SOCIAL NETWORKS AND ARCHIVAL CONTEXT



ICA /ALA Meeting. Mexico City, 2017

Daniel Pitti, University of Virginia | Kelly Spring & Jimmy Zavala, University of California, Irvine



Session Overview

- Description and history of SNAC Cooperative (Daniel)
- Governance (Kelly)
- History Research Tool (Jimmy)
- Editing User Interface (Kelly)

SNAC Cooperative

Archival Records

- Record: Linguistic, symbolic, or graphic information represented in any persistent form, on any durable carrier, by any method, by an Agent in the course of life or work events and Activities.
- Records document (are evidence of) people living and working together
- The lives and work of people, and the records created by them constitute a vast network of interrelated people and documents: a social document network

Description of Archival Records

- Archival descriptive dominated by using one provenance-based apparatus that combines/intermixes context description and record description: the finding aid
 - Creators are documented in detail
- Many individuals and groups documented in the records are also referenced in description (as access points or informally)
- Archival description documents interrelations among people and records
- A vast social-document network connecting the past to the present to the future
- To date, interrelations have been implicitly rather than explicitly documented
- SNAC makes the social and document interrelations explicit



Rationale for Cooperative

- The corporate body, person, or family (CPF) document in the holdings of one repository is often documented in the holdings of another repository
- Research, identify, and describe the CPF entity once, and shared by all: economy
 of cooperatively curating the data
- Link the people descriptions to one another and to record descriptions: build the social-document network
- An international, Internet-based "linked archival authority system"
- Research economy
 - Integrated access to distributed historical resources
 - Expand the context and understanding of the records by revealing the social networks



Cooperative

- Cooperative host: University of Virginia Library
 - Administration and governance coordination
 - Technology infrastructure
 - -Cooperative members ...



Cooperative Members

- American Institute of Physics
- American Museum of Natural History
- Archives, National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore, India
- Archives nationales de France
- Brigham Young University
- · California Digital Library
- Cecilia Preston (individual scholar)
- George Washington University
- Getty Research Institute
- Harvard University
- Indiana University
 –Purdue University
 Indianapolis
- Jane Addams Papers (documentary editing)
- Library of Congress
- Mojave Desert Archives

- U.S. National Archives and Records Administration
- New York Public Library
- Princeton University
- Smith College
- Smithsonian Institution
- Tufts University
- University of California, Irvine
- University of Miami
- University of Nebraska
 - Library
 - Walt Whitman Archive (documentary editing)
- University of North Carolina, Chapel Hill
- University of Oregon
- University of Virginia
- Utah State Archives
- Yale University

History: Research & Demonstration 2010-2015

- From R&D to Cooperative Program
- Funding: NEH (2010-12), IMLS (2011-2014) and Mellon (2012-2015)
- R&D Partners
 - University of Virginia, Institute for Advanced Technology in the Humanities
 - University of California, Berkeley School of Information
 - California Digital Library (University of California)

Objectives

- Demonstrate that data describing people in existing archival description can be used to ...
 - –Address the challenge of finding/discovering/locating/understanding distributed historical resources and ...
 - –Reveal the social-document network implicit in the description and ...
 - Lay the foundation for an international cooperative for centrally maintaining the biographical/historical descriptions

Data Sources

- 2.25M WorldCat archival descriptions (MARC21)
- Nearly 190,000 EAD-encoded finding aids, primarily from US and UK, though some French
- 300,000 British Library authority records
- U.S. National Archives and Records Administration authority records
- Agency descriptions from Smithsonian Institution Archives/New York State Archives
- And more ...

Methods and Processing

- Extract/Assemble/Migrate EAC-CPF (archival standard) records from existing archival description
 - -Extracting both creators and referenced CPF names
- Match EAC-CPF records against one another and against existing authority records (Virtual International Authority File (VIAF))
 - Enhance EAC-CPF by normalizing entries, adding alternative entries, titles, "same as" links (VIAF)
- Create a prototype historical resource and access system
 - Social networks in which people lived and worked
 - -Integrated access to distributed archival resources
 - Access to other resources by and about (publications, other artifacts)

The Identity Resolution Challenge

- Different names for the same person
- Different people with the same names
- Names are weak identifiers
- A challenge for computers
- A challenge for people

 Identity must be based on evidence and as much as possible: names, dates, places ...

Extraction Results

- Original Source Records: 6,719,064
 - -4,653,365 Persons
 - -1,868,448 Corporate Bodies
 - -197,251 Families
- Merged Records: 3,741,262
 - -2,466,425 persons
 - -1,077,588 corporate bodies
 - -197,249 families



Establishing the Cooperative: 2015-2019

- Funding: Mellon Foundation 2015-2019
- Two primary objectives
 - Social
 - Administration
 - Community governance
 - A growing team of trained editors
 - Business model
 - Technological
 - Transform R&D platform into maintenance platform
 - An editing interface for manually creating, revising, linking descriptions
 - A History Research Tool for researchers
 - Collaborative Ingest Tool for importing large sets of data



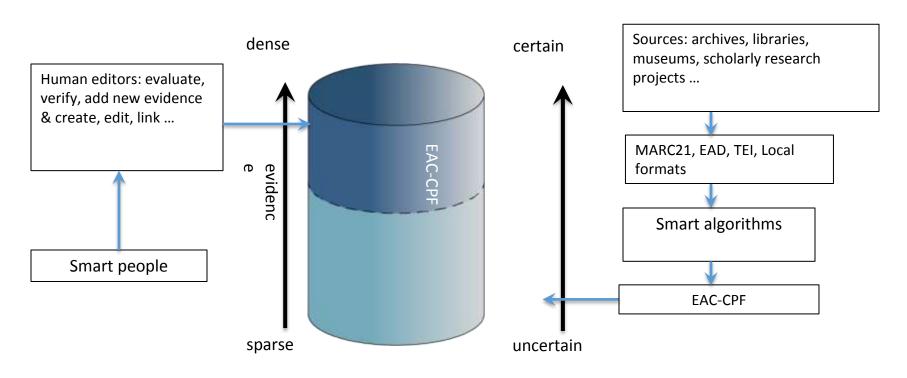
Current Focus of Work

- Developing capacity to increase membership
- Training editors, training trainers of editors:
 - Real and virtual classroom
 - Self-paced online modules
- Collaborative Ingest Tools
 - Enable new data contributors to assist in improving the quality of the data before import
 - Enable new contributors to assist in Identity Resolution
 - Enable new data contributors to integrate persistent identifiers into local description of archival records

Going Forward: 2017-2019

- Continue both the social and technological development, including making both the editing and research interfaces available in other languages
- As we build capacity, recruit new members
- Increase members of all kinds, including international members

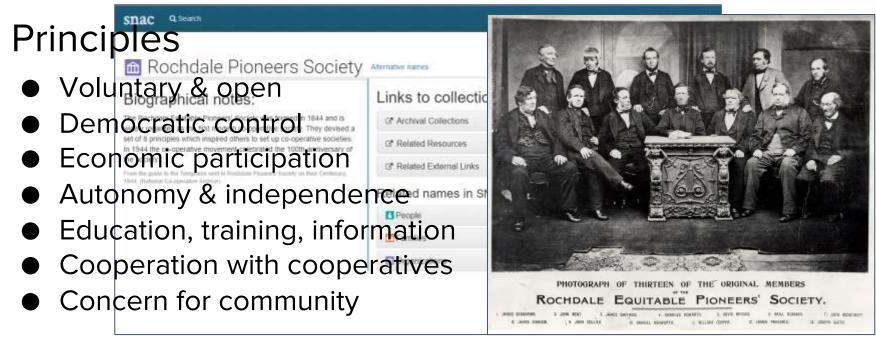
SNAC Cooperative Identities



The SNAC Cooperative Community



The SNAC Cooperative Community





The SNAC Community Structure



Operations Committee



Technology Infrastructure



Editorial Policy & Standards



SNACSchool



Communications

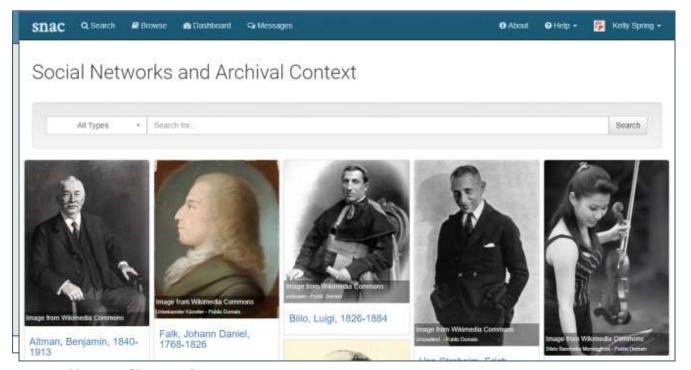
Technology Infrastructure



Technology Infrastructure Highlights

- History Research Tool (HRT) transition
- HRT Advanced Search Feature
- User Interface development for maintenance system

History Research Tool transition



SINVACENHA is Board to the a Troot of Teater of the Control of the

Phase II Goals

- Feedback and development
- Batch data refinement
- SNAC API integration



Technology Infrastructure

Technology Infrastructure Phase II Projects













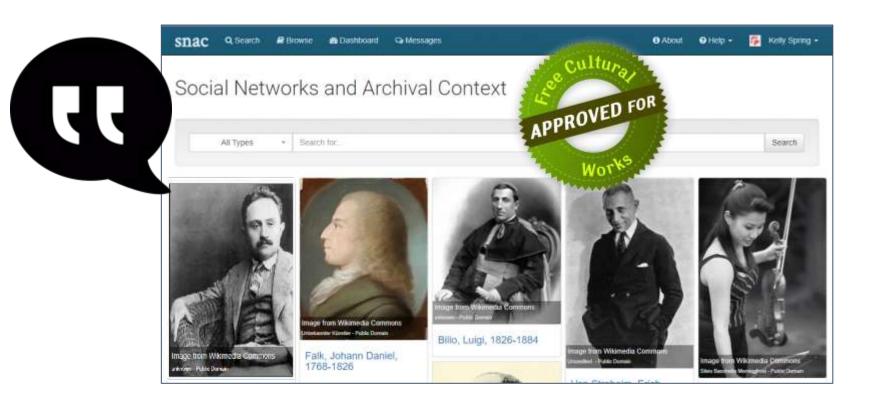
Editorial Policy & Standards



Editorial Policy & Standards Highlights

- Defined roles and permissions
- Developed editorial practice
- Determined ownership policy

Constellation Ownership Policy



Phase II Goals

- Demographic classification
- Function & Occupation policy
- Geographic names
- Subject terms



Editorial Policy & Standards

Editorial Policy & Standards Projects Include















SNACSchool



SNACSchool Highlights

- Developed training program
- Created SNACSchool Lite
- Hosted live training events

Phase II Goals

- Enhance training material
- Effective training methods
- More training!



SNACSchool

SNACSchool Projects Include













Communications



Communications Highlights

- Established regular information exchange
- Evaluated existing channels of internal communication
- SNAC Cooperative central portal



Cooperative Members Portal



Phase II Goals

- Develop the "SNAC brand"
- Cooperative portal content
- Outreach efforts



Communications

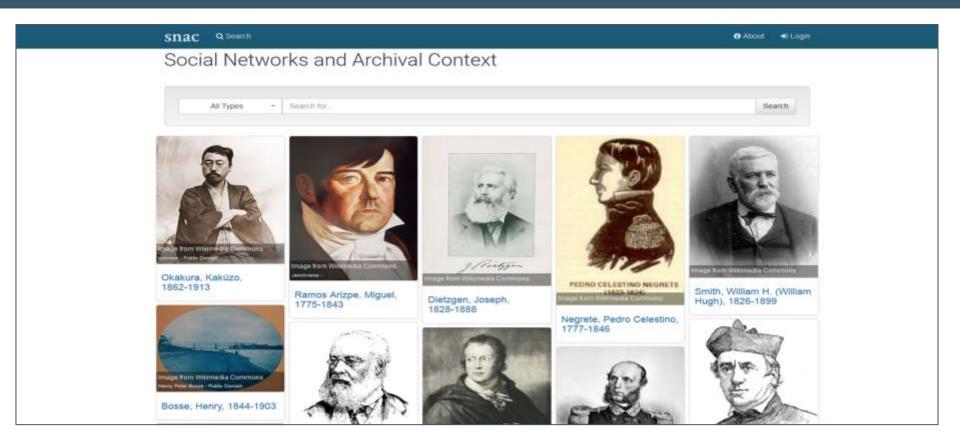
Communications Projects Include

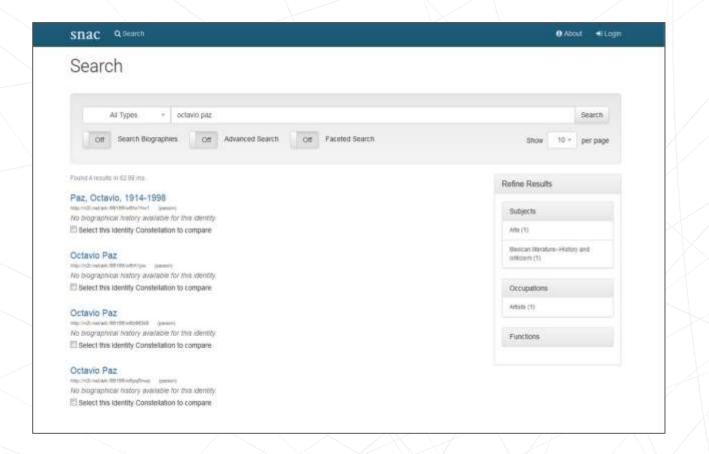


History Research Tool Demonstration



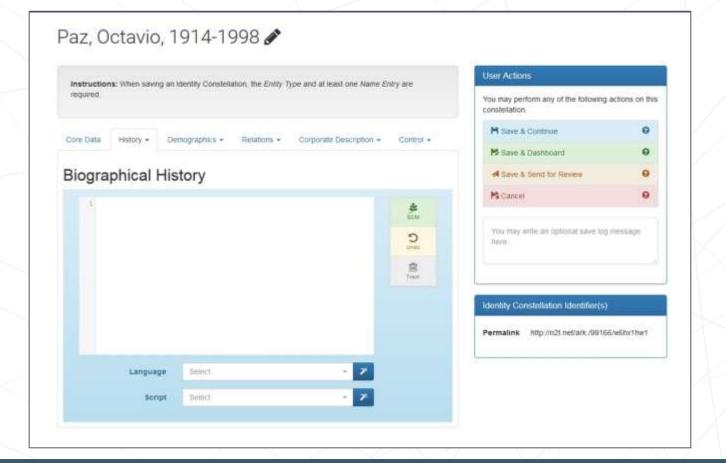
SNAC R&D Technical Highlight: Front-end history research tool

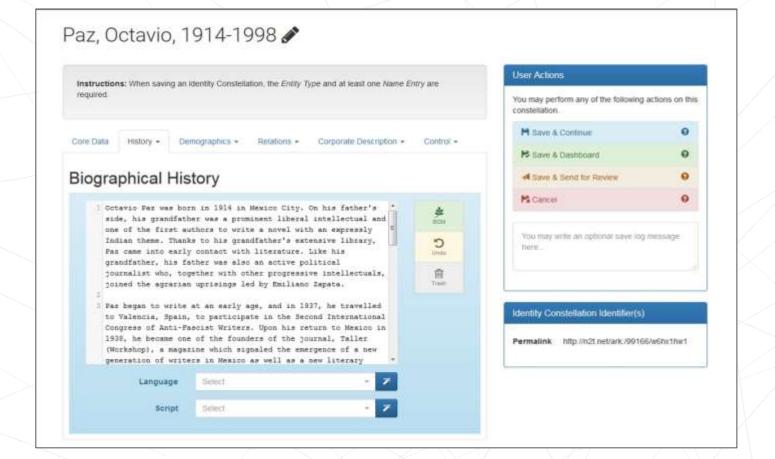




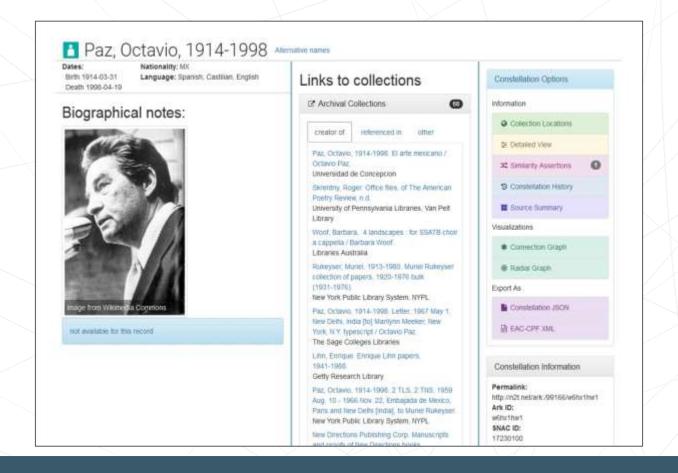
















Collection Locations



Lewfet | © OpenStreetHap

Auraria Library, Colorado University.

Beinecke Rare Book and Manuscript Library

Benson Latin American Collection, The University of Texas at Austin

Biblioteca Nacional de España, Biblioteca Nacional - Madrid

British Library: Western Manuscripts

Campbell University, Wiggins Memorial Library

Columbia University in the City of New York, Columbia University Libraries

Columbia University. Rare Book and Manuscript Library,

David M. Rubenstein Rare Book and Manuscript Library, Duke University

Emory University Library, Special Collect Department

Getty Research Institute

Getty Research Library

Harry Ransom Humanities Research Center (HRC); University of Texas at Austin,

University of Texas Libraries

Harry Ransom Humanities Research Center, University of Texas at Austin.

Harvard University Archives.

Harvard University, Houghton Library

Haverford College Library

Houghton Library, , Harvard College Library, Harvard University

Libraries Australia

Library of Congress

Manuscript Division, Library of Congress

Miami University

Miami University Libraries, King Library

Snac Social Networks & Archival Context



Ashton, Dore Yevtushenko, Yevgeny Aleksandrovich, 1933-

Guillén, Jorge, 1893-1984

Levin, Harry, 1912-1994 Bronk-William

Gascoyne, David, 1916-2001

Cummings, E.E. (Edward Estlin), 1894-1962

Paz, Octavio, 1914-1998

Amichai-Yehuda

Peden, Margaret Sayets; Munel, 1913-1980

Bishop, Elizabeth, 1911-1979

Levertov, Denise, 1923-1997

Rodman, Seiden, 1909-2002

Snac Q.Search Ø About • Login

Alternative names

🚹 Guillén, Jorge, 1893-1984

Dates: Birth 1893-01-18 Death 1984-02-06 Nationality: ES

Language: Spanish, Castilian

Biographical notes:



Professor of Spanish literature in both Europe and America and an award-winning poet. He went into exile in 1938 and settled in the United States in 1940.

From the diseptytion of Carrespondance with Brush Epilysuskalds / Jurge Guillen 1954-1963 (University of Wascaman - Madalon, General Larary System), WorldCaf record of 29990573

Guillén (1893-1964) was a Spanish lyric poet.

1924-1905 (Inclusive) 1940-1965 (Inclusive) University) Viertificat recent of 81017346

Jorge Gullén, a Spanish poet, published Cantico

in 1964, his first book to appear in English. The bilingual edition contained translations by Norman Di Glovanni, Richard Wilbur, W.S. Merwin, Alastair Reed, Alan Dugan, Barbara Howes, and others.

From the Interception of Judge Guillion embruses, 1961-1964 (Judy 1965-1964), (Pannaylvania State Household, Ministrative Control of THE 1969).

Biographical Note: Jorge Guillén was a Spanish poet and teacher. He was born in 1893 and taught in the United States from 1938-1958.

His opposition to Franco during the Spanish Civil War led him into self-imposed exile. While in the U.S., he taught at Middlebury, Wellesley, Harvard, Yale, Berkeley, and Ohio State. He died in 1984 at age 90 in

Links to collections Archival Collections Related Resources Related External Links Related names in SNAC People Families Organizations





Subjects:



Editing Interface Demonstration

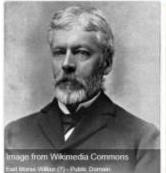


Social Networks and Archival Context

All Types

Search for ...

Search



Eliot, T. L. (Thomas Lamb), 1841-1936





Bigot de Morogues, Marie, 1786-1820





Ehrenberg, Herman, 1816-1866





Brooks, Alfred H. (Alfred Hulse), 1871-1924

Image from Wikimedia Commons

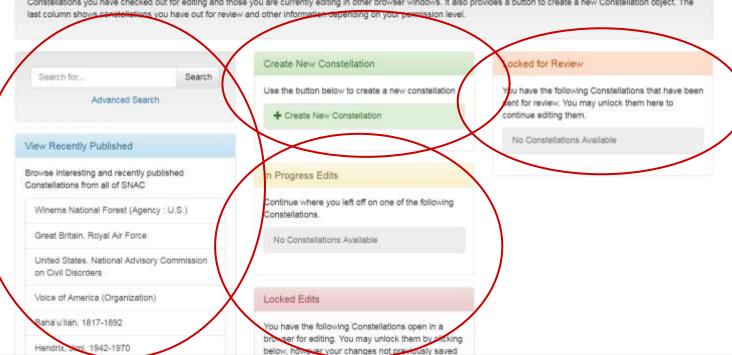
Unknown (US Geological Survey) - Public Domain





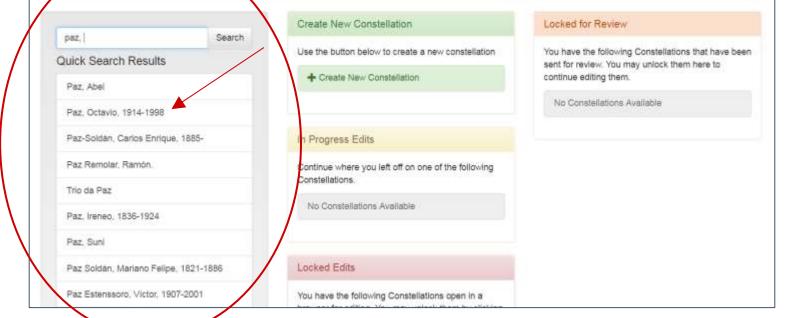
Dashboard

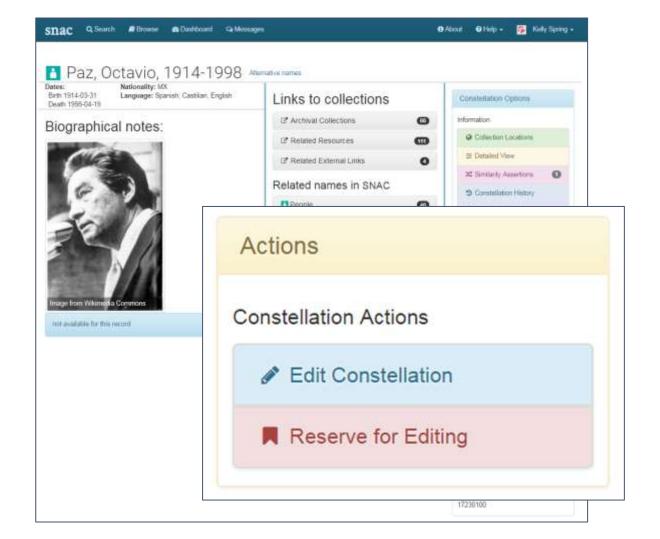
Welcome, Kelly, this is your dashboard. The left column shows the most recently published Constellations across SNAC and provides searching. The second column shows Constellations you have checked out for editing and those you are currently editing in other browser windows. It also provides a button to create a new Constellation object. The last column shows centralistics you have out for review and other information opening on your personal revel.

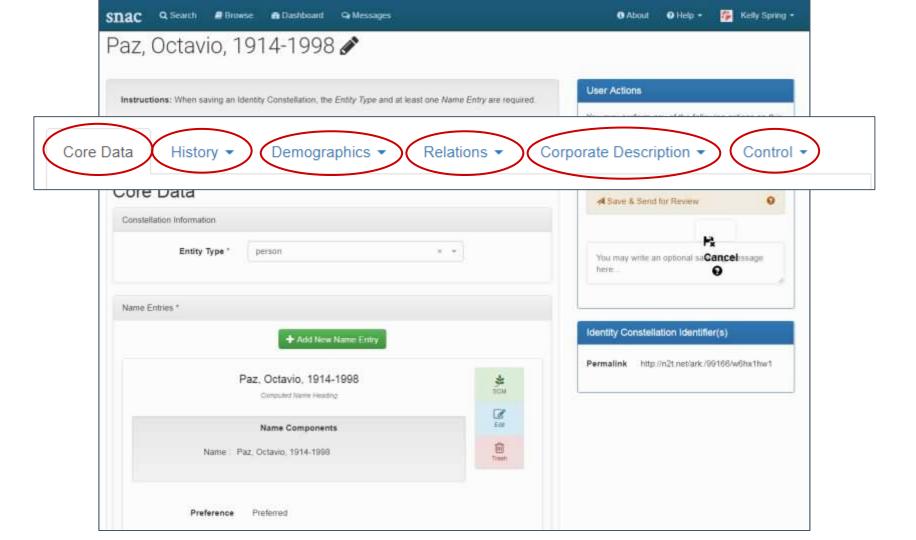


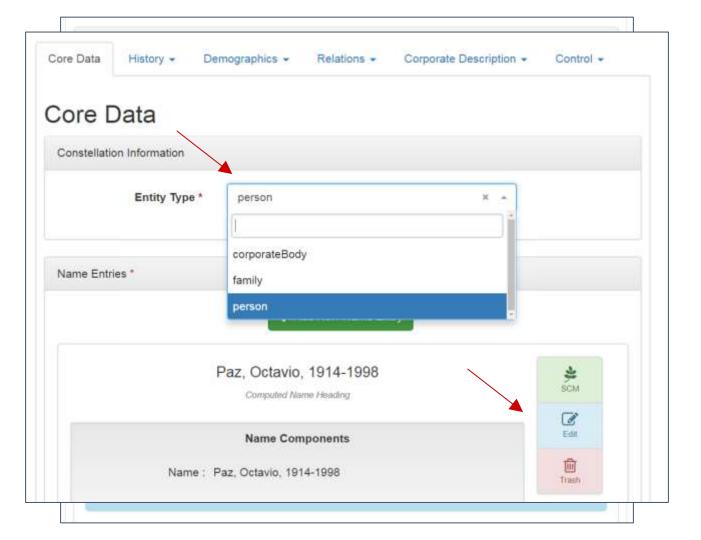
Dashboard

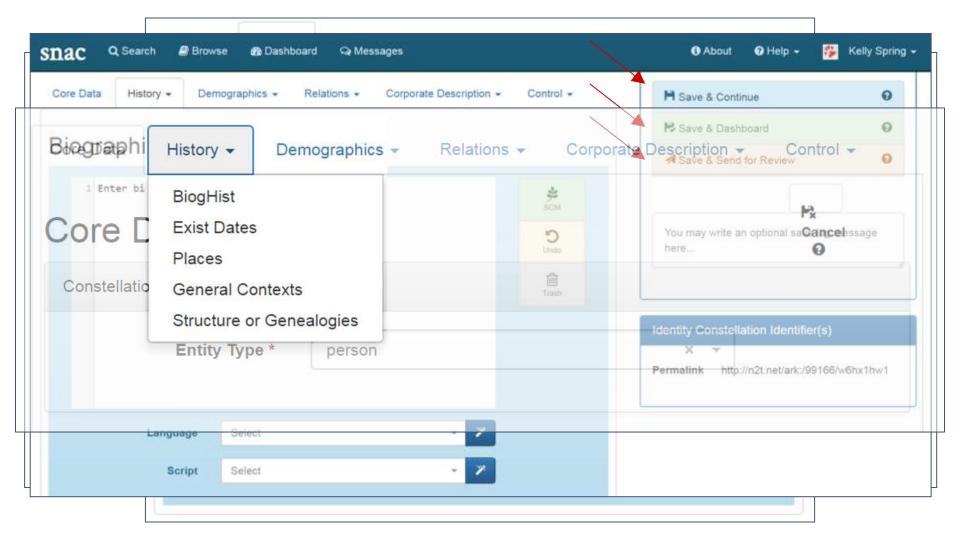
Welcome, Kelly, this is your dashboard. The left column shows the most recently published Constellations across SNAC and provides searching. The second column shows Constellations you have checked out for editing and those you are currently editing in other browser windows. It also provides a button to create a new Constellation object. The last column shows constellations you have out for review and other information depending on your permission level.

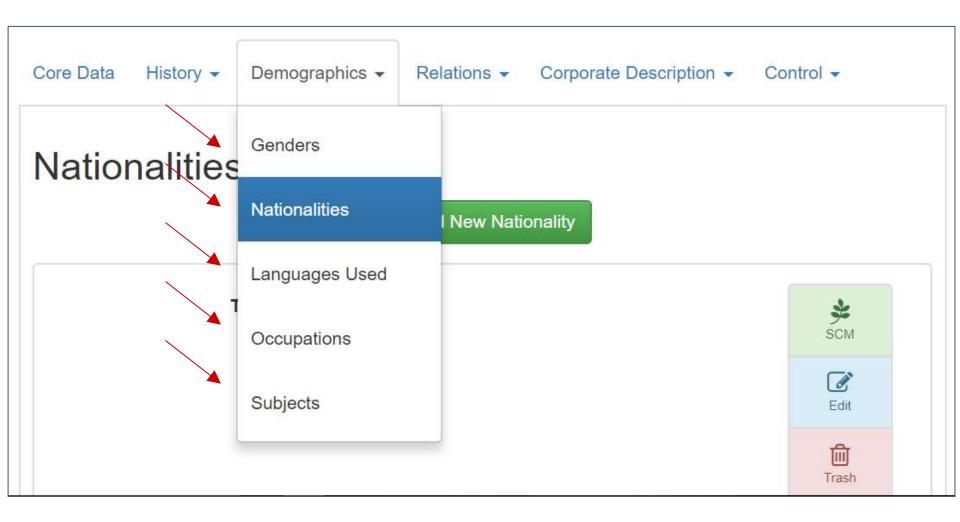


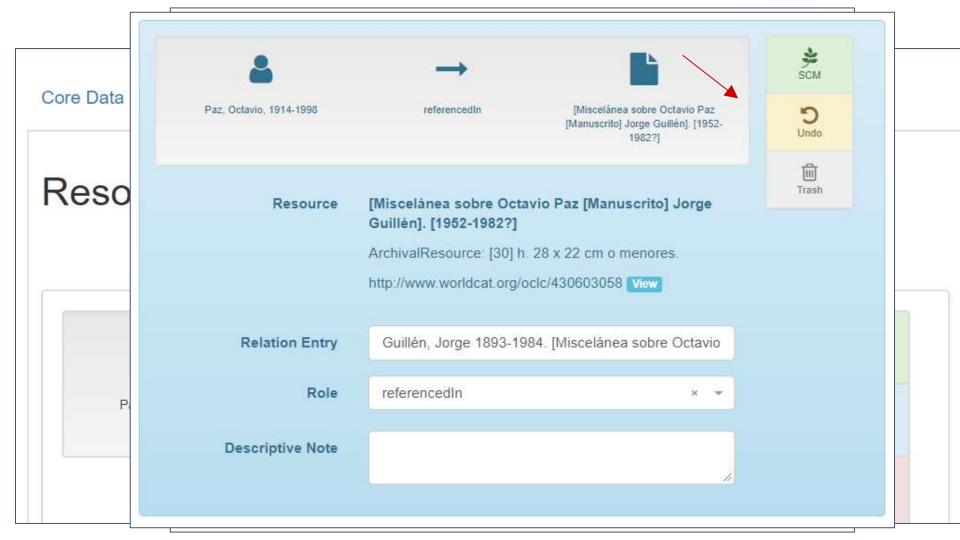


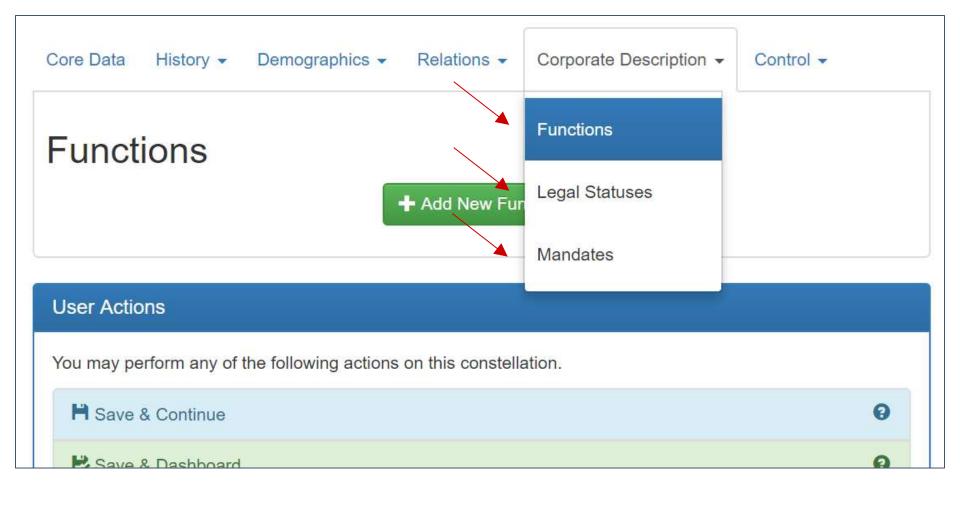


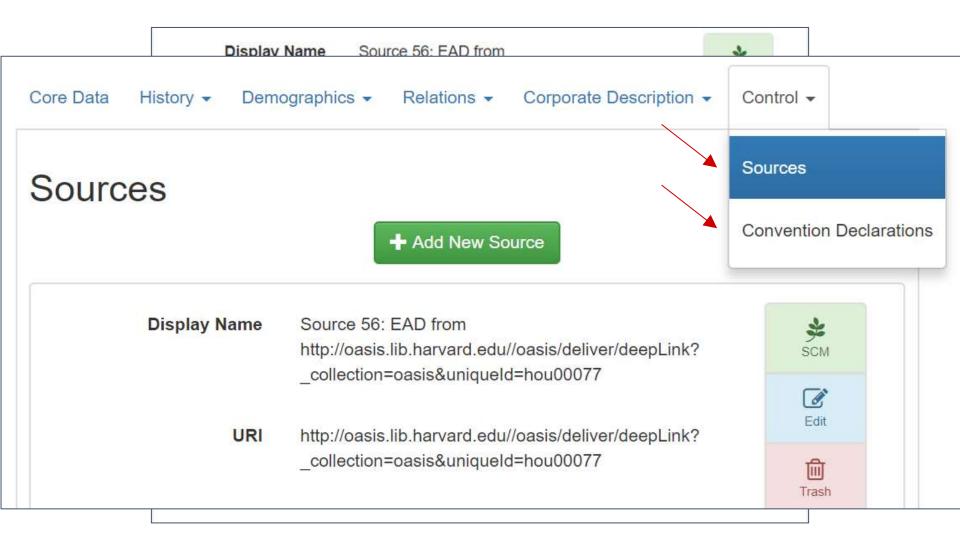






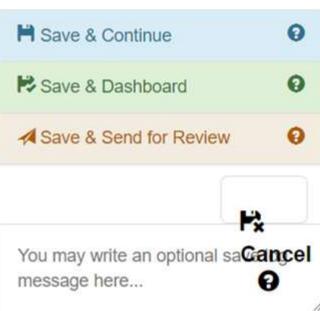






User Actions

You may perform any of the following actions on this constellation.

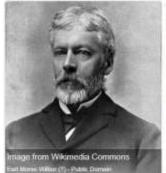


Social Networks and Archival Context

All Types

Search for ...

Search



Eliot, T. L. (Thomas Lamb), 1841-1936





Bigot de Morogues, Marie, 1786-1820





Ehrenberg, Herman,

1816-1866











Session Review

- Description and history of SNAC Cooperative
- Governance
- History Research Tool
- Editing User Interface

Thank You!