

Smithsonian Libraries and Archives Wikidata & Wikibase Projects

2021 LD4



July 21 2021



bit.ly/2021LD4SMI0721



Smithsonian
Libraries and Archives

How do we do this?

GLAMs - Modeling out data from the *Artists Files*
to the Zs: *Zebra Genomes*, 中國人 (*Zhongguo ren*), *Zulu Kingdom*, in the Smithsonian Institution



Smithsonian
Libraries and Archives

Who we are



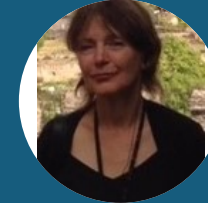
Jackie Shieh

Descriptive Data Management
Librarian
[*ShiehJ*](#)



Anne Evenhaugen

Head, Art Libraries
[*aceven*](#)



Kathryn Phillips

Librarian
[*JuliaAgricola*](#)



Diane Shaw

Special Collections Cataloger
[*Uncommon frittillary*](#)



Heidi Berthoud

Head of Resource Description
[*Berthoudh*](#)



Richard Naples

Technical Information Specialist
[*Drastician*](#)



Suzanne Pilsk

Head of Metadata Department
[*Pilsk*](#)

*...plus many other staff members at the
Smithsonian Libraries and Archives!*



Smithsonian
Libraries and Archives

Topics that we will cover today:

- Project genesis: Smithsonian and its world of data
- Why Wikidata platform
- Our PCC Projects
- Wikidata challenges and solutions
- Wikibase journey

How did we get started?

- Authority control need
- Starting point
 - Artists files—Libraries and museums
 - Non-Latin collections
 - Research outputs



PCC Wikidata Smithsonian Projects

- African Ethnic Groups
- Artists' Files
- Chinese Ancestors Portraits
- Dibner Scientific Portraits
- Smithsonian Research Online

*These projects all centered around dealing with **names** (people, organizational units), but also have challenges about **multilingual scripts**, concerns about **ethics**, and questions on how we can make **inferences***

This is what we did

Property

- Application & constraint
- Policies
- Best practices
- Ethical question
- Political issue
- SPARQL impact

Issues and resolutions

- ...

Properties

Example 1: Names

What's in a Name?

- Family (P53)
- Family name (P734)
- Name of clan (P???)
- Lineage name (P ???)

Issues and resolutions

- Naming conventions
- Nominal kinship
- Given name constraints
- Inference

Names: Chinese Ancestor Portraits



Gong Family (Q103537674)

Li clan of Longxi (Q15914504)

Xing (Q61942409) and Shi (Q11128762)

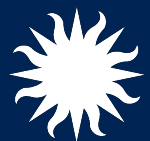
Lady Wanyan (Q97330058)

What would you have done?

Slack questions

- Inference issue, same clan, not of same family name
- Women's name, <wife of>, <Mrs. ***>

Sex or Gender



Smithsonian
Libraries and Archives

Example 2: Sex or Gender

Sex or gender (P21)

Issues and resolutions

- Ethical questions
- Conceptualization
- How to operationalize

What would you have done?

Slack questions

- How to get around the problem of retrieving an item based on data at birth and publicly available, but the named subject objects to this approach.
- Should the property reflect the latest gender status of the subject for the living vs historical person?
- In the absence of known gender, if this is a <core> property, how will you reconcile the requirement and practice?
- How non-human, e.g., Lassie? A <she> character but in reality a <he> actor?!
- Lack of control vs. the “responsibility” of the Smithsonian name?

Ethnic Groups

Example 3: Ethnic Groups

Ethnic group (P172)

- ...

Issues and resolutions

- Ethnic vs. indigenous group
- Historical and present
- Multi-ethnic group identification

What would you have done?

Slack questions

- How do we decide the name of an ethnic group when it has changed across the history?
- What people call themselves vs. how others call them? That could be also a multilingual concern especially during the colonial period.
- Describing people of variety of ethnic groups, e.g., Qing dynasty from Han ethnic to becoming Manchu ethnic due to honorific promotion; or self-promotion for social status, changing from Christian, Jewish to Muslim, vice versa?

Floruit (floru-what?!?)

Example 4: Active Period of an Artist

Active period of an artist

- Floruit (P1317)
- work period (start) (P2031) & work period (end) (P2032)

Issues and resolutions

- How *wikidata* defines these properties vs. How *we* define in the field
- Lack of Consistency of Use
- How do you SPARQL these?

What would you have done?

Slack questions

- What is the best practices for recording an active period of an artist, writer, etc.?
- How does the period of “best known period” of an artist aside from the start of their work?
- What is the criteria of an editor to use in terms of the research for adding a property statement?

Wikibase: an SI-centric Wikidata



Smithsonian
Libraries and Archives

Wikibase: an SI-centric Wikidata

- What to emulate from Wikidata?
- Data modeling
 - Greater local control over data description and quality, but at the cost of interrelatability?
 - PII concerns
- Documentation: best practices
- System customization
 - Wikidata: Schemas in place of <property—constraint>
 - Hierarchical user account setup
 - UI customization

What will you do?

- Slack questions
 - Data modeling needs
 - User account for various access/edit needs

References and Credits

PCC Wikidata Pilot: Smithsonian Libraries

wikidata.org/wiki/Wikidata:WikiProject_PCC_Wikidata_Pilot/Smithsonian_Libraries

Wikibase yearly summary 2021

lehir.net/wikibase-yearly-summary-2021/

Landsdowne Portrait by Gilbert Stuart (1796) National Portrait Gallery, Washington D.C.
commons.wikimedia.org/entity/M33579

The Hope Diamond
Photo courtesy of the Smithsonian Institution
www.si.edu/newsdesk/photos/hope-diamond-0

Dorothy's Ruby Slippers
americanhistory.si.edu/press/fact-sheets/ruby-slippers

Mei Xiang (Meghan Murphy photos)
insider.si.edu/2010/04/national-zoos-giant-panda-mei-xiang-is-not-pregnant/

1909 Wright Military Flyer
airandspace.si.edu/collection-objects/1909-wright-military-flyer/nasm_A19120001000

Smithsonian Institution Building (aka the Castle)
commons.wikimedia.org/entity/M96312914



Smithsonian
Libraries and Archives

Staff Credit Slide

Heidy Berthoud, Head of Resource Description. [Berthoudh](#)

Rebecca Bruner, Library Technician, Hirshhorn Museum and Sculpture Gallery Library. [Rebeccabruner](#)

Anne Evenhaugen, Head Librarian, Smithsonian American Art and Portrait Gallery Library. [aceven](#)

Nilda Lopez, Library Technician, Cooper Hewitt Smithsonian Design Library. [NeoNilz](#)

Richard Naples, Technical Information Specialist, Metadata Services. [Drastician](#)

Kathryn Phillips, Librarian for South, Southeast and Middle Eastern Art and Architecture. [JuliaAgricola](#)

Suzanne Pilsk, Head, Metadata Services. [Pilsks](#)

Alexandra Reigle, Librarian, Smithsonian American Art and Portrait Gallery Library. [AReigle](#)

Keala Richard, Conservation Technician, Smithsonian Libraries and Archives Preservation Department. [Km31284](#)

Diane Shaw, Special Collections Cataloger. [Uncommon fritillary](#)

Jackie Shieh, Descriptive Data Management Librarian. [jshieh](#)

Mike Smith, Librarian for Korean art, American art, Conservation and Scientific Research, and serials. [78smith](#)

Janet Stanley, African Art Museum Librarian.

Amy Watson, Resource Description.

Yue Shu, Librarian for Chinese language. [Shuyuex0320](#)

Thank you!



Smithsonian
Libraries and Archives