



Smithsonian
National Museum of Natural History

Digitization: Past, Present, and Future

Paul Kimberly and Rebecca Snyder

Overview

- The Smithsonian Institution
- The National Museum of Natural History (NMNH)
- NMNH Collections
- NMNH Digitization

The Smithsonian Institution

- Established in 1846 "for the increase and diffusion of knowledge"
- 19 museums and the National Zoological Park
- 168 affiliated museums throughout the U.S., Panama, and Puerto Rico



The NMNH

NMNH Consists of:

- 2 buildings: NHB & MSC
- 7 scientific departments
- National Anthropological Archives & Human Studies Film Archives (NAA & HSFA)
- Education & Outreach (Q?rius)
- Onsite affiliate organizations
- Marine station in Ft. Pierce, FL
- Artic Studies Center in Anchorage, AK
- Laboratories of Analytical Biology
- Onsite libraries with ~20 special collections



Staffing



- 470 staff (350 science staff)
- 50 contractors (non-construction)
- 42 staff from Affiliated Agencies
- 750 Fellows and Academic Appointments
- 11,000 short-term visiting researchers (<30 days)
- 600+ volunteers and interns

Holdings

SD600* collections

- Collection records
- Descriptive records
- Digital Surrogates
- Finding aids
- Specimens/objects
- Taxonomic inventories
- Documentation such as field notebooks and ledger books
- Illustrations
- Archival materials

Non-SD600 materials

- Conservation records
- Illustrations
- Maps
- Publications and manuscripts
- Exhibit materials

* Smithsonian Directive 600 is the SI's official documentation for collections care and preservation.

How Much We Have

Our best guess is

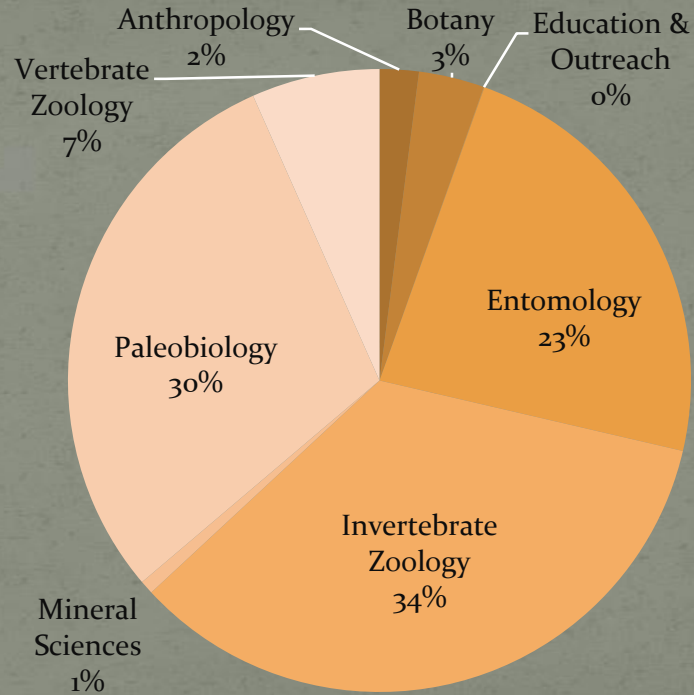
144.6 million

Specimens, Genetic Samples & Objects

What Does 144 million look like?



Collections



What is Digitization?

Digitization is:

- Creation of digital descriptive records (databasing)
- Creation of digital surrogates (images representing collection items)
- Transcription of printed data, metadata, field books, and other supporting materials including collection item labels
- Transferring analog information to digital (35 mm slides, audio recordings, motion picture film, etc.)

What is Digitization?

Digitization is:

- Creation of digital descriptive records (databasing)
- Creation of digital surrogates (images representing collection items)
- Transcription of printed data, metadata, field books, and other supporting materials including collection item labels
- Transferring analog information to digital (35 mm slides, audio recordings, motion picture film, etc.)

Digitization is not:

- Simply taking digital imagery

How Much Is Digitized

6.8 Million Database Records

38.5 Million Objects Represented

~900,000 Surrogates (images of objects)

6.4 Million Web Accessible Records

THE PAST



Paleobiology Laboratory
(date unknown)

How We Got Here

- Early 1960's computers begin to appear at NMNH
- 1963 - Creation of Automated Data Processing Committee (ADP)
- 1967 - SI purchases a series of mainframes, first database online, 10,000 specimen records created
- 1970 - NMNH creates custom database – SELGEM; Early scanning equipment is purchased
- 1979 - Begin the “Great Inventory”
- 1985 - Begin migration from SELGEM to INQUIRE
- 1992 - NMNH and parts of SI join the internet; Digital cameras begin to appear
- 2002 - Begin migration to EMu, ~1 million INQUIRE records; Hundreds of scanners, cameras and digitizing equipment onsite
- 2011 - Transaction Management integrated with EMu
- 2013 - SI Transcription Center goes live
- 2014 - Bumble Bee Rapid Capture Pilot Project

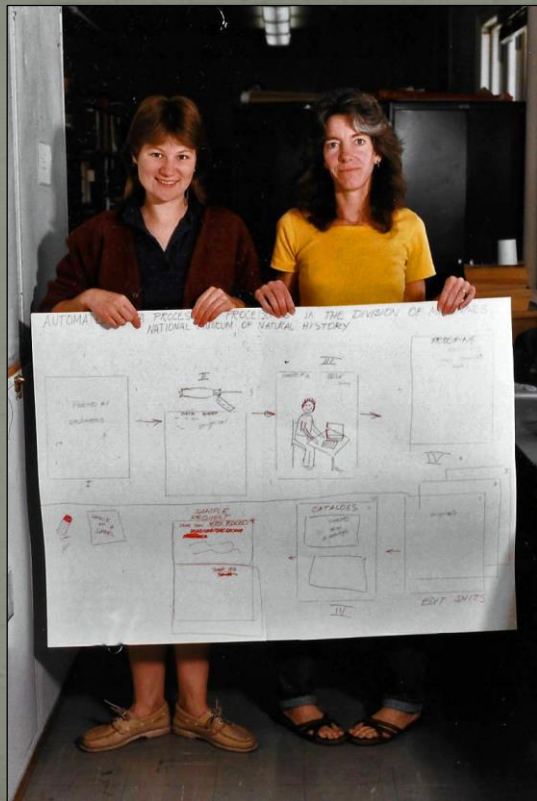
The Great Inventory 1979-1985

- Federal government mandated the Smithsonian to take an inventory of its holdings
- Approximately 20 technicians were hired temporarily for inventory
- The final specimen count was the foundation for our current totals
- It is accepted that that number is an approximation of what we held at that time
- I'm still searching for the final number that was generated
- Since then we have built on that base by tracking transactions in our system and then adding and subtracting these data each year

The Great Inventory 1979-1985



The Great Inventory 1979-1985



THE PRESENT



NMNH Digitization Steering Committee

(formed 2012)



Aka: DigiComm

DigiComm Looking Both Ways

The Castle

OCIO

DPO



DigiComm

Anthropology

Botany

Entomology

Invertebrates

Mineral Sci

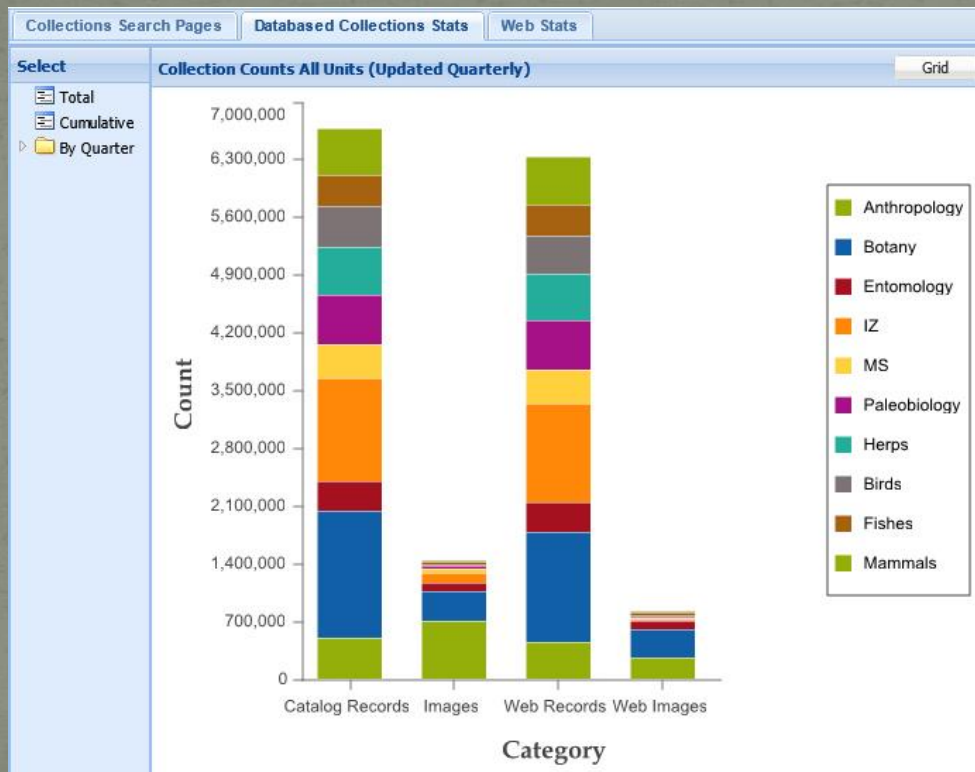
Paleobiology

Vertebrates

Education

March 2015 Statistics

6.8
million
catalog
records



~900 K

digital
surrogates

1.2

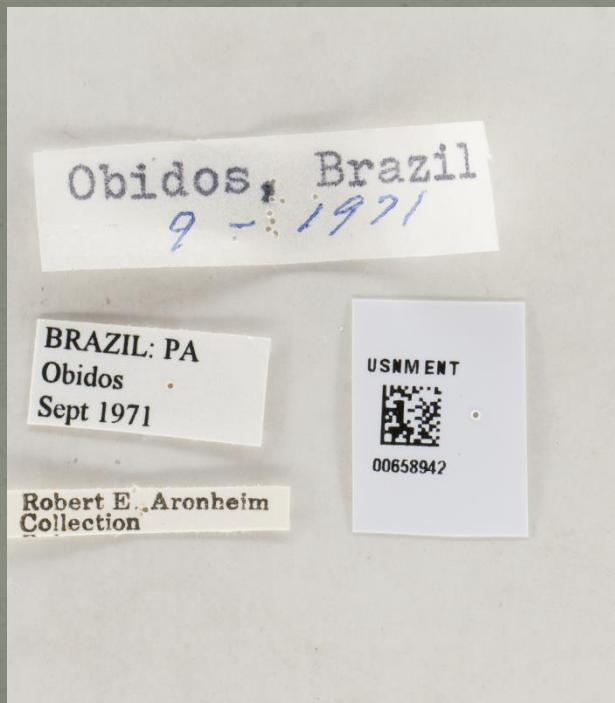
million
multimedia
assets

Thinking by type not department





Pinned Specimen Labels







Dry Specimen Label

Taxonomy (1) - Display

File Edit Select View Tools Tabs Hierarchy Multimedia Window Help

Paronuphis antarctica (Monro, 1930) : Onuphidae : Polychaeta : Annelida

Classification

Parent:

Family & No: Onuphidae

Genus & No: Paronuphis

Subgenus:

Species: antarctica

Subspecies:

Other: Other Rank Other Value

Seq. Num.:

Hybrid & Parentage

Hybrid?: Rank:

Sex: Parent:

Cultivar Name

Cultivar?: Name:

Controls

Applicable Code: ICZN Rar

Scientific Name

Automatic?: Yes Name: Paronuphis antarctica (Monro, 1930)

Currently Accepted Name

Currently Accepted?: Unknown Current Name:

Classification Citations Primary Citation Authors Availability Validity Homotypics All Sys

Display Taxon 1 of 1 snyder Admin nmrhive

Paronuphis antarctica (Monro) Cat. No. **58434**

Stack Shelf Cont. **ALC.** Spms. **2**

Locality **Antarctic Peninsula 61°14'to 61°16'S.;
56°14'to 56°12'W.**

Station **407** Depth **232-247 meters**

Coll. by **USNS ELTANIN** Date **31 December 1962**

From **Transfer SOSC** Date

Remarks

Id. by **Olga Hartman** Acc. No. **330636**

SI-179
Rev. 8-9-73 **INVERTEBRATE ZOOLOGY** **Onuphidae**



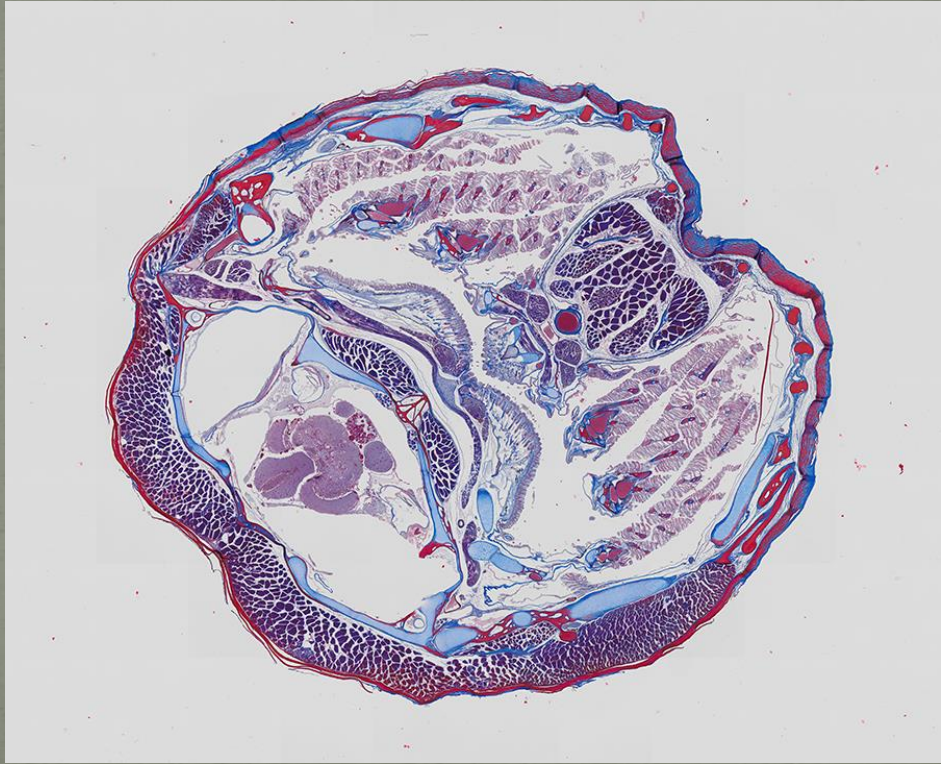




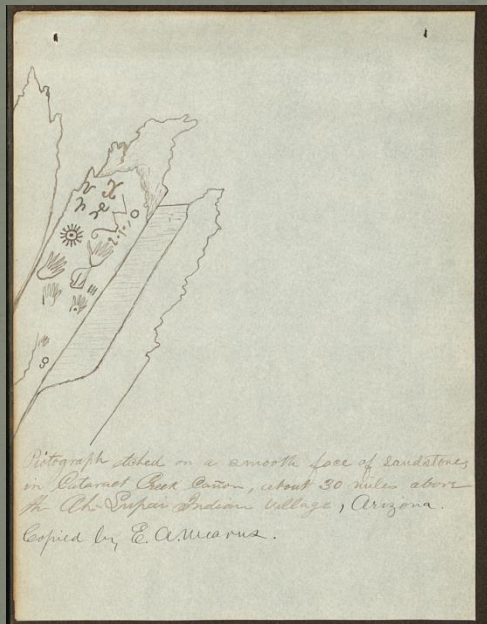


Microscope and Histology Slides

- Upwards of 5 million slides throughout the museum
- Currently testing scanning equipment
- The first scan pass captures the slide in reflected light to capture labels



Field Notes



(29)

Birds seen:-

1. *Melospiza propinqua*. - One seen on the march, while crossing a wide desert prairie, on which only, gramma grass and sage brush and a few weeds were growing. The Robin was seen on one of a couple of low ridges upon which there was a very scanty growth of cedars.
2. *Mysadestes townsendi*. - One was shot at the bottom of Black Tanks Cañon, and another was seen on the ridge of cedars in the desert.
4. *Psittopsittacus plumbeus*. - As we approached the deep cañon of the Cataract Creek a few grasshoppers and cedars, agaves and cactuses were found - also some holly, and hares, just on the verge of the cañon. I saw some of these delicate. I think and a lot of mice also.

Some painted on the red sandstone rocks about 25 miles from the Ah-Supai village in the cañon of Cataract Creek, Arizona.

(12)

7. *Sayornis saya*. - A few observed.
16. *Colaptes auratus mexicanus*.
17. *Peucaea aedon*. - One seen at Hall Cañon.
20. *Contopus richardsonii*.
21. *Archibuteo ferrugineus*. - I shot one, three days ago, as it sat on a cliff of volcanic rock, but it got away, wounded.
22. *Empidonax griseus*.
23. *Geothlypis trichas*. - One shot by Dr. Reichel at Ash Fork.
24. *Geothlypis gambeli*. - Very abundant at Hall Cañon where I shot this 15.
25. *Junco hiemalis*. - Several were collected. They are the first I have seen in Arizona.

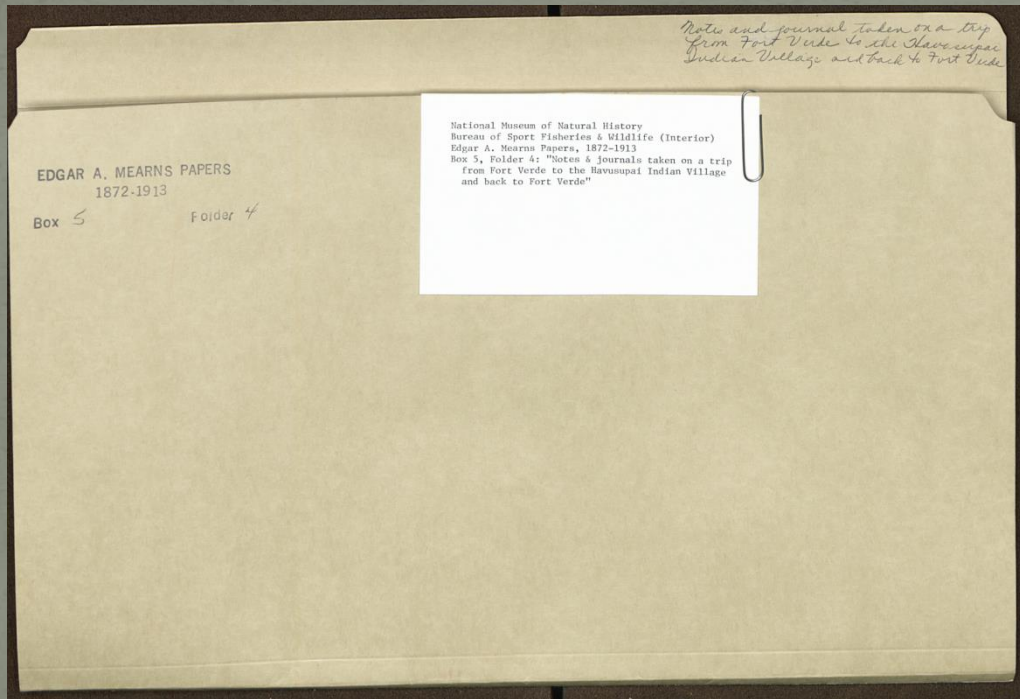
Specimens Collected:

10. 1900. *Accipiter velox* 2 ad. in the same. November 6, 1904. Ash Fork, Ariz. (Hall's 20000)
11-80; 22-70; 70-67; cul. 25; 25; 25; 25;
Cul. 40; 25; 62; tarsus, 2-10; wings, 150; claws, 150. In the plumage with iris yellow, all greenish. Bill plumbeous at base. Maxillary base brownish to tip. Feet black, yellow, claws black. Reported to me by Dr. Reichel, who shot it, shortly after. I borrowed a 2-foot scale from Dr. Reichel. All measurements are given in inches and decimal fractions thereof.

Ledgers

Chart No.	Chart No.	Original No.	Spec.	Sex.	Locality.	When Collected.	Status of Object.	Preserving Method.	Mounting.	Exhibit No.	Collected by.	Oct.	1891.	No. of Spec. Exam.	Remarks.
1.2	5				Donalaska Alaska						W. H. Bull	W. H. Bull	Dec 26	1	on hydrozoa
2.2	"				"						"	"	"	1	on Pteropod
3.2	"				"	1878		Seash.			"	"	"	1	on Chamaeum
4.2		208	<i>Halysaetia (Halysaetia)</i>		Cowley Basin	1879			W. S. F. Comm		Pearl and	"	"	1	from tubes
5.2		747							"		"	"	"	1	from tubes
6.4			<i>Nereis virens</i>		Newport, R.I.	1860	Alc.				U.S.F.C.		102	29/100	1 of 22
7.4	7	2	<i>Polysia</i>		Vineyard Sound	1870					"	"	85	1, bottle	2 of 100
8.4	8	3	<i>Glymonella triquetra</i>		Newport, R.I.	1880					"	"	85	1, bottle	2 of 100
9.4		4	<i>Hydrobia ulter</i>		Loc. 849			without tubes			"	"		2 of 100	2 of 100
10.4	6	5			Loc. P. 99, 100			with tubes			"	"	100	2, 2 of 100	2 of 100
11.4		6			Off Newport, R.I.		Day	tubes only			"	"		2 of 100	2 of 100
12.4		7	<i>Phaeosoma carmentisium</i>		Loc. 792		Alc.				"	"		1 of 100	1 of 100
13.4		8	<i>Goullii</i>		Newport, R.I.			Shore			"	"		1 of 100	1 of 100
14.4		9	<i>Rhyndobolus thomasi</i>		"			.. sand			"	"		1 of 100	1 of 100
15.4		10	<i>Loxitta elegans</i>		Vineyard Id.						"	"		1 of 100	1 of 100
16.4		201	<i>Nereis virens</i>		New Haven Conn	1890			Shore		"	"		1 of 100	1 of 100
17.4		204			"						"	"		1 of 100	1 of 100
18.4		205			Tan Pond, New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
19.4		206	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
20.4		207	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
21.4	7	70	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
22.4		71	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
23.4		72	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
24.4		73	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
25.4		74	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
26.4		75	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
27.4		76	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
28.4		77	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
29.4		78	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
30.4		79	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
31.4		80	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
32.4		81	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
33.4		82	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
34.4		83	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
35.4		84	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
36.4		85	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
37.4		86	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
38.4		87	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
39.4		88	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
40.4		89	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
41.4		90	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
42.4		91	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
43.4		92	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
44.4		93	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
45.4		94	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
46.4		95	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
47.4		96	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
48.4		97	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
49.4		98	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
50.4		99	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100
51.4		100	<i>Loxitta virens</i>		New Haven Conn	1890	Alc.				"	"	1892	1 of 100	1 of 100

Archive Documentation and A/V



Office of the Registrar Digitization

VERTEBRATE ZOOLOGY

COMMON NAME COOPER'S HAWK		NC CATALOG NUMBER(006/003) 64 375
GENUS(071) ACCIPITER COOPERII		
SPECIES(076)		
SUBSPECIES(079)		
ORDER FALCONIFORMES	FAMILY ACCIPITRIDAE	
DATE COLLECTED DAY/MON/YEAR 11/16/83		
COUNTRY (100) U.S.		
STATE/PROV/INDE (103) PA		
COUNTY (104) Snyder County		
SPECIFIC LOCALITY (NWP REFERENCE)(104)		
MODIFIER OF SP., LOC.(108)		
ELEVATION (112)		DEPTH (114)
OCEAN (101)		
SEA (102)		
SEA-LEVEL (109)		
LATITUDE-LONGITUDE (110)		
COLLECTOR'S NAME (125) LESLIE F. SHAWY		COLLECTOR'S NUMBER (126)
SHIP-ORIGIN-STATION NUMBER(127)		
NUMBER OF SPECIMENS (130)	COLLECTION METHODS	
DONOR'S NAME (180)		
ACCESSION NUMBER (155)	DATE CATALOGED IN NC (187) 10 FEB 1983	
OLD USNM NUMBER, REMARKS (200) Snyder County, PA.		
SEX (401) MALE <input type="checkbox"/> FEMALE <input checked="" type="checkbox"/> ? <input type="checkbox"/>		

<input type="checkbox"/> A SKIN ONLY	<input type="checkbox"/> F OTHER EXPLANATION	<input type="checkbox"/> J SKIN, SKULL & SKELETON
<input type="checkbox"/> B SKULL & SKELETON		<input type="checkbox"/> K ALCOHOLIC, SKULL REMOVED
<input type="checkbox"/> C ANATOMICAL	<input type="checkbox"/> G SKULL ONLY	<input type="checkbox"/> L POINTED HEAD
<input type="checkbox"/> D ALCOHOLIC	<input type="checkbox"/> H BODY SKELETON ONLY	<input type="checkbox"/> M ENTIRE (EXHIBIT TYPE) POINT
<input type="checkbox"/> E PARTIAL SKELETON	<input type="checkbox"/> I SKIN & SKULL	

CONTRIBUTION TO THE NATURALIST CENTER

I am contributing the following items to The Naturalist Center, National Museum of Natural History. I understand that if these items, or any part of them, cannot be used by the Center, its staff is authorized to dispose of them as it sees fit.

(List items below.)

Hawk - Cooper's Hawk - Snyder Co. Pa. - No. 64375
Black Bird - Maryland - No. 64376
Blackbird - Mountain Island, N.C. - No. 64377
2 - House of Wicks
Platystrophia - Maryland, Simpson Bay - No. 64379

I would like to receive a letter acknowledging this contribution.

Name: Leslie F. Shawy

Address: 2302 Dawson St.

City: Hillsport, Md. State: Md. Zip: 21003

~500,000 documents

Office of the Registrar Digitization

From: To:	U. S. N. M. INDEX	ACC 184488
Date	C. & R. No.	
10-18-1949.	Sato, Dr. M.,	
	College of Agriculture,	
	Tsuruoka-city,	
	Yamagata-pref., Japan.	
7 specimens of lichens from Japan.		
GIFT		
Ref. to	Ans.	October 18, 1949.
	Inv. Sh. Off.	
	Distribution	
U. S. GOVERNMENT PRINTING OFFICE 16-58853-1		

Index to U. S. NATIONAL MUSEUM Correspondence.	
Letter Received from	Dearmont, J. B. No. 2406
Date of Letter,	Address: 89 St. Peter St., Montreal, Canada. Returnable to Office. March 20
2-21, 1903	Subject: Req. S. I. Rep. for 1895
Order made	7. M. Rep. for 1897, etc.
Answered	etc. - various other rept's named.
2-25	Temporarily Ans'd
	Referred to Mrs. Brock March 13
FILE	Notes (Address to S. I.)
	LIBRARY BUREAU 49787C

>350,000 index cards

Crowd Sourcing

The screenshot shows the Smithsonian Digital Volunteers Transcription Center website. The browser address bar displays <https://transcription.si.edu/>. The page features a teal header with the Smithsonian logo and navigation links: HOME, PROJECTS, ABOUT, TIPS. Below the header is a dark navigation bar with "SMITHSONIAN DIGITAL VOLUNTEERS: TRANSCRIPTION CENTER" and "SIGNUP LOGIN".

The main content area features a "FEATURED" section for "THE BUMBLEBEE PROJECT" with a "VIEW PROJECTS" link. The featured image shows a bumblebee specimen with handwritten labels: "Winchester, Va. May 17, 1936 G.A. Sandhouse", "Bombus affinis Cresson Det. by S. CZAPENIK", and "USNM ENT". Below the featured image are four smaller project thumbnails: a bee specimen, a document, an American flag, and a record book.

Navigation arrows are present on the left and right sides of the featured image area. A vertical "feedback" button is located on the right side of the page.

JOIN US!
LEARN HOW TO TRANSCRIBE
Become a Smithsonian Digital Volunteer and

BROWSE PROJECTS
Select a category below to begin browsing projects.

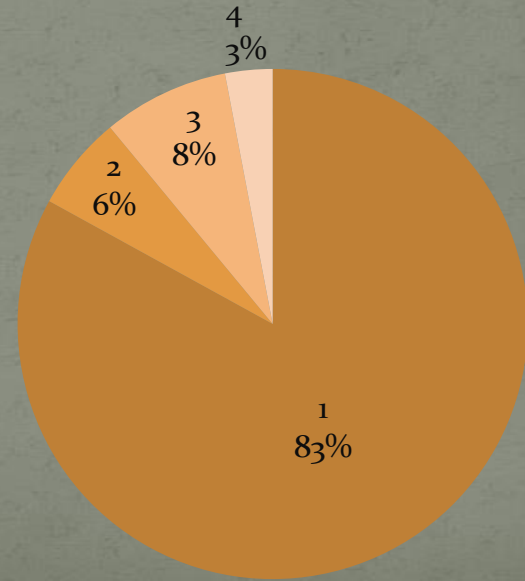
LATEST UPDATES
penguinfun marked for review a page from Harvard-Smithsonian Center for

Statistics

- Yearly reporting to SI of Collection and Digitization Assessment
- Rank database records on a scale from 1-4 (Record Quality)

Record Quality

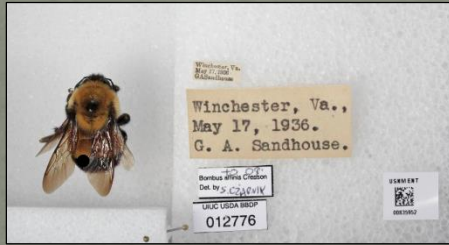
- 1 = No database record 32 million
- 2 = Below standard record 2 million
- 3 = Standard record 3 million
- 4 = Above standard record 1 million



What's Left To Do

- We estimate it will take 35.5 million catalog records to represent our current collections
- That means we need at least another 26 million records... (*EEK!*)
- It has taken 50 years to get to 6.8 million, though our rate has dramatically improved over the years!
- So we still have a long way to go!!!

THE FUTURE



THE FUTURE

- Prioritized projects based on research question(s)
- Rapid capture – Improved workflows
- Greater automation
- Digitization and cataloging will become more dependant upon each other
- Create a digitization center
- Increased crowd sourcing – transcription centers and beyond

You Say Potato!

You Say Potato!
I Say Potato!

You Say Potato!
I Say Potato!
You Say Digitise!

You Say Potato!
I Say Potato!
You Say Digitise!
I say Digitize!

Thanks

- Rebecca Snyder
- NMNH Staff
- CSIRO
- iDigBio
- Paul Flemons