

Wikidata vs. custom Wikibases: community history case studies

Amy Ruskin
Data Engineer, Digital Scholarship Group
Northeastern University

Boston Research Center

- Digital community history and archives lab based in the Northeastern University Library
- Projects in partnership with organizations in Chinatown, Roxbury, East Boston, and the South End
- Our role: provide **technical infrastructure** for **local history**



<https://bostonresearchcenter.org/>



Wikibase: Chinatown Collections Survey



“Mural on Oak Street in Boston's Chinatown”, [ca. 1976-2008], <http://hdl.handle.net/2047/d20180779>

Background and goals

- Initial meetings between BRC and community organizations in Chinatown in spring 2021
- Lots of digital history projects already happening in community
 - BRC could help make those more findable
 - Launching point for neighborhood research
- Project goal: create **searchable** and **browsable multilingual directory** of collections

Boston Chinatown Collection Survey

- Develop survey
- Review survey with partners
- Translate survey (English, simplified Chinese, traditional Chinese)
- Publicize survey
 - Blog posts on BRC and Boston Public Library (BPL) websites
 - Direct outreach to people/organizations
- Gather initial responses in spreadsheet
- Re-invite non-respondents
- Update community partners
- Republicize survey via community partners

Boston Chinatown Collection Survey (English)

What is your name?

What is your email address?

What is your phone number?

Name of repository, organization, or individual owner of collections:

Please describe the collection(s) that pertain to Boston's Chinatown. For example, what topics, people, events, places, and organizations are represented? What are the format or formats of these materials (print, manuscript, publications, photographs, cassettes, CDs, digital files, etc.)?

To what degree have these collections been processed (that is, reviewed, and organized or cataloged)?

How have they been made accessible to researchers and/or the public?

To what extent have they been digitized?

What are your hopes, dreams, and concerns for these collections?

As we think about bringing these collections together in an online catalog, what features do you think would be most helpful to users? For example, ability to search for keywords, filter by date, etc.

Do you have items that are specifically relevant to the immigration history of Chinatown?

As we think about bringing these collections together, what else should we consider?

Would you like to participate further in this inventory project? If so, how?

Would you like to receive monthly email updates about community history projects and events?

Initial survey responses

Name of repository, organization, or individual...	Please describe the collection(s)	Processed?	Accessible?	Digitized?	Hopes, dreams, cor
Cambridge Historical Commission	One research file containing a copy of the National Register of Historic Places Multiple Property Documentation Form for Boston's Chinatown (dated June 1, 2016)	Fully processed	Yes	Not digitized	That researchers m
Moakley Archive, Suffolk University	Harry Hom Dow Papers: Dow, a son of Chinese immigrants, was raised in Massachusetts during a time when the Chinese Exclusion Act of 1882 was in place and anti-Chinese sentiment and discrimination were common. In spite of those challenges, the Suffolk Law School graduate became the first Chinese-American to pass the Massachusetts Bar exam, served in the military, opened a private law practice, and was a noted community activist. Dow, a son of Chinese immigrants, was raised in Massachusetts during a time when the Chinese Exclusion Act of 1882 was in place and anti-Chinese sentiment and discrimination were common. In spite of those challenges, the Suffolk Law School graduate became the first Chinese-American to pass the Massachusetts Bar exam, served in the military, opened a private law practice, and was a noted community activist. Consult the Dow Paper's to learn more about his collection.	Partially processed. We are dealing with a few challenges - some of files are Dow's client files from his law practice which cannot be made public without the consent of the clients and there are other materials solely written in Chinese that we don't currently have resources to translate into English.	yes, but only a small portion	a small portion have been digitized.	Hoping to have all t
Schlesinger Library, Radcliffe Institute, Harvard University	1. Chinese American Women Oral History Project (http://id.lib.harvard.edu/alma/990118374940203941/catalog) - audio cassettes, videotapes, paper and electronic transcripts 2. Cheng Imm Tan papers (http://id.lib.harvard.edu/alma/99153757003203941/catalog) - paper (manuscript), photographs, audio cassettes, videotapes, digital files	Both collections have basic descriptions in the Harvard Library catalog, but are not yet fully cataloged or inventoried.	Some clips of the Chinese American Women Oral History Project are available on a website from an exhibit: https://projects.iq.harvard.edu/chineseamericanwomen/oral-history-project	Not at all.	I'd like them to get p
Norman B. Leventhal Map & Education Center at the	Maps of Boston from 19th century to present that show Chinatown, atlases showing buildings and often owners from 1870s to 1980s, plans/maps from the		Can be searched in Boston Public Library online catalog, can be viewed by visitors to	Most maps that are not under copyright	

Data models and controlled vocabularies

- What types of entities do we need to represent?
- What fields do we use to represent each type of entity?
 - What controlled vocabularies do we need to populate those?
- What relationships between entities should we capture?
- [Meta: is this possible given the information in the survey responses? What user needs does this serve?]

Key entity types to model

- Collections

- Archival collections
 - Fonds/series/files with specific relevance to Boston's Chinatown (e.g., Chinatown Progressive Association records)
 - Fonds/series/files that happen to include materials related to Boston's Chinatown (e.g., the FayFoto Boston photographs)
- Digital projects/websites
- Personal collections

- Organizations

- Maintainers of collections
- Donors of records in collections
- Creators of records in collections

- People

- Maintainers, donors, creators, and/or contact points of collections
- Founders of organizations

Controlled vocabularies

- Type of collection
 - archival collection, website, personal collection
- Type of materials included (general)
 - National Archives General Records Type Authority List
- Type of materials included (specific)
 - Getty Art and Architecture Thesaurus, Information Forms hierarchy
- Access status
 - National Archives Access Restriction Status Authority List
- Digitization status
 - born-digital, fully digitized, partially digitized, not digitized
- Subjects
 - Student coding survey responses assigned tags
- Type of organization
 - community organization, historical society, municipal archive, academic archive

Survey responses coded in Airtable (tabular view)

Chinatown Collections						
Data Automations Interfaces						
Collections Organizations People + Add or import Extensions						
Views Shareable view 10 hidden fields Filter Group Sort Color Share view						
	Title	Description (modified)	Collection type	Maintainer (organization)	Maintainer (person)	Material types (general)
1	[Carmen Chan's Sampan photos]	This collection contains photographs taken by Carmen Chan...	personal collection		Carmen Chan	Photographs and other Grap...
2	Asian Community Development Corporation (Jac...	The Asian Community Development Corporation's recor...	archival collection	Northeastern University Archives and Special Collections		Textual Records
3	Boston Greenspace Alliance records	This collection includes one folder of materials titled Chl...	archival collection	University Archives and Special Collections, Joseph P. Healey		Textual Records
4	Boston Urban Gardeners records	This collection includes folders of materials on Chinato...	archival collection	University Archives and Special Collections, Joseph P. Healey		Maps and Charts Textual F...
5	Caroline Chang collection	This is one of CHSNE's many collections. It currently ha...	archival collection	Chinese Historical Society of New England		
6	Cheng Imm Tan papers	This collection contains audio, videos, pictures, and tea...	archival collection	Radcliffe Institute, Harvard University		Photographs and other Grap...
7	Chinatown Branch, Boston Public Library	This is one of the 25 Boston Public Library branches. It i...				
8	Chinatown Story Cart	This is an art project created 2020-2021. In collaboratio...	website	Asian Community Development Corporation	Lily Xie Crystal Bi Wegner	Photographs and other Grap...
9	Chinese American Experiences Mass. Memories ...	Chinese American Experiences Mass. Memories Road S...	archival collection	University Archives and Special Collections, Joseph P. Healey		Moving Images Photograp...
10	Chinese American Women Oral History Project	Collection of oral histories that center around life in Chi...	archival collection	Radcliffe Institute, Harvard University		Sound Recordings Moving...
11	Chinese Consolidated Benevolent Association of ...	Collection of primarily physical documents related to Ch...	archival collection	Chinese Consolidated Benevolent Association of New England		Photographs and other Grap...
12	Chinese Progressive Association records	The Chinese Progressive Association records document...	archival collection	Northeastern University Archives and Special Collections		Textual Records Photograp...
13	FayFoto Boston photographs	Collection of photographs of Boston, including aerial ph...	archival collection	Northeastern University Archives and Special Collections		Photographs and other Grap...
14	Harry Hom Dow Papers	Dow, a son of Chinese immigrants, was raised in Massac...	archival collection	Moakley Archive & Institute		Textual Records
15	Haymarket People's Fund records	This collection includes material on the Chinatown Peop...	archival collection	University Archives and Special Collections, Joseph P. Healey		Textual Records
16	Hudson Street Chronicles	This is a collection of photos, stories, and essays on Cy...	website		Cynthia Yee	Photographs and other Grap...
17	International Society Records	The International Society originally grew out of the Chin...	archival collection	Northeastern University Archives and Special Collections		Textual Records Photograp...
18	Newspapers at Northeastern University Archives ...	This collection contains issues from Sampan, and other ...	archival collection	Northeastern University Archives and Special Collections		Textual Records Photograp...
19	Jeffrey Wong type sets	This is a personal collection of Yu-Wen Wu, who is in po...	personal collection		Yu-Wen Wu	Artifacts
20	Judy Branfman papers and "Standing Our Ground...	This collection includes a Chinatown user survey from 1...	archival collection	University Archives and Special Collections, Joseph P. Healey		Photographs and other Grap...
21	Kwong Kao Chinese School, Gung Ho Club	One of CHSNE's many collections. This collection has n...	archival collection	Chinese Historical Society of New England		Textual Records
22	Norman B. Leventhal Map Center Collection	Maps of Boston (including Chinatown) from the 1800s t...	archival collection	Norman B. Leventhal Map & Education Center at the Boston P...		Maps and Charts
23	Margaret Wong Collection [restaurantur]	One of CHSNE's many collections. This collection has n...	archival collection	Chinese Historical Society of New England		
24	Menu Collection	Collection of menus from Chinese restaurants in Boston'...	archival collection	Chinese Historical Society of New England		Photographs and other Grap...
25	Mong Q. Lee collection	This collection contains most of Mong Q. Lee's belongin...	archival collection	Chinese Historical Society of New England		Photographs and other Grap...
+ Add...		37 records				

Survey responses in Airtable (form view)

A Title ⓘ

Chinese Progressive Association records

Description ⓘ (modified) ⓘ
The Chinese Progressive Association records document activism and advocacy in Chinatown and in Greater Boston's Chinese and Asian American community from the 1970's through the 2000's. Topics include urban development, immigrants' rights, workers' rights, race-related violence, tenants' rights, and Chinese cultural activities. Includes administrative records, photo, video, and clippings.

Collection type ⓘ
archival collection

Maintainer ⓘ (organization)
Northeastern University Archives and Special Collections
ALTERNATE NAMES SHORT DESCRIPTION WIKIDATA URI
NUASC university archives f... http://www.wikidata.org/...
+ Add record

Maintainer ⓘ (person)
+ Add record

Material types ⓘ (general)
Textual Records ✕ Photographs and other Graphic Materials ✕ +

Material types ⓘ (specific)
correspondence ✕ memorandums ✕ booklets ✕ programs ✕
handbooks ✕ fact sheets ✕ fliers ✕ brochures ✕
pamphlets ✕ newspaper clippings ✕ reports ✕ interviews ✕
posters ✕ VHS ✕ photographs ✕ slides ✕ minutes ✕
agendas ✕ +

Access Restriction Status ⓘ
Restricted - Partly

Digitization status ⓘ
Partially digitized

Access policy notes ⓘ

Project status ⓘ
Further accruals expected

Project start year ⓘ
2008

Project end year ⓘ

Materials start year ⓘ
1976

Materials end year ⓘ
2006

Subjects ⓘ
urban development ✕ unionization ✕ unemployment rights ✕
immigrants' rights ✕ workers' rights ✕ race-related violence ✕
tenants' rights ✕ bilingual education ✕
Chinese cultural activities ✕ strike organizing ✕
retraining of adult workers ✕ +

Language of materials ⓘ
English ✕ Chinese ✕ +

Collection URL ⓘ
<http://hdl.handle.net/2047/D20232667>

Collection description URL ⓘ
<https://archivesspace.library.northeastern.edu/repositories/2/resources/946>

Records creator ⓘ (organization)
Chinese Progressive Association
ALTERNATE NAMES SHORT DESCRIPTION WIKIDATA URI
organization based ... http://www.wikidata.org/...
+ Add record

Records creator ⓘ (person)
+ Add record

Records donor ⓘ (organization)
+ Add record

Records donor ⓘ (person)
Lydia Lowe
WIKIDATA URI PUBLIC FIG... CREATOR OF RECORDS... CONTACT FOR
http://www.wikidata... ✓ Chinese Women Worl Chinese Wo
+ Add record

Contact ⓘ
Molly Brown
WIKIDATA URI PUBLIC FIG... CREATOR OF RECORDS... CONTACT FOR
Chinese Pro
+ Add record

Database needs and nice-to-haves

Needs:

- Multilingual capabilities—represent all field names and values in both English and Chinese
- Represent collections, people, organizations, and the relationships between them

Nice-to-haves:

- Lightweight/easy to implement
- Editing interface that allows project participants to add/edit data directly
- Flexibility in schema to expand project scope, deal with idiosyncratic entities

Why Wikibase over Wikidata?

- Not all of the items are appropriate for inclusion in Wikidata
 - Notability guidelines
 - Privacy concerns
- More control over modeling practices
 - Don't have to inherit Wikidata modeling efforts that don't make sense in local context
 - Can add properties specific to our usage
- More control over editing access
- (Pilot project for broader use of Wikibase in the library)

Installing Wikibase suite

- Full Wikibase suite deployed in an EC2 instance on AWS with Docker
- Process of installing/configuring/testing Wikibase installation—not completely straightforward
 - Contact Rob Chavez (r.chavez@northeastern.edu) for more information
- More customization to do before making public
 - Add logo
 - Set ordering of statements on entity pages
 - Set formatURI property so identifier values appear as hyperlinks

Populating Wikibase

Automated:

- Copying labels/descriptions/aliases from Wikidata entities
- Bulk-creating entities from CSV files with monolingual labels and descriptions
- (See: [NEU-DSG/wikibase-utilities](https://github.com/NEU-DSG/wikibase-utilities) GitHub repository for the Python scripts used here)

Manual:

- Creating statements on items
- Adding in text from professional translator

Wikibase demo

Chinese Progressive Association records (Q40)

archival collection held at the Northeastern University Archives

 edit

[↗ In more languages](#)

[Configure](#)

Language	Label	Description	Also known as
English	Chinese Progressive Association records	archival collection held at the Northeastern University Archives	
Chinese	华人进步会记录	<i>No description defined</i>	


[Fewer languages](#)



Statements

title	 Chinese Progressive Association records (English)	
	↗ 0 references	+ add reference
	 华人进步会记录 (Chinese)	
	↗ 0 references	+ add reference
		+ add value


language	 English	
	↗ 0 references	+ add reference
	 Chinese	
	↗ 0 references	+ add reference
		+ add value

access restriction status	 partly restricted access	
	↗ 0 references	+ add reference
		+ add value

materials start date	 1976	
	↗ 0 references	+ add reference
		+ add value

materials end date	 2006	
	↗ 0 references	+ add reference
		+ add value

corresponding Wikidata item	 Q109364280	
	↗ 0 references	+ add reference
		+ add value

instance of	 Chinatown Collection	
	↗ 0 references	+ add reference
		+ add value

records creator	 Chinese Progressive Association	
	↗ 0 references	+ add reference
		+ add value

maintainer	 Northeastern University Archives and Special Collections	
	↗ 0 references	+ add reference
		+ add value

description	 The records in this collection document activism and advocacy in Chinatown and in Greater Boston's Chinese and Asian American community from the 1970's through the 2000's. Materials include administrative print records, public relations records, photo, video, and clippings. (English)	
	↗ 0 references	+ add reference
	 华人进步会记录记载了从 1970 年代到 2000 年代在华埠和大波士顿华人和亚裔美国人社区的盛行主义和倡议活动。主题包括城市发展、移民权利、工人权利、种族暴力、租户权利和中国文化活动，包含行政记录、照片、视频和剪报。 (Chinese)	
	↗ 0 references	+ add reference
		+ add value

project status	 Further accruals expected (English)	
	↗ 0 references	+ add reference
	 预计会增加更多藏品 (Chinese)	
	↗ 0 references	+ add reference
		+ add value

Building front-end

- Currently using [Snowman](#)
 - (See: [NEU-DSG/chinatown-collections-site](#) GitHub repository for specific implementation)
- SPARQL queries + HTML/Go templates => static site to display data from Wikibase in a more user-friendly manner
- Functionality for:
 - Seeing full lists of collections, organizations, and people in the Wikibase in separate, sortable tables
 - Viewing individual entities and the information we have about them
 - Browsing collections organized by various categories
 - Reading longer form text about the project, data models, etc.
 - Switching between English and simplified Chinese for all of that

Site demo

Chinatown History Directory

An index of sources on Boston's Chinatown

Boston's Chinatown is home to many community-oriented cultural and activist organizations. These organizations, along with many longtime neighborhood residents, hold a wealth of historical collections of photographs, newsletters, flyers, recorded interviews, maps, meeting minutes, and more. In addition, many area libraries, archives, and museums host collections pertaining to Chinatown's history. In collaboration with Chinatown community partners, this website aims to document archival collections drawn from a diverse range of sources.

This directory is intended as a launching point for neighborhood research, presenting indexes of these collections and the individuals and organizations connected to them.

Indexes by entity type

[Collections](#)

Archival collections, websites, and personal collections

[Organizations](#)

Groups or institutions connected to the collections as maintainers, donors, or creators

[People](#)

Individuals connected to the collections as maintainers, donors, or creators

Browse collections by...

[Maintainers](#)

People or organizations responsible for hosting and managing the collection

[Subjects](#)

Topics covered in the collection materials

[Material types](#)

Types and formats of records included in the collection

This is a project of the [Boston Research Center](#), a digital community history and archives lab based at Northeastern Library, in partnership with the Boston Public Library.

Questions or suggestions? Reach us at brc@northeastern.edu.

Funded in part by generous support from the Andrew W. Mellon Foundation and the National Endowment for the Humanities.

Collections

zh

The types of collections listed here include archival collections, websites, and personal collections, and cover a variety of topics related to the Chinatown neighborhood and Boston's Chinese community. Listed below is a brief description of each collection and its contents, as well as its digitization status.

Show 10 ▾ entries

Search:

title	collection type	digitization status	description
Asian American Resource Workshop records	archival collection	partially digitized	The records in this collection document the AARW's civil rights and anti-violence efforts for Asian Americans, and their creation of Asian American educational and cultural resources. It also includes documentation on the SafetyNet Violence Prevention Project and the Sticky Rice Project. Types of documents include minutes, calendars, grant applications, budgets, photographs, comics, etc.
Asian Community Development Corporation (Jacquie L. Kay) records	archival collection		The Asian Community Development Corporation's records document the development of affordable housing and assistance for small business development, education, and job training for Asian-Americans in urban areas. The majority of the collection documents two ACDC development projects, Oak Terrace and The Metropolitan, including funding reports, project status reports, and design specifications.
Boston 200 records	archival collection	partially digitized	Collection of oral histories documenting Chinatown and South End. Oral histories labeled under South End often address individuals and activities in Chinatown. Many of the transcripts of the oral histories are partial and appear to have been edited.
Boston Greenspace Alliance records	archival collection	not digitized	This collection includes one folder of materials titled Chinatown, 2004 (box 3, folder 146).
Boston Urban Gardeners records	archival collection	not digitized	This collection includes folders of materials on Chinatown from the 1980s, such as surveys (business, housing, user), reports (land use), plans, maps (Downtown Crossing and Chinatown), and a publication ("Chinatown 1988: Urban Design for Community Growth")
Carmen Chan's Samoan photos	personal collection	not digitized	This collection contains photographs taken by Carmen Chan, a reporter for Chinatown's local newspaper Sampan. The collection contains a large number of black and white photos.
Caroline Chang collection	archival collection	not digitized	This is one of CHSNE's many collections. It currently has not been processed yet, so enquiries should be made to CHSNE for research availability.
Cheng Imm Tan papers	archival collection	not digitized	This collection contains audio, videos, pictures, and teaching documents of Gund Kwok, an all-women Asian Lion Dance troupe based in Boston's Chinatown. It also contains some of Cheng Imm Tan's legal and personal documents, such as immigration and travel documents, awards, and family photographs.
Chinatown Story Cart	website	fully digitized	This is an art project created 2020-2021. In collaboration with local youth groups like BCNC, CSTQ, and ACDC, interviews were collected and visual media was created about the Chinatown community's lives and their experience of the pandemic. It also includes story sets with visual art and reflection activities sent to and collected from the residents. The majority of the collection is digitized.
Chinese American Experiences Mass. Memories Road Show	archival collection	fully digitized	Chinese American Experiences Mass. Memories Road Show is a digital collection of items (photographs, documents, video interviews) contributed by community members. These materials were collected during a participatory archiving event held at the Pao Arts Center in Chinatown on June 2, 2018.

Showing 1 to 10 of 36 entries

Previous **1** 2 3 4 Next

This is a project of the [Boston Research Center](#), a digital community history and archives lab based at Northeastern Library, in partnership with the Boston Public Library.

Questions or suggestions? Reach us at brc@northeastern.edu.

Funded in part by generous support from the Andrew W. Mellon Foundation and the National Endowment for the Humanities.

Chinese Progressive Association records

zh

The records in this collection document activism and advocacy in Chinatown and in Greater Boston's Chinese and Asian American community from the 1970's through the 2000's. Materials include administrative print records, public relations records, photo, video, and clippings.

material types (general)	photographs and other graphic materials; textual records
material types (specific)	interviews; newspaper clippings; programs; booklets; agendas; slides; photographs; VHS; posters; reports; pamphlets; brochures; fliers; handbooks; memorandums; minutes; correspondence
materials start date	1976
materials end date	2006
subject	
language	Chinese; English
official website	
described at URL	https://archivesspace.library.northeastern.edu/repositories/2/resources/946
collection type	archival collection
records creator	Chinese Progressive Association
donor	Lydia Lowe
maintainer	Northeastern University Archives and Special Collections
access restriction status	partly restricted access
digitization status	partially digitized
access policy notes	
project status	Further accruals expected
project start year	2008
project end year	

Collections

This is a project of the [Boston Research Center](#), a digital community history and archives lab based at Northeastern Library, in partnership with the Boston Public Library.

Questions or suggestions? Reach us at brc@northeastern.edu.

Funded in part by generous support from the Andrew W. Mellon Foundation and the National Endowment for the Humanities.

华人进步会记录

en

华人进步会记录记载了从 1970 年代到 2000 年代在华埠和大波士顿华人和亚裔美国人社区的倡行主义和倡权活动。主题包括城市发展、移民权利、工人权利、种族暴力、租户权利和中国文化活动，包含行政记录、照片、视频和剪报。

一般记录类型	照片和其他图形材料; 文字记录
特定记录类型	访谈; 剪报; 计划; 小册子; 议程; 幻灯片; 照片; 录像带; 海报; 报告; 资料手册; 传单; 手册; 备忘录; 会议纪要; 通信
所涉期间开始日期	1976
所涉期间结束日期	2006
作品主题	
作品或名称语言	中文; 英文

官方网站	
详细描述网址	https://archivesspace.library.northeastern.edu/repositories/2/resources/946

馆藏类型	档案馆藏
馆藏创建者	中国进步会
捐赠者	Lydia Lowe
维护者	东北大学档案与特藏
访问限制状态	部分限制访问
数字化状态	部分数字化
访问政策说明	
项目状态	预计会增加更多藏品
项目开始年份	2008
项目结束年份	

馆藏

这是设于东北大学图书馆数字社区历史和档案实验室的[波士顿研究中心](#)与波士顿公共图书馆合作的一个项目。

有问题或建议吗？电邮 brc@northeastern.edu 与我们联系。

部分资金来自安德鲁•梅隆基金会和国家人文基金会的慷慨支持。

Future work

- Improve UI
- Add support for traditional Chinese
- Generate visualizations (maps? timelines?)
- Develop Wordpress plug-in to interact with Wikibase, replace Snowman site
- Include more types of entities (e.g., events)

Thank you!

Contact: asruskin@gmail.com