的年代 文物 514 (1999:3) 172-198
 - 张培瑜 日月食卜辞的证认与殷商年代 中国社会科学 119 (1999:5) 56-63
 - 王宇信等 甲骨文月食刻辞的考辨 甲骨学一百年 北京 1999.9 636-647
 - 彭裕商 宾组卜辞五次月食的先后次序 中原文物 (2003:3) 111 44-47,84
 - Espenak & Meeus *Five millennium canon of lunar eclipses* NASA 2009

XML Metadata template (TEI standard) for cataloging rare books

Tag Name (TEI)	Field displayed in CUDL	Content	Tag and content displayed in CUDL
<title>太平天日</title>	Uniformed Title	太平天日	brief view in the search result)
<title type="alt">Tai ping tian ri</title>	Alternative Title(s)	Tai ping tian ri	Alternative Title(s): Tai ping tian ri
<author> <persName key="person_v72194171" ref="htl Hong, Xiuquan, 1814-1864</persName> </author>	Author(s)	洪秀全 Hong, Xiuquan, 1814-1864	Author(s): 洪秀全 Hong, Xiuqu
<editor key="person_v74620220" ref="http://viaf.org/viaf/74620220" role="fmo">Wade, T. F. (Thomas Francis), 1818-1895</editor>>	Former Owner(s)	Wade, T. F. (Thomas Francis), 1818-1895	Former Owner(s): Wade, T. F. (Thomas Francis), 1818-1895
<textLang mainLang="chi">Chinese</textLang>	Language(s)	Chinese	Language(s): Chinese
<note> 框20.8 x 13公分, 8行20字, 白口, 四周單邊, 單黑魚尾, 版心上鑄書名。 </lb></note>	Note(s)	框20.8 x 13公分, 8行20字, 白口, 四周單邊, 單黑魚尾, 版心上鑄書名。	Note(s): 框20.8 x 13公分, 8行20字, 白口, 四周單邊, 單黑魚尾, 版心上鑄書名。



Metadata:

1. VIAF for name authority
2. LCSH for subject headings
3. Local SH/NAF for internal search





[滿漢雙文雍正特恩誥
命碑]

Classmark:

(FH.990.137(1)-(4))



TEI/XML for rubbing collection

Examples of various features we use:

1. Decoration: e.g. linkable seals. `<DecoDesc>` and `<decoNote>`
2. Multi-parts for one item (classmark): `<msPart>`
3. Table of Contents (linkable): `<msContents>`, `<msItem>`
4. Thumbnails (linkable): `<facsimile><Graphic><surface>` (automatically generated in CUDL)
5. Transcription: `<text><div>`, `<pb>`, `<p>`, `<lb/>`

Joint Digital Collections Webpage

Sally T. WongAvery Digital Collections: A Three-Year Collaboration at UC San Diego and Cambridge University

Search this Guide Search

Sally WongAvery Digital Collections at UC San Diego and Cambridge University - A Three-Year Collaboration

The Sally T. WongAvery Digital Collection grew out of the WongAvery East Asian Collections Collaboration between Cambridge University Libraries and UC San Diego Library (June 2023 to June 2026).

This digitisation project has been made possible through the generous sponsorship of the Avery-Tsui Foundation in honour of Sally T. WongAvery. With their support, we have been able to preserve and share digitally an array of items from the two university libraries Chinese collections, which are now accessible for the first time to both scholars and the public worldwide. We believe that this project will have a profound impact on research in fields such as history, art, art history, anthropology, sociology, archaeology, palaeography, linguistics, religion and geography, while simultaneously expanding opportunities for scholarship, education, and engagement with collections.

UC San Diego Library

Chinese News

The Chinese News, founded by Mr. Stanley Ting in 1983, distributed news and advertising to the Chinese-American population in San Diego until 1998. The weekly newspaper contains local and international news articles written in Chinese in addition to advertising in English and Chinese.

We Chinese

Founded in 2001, We Chinese is a monthly Chinese-language magazine published in San Diego, featuring news concerning Chinese Americans in Southern California and the San Diego area. The magazine highlights prominent figures of the community and provides information about local commerce and services.

Carol Fletcher Collection on the U.S. Women Scientists China Study Tour, 1981

Papers and photographs documenting Carol Fletcher's trip with the U.S. Women Scientists China Study Tour in 1981, including her travel journal. The tour group met with professional Chinese women, were introduced to officials of the People's Republic who formulated science and technology policy, and travelled to Shanghai, Wuji, Nanjing, Sian, and Beijing.

Guardian Newspaper Tour

The National Guardian Newspaper was founded in 1948 in New York City as an independent left-wing weekly publication. With a change in editorial leadership in the 1970s, the newspaper was rebranded as The Guardian and intensified their focus on reporting in Southeast Asia and China in particular. The newspaper sponsored several tours in China during the years 1972 and 1973. The tour participants were recruited from newspaper subscribers and by word of mouth.

Anti-Confucius Book Collection

The Anti-Confucius Book Collection consists of 747 titles donated by UCSD PhD alumnus Matthew Wills. They were largely political propagandas produced during 1960s-1970s, also known as the Cultural Revolution period in China. Many of the titles are compiled by local propaganda entities that discontinued after 1976. Materials in this collection reflect a 1970s campaign to reinforce the power of political elites and

Cambridge University Library

Oracle Bone Collection

Including 600+ Hopkins Oracle Bone collection in 2D format, 50 bones in 3D format and 70+ additional oracle bones in 2D format.

Rubblings

The group of rubbings particularly present the bilateral nature of materials including Chinese and Phags-pa, Syriac, Uyghur, traditional Mongolian script, Manchu script as well as Quasi-archaic Chinese script, and various religions including Confucianism, Daoism, Buddhism, Islam, Catholicism, Nestorianism, Daoism, and Buddhism.

Rare Books

Selected rare books from Chinese collections which highlights the University's long-standing scholarly strengths in philology, Classical Chinese studies, printing history, dialect studies, and folk religion. Selection criteria have included rarity, uniqueness, research and aesthetic value, as well as relevance to the history and communities of the University of Cambridge.

Taijping Heavenly Kingdom Collection

Part of the collection of Sir Thomas Francis Wade and contains a range of propaganda publications produced by the Taijping Rebellion's insurgent regime. These materials offer rare insight into the political ideology, religious rhetoric, and communication strategies of the Taijping Heavenly Kingdom.

Manuscripts

Comprise a small but significant group of handwritten Chinese texts preserved at Cambridge University Library. These manuscripts include unique or rare copies of historical, literary, and religious works. Manuscripts such as *Tianyuan yu li xiang yi fu* containing hand-drawn illustrations or annotations, offering valuable evidence for the transmission and material culture of Chinese texts.

Painting and Calligraphy

Including a long calligraphy handscroll, rubbing calligraphy and paintings, reflecting the central role of brushwork in Chinese artistic and scholarly culture. Produced by scholars and artists across different periods, the pieces demonstrate

https://ucsd.libguides.com/ucsd_cambridge_wongavery

Anti-Confucius Book Collection

The Anti-Confucius Book Collection consists of 747 titles donated by UCSD PhD alumnus Matthew Wills. They were largely political propagandas produced during 1960s-1970s, also known as the Cultural Revolution period in China. Many of the titles are compiled by local propaganda entities that discontinued after 1976. Materials in this collection reflect a 1970s campaign to reinforce the power of political elites and affirm the absolute correctness of Maoist socialism.

Carol Shankel Photo Slides

The 740 slides document six trips taken by a group of faculty, students, and staff from the University of Kansas to China in the early 1980s. The delegation was among the first groups of U.S. visitors to China. Their first two trips, in 1980 and 1982, are especially significant. The sites they visited are rich in archaeological value, reflecting the academic interests of the group.

Thea Berggren Photographs

A photographer and visual artist, Thea Berggren traveled to China in 1980 and took more than 800 photos. They capture and tell the story of the Chinese people in Beijing, Luoyang, Kunming, Xi'an, Chengdu and Guangzhou.

as *Tianyuan yu li xiang yi fu* containing hand-drawn illustrations or annotations, offering valuable evidence for the transmission and material culture of Chinese texts.

Painting and Calligraphy

Including a long calligraphy handscroll, rubbing calligraphy and paintings, reflecting the central role of brushwork in Chinese artistic and scholarly culture. Produced by scholars and artists across different periods, the pieces demonstrate the aesthetic principles and cultural significance of Chinese literati art, where painting and calligraphy were closely intertwined forms of expression.

Painting Manual Collection

Including *Jiziyuan hua pu* in two different editions and *Ba zhong hua pu* published in Japan. It consists of illustrated manuals used for the teaching and transmission of traditional Chinese painting techniques. These works typically present step-by-step visual models of subjects such as flowers, birds, landscapes, and figures, serving as practical guides for artists and students. Many of these manuals belong to the rich tradition of woodblock-printed painting guides that circulated widely from the Ming and Qing periods onward.



Last Updated: Mar 6, 2025 4:21 PM | URL: https://ucsd.libguides.com/ucsd_cambridge_wongavery | [Print Page](#)

Login to LibApps
Report a problem.

Q&A

謝謝

Thank you!

