

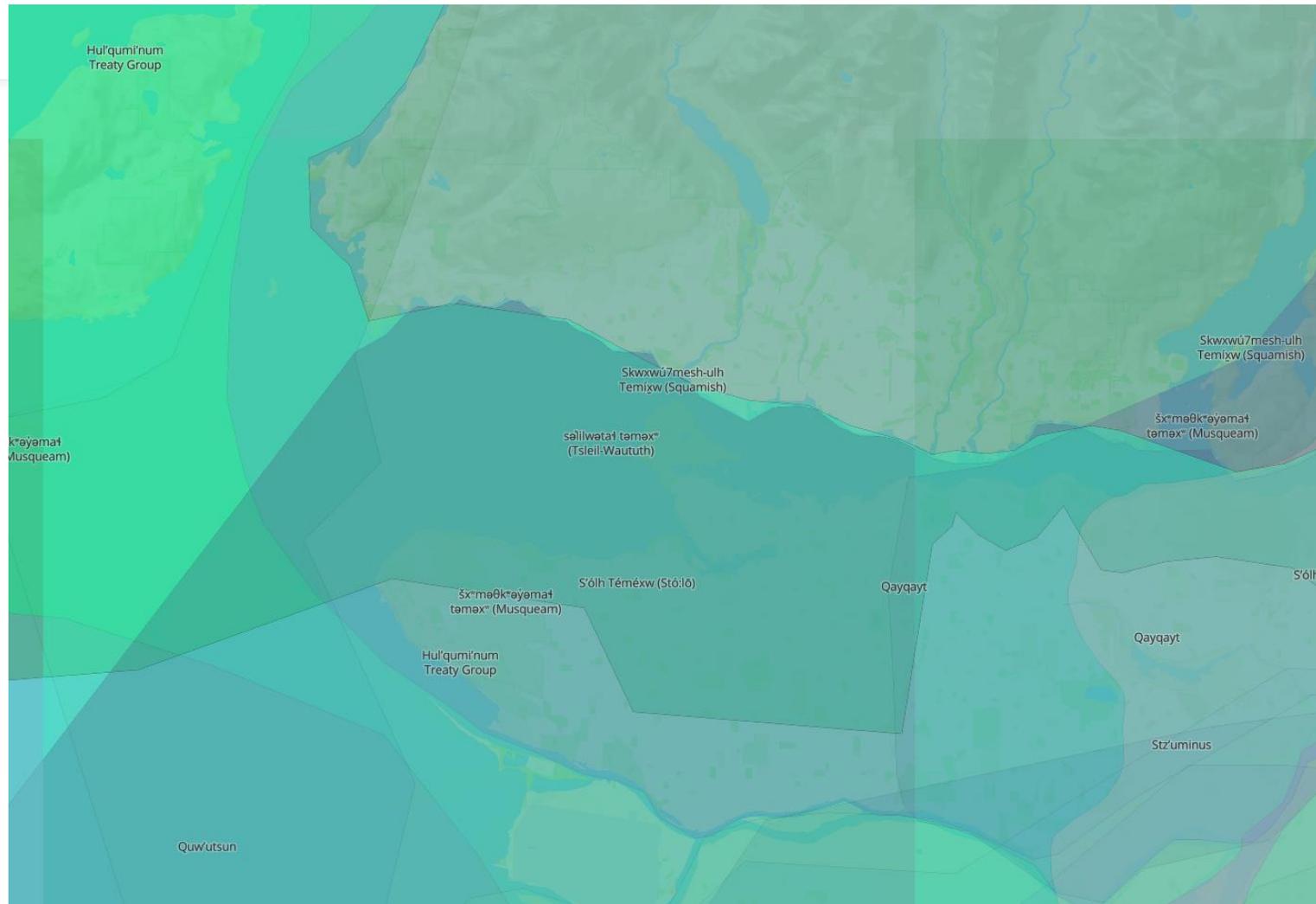


A Legacy of Race and Data: Mining the History of Exclusion

Sarah Zhang

Simon Fraser University, Canada

Land acknowledgement





Outline

- Historical background: Chinese Canadian Head Tax
- Register of Chinese Immigrants to Canada, 1886 to 1949: digitization and normalization of origin place names
- Hack the data: data analysis and visualizations
- Future directions for the data

Historical background

- 1885: Chinese Immigration Act was enacted
- Head tax: started with \$50, increased a few times to its maximum of \$500 in 1903
- Intended to constrain Chinese immigration
- No other ethnic group in British Columbian history suffered such formally sanctioned mistreatment of its members on entering Canada over such an extended period (source: [Federal Head Tax, Government of British Columbia website](#)).



Image above: Headtax receipt

Digitization project's output: Library and Archives Canada

The screenshot displays the Library and Archives Canada website interface. At the top, the logo 'Library and Archives Canada' is on the left, a red maple leaf is in the center, and the 'Canada' logo is on the right. A search bar contains the text 'Search BAC-LAC.gc.ca' and a 'Search' button. Below the header is a navigation menu with 'Discover the Collection', 'Search the Collection', 'Services for the Public', and 'Services and programs'. A breadcrumb trail reads: 'Home → Discover the Collection → Immigration → Immigration Records → Immigrants from China, 1885-1949 → Search: Immigrants from China'. On the left side, a sidebar menu lists: 'Immigrants from China, 1885-1949', 'The records', 'About the database', 'Search screen', 'How to interpret the results', 'How to obtain copies', and 'Other resources'. The main content area is titled 'Search: Immigrants from China' and includes the instruction 'Enter one or more search terms. (eg. Whiteaves, White*)'. There are three input fields labeled 'Name', 'Date of Registration (1938 or 1938-11 or 1938-11-30)', and 'Certificate Number'. At the bottom of the form are 'Search' and 'Clear' buttons.

A searchable database

Digitization project's output: the dataset

The screenshot shows the UBC Library website interface. At the top, the UBC logo and 'THE UNIVERSITY OF BRITISH COLUMBIA' are displayed. Below this is a navigation bar with links for 'Library Home', 'Search Collections', 'Hours & Locations', 'Use The Library', 'Get Research Help', and 'About Us', along with an 'Ask Us!' button and a 'LOGIN' dropdown. The breadcrumb trail reads: 'Library Home / Search Collections / Open Collections / Browse Collections / UBC Faculty Research and Publications / Register of Chinese Immigrants to Canada, 1886-1949'. The main heading is 'Open Collections', followed by a sub-heading 'UBC Faculty Research and Publications'. The collection title is 'Register of Chinese Immigrants to Canada, 1886-1949', with author 'Ward, W. Peter; Yu, Henry' and year '2008'. A description of the register is visible below the title.

Library Home / Search Collections / Open Collections / Browse Collections / UBC Faculty Research and Publications / Register of Chinese Immigrants to Canada, 1886-1949

Open Collections

UBC Faculty Research and Publications

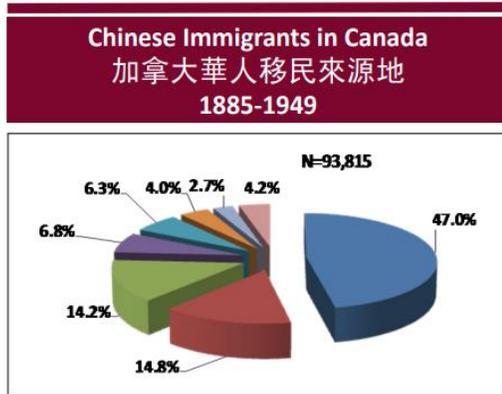
Register of Chinese Immigrants to Canada, 1886-1949

Ward, W. Peter; Yu, Henry
2008

The Register lists all immigrants of Chinese origin who entered Canada under the so-called Chinese Head Tax Act and subsequent legal restrictions which, between 1885 and 1922 attempted to limit immigration from China by means of an escalating entry tax, and between 1923 and 1949 prohibited all Chinese immigration apart from a small number of individuals in highly restrictive categories. The Register also lists those who came to Canada prior to 1885 who, by registering with immigration authorities for a nominal sum,

<https://open.library.ubc.ca/cIRcle/collections/facultyresearchandpublications/52383/items/1.0075988>

Where did the immigrants come from?



- Taishan 台山
- Xinhui 新會
- Kaiping 開平
- Panyu 番禺
- Zhongshan 中山
- Enping 恩平
- Heshan 鶴山
- Other 其他



Normalizing transliterations of origins

On the county level:

Challenge to Researchers
Spelling Diversity
 新 (San) 會 (Voy)

• San Ooi	• Sen Way	• Sin Wai	• Sun Wae
• San Woi	• Sen Wey	• Sin Way	• Sun Wai
• San Wui	• Sen Weyh	• Sin Woi	• Sun Way
• San Wae	• Sen Woy	• Sin Woy	• Sun Wey
• San Wai		• Sin Wey	• Sun Woi
• San Way		• Sin Wui	• Sun Woy
• San Wey		• Sing Wae	
• San Woy		• Sing Woi	
• Sang Wol			

Gold Mine for Linguists!

Source: Chiang, R., 2016

On the village level:

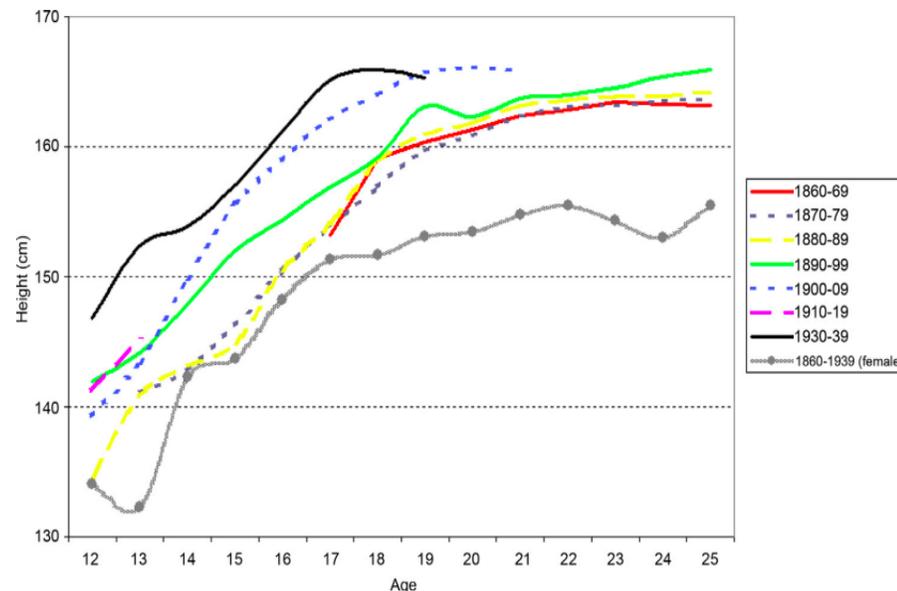
Variant Forms of Romanization 白沙 (Taishan 台山)		
Back Sa	Bak Shor	Pack Shan
Back Saer	Bak Shr	Pack Shar
Back San	Bak Shuk	Packshar
Back Sar	Bak Shut	Pack Shu
Back Ser	Bak Suck	Pah Sha
Back Shar	Bak Sun	Pak Cha
Backshar	Bak Thut	Pak Sah
Bak La	Bark Sar	Pak San
Bak Lai	Bark Sun	Pak Sar
Bak Sa	Bat Shar	Pak Sha
Bak Sack	Bok Sha	Pak Sha Une
Bak Sah	Bok Shar	Pak Shar
Bak Sahr	Buck Sa	Pak Sue
Bak Sai	Buck Sah	Pang Sa
Bak Sam	Buk Sah	Pang Sar Hang
Bak San	Buk Shai	Pang Sha
Bak Sar	Buk Shar	Pang Shar
Bak Sha	Packhar	Par Sar Long
Bak Shai	Pack Sam	Par Shar
Bak Shar	Pack San	Park San
Bak Shat	Packsar	Park Sar
Bak Shen	Pack Sha	
Bak Sheuk		
Bak Sar (or Nam Long)		

Hack the data

- Inquiries:
 - To what extent this dataset has been utilized by scholars?
 - What new possibilities for the study of early Chinese immigrants to Canada or head tax could this dataset offer if new computational/digital tools are employed?
- Data cleaning, interpretation and conversion
- Statistical analysis and network analysis- leveraging the normalized place names in the data

Previous scholarly work: Peter Ward's study on the immigrants' wellbeing

- Peter Ward's data revealed an increase of over 4 cm in the adult heights of both sexes between 1850 and 1930.
- One likely cause of improved well-being is the influence of the migrants' remittances on socioeconomic change in south China.



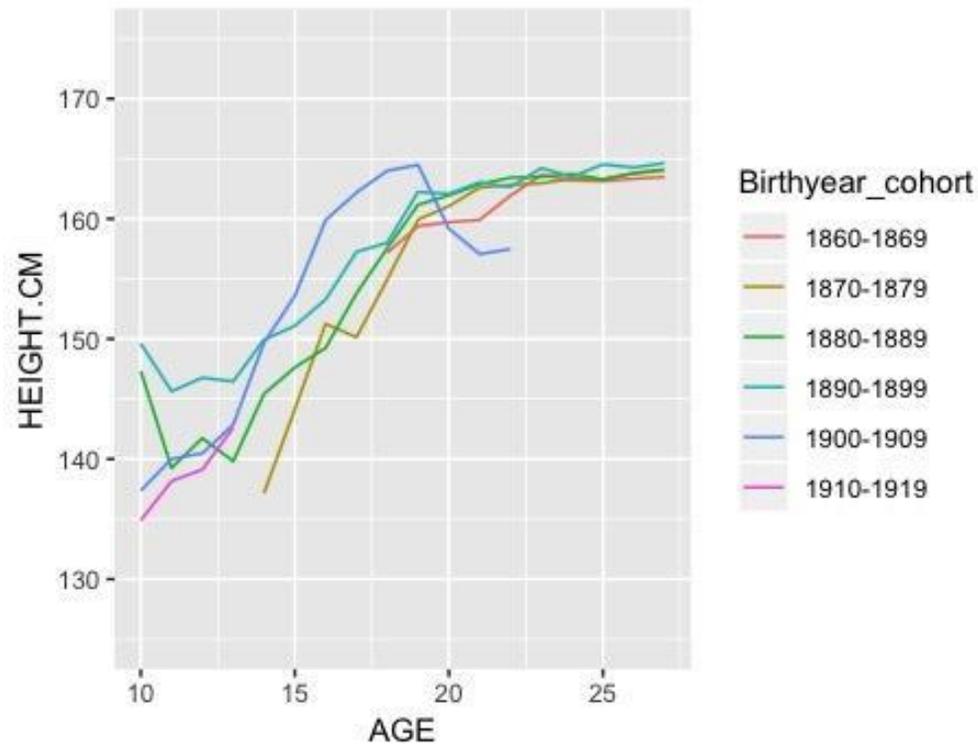
(Peter Ward, 2013)

Statistical analysis: down to the county level

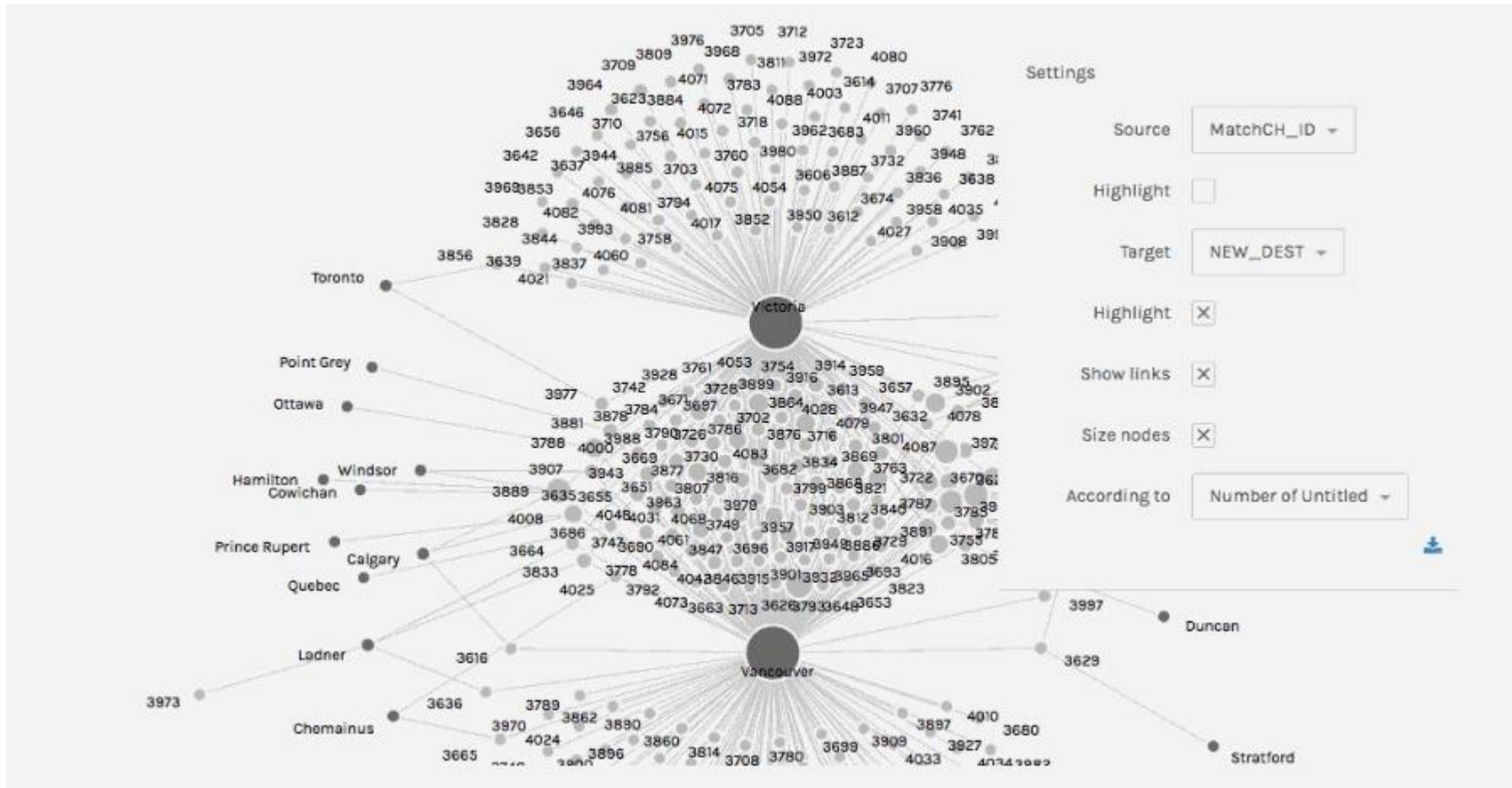
When analyzing only the immigrants from Xinhui, however, very interestingly different patterns are revealed

Nuanced patterns: less evident rise in mean height over the years

Why is the **mean height** for adolescent years grow so quickly, while **plummeting** in following for the 1900-09 cohort?



Network visualization: down to the village level



Future directions for the dataset: from the data stewardship perspective

- FAIR (Findable, Accessible, Interoperable, Reusable) principle
- Library digital collections vs research data repositories
- Spotlighting **Reusable**
- What need be improved?
 - Enrich the file-level and column-level metadata, recreated a codebook
 - A more unrestricted license

Future directions for the dataset: from the research perspective

- **Historical GIS:**

using spatial data to

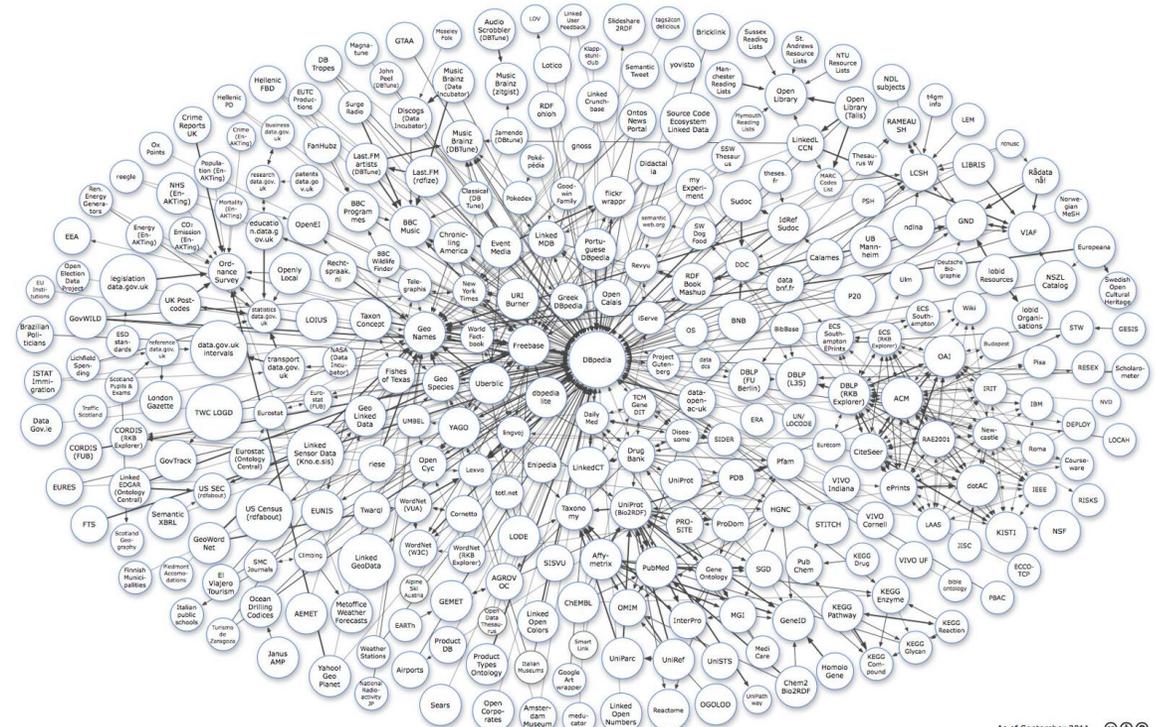
discover a unique perspective
on past events



Image: "[#BLMcareers](#)" by [mypubliclands](#)

Future directions for the dataset: from the research perspective

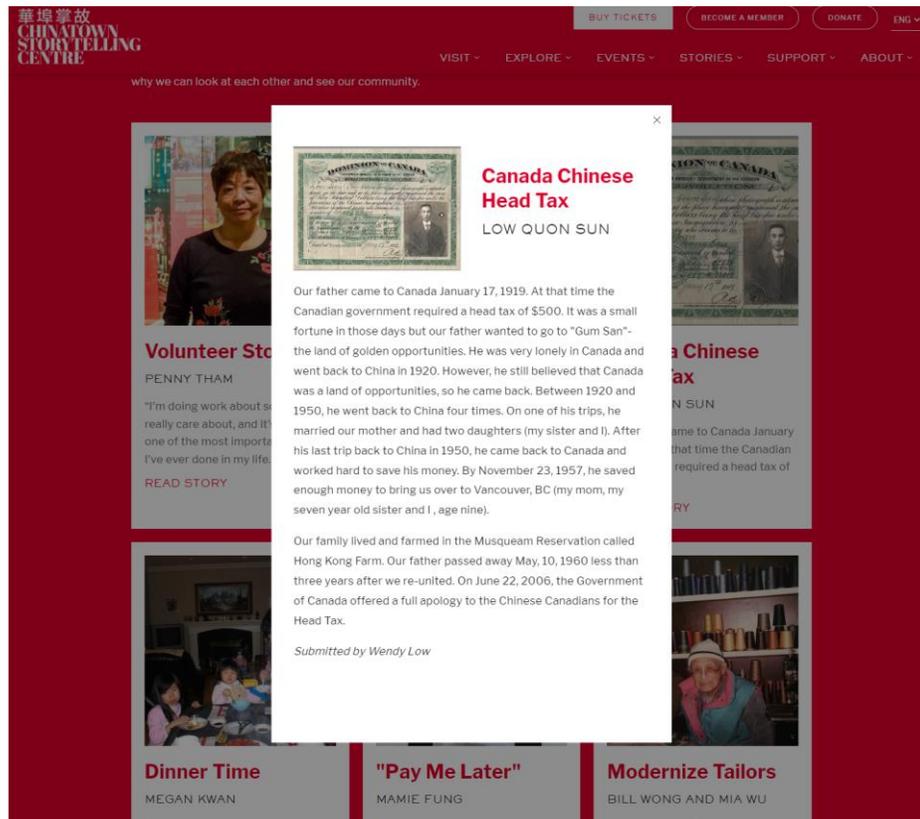
- **Linked Open Data:**
linking the data to web resources
outside of the dataset (Smith, 2016)



As of September 2011 © I I I

Image: "[State of the Linked Open Data \(LOD\) Cloud in September 2011](#)"
by [dullhunk](#)

Future directions for the dataset: from the research perspective



- **Open Science and Community Engagement:**

What kind of story should be told? What questions do they want to see answered?

Source: Chinatown Storytelling Center <https://www.chinatownstorytellingcentre.org/stories/#community>

References

- Chiang, Rudy. “Canada Immigration Chinese Head Tax Record Sun Woy District Part 1: 1885-June, 1903.” UBC Asian Library. 2012. https://branchasian.sites.olt.ubc.ca/files/2011/09/Head_Tax_Record_SunWoy_RC.pdf
- Hermansen, Sally, and Henry Yu. “The Irony of Discrimination: Mapping Historical Migration Using Chinese Head Tax Data.” In *Historical GIS Research in Canada*, edited by Jennifer Bonnell and Marcel Fortin, 225-237. Calgary, AB: University of Calgary Press, 2014.
- Ward, W. Peter. “Stature, Migration and Human Welfare in South Shina, 1850-1930.” *Economics and Human Biology* 11, no. 4 (2013): 488-501.
- Smith, J. (2016). Working with the Semantic Web. In *Doing Digital Humanities* (pp. 309-324). Routledge.
- 'Mapping the Villages & Towns Recorded in the Register of Chinese Immigration to Canada from 1885 to 1949': <https://asian.library.ubc.ca/files/2012/01/Head-Tax-brochure2.pdf>.

To learn more about our work

- OSF project: Hacking for the Register of Chinese Immigrants to Canada, 1886-1949. <https://osf.io/9zr6f/>
- Upcoming book chapter: Cho, Allan, & Sarah Zhang. "A Legacy of Race and Data: Mining the History of Exclusion."
Future Horizons: Canadian Digital Humanities. Eds. Sarah Roger and Paul Barrett.