

20. Ricc Ferrante (*SI*):

Re-purposing Crowdsourced Projects and
Computational Data

Repurposing Crowdsourcing Projects as Computational Data

Riccardo Ferrante
april 2019

Smithsonian Institution Archives
@raferrante

04

SI Archives crowdsourcing history

Response to researchers

Sprung out of a 2009 project to expose hidden collections material

Let us see this material online!

Online delivery of handwritten materials underscored the need for transcription

Transcription output is ripe for data mining and other computational analysis

SUBPROJECTS



LAND AND FRESHWATER MOLLUSCA
FOUND IN GREAT BRITAIN AND IRELAND
BY ARTHUR STELFOX, 1911-1917



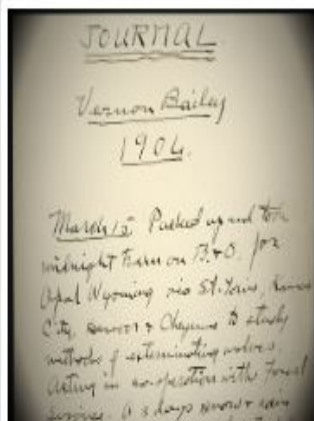
ARTHUR W. STELFOX - DIARY 1:
HYMENOPTERA



ARTHUR W. STELFOX - VOL. 2, DAILY
LISTS



ARTHUR STELFOX - DIARY OF INSECTS,
CHIEFLY HYMENOPTERA, 1929-31, VOL. 3



100% COMPLETE

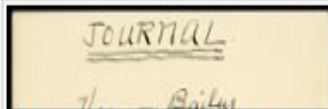
1,720 Total Pages

250 Contributing Members

VERNON ORLANDO BAILEY PAPERS, 1889-1941 AND UNDATED

These projects are completed volunteer transcriptions of Vernon Bailey's field notes from Smithsonian Institution Archives collections documenting his career as a field naturalist between 1889 and 1941. Vernon Orlando Bailey (1864-1942) was born in Manchester, Michigan. At an early age his family moved to Elk River, Minnesota, where he developed an interest in natural history. Around 1885, Bailey began sending collections of birds and mammals to Clinton Hart Merriam, Chief of the newly created Division of Economic Ornithology and Mammalogy of the United States Department of Agriculture (in 1896 the name was changed to the Bureau of Biological Survey). In 1887, Bailey was appointed Special Field Agent for the Division of Economic Ornithology and Mammalogy. His title was changed to Chief Field Naturalist in 1890, and he remained with the Biological Survey until his retirement in 1933. Bailey's chief biological interest was the study of the life history and distribution of mammals.

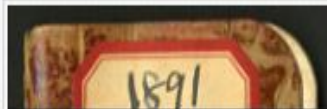
SUBPROJECTS



JOURNAL KEPT BY BAILEY ON FIELD TRIP
TO WYOMING AND NEW MEXICO, MARCH
15-JUNE 1906



VERNON BAILEY - FIELD NOTES, UTAH,
ARIZONA, AND NEVADA, JANUARY 1-
FEBRUARY 11, 1889



VERNON BAILEY - FIELD NOTES,
OCTOBER - NOVEMBER 1891 AND
JANUARY 1892



VERNON BAILEY - FIELD NOTES,
DECEMBER 11-16, 1891 AND JUNE 20-
JULY 19, 1892

Finding relationships

Secretary Baird - The first US National Museum

Hundreds of correspondents

Collections of the correspondents' personal papers

Surfacing additional figures, their social and scientific influence

Events, field of study development

The screenshot shows the Smithsonian Digital Volunteers Transcription Center interface. At the top, the Smithsonian logo and navigation links (HOME, PROJECTS, SEARCH, ABOUT, TIPS, BLOG, MY WORK) are visible. Below the navigation bar, the page title is "SMITHSONIAN DIGITAL VOLUNTEERS: TRANSCRIPTION CENTER". The main content area features a breadcrumb trail: "HOME > PROJECTS > SMITHSONIAN INSTITUTION ARCHIVES" followed by the project title "SPENCER FULLERTON BAIRD INDEX OF CORRESPONDENCE, 1850S-1870S".

The "About the Project" section includes the following text: "Who corresponded with the Smithsonian in its early years? From the 1850's to the 1870's, second Smithsonian Secretary Spencer F. Baird reached out to naturalists and collectors across the globe in pursuit of his dream to establish a national collection. Writing thousands of letters a year, Baird maintained a list of these correspondents in an atlas, identifying their names and organizing them by where they lived." It also contains a call to action: "** Thank you for helping us make future research possible on this network of 19th century scientists and others who played a part in the first two decades of Smithsonian history. Would you be interested in trying your hand at the [personal diary of one of Baird's correspondents, William H. Dall](#) of the Western Union Telegraph Expedition?" and a "Read More" link.

On the right side, a "Completed!" box displays "Project Progress (details)" with a green progress bar. Below this, two statistics are shown: "58 CONTRIBUTING MEMBERS" and "85 TOTAL PAGES".

At the bottom of the page, there is a "Show pages needing (scroll down to load more):" section with filters for "Transcription", "Review", "Notes", and "All". Below this, a grid of eight thumbnail images shows open books with handwritten text, each labeled "Complete" at the bottom.

A vertical "feedback" button is located on the left side of the page.

Challenges

Volume requiring transcription & transcription accuracy

Platform, a.k.a. exposing this for computational analysis and machine learning

Staff skill development, perceived value & relevance

Research community awareness

Finding executive champions