

# Emerging Pathways for the Transcription of Scientific Field Notes

CLIR DHC Symposium

October 12, 2022 | Baltimore, MD

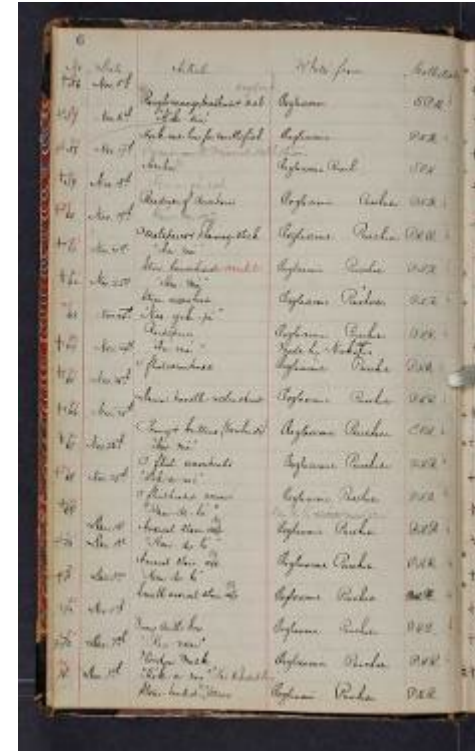
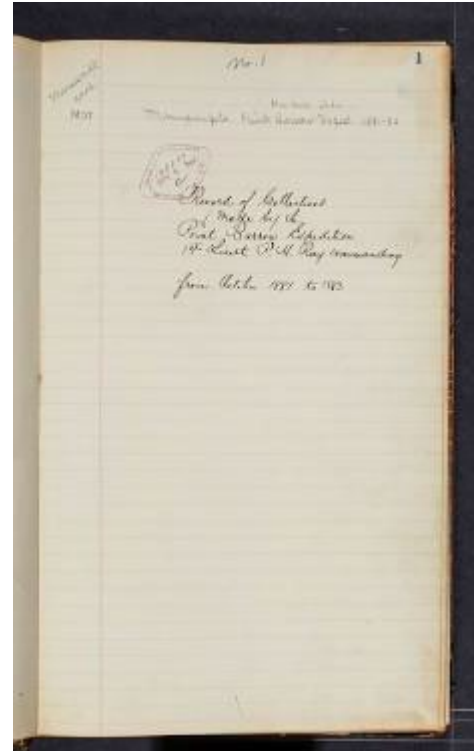


Smithsonian  
*Libraries and Archives*



HARVARD  
LIBRARY

# Field Notes: Invaluable Primary Resources



[Smithsonian Institution Archives](#), SIA RU007203, Created by Murdoch, John, 1852-1925, "Record of collections made by the Point Barrow Polar Expedition, 1881-1883", SIA2015-001147





# Our Panel

## **Riccardo Ferrante**

Associate Director of Information Systems, Digital Lifecycles and User Experience  
Smithsonian Libraries and Archives

## **Sonoe Nakasone**

Data Specialist  
American Women's History Initiative and Smithsonian American Art Museum

## **Katie Mika**

Data Services Librarian  
Harvard Library and Institute for Quantitative Social Science (IQSS)





# A good start

Making scientific field notes accessible  
with The Field Book Project

By Riccardo Ferrante, he/him



**Smithsonian**  
*Libraries and Archives*



**HARVARD**  
LIBRARY

# Building a road of discovery



- First things first - where are they?
  - In 2009, the CLIR *Cataloging Hidden Collections* program funded the Smithsonian's *Exposing Biodiversity Fieldbooks and Original Expedition Journals* with two objectives
    - Catalog an estimated 6,000 fieldbooks and journals whose only online presence, if any, was as a line in an online archival finding aid.
    - Build and launch an online catalog of the field books making it accessible to researchers worldwide.
    - Utilize standards such as MODS and EAC-Cpf to facilitate data sharing with biodiversity and natural history organizations, libraries, archives and aggregators.



# Consequences of a good thing



- Along the way
  - Item-level cataloging revealed a **pressing need for conservation** and preservation treatment. Working directly with decades- and century-old field journals and notes, project catalogers were trained to make preliminary preservation assessments and alert conservators when further handling of a field book posed such risk to its condition.
  - **“Have you digitized this?”**  
followed soon after by **“Do you have a transcript of this?”**



# Pressing forward



- Preservation and Conservation
  - The National Park Service's Save America's Treasures grant program allowed us to stabilize at-risk field books.
- Digitization
  - Small internal grants funded internships to continue small project-based digitization.
- Transcription
  - An inaugural contributor to the Smithsonian's crowd-sourced transcription platform provided a way to respond despite lack of funds



# Pressing forward (more)



- Preservation and Conservation
- Digitization
- Rapid-capture digitization and more cataloging
  - A large grant from the Arcadia Foundation enabled the cataloging of over 3,000 additional scientific biodiversity-related field books, growing the catalog by 50%, and enabling us to implement rapid capture digitization of these materials.
- Access via more platforms, at a deeper level
  - Digitized field books were added to the Field Book Registry, the Biodiversity Heritage Library, and the Digital Public Library of America.
  - Crowd-sourced transcriptions create a full-text searching option on those platforms, also Google.





# Even better together

In 2015, CLIR Launches the *Digitizing Hidden Collections* program.

Ten BHL partners came together to further expand the scientific field notes collections in the Biodiversity Heritage Library.

## **Biodiversity Heritage Library Field Notes Project**

- Smithsonian Institution
- Missouri Botanical Garden, Peter H. Raven Library;
- American Museum of Natural History;
- Yale Peabody Museum;
- Harvard University, Herbaria Botany Libraries;
- Harvard Museum of Comparative Zoology, Ernst Mayr Library;
- University of California, Berkeley Museum of Vertebrate Zoology;
- New York Botanical Garden, LuEsther T. Mertz Library;
- The Field Museum;
- Internet Archive



# It's about access, right?

- The access-based empowerment we expected
  - From cataloging
    - Taxonomic names, geolocations, subjects, authors
  - From digitization
    - Viewing the item
    - If printed, OCR is possible for full text and specialized indexing
- Increased use and downloads globally



# Empowerment surprises

- Transcription impact on **handwritten** field notes
  - Paleography is a dying art, but far from dead yet. Thousands of volunteers worked together, producing high quality transcripts when there were no funds to be found. (thank you, Internet crowd!)
  - Full text searching became possible, including searching across handwritten and printed material, i.e., primary and secondary sources together or apart.
  - Taxonomy links between historic and current names, both scientific and common
  - Hidden or misrepresented collectors



09.252	<i>Opuntia</i>	9310	Collected by A.K. Brattie at request of C. V. Piper on Walbron Island, San Juan Co., Washington.	Aug, '09
09.253	<i>Hylcoereus*</i>		E. L. Merrill, Maunabo, P.I.	Aug. 09.
09.254	<i>Rhipsalis coccinea</i>			Aug 6, 09
09.250	<i>Hylcoereus triangularis</i>			
09.256	<i>Phyllanthus</i> <i>Epiphyllum Gaillardas</i>		→ Mrs. D. D. Gaillard Canal Zone.	
09.257	<i>Opuntia ballii</i> Rose		C. R. Ball, Pecos, Texas #1506	Aug. 16, '09
09.258	<i>Opuntia kuntzei</i> Rose		Dr. Carl Schomburgk, Brownell, Pima Co., Arizona.	Aug. '09
09.259	<i>Anacardium</i> (seed) 11/09	3	Dr. Y. B. McCoy Mexico	Aug. '09
		(seed) 10/10/09	" "	

# Empowerment surprises


Transcription revealed hidden or misrepresented collectors, especially women

- Example: Joseph Nelson Rose's 1851 catalog of specimens documented over 25 women contributors
- Meet Mrs. Katherine Ross Gaillard. Misattributed in several catalogs as "D. D. Gaillard," these entries can now be properly attributed.



# Untapped potential

- **Linked data**



[Main page](#)  
[Community portal](#)  
[Project chat](#)  
[Create a new Item](#)  
[Recent changes](#)  
[Random Item](#)  
[Query Service](#)  
[Nearby](#)  
[Help](#)  
[Donate](#)

[Lexicographical data](#)  
[Create a new Lexeme](#)  
[Recent changes](#)  
[Random Lexeme](#)

[Tools](#)  
[What links here](#)  
[Related changes](#)

Item [Discussion](#)

## Katherine Ross Gaillard (Q81115987)

American botanical collector [edit](#)

[Mrs David DuBose Gaillard](#) | [Mrs D. D. Gaillard](#) | [Katherine Ross Davis](#) | [K. R. Davis](#) | [K. R. Gaillard](#) | [K. Gaillard](#)

[In more languages](#)  
[Configure](#)

Language	Label	Description	Also known as
English	Katherine Ross Gaillard	American botanical collector	<a href="#">Mrs David DuBose Gaillard</a> <a href="#">Mrs D. D. Gaillard</a> <a href="#">Katherine Ross Davis</a> <a href="#">K. R. Davis</a> <a href="#">K. R. Gaillard</a> <a href="#">K. Gaillard</a>
Spanish	Katherine Ross Gaillard	No description defined	
Traditional Chinese	No label defined	No description defined	
Chinese	No label defined	No description defined	

[All entered languages](#)

# Untapped potential

- Extracting data sets from handwritten sources

F.W. True and Daniel W. Prentiss, Jr.  
 "Maine, 1897."  
 Pages 6 and 7 of this field catalog

1897  
July

No.	Name	Sex	Locality	Date	Measurements	Remarks	
116	<i>Microtus pennsylvanicus</i>	♂	Tommy's Island, Hancock Co., Me.	19	113 33 21	Grassy meadow	cheese
17	"	♀	"	"	" " "	Skull	"
18	"	♂	"	"	" " "	"	"
19	"	♂	"	"	" " "	"	"
120	<i>Blarina brevicauda</i>	♀	Brookline, Hancock Co., Me	20	192 29 16	Caught in wet meadow	skull
21	<i>Eutamias gapperi</i>	♂	"	21	131 36.5 18	"	cheese
22	<i>Blarina brevicauda</i>	♀	"	21	115 29 16.5	"	"
23	Sorex	♂	"	23	- - 12	Sphagnum swamp Tail broken off	meat
24	<i>Microtus pennsylvanicus</i>	♂	"	23	163 43 21	"	meat
25	<i>Blarina brevicauda</i>	♀	"	24	130 28 15	" many specimens	"
26	"	♀	"	"	128 29 15	"	"
27	"	♀	"	25	121 26 17	"	"
28	<i>Sciurus hudsonicus</i>	♂	"	26	- - 46	Spruce woods near clearing, soil defective	meat
29	<i>Microtus pennsylvanicus</i>	♀	"	28	- 185 46	Sp. High thicket	meat
30	"	♂	"	29	174 22.5 15.5	"	meat
31	"	♀	"	1	177 53 20	High hillside	"
32	Sorex	♀	"	1	104 46 25	"	"
33	<i>Mephitis mephitis</i>	♂	"	1	530 78 73	"	meat
34	<i>Microtus pennsylvanicus</i>	♂	"	2	131 38 18.5	"	"
35	"	♂	"	2	103 28.5 18	"	"
36	<i>Blarina brevicauda</i>	♂	"	2	104.5 26 15.3	"	"
37	<i>Mephitis mephitis</i>	♂	"	5	300 192 71	Kindness of Lawrence Jannay	meat
38	<i>Blarina</i>	♀	Walker's Pond Me.	10	118 28 16	swampy ground	meat

**Thank you.**





# Legacy Data to Linked Data

How Linked Data Could Change Field Book Metadata

By Sonoe Nakasone, she/they



**Smithsonian**  
*Libraries and Archives*



**HARVARD**  
LIBRARY





# Vision of 2009 Field Book Project

- Repository of metadata about field books
- Detailed item level descriptions of field books
- Save time and money on travel for research





'Linked Data' = 2006 | Our project = 2009

- Field Book Project used XML schemas: MODS, NCD, and EAC-CPF
- 'Linked data' was around, but new and unfamiliar to many in libraries





# If we did it LD

**Pearl and Hermes Reef --> Honolulu --> Hawai'i --> United States**

Location subjects in field book record

Pearl and Hermes Reef  
Hawai'i  
United States

EVERY TIME, in all 49 items matching "**Pearl and Hermes Reef**"

Works for queries: 1) Show me all "**Pearl and Hermes Reef**" field notes; 2) all "**Hawai'i**" field notes; and 3) all "**United States**" field notes

# If we did it LD

**Pearl and Hermes Reef** --> Honolulu --> **Hawai'i** --> **United States**

As triples...

"Field book X" "depicts location" "**Pearl and Hermes Reef**". (EVERY TIME)

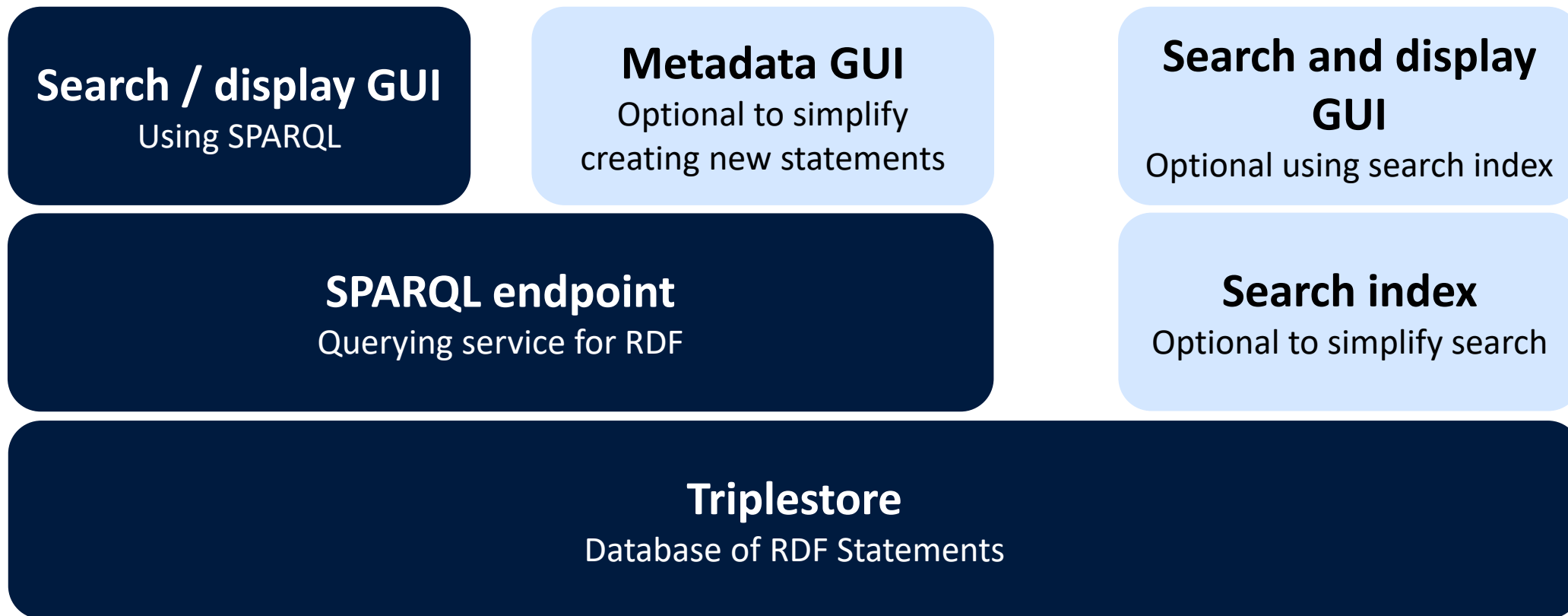
"Pearl and Hermes Reef" "is part of" "**Hawai'i**". (ONCE)

"Hawai'i" "is part of" "**United States**". (ONCE)

Works for queries: 1) Show me all "**Pearl and Hermes Reef**" field notes;  
2) all "**Hawai'i**" and its subsidiaries field notes; and 3) all "**United States**" and its subsidiaries field notes



# Example components of a Linked Data stack





# Wikibase, Wikidata, or similar

- Ready to use linked data infrastructure
- Tools that lower barrier for sparql queries like Query Builder by Jakob Warkotsch (WMDE): <https://bit.ly/3V1pLR5>
- Connectivity to other data on the web and search engine users



# (Truncated) record for a field book

Title: Pearl and Hermes Reef, September 1966, Lewis  
Creator: LEWIS, T. JAMES  
Repository location: Smithsonian Institution Archives  
Places: PEARL AND HERMES REEF  
HAWAII  
UNITED STATES  
Start date: 1966-09-25  
End date: 1966-09-27  
Topic: PACIFIC OCEAN BIOLOGICAL SURVEY PROGRAM  
Ornithology  
Accession Number: SIA RU000245  
Collection: National Museum of Natural History, Pacific Ocean Biological Survey Program, records, circa 1961-1973, with data from 1923

See an example of  
a different field book in  
Wikidata:

[https://bit.ly/  
3rsptVV](https://bit.ly/3rsptVV)



# Looking ahead toward LD for field books

- Field book metadata is already structured data
- Some data shouldn't be open
- Shared data model to aide interoperability







# Thank you

Nakasones@si.edu



# Extracting Datasets from Digital Collections

Katie Mika

Data Services Librarian

Harvard Library & Institute for Quantitative Social Science



Smithsonian  
*Libraries and Archives*



HARVARD  
LIBRARY



# Kinds of data

## How to extract data & datasets from items

## Collections as Data as a distinct service model





# Kinds of Data





*“Data refers to entities used as evidence of phenomena for the purposes of research or scholarship.”*

*“Data are not pure or natural objects with an essence of their own. They exist in a context, taking on meaning from that context and the perspective of the beholder.”*

Borgman, C. (2015). Big data, little data, no data: scholarship in the networked world. Cambridge, Massachusetts: MIT Press.



## LAPEYROUSIA CRUENTA

Blood-spotted Lapeyrouisia

Native of South Africa

Family IRIACACEAE

Iris Family

*Anomatheca cruenta* Lindl. Bot. Beg. 18: pl. 1369. 1830.  
*Lapeyrouisia cruenta* Baker, Hamilt. Ind. 373. 1850.

*Lapeyrouisia* is a genus of some forty-five species, mostly African, inhabiting the dry hilly Cape regions, and one of a series of most beautiful garden subjects from that country, whence came our standard garden flower gladiolus. From the same regions are also produced the fragrant freesia of greenhouses, the belladonna lily, and such other tender "bulbs" as *Watsonia*, *Antholyza* or wand-flower, *Tritonia*, *Crocus*, *Mandragora* and *Ixia*, all brilliantly colored. Like the gladiolus, *Lapeyrouisia* has a corm ("bulb"), and can be propagated in much the same way, so it is treated in the same manner as that popular flower. The corms can be taken up and dried off over winter. South of New York it is probably hardy, needing a covering only during the most severe weather.

Under such conditions it is said to require frequent division of the corms. It is cultivated in Bermuda occasionally and probably in other West Indian islands.

Our illustration was made from greenhouse-grown seedlings, the seed having been sent from the Botanic Garden at Cambridge, England, last summer. In England it is grown in rockeries, and other sunny situations out doors.

*Lapeyrouisia cruenta* is a perennial herb from a small ovoid corm with fibrous reticulated brown coat. The smooth stem is about one foot high, with many narrowly sword-shaped leaves in two ranks. The lower leaves are about six inches long, the upper shorter and narrower. The two to six flowers are in loose terminal spikes, each flower subtended by two ovate-acute herbaceous bracts, one larger than the other. The perianth-tube is slender, straight, about an inch and a half long, the six segments bright red, about a half of an inch long, oblong to lance-shaped and obtuse; three of them plain red, and three larger with a conspicuous dark red spot at the base. The stamens are about one quarter of an inch long, inserted in the throat of the perianth-tube, the anthers two-lobed, the stigma red, two-lobed, on a long slender style. The fruit is a rough, rounded, three-celled capsule, with many round seeds in two series.

KENNETH R. ROYSTON.

Addisonia 11

(Plate 358)

LAPEYROUSIA CRUENTA

Blood spotted Lapeyrouisia

Native of South Africa

Family Iridaceae Iris Family

*Anomatheca cruenta* Lindl. Bot. Beg. 18: pl. 1369. 1830.

*Lapeyrouisia cruenta* Baker, Hamilt. Ind. 373. 1850.

*Lapeyrouisia* is a genus of some forty-five species, mostly African inhabiting the dry hilly Cape regions, and one of a series of most beautiful garden subjects from that country, whence came our standard garden flower gladiolus. From the same regions are also produced the fragrant freesia of greenhouses, the belladonna lily, and such other tender "bulbs" as *Watsonia*, *Antholyza* or wand-flower, *Tritonia*, *Crocus*, *Mandragora* and *Ixia*, all brilliantly colored, like the gladiolus. *Lapeyrouisia* has a corm ("bulb"), and can be propagated in much the same way, so it is treated in the same manner as that popular flower. The corms can be taken up and dried off over winter. South of New York it is probably hardy, needing a covering only during the most severe weather.



LAPEYROUSIA CRUENTA

"Details - Addisonia : Colored Illustrations and Popular Descriptions of Plants - Biodiversity Heritage Library." Accessed October 5, 2022. <https://www.biodiversitylibrary.org/bibliography/5044>.



Smithsonian  
Libraries and Archives



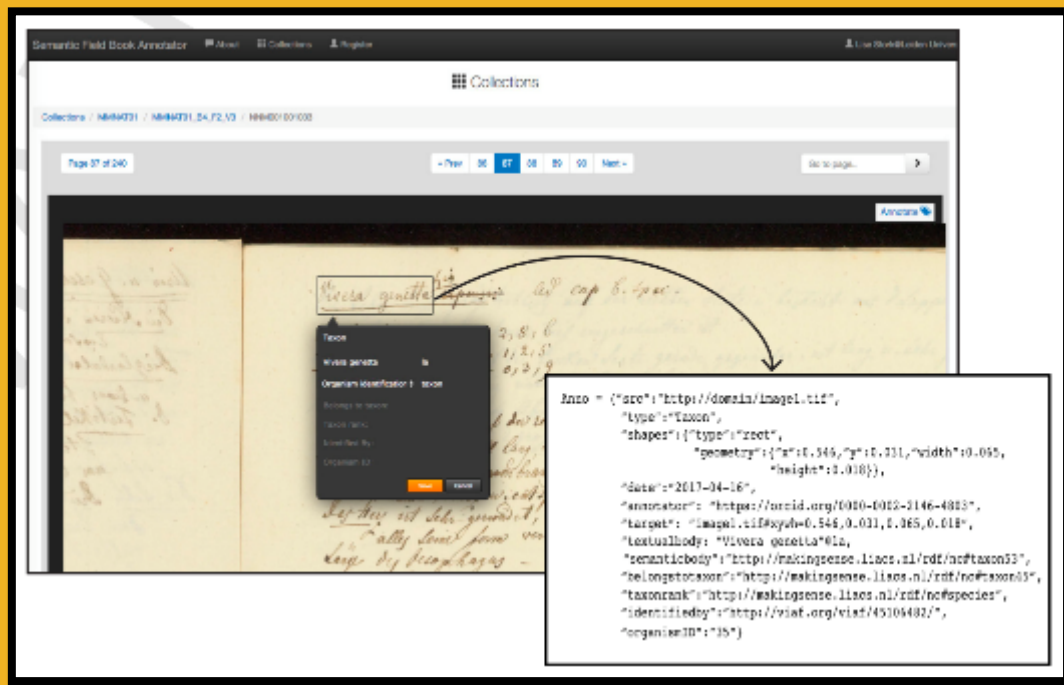
HARVARD  
LIBRARY

```

1 <desc> online.html</https://www.biodiversitylibrary.org/view/11171?url=http://www.zo.org/2261/49/Science-Instance%
2 </html>
3 <html:Abstracts> : colored illustrations and popular descriptions of plants</html:
4 </html:Info>
5 <html:Info type="Abstracts">
6 <html:Abstracts/html:
7 </html:Info>
8 <name type="personal">
9 <html:Publication Mary E. Wilson</html:
10 </html:
11 </html:
12 </html:
13 </html:
14 </html:
15 <name type="corporate">
16 <html:PartNew York Botanical Garden</html:
17 </html:
18 </html:
19 </html:
20 </html:
21 </html:
22 </html:
23 </html:
24 </html:
25 </html:
26 </html:
27 </html:
28 </html:
29 </html:
30 </html:
31 </html:
32 </html:
33 </html:
34 </html:
35 </html:
36 </html:
37 </html:
38 </html:
39 </html:
40 </html:
41 </html:
42 </html:
43 </html:
44 </html:
45 </html:
46 </html:
47 </html:
48 </html:
49 </html:
50 </html:
51 </html:
52 </html:
53 </html:
54 </html:
55 </html:
56 </html:
57 </html:
58 </html:
59 </html:
60 </html:
61 </html:
62 </html:
63 </html:
64 </html:
65 </html:
66 </html:
67 </html:
68 </html:
69 </html:
70 </html:
71 </html:
72 </html:
73 </html:
74 </html:
75 </html:
76 </html:
77 </html:
78 </html:
79 </html:
80 </html:
81 </html:
82 </html:
83 </html:
84 </html:
85 </html:
86 </html:
87 </html:
88 </html:
89 </html:
90 </html:
91 </html:
92 </html:
93 </html:
94 </html:
95 </html:
96 </html:
97 </html:
98 </html:
99 </html:
100 </html:

```

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52
53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78
79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104



"Details - Addisonia : Colored Illustrations and Popular Descriptions of Plants - Biodiversity Heritage Library." Accessed October 5, 2022. <https://www.biodiversitylibrary.org/bibliography/5044>.

Stork, Lise, Andreas Weber, Eulàlia Gassó Miracle, Fons Verbeek, Aske Plaat, Jaap van den Herik, and Katherine Wolstencroft. "Semantic Annotation of Natural History Collections." *Journal of Web Semantics* 59 (December 2019): 100462. <https://doi.org/10.1016/j.websem.2018.06.002>.

Title table of BHL hosted material. Accessed October 5, 2022. <https://about.biodiversitylibrary.org/tools-and-services/developer-and-data-tools/#Data%20Exports>.



Favorite, Capt. Keen, of Portland, arrived and anchored in the Western anchorage. Reported the Jamestown launch as forty miles this side of Chilkah River a week ago, bound up and all well. Drew large seine in the lake for trout with poor success being just the commencement of the season. Small messes are caught with a hook.

Tuesday, June 1st 1880

[[page includes tables with 5 columns - "\*" used to better distinguish separate column fields for readability.]]

[[table]]  
 [[table headings:]]  

Time	Barom.	Air.	Water.	Wind.
6 A.M.	29.64	46	47	Variable
12 M.	50	50	48	[[ditto for: Variable]]
6 P.M.	70	50	48	[[ditto for: Variable]]

 [[/table]]

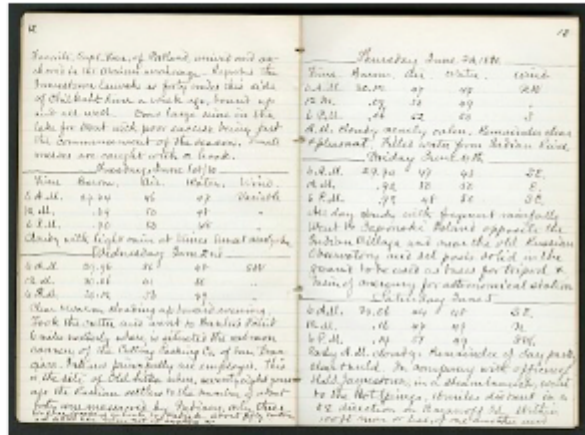
Cloudy with light rain at times sunset nearly clear

Wednesday June 2nd

[[table]]  

Time	Barom.	Air.	Water.	Wind.
6 A.M.	29.86	56	48	SW
12 M.	30.06	61	50	[[ditto for: SW]]
6 P.M.	30.12	53	49	[[ditto for: SW]]

 [[/table]]



Coll. No.	Species	Sex	Date	Loc.	Altitude	Time	Wind	Direction	Remarks	Locality	Other
2059	Passer domesticus	♀	July 1886	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2060	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2061	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2062	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2063	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2064	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2065	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2066	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2067	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2068	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2069	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2070	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2071	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2072	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2073	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2074	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2075	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2076	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2077	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2078	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2079	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2080	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2081	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2082	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2083	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2084	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2085	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2086	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2087	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2088	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2089	"	♀	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft
2090	"	♂	"	10° 25' N 140° 15' W	170	1:30	light	SE	nest with 3 eggs	1000 ft	1400 ft

"William H. Dall - Field Notes, 1880 (1 of 2) | Smithsonian Digital Volunteers." Accessed October 7, 2022. <https://transcription.si.edu/project/6980/>

At-sea, 1963-1966, 1968, part 3 : July - August 1966. Pacific Ocean Biological Survey Program - 1966. <https://www.biodiversitylibrary.org/bibliography/148243>





Wikisource

Page: Field Notes of Junius Henderson, Notebook 1.djvu/3

This page has been proofread.

Boulder, Colo.

- July 29, 1905. Saw Say Phoebe and robin, America Flicker, Northern Flicker.
- July 29. Saw Say Phoebe, Heard a America Robin at 11:50 last night.
- July 30. Saw Say Phoebe
- July 31. Heard Pine Siskin at noon at Court House, have seen none for a week. Expense University Professor trip, 2 tickets to Denver Dr. Ramsay and I — \$2.00. Saw a Kingbird and America Robin on way to depot. Left Boulder with Dr. Ramsay at 5:15 p.m., reached Denver about on time. Stayed at Oxford Hotel to be near depot in the morning. Went to City Park and heard band and saw moving pictures including "Stage Robbery" which, to say the least, was not an elevating spectacle, nor helpful to venturesome boys, apt to be carried away with the wildness of such a life.

Category: Proofread

*1905*  
 July 29. *Boulder, Colo.*  
 Saw say Phoebe and flicker  
 Robin, flicker

July 29. *Saw Say Phoebe, Heard a robin at 11:50 last night.*

July 30. *Saw Say Phoebe*

July 31. *Heard siskin at noon at court house. Have seen none for a week. Expense University Professor trip 2 tickets to Denver Dr. Ramsay & I — \$2.00. Saw a Kingbird and robin on way to depot. Left Boulder with Dr. Ramsay at 5:15 p.m. Reached Denver about on time. Stayed at Oxford Hotel to be near depot in the morning. Went to City Park and heard band and saw moving pictures, including "Stage Robbery," which, to say the least, was not an elevating spectacle, nor helpful to venturesome boys, apt to be carried away with the wildness of such a life.*

Show me all "Pearl and Hermes Reef" field notes; all "Hawai'i" and its subsidiaries field notes; and all "United States" and its subsidiaries field notes

Show me all "*Puffinus pacificus*" field notes; where "date" is "July"

Thomer, Andrea, Gaurav Vaidya, Robert Guralnick, David Bloom, and Laura Russell. "From Documents to Datasets: A MediaWiki-Based Method of Annotating and Extracting Species Observations in Century-Old Field Notebooks." *ZooKeys* 209 (July 20, 2012): 235–53. <https://doi.org/10.3897/zookeys.209.3247>.



The following text is generated from uncorrected OCR or manual transcriptions.

[Begin Page: Page [2],Page [3]]

\  
 ^ 2 S'  
 t o \ V  
 s'y  
 v w - 3 V^ V^ A, A^  
 D x V i v ^ A ^ l y V I K A y A  
 I  
 /  
 I' kv F D^  
 V^ i i  
 I A ^ ^ U S  
 ' i  
 30 / V.  
 c 4 ^ L c^ A  
 I t  
 0 A / C  
 L O S V ( V ) 1 / C w . I ? i r  
 f i .  
 / A ;  
 f i P p j | ^ C ^ . j j ) S w i t s  
 4 y . i . ^ I A r > ^ I S  
 I r  
 I f  
 I 1  
 A - ^ r u s o A 4 i  
 K  
 I 1  
 A j ^ 8 ( j j )  
 C^ - C I C I  
 I a / 1 S ( j A c - V j E U v E )  
 S - i V I t  
 I 1  
 S - ^ ^ r o A ^ I A S u w i  
 I t v - i  
 7  
 V^ c - c  
 c  
 V j i A  
 r A  
 /  
 to 4 .  
 7  
 M  
 X  
 I f  
 I A  
 V 0 C A V ' 0 t O  
 \ o s' f o l s t a ^ l o' u j  
 I f a : 1 6 3 " h u r  
 i s " o s' i C s ^ u i W  
 1 7 3 k V I U "

Field No.	Sp. No.	Date	Locality	Sex	Tail Len	Other Measurements	Remarks
2087	2087	July 1962	1st. 10000' 10' N	♂	120	...	Huber 836
2088	2088	July 1962	1st. 10000' 10' N	♂	120	...	Chandler 9
2089	2089	July 1962	1st. 10000' 10' N	♂	120	...	BAComia 138
2090	2090	July 1962	1st. 10000' 10' N	♂	120	...	21
2091	2091	July 1962	1st. 10000' 10' N	♂	120	...	Huber 837
2092	2092	July 1962	1st. 10000' 10' N	♂	120	...	PEARSON 23
2093	2093	July 1962	1st. 10000' 10' N	♂	120	...	Huber 838
2094	2094	July 1962	1st. 10000' 10' N	♂	120	...	839
2095	2095	July 1962	1st. 10000' 10' N	♂	120	...	846
2096	2096	July 1962	1st. 10000' 10' N	♂	120	...	PEARSON 24
2097	2097	July 1962	1st. 10000' 10' N	♂	120	...	Bulmer 955
2098	2098	July 1962	1st. 10000' 10' N	♂	120	...	BAComia 134
2099	2099	July 1962	1st. 10000' 10' N	♂	120	...	Chandler 10
2100	2100	July 1962	1st. 10000' 10' N	♂	120	...	PEARSON 25
2101	2101	July 1962	1st. 10000' 10' N	♂	120	...	5 Skelton
2102	2102	July 1962	1st. 10000' 10' N	♂	120	...	Hub 53
2103	2103	July 1962	1st. 10000' 10' N	♂	120	...	Huber 840
2104	2104	July 1962	1st. 10000' 10' N	♂	120	...	851
2105	2105	July 1962	1st. 10000' 10' N	♂	120	...	852
2106	2106	July 1962	1st. 10000' 10' N	♂	120	...	5 Skelton
2107	2107	July 1962	1st. 10000' 10' N	♂	120	...	Hub 54
2108	2108	July 1962	1st. 10000' 10' N	♂	120	...	52
2109	2109	July 1962	1st. 10000' 10' N	♂	120	...	Huber 853
2110	2110	July 1962	1st. 10000' 10' N	♂	120	...	68
2111	2111	July 1962	1st. 10000' 10' N	♂	120	...	Huber 841
2112	2112	July 1962	1st. 10000' 10' N	♂	120	...	804
2113	2113	July 1962	1st. 10000' 10' N	♂	120	...	Hub 53

E	F	G	H	I	J	K	L	M	N
1966	LOCATION	FAT	CON.	REproduction	Remarks				
	180 05' N 161 67' W			QUA A' IL 5/ Color wabes					Huber 836
	150 0' N 163	473.7		I. ovum / mm					Chandler 9
	15 05' N 163	420				FSB			BAComia 138
	110 38' N 166 0 23 W	very light		Dv Ary 8x 44-1					21
	110 34' N 16	134.7	Light - mad	testes 2x1 mm (black)	BES				Huber 837
	090 56' N 168 06' W	1154+		JonAcs Nat f color Notes					PEARSON 23
		360	Light - med.	ROO Ary 7x 5-2m					Huber 838
		366	med.	QUA MINUTE					839
	600 18' N 166	649	Very halt	4 Fat's 12 x 6 Color Watts 9-007	FS				846
	06 0 10' N 17	422		ovg granular color Natis 4-4-00 9					PEARSON 24
	050 33' N 172 0 09' W	Very light		Ov Ary 6 x 2r color wates	4-011 BES				Bulmer 955
	05 22' N 172	410		L 7. 4y/mm color Natis					BAComia 134
		442		Ovary 5x4mm color wates			SB		Chandler 10
	230 44' N 17	315		LT 2x1mm RT 1X1mm					PEARSON 25
	00 47' N 176	318.7	Light	Color Nat Es			5 Skelton		Hub
	0 0 47' N 176	207.5		1+ 3x2mm R color wates	Plus **		5 dN		Hub 53
	00 47' N 176	160.5							Huber 840
	0 46' N 177.3	170							851
		186		NOB P	molt				852
	0 0 46' N 177	364	May	120765 3X2r Color Notes		5	SkElston		Hub 54
	10 00' N 176	175		Ovary 5V64.3 color wates					52
	10 05' N 176 0 56 W								Huber 853
	0 50' N 176.4	371		Fat's 4x 2 n color wates		FSB			68
	0 1 58' N 176	157	NO	Sram Ar		BS	Skin		Huber 841
	1 0 03' N 176 0 36' W								804
	100 1' N 1760	172	light	1. 604 / mm.					Skeleton
									Hub 53



# How to extract data and datasets from items



Field number	Species	Sex	Date	1966	Location	Wgt.	Tot. Len.	Measurements	Remarks			
2259	<i>Phaethon rubricauda</i>	♀	6 July	18° 05' N 148° 30' W	670		very light	very light	color notes	S	SRW	Huber 836
60	<i>P. externa externa</i>	♀	7 July	15° 1' N 148° 19' W	473.2		very light	very light	color notes	BS	"	Huber 837
61	"	♂	"	15° 08' N 148° 26' W	429		light	L.T. 42.2mm	color notes	BSB	"	Huber 838
62	"	♀	8 July	15° 58' N 148° 23' W	472		very light	very light	color notes	ES	"	Huber 839
63	<i>Pt. hypoleuca nigripennis</i>	♂	9 July	15° 34' N 148° 27' W	134.7		light-mid	Feet 22mm (dark)	color notes	0 ES	"	Huber 840
64	<i>Sterna fuscata</i>	♀	"	08° 06' N 148° 06' W	157		light	quarry 100mm	color notes	S	"	Huber 841
65	<i>Bulweria pacifica</i>	♀	"	09 36' N 148° 25' W	367		very-mid	quarry 70mm	molt	S	"	Huber 842
66	"	♀	"	"	365		mid	quarry 70mm	molt	ES	"	Huber 843
67	"	♀	"	09 38' N 148° 21' W	581		mid	quarry 70mm	molt 4000	ES	"	Huber 844
68	<i>Phaethon rubricauda</i>	♂	"	09° 18' N 148° 25' W	844		very light	Feet 22.5mm	color notes	ES	"	Huber 845
69	Last number											
70	<i>Bulweria galapagensis</i>	♀	10 July	06° 15' N 170° 24' W	422		very light	quarry 70mm	molt	ES	"	Huber 846
71	<i>Pt. hypoleuca nigripennis</i>	♀	"	07° 36' N 170° 05' W			very light	quarry 60mm	color notes	4-100 ES	"	Huber 847
72	<i>Bulweria pacifica</i>	♂	"	05° 22' N 170° 21' W	710			L.T. 49mm dark	color notes	BS	"	Huber 848
73	"	♀	"	"	442		light	quarry 60mm	color notes	SB	"	Huber 849
74	"	♂	11 July	08° 44' N 170° 54' W	315		no	L.T. 30mm AT 11mm	molt	S	"	Huber 850
75	"	♂	13 July	0° 47' N 170° 57' W	315.7		light	Feet 22.5mm	color notes	S	Skull	Huber 851
76	<i>Sterna fuscata</i>	♂	"	0° 47' N 170° 58' W	202.5			L.T. 30mm AT 11mm	color notes	S	Skull	Huber 852
77	"	♀	"	0° 47' N 171° 14' W	160.5					B	Skull	Huber 853
78	"	♀	"	0° 46' N 171° 51' W	179					"	"	Huber 854
79	"	♀	"	"	186					"	"	Huber 855
80	<i>Bulweria pacifica</i>	♂	"	0° 46' N 171° 36' W	364		very	quarry 30mm	molt	S	Skull	Huber 856
81	<i>Sterna fuscata</i>	♀	14 July	1° 08' N 170° 52' W	175		light	quarry 50mm	color notes	BS	"	Huber 857
82	"	♀	"	1° 05' N 170° 53' W				quarry 20mm AT 11mm		BS	Almond	Huber 858
83	<i>Bulweria pacifica</i>	♂	"	0° 50' N 170° 43' W	371		light	Feet 22.5mm	molt	0 BS	Skull	Huber 859
84	<i>Sterna fuscata</i>	♀	15 July	0° 58' N 170° 37' W	157		no	quarry 50mm		BS	Skull	Huber 860
85	"	♀	"	1° 03' N 170° 30' W				quarry 50mm			Almond	Huber 861
86	"	♀	"	1° 01' N 170° 26' W	172		light	quarry 50mm	molt	S	Skull	Huber 862

Field Number	SPECIES	SEX	DATE	1966	LOCATION	WGT.	FAT CO.	REproduction	Remarks		
2259	Phaethon rubricauda	♂	6 July	1956	18°05'N 160°40'W	670		OV 11x14mm OVI 5/2 - im. adme	Colcwates	SKIN	Huber 836 13
60	Pt. Externa Externa	♀	7 July		150°11'N 163°19'W	473.7	Very light	NRP OVI 12x5mm OVI 10x7mm	molt, color wates	=	Chancl 9
61	"	♀	"		15°03'N 163°26'W	429	light	IVBP	molt color Notes	ES	BAL con 103
62	"	♂	8 July		11°38'N 160°23'W	472	very light	DVA 844-1 NRP	molt color Notes	ES	PEANS 21
63	Puffinus puffinus	♂	9 July		11°34'N 166°27'W	131.7	Light med	Testes 2x mm (black)	color Notes	BS	Huber 837
64	Sterna fuscata	♂	"		09°56'N 160°05'W	157	11s 1+	Jon Ar Not found.	color Notes	S	PEARS 23
65	Puffinus pacificus	♂	"		09°36'N 168°25'W	360	Light-med.	ROVAR 7x5.2mm NRP OVI 10x7mm	molt	S	Huber 838
66	"	♂	"		11	365	med.	VcB MINUTE	molt	S	" 839
67	"	♂	"		09°38'N 168°23'W	383	med.	DoA/1x4m, 01	molt 4 00-25	S	PEARS 22
68	Phaethon rubricauda	♂	"		60°18'N 168°25'W	849	very light	+Es: F12x6mm	Molt 11006 Colcwates 007 ES	S	Haber 846
69	Lost Number										
70	Puffinus puffinus	♂	10 July		06°19'N 171°29'W	422	very light	NRP OVI 12x5mm OVI 10x7mm	molt 4 008 color Notes 4-01 EBS	S	PEARS 24
71	Pt. hypoleuca	♂	"		05°03'N 172°09'W	410	Very light	OVI 12x5mm	color Notes 4-01 EBS	S	Bulmer 935
72	Puffinus pacificus	♂	"		05°22'N 172°21'W	410	light	OVI 12x5mm	color Notes	BS	BAL con 104
73	"	♂	"		"	442	light	OVI 12x5mm	color Notes	SB	Chancl 10
74	"	♂	"		2°34'N 173°05'W	315	NO	LI 2x1mm RT 1x1mm	molt	S	PEARS 25
75	"	♂	12 July		00°47'N 176°57'W	318.7	light	NRP	molt	S	SKEELTON Hoff 67
76	Sterna fuscata	♂	"		0°01'N 176°58'W	207.5		LI 3x1mm RT 2x1mm	color Notes 793-AR 145	SKIN	Hoff 63
77	"	♂	"		00°47'N 176°14'W	160.5			color wates Flu	S	Alcoholic Huber 890
78	"	♂	"		0°46'N 177°34'W	179				S	" 891
79	"	♂	"		"	186				S	" 892
80	Puffinus pacificus	♂	"		00°16'N 177°05'W	364	heavy	NOBP 120 (3x2mm)	molt color Notes	S	SKEELTON Hoff 64
81	Sterna fuscata	♂	1 July		10°00'N 176°52'W	175	light	B. bare Ovar 5x4.1mm B. bare	colcwates	S	" 52
82	"	♂	"		10°05'N 176°55'W					BS	Alcoholic Huber 893
83	Puffinus pacificus	♂	"		0°50'N 176°43'W	371	light	NRP Testes 4x2mm	Molt color wates	EBS	SKEELTON Hoff 68
84	Sterna fuscata	♂	"		0°58'N 176°37'W	157	NO	veral BP Ovar 5x4.1mm B. bare		BS	Skin Huber 841
85	"	♂	"		10°03'N 176°36'W					S	Alcoholic " 894
86	"	♂	"		10°01'N 176°36'W	172	light	B. from plate OVI 60x7mm	Molt Colmites	S	SKEELTON Hoff 53



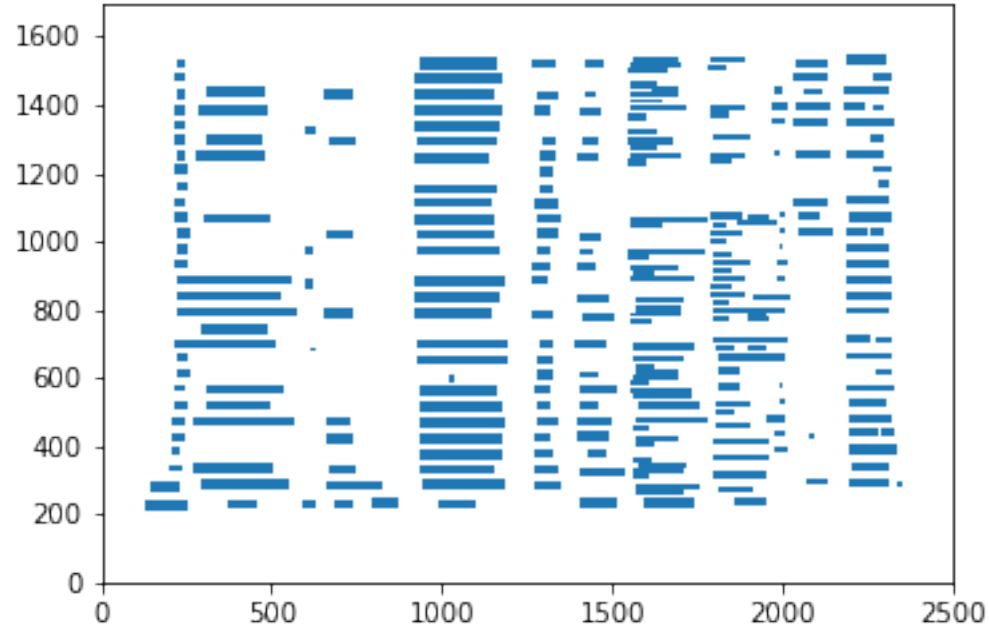
Microsoft

Spec Number	Species	SEX	DATE	LOCALITY	ALTITUDE	ALCO	Preparation	Remarks	Collector	Number
60	Phyllanthus leucostachyus	♂	6 July 1970	10022N 100° 40' W	427.0	Weight	Herbarium	collected	W. S. Gentry	444886
61	Phyllanthus leucostachyus	♀	7 July	10022N 100° 40' W	427.0	Weight	Herbarium	collected	W. S. Gentry	444887
62	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	422	Weight	Herbarium	collected	W. S. Gentry	444888
63	Phyllanthus leucostachyus	♂	9 July	10024N 100° 27' W	424.7	Weight	Herbarium	collected	W. S. Gentry	444889
64	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	427	Weight	Herbarium	collected	W. S. Gentry	444890
65	Phyllanthus leucostachyus	♂	9 July	10024N 100° 27' W	424.7	Weight	Herbarium	collected	W. S. Gentry	444891
66	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	427	Weight	Herbarium	collected	W. S. Gentry	444892
67	Phyllanthus leucostachyus	♂	9 July	10024N 100° 27' W	424.7	Weight	Herbarium	collected	W. S. Gentry	444893
68	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	427	Weight	Herbarium	collected	W. S. Gentry	444894
69	Phyllanthus leucostachyus	♂	9 July	10024N 100° 27' W	424.7	Weight	Herbarium	collected	W. S. Gentry	444895
70	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	427	Weight	Herbarium	collected	W. S. Gentry	444896
71	Phyllanthus leucostachyus	♂	9 July	10024N 100° 27' W	424.7	Weight	Herbarium	collected	W. S. Gentry	444897
72	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	427	Weight	Herbarium	collected	W. S. Gentry	444898
73	Phyllanthus leucostachyus	♂	9 July	10024N 100° 27' W	424.7	Weight	Herbarium	collected	W. S. Gentry	444899
74	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	427	Weight	Herbarium	collected	W. S. Gentry	444900
75	Phyllanthus leucostachyus	♂	9 July	10024N 100° 27' W	424.7	Weight	Herbarium	collected	W. S. Gentry	444901
76	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	427	Weight	Herbarium	collected	W. S. Gentry	444902
77	Phyllanthus leucostachyus	♂	9 July	10024N 100° 27' W	424.7	Weight	Herbarium	collected	W. S. Gentry	444903
78	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	427	Weight	Herbarium	collected	W. S. Gentry	444904
79	Phyllanthus leucostachyus	♂	9 July	10024N 100° 27' W	424.7	Weight	Herbarium	collected	W. S. Gentry	444905
80	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	427	Weight	Herbarium	collected	W. S. Gentry	444906
81	Phyllanthus leucostachyus	♂	9 July	10024N 100° 27' W	424.7	Weight	Herbarium	collected	W. S. Gentry	444907
82	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	427	Weight	Herbarium	collected	W. S. Gentry	444908
83	Phyllanthus leucostachyus	♂	9 July	10024N 100° 27' W	424.7	Weight	Herbarium	collected	W. S. Gentry	444909
84	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	427	Weight	Herbarium	collected	W. S. Gentry	444910
85	Phyllanthus leucostachyus	♂	9 July	10024N 100° 27' W	424.7	Weight	Herbarium	collected	W. S. Gentry	444911
86	Phyllanthus leucostachyus	♂	9 July	10022N 100° 43' W	427	Weight	Herbarium	collected	W. S. Gentry	444912

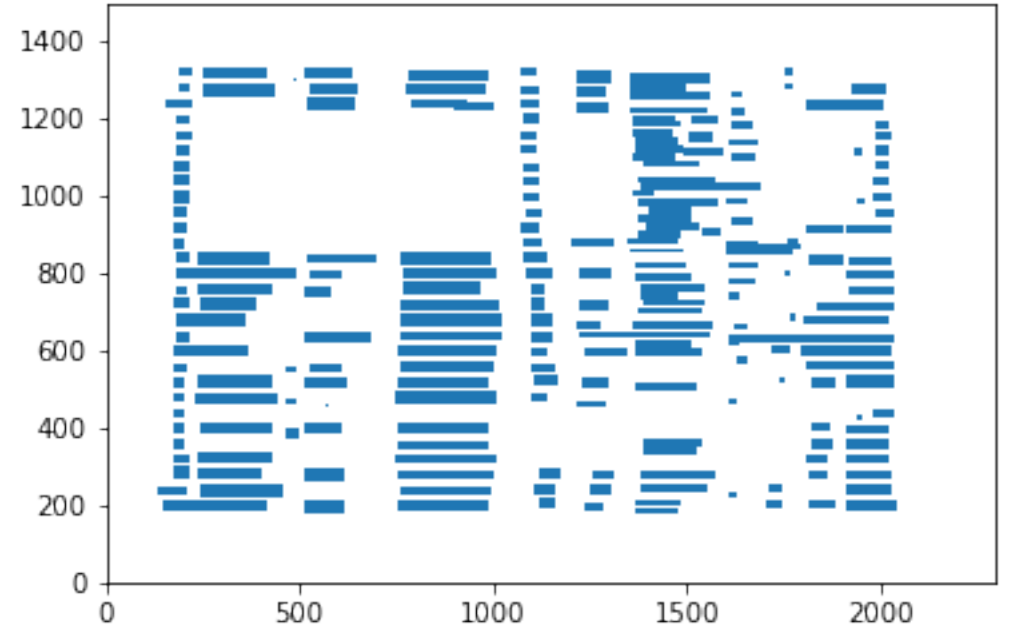
```

    "confidence": 0.986,
    "text": "Alco"
  }
}
},
{
  "appearance": {
    "style": {
      "confidence": 0.909,
      "name": "handwriting"
    }
  },
  "boundingBox": [
    2468,
    254,
    2613,
    253,
    2613,
    281,
    2468,
    283
  ],
  "text": "Haber 872",
  "words": [
    {
      "boundingBox": [
        2469,
        256,
        2540,
        255,
        2540,
        283,
        2469,
        282
      ],
      "confidence": 0.421
    }
  ]
}

```



Page 4



Page 7

## Creating data tables from images of manuscript field notes

This script uses the bounding box coordinates to generate a pandas DataFrame

Bounding box notes:

Bounding boxes are made up of 8 coordinates that identify points along the x and y axes and create a box. The coordinates are as follows:  $[0,1, 2,3, 4,5, 6,7] == [x1,y1, x2,y2, x3,y3, x4,y4]$ . This information helps to create the dictionaries of `x_coords` and `y_coords` used throughout the script.

### 1. Setup

- Import libraries and data
- Setup vars needed later
- Plot boxes (just helpful to see)

```
In [71]: # Import libs and data
import json
import pandas as pd
import itertools
import collections
import statistics
import numpy as np

f = open('HP_out/sea18631966196800natlib_0006.handprint-microsoft.json')
data = json.load(f)
```

```
In [72]: #list of lines
lines = []
for d in data['analyzeResult']['readRegular'][0]['lines']:
    lines.append(d)
```

```
In [73]: #dict of all x_coords of lines
x_coords_lines = {}
for ln in lines:
    temp_list = []
    temp_list.append([ln['boundingBox'][0], ln['boundingBox'][2], ln['boundingBox'][4], ln['boundingBox'][6]])
    text = ln['text']
    x_coords_lines[text] = temp_list[0]
```

```
In [74]: #dict of y_coords of lines
y_coords_lines = {}
for ln in lines:
    temp_list = []
    temp_list.append([ln['boundingBox'][1], ln['boundingBox'][3], ln['boundingBox'][5], ln['boundingBox'][7]])
    text = ln['text']
    y_coords_lines[text] = temp_list[0]
```



[https://github.com/kmika11/bhl\\_unlocking\\_datatables/blob/main/creating\\_tables\\_using\\_bounding\\_boxes.ipynb](https://github.com/kmika11/bhl_unlocking_datatables/blob/main/creating_tables_using_bounding_boxes.ipynb)





col_1	1	2	3	4	5	6	7	8	9	10	11	12	13
Field Number	SPECIES	SEX	DATE	1966	LOCATION		FAT CON .	REproduction		Remarks			
2259	PhAsthew rubricAndA	#NAME?	6-Jul-56		180 05'N 160 40' W	670"		OUA A' IL 5/225 -pota4me	Color watss				Huber 836
60	Pt. ExterNA ExterNA		7-Jul		150 II'N 1830 19'W	473.7		L . ovum / mm					Chandler 9
61					15 03'N 1630 26' W	429					ESB		BALcomb 103
62		0	8-Jul		110 38'N 166 0 23 W		very light	DvAry 8X 44-1					21
63	it hypoluca NigrIDENNIS	No	9-Jul		110 34' N 1660 27'W	134.7	Light - mad	testss 2x1 mm (black )		BES			Huber 837
64	SterNA tuscatA				090 56N 1680 06'W		1154+	JonAds Not found.	color Notes				PEARSON 23
65	Puffinus pacificus					360	Light - med.	RODAry 7x 5-2m					Huber 838
29						365	med.	OUA MINUTE					839
68	PhaEthon rubricandA	KO			690 18'N 168025 W	849	Very halt	+EstEs 12 x 6mm	Color Watts 9 -007		ES		846
70	Puffinus puffinus NEWEilli		10-Jul		06 0 19'N 1710 29' W	422		oug granular	color Notss 4-010 EBS	H-00 9			PEARson 24
71	Pt. hypoluca Nigri PENNIS				050 35'N 1720 09'W		Very light	Ov Ary 6 x 2mm	colarvates	4-011 BES			Bulmer 935
72	Puffivins pacificus	31			05 22'N 1720 21' w	410		L. 7. 4x/mm dark,	colorNotis				BALcom5 104
23						442		Ovary 5x4mm	color vates		SB		Chandler 10
74		8			2:30 44'N 173059 W	315		LT 2x1mm RT 1Xlmm					PEARSON 25
75			13 tily		00 47'N 1760 57W	318.7	Light		ColorNatEs		5	SkElEton	Hoff
76	SterNA EuscatA				0 0 47' N 1760 58' W	207.5		L+ 3X/mM Rt 2xl --	colorwatss	Flus **		SKIN	Hott 63
77					00 47'N 176 14 W	160.6							Huber 890
78					0 46'N 177 34'W	179							891
79						186		NOBP	molt				892
80	Puttivis pacificus				0 0 46'N 1770 36 'W	364	klavy	120765 3X2mm	Color Notes		S	SkElston	Hoff 64
81	StERNA fuscato		14-Jul		10 00N 176 520	175		Ovary 5x4.1m	color wotes				52

[https://github.com/kmika11/bhl\\_unlocking\\_datatables/blob/main/sea19631966196800natib\\_0004.csv](https://github.com/kmika11/bhl_unlocking_datatables/blob/main/sea19631966196800natib_0004.csv)



# What we learned

- Can work for optimal items
- Contribute datasets to Open Data landscape
- Validation is iterative
- “Low hanging fruit”





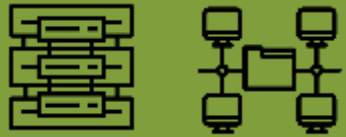
# Collections as Data as a distinct service model



# Granularity Curation Publication

Field #	Name	Sex	Date	Locality	Elev.	Wt.	T.L.	T.V.	M.F.	E.	Notes	
11741	MINIOPTEBUS		1953 APRIL 5	CAPE VOGEL PENINSULA, MANUA	S.L.	-	-	-	-	-	IN FORMALIN. SHOT AFTER DARK - BEST HOUSE.	
42	"		"	"	S.L.	-	-	-	-	-	"	
43	"		"	"	S.L.	-	-	-	-	-	"	
44	"		"	"	S.L.	-	-	-	-	-	"	
11745	RATTUS	♂	"	"	15 M.	33 oz.	(216)	150	40	15	SKIN + SKULL	RAIN FOREST
* 46	"	♀	"	"	15 M.	47 oz.	(306)	183	39	14	"	"
* 47	MELOMYS	♀	"	"	15 M.	53 oz.	249	137	30	11	"	"
* 48	RATTUS	♂	"	"	S.L.	7.6 oz.	265	167	37	15	"	1° SHORT GRASS
* 49	MELOMYS	♀	APRIL 6	"	15 M.	2.3 oz.	268	141	27.5	12	"	RAIN FOREST
11750	"	♂	"	"	15 M.	1.0 oz.	209	112	25.5	10	"	"
* 51	RATTUS	♂	"	"	S.L.	1.2 oz.	(200)	26	12	"	"	BANANA TREE
* 52	DORSONIA	♀	"	"	S.L.	11.8 oz.	260	91	45	31	"	"
* 53	"	♀	"	"	S.L.	11.3 oz.	248	88	37	30	"	"
* 54	"	♀	"	"	S.L.	11.8 oz.	254	26	44	31	"	"
11755	"	♀	"	"	S.L.	11.3 oz.	247	27	38	32	"	"
56	MOUSE	IM	APRIL 7	"	S.L.	-	-	-	-	-	IN FORMALIN	HOUSE IN VILLAGE
57	"	IM	"	"	S.L.	-	-	-	-	-	"	"
* 58	RATTUS	♀	"	"	S.L.	1.2 oz.	233	117	24	12	SKIN + SKULL	BANANA TREE
* 59	PETAURUS	♀	"	"	S.L.	3.1 oz.	353	177	28.5	19	"	"
11760	MELOMYS	♂	"	"	15 M.	1.4 oz.	(215)	(91)	26.5	11	"	"
* 61	RATTUS	♀	"	"	15 M.	5.4 oz.	(317)	144	35.5	15	"	"
* 62	PETAURUS	♂	APR. 8	"	S.L.	3.8 oz.	338	182	28	17	"	"
* 63	RATTUS	♂	"	"	S.L.	8.7 oz.	(354)	172	42.5	14	"	"
* 64	MELOMYS	♀	"	"	S.L.	2.6 oz.	288	148	28.5	-	"	"
11765	MINIOPTEBUS	♀	"	"	S.L.	63	743	8.5	7	6	"	"
66	Small, shrew-like	♀	"	"	S.L.	-	-	-	-	-	IN FORMALIN	"
67	MINIOPTEBUS	♀	"	"	S.L.	-	-	-	-	-	"	"
* 68	DORSONIA	♂	APR. 9	"	65 M.	17.2 oz.	280	30	40	30	SKIN + SKULL	IN LIMESTONE CAVES
* 69	"	♂	"	"	"	17.5 oz.	281	29	42.5	32	"	"
11770	"	♂	"	"	"	17.7 oz.	285	31	44	30	"	"

Van Deusen, Hobart M., American Museum of Natural History, and Archbold Expedition to New Guinea (1953). 4th Archbold Expedition to New Guinea, 1953 Papua -- Cape Vogel, Maneau Range (Mt. Dayman), Goodenough Island, 1953. <https://doi.org/10.5962/bhl.title.133896>.



Collections librarians  
 Archivists  
 Research librarians  
 Researchers

Systems librarians  
 IT  
 Digitization  
 Digital preservation

Data curators  
 Research librarians

Cataloging librarians  
 Metadata librarians  
 Systems librarians  
 External communities



**Thanks!**

