

# NMNH Rapid Capture



**DESIGNING WORKFLOWS TO INCREASE  
THE RATE OF DIGITIZATION**

**PAUL KIMBERLY, PATRICIA GENTILI-POOLE, KEN  
RAHAIM**



**Smithsonian Institution**

# Overview



- Bumble Bee Rapid Capture Pilot
- NMNH Herbarium rapid capture initiative
- NMNH rapid capture tests and future projects

# What is Rapid Capture?



- A faster more efficient way to digitally capture specimens and associated data and metadata
- New streamlined workflows
- Innovative new equipment
- Higher resolution photographic equipment



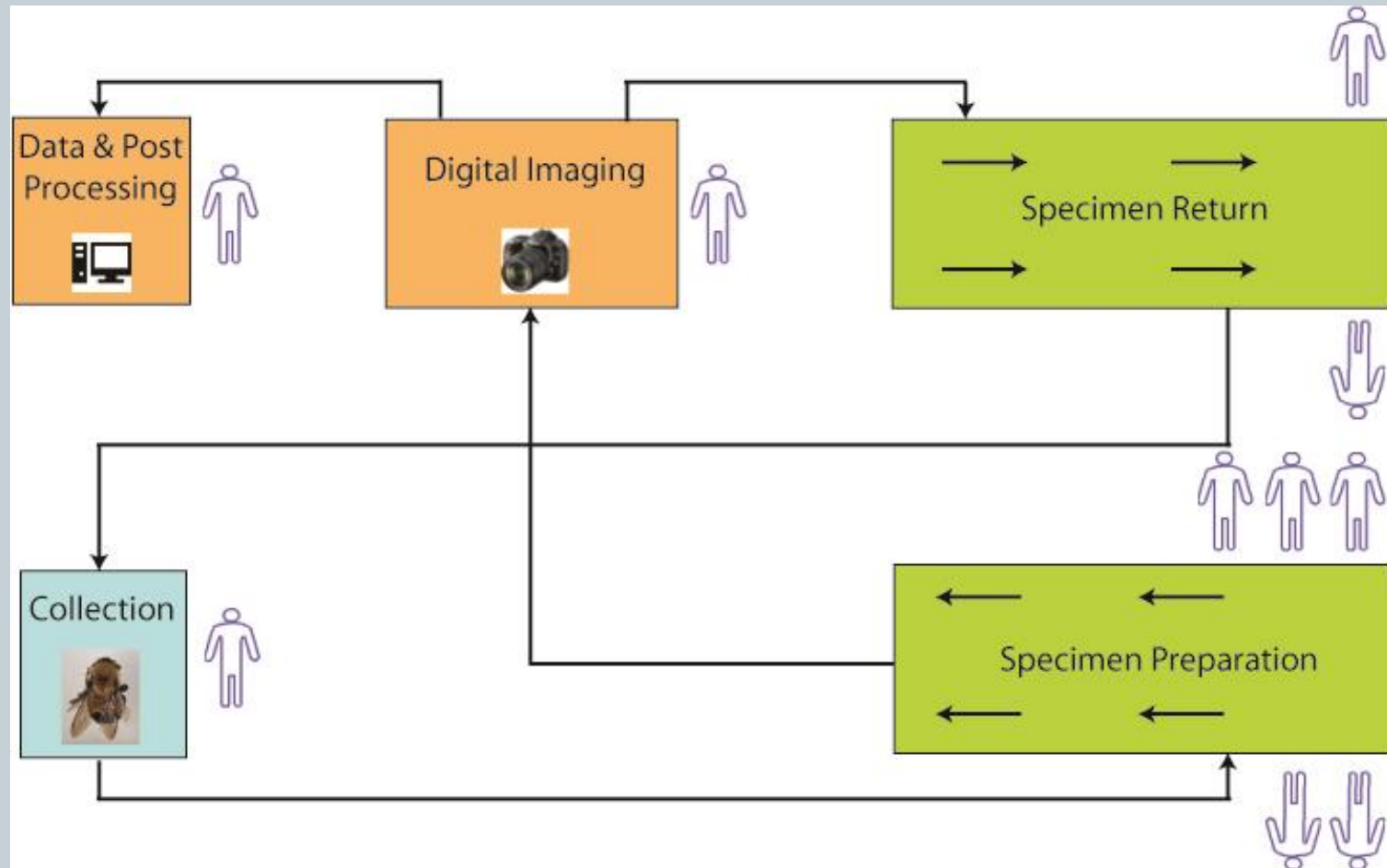
# Bumble Bee Rapid Capture Pilot Project Goal



**Develop a Workflow** to speed the process of digitization of pinned collections in Entomology by capturing digital of specimens including labels to create complete records for 14,400 bumble bees.

We passed the 14,400 in 3 weeks, and hit 44,047 in the 8 week pilot.

# Bumble Bee Rapid Capture Workflow



# Specimen Processing



Collection



Adding barcode labels



Removing labels from pins



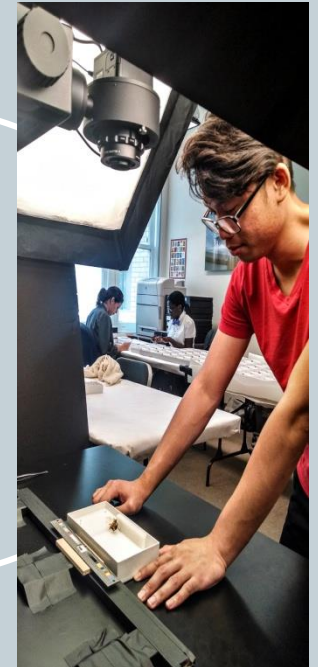
Staging specimen



Replacing specimens in trays and drawers



Replacing labels on pins

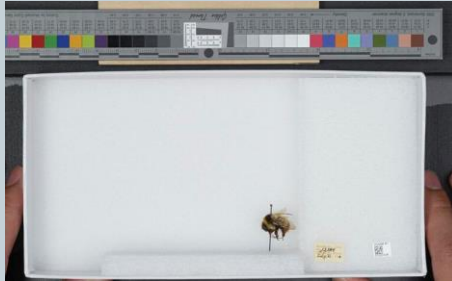


Imaging

# Specimen Processing



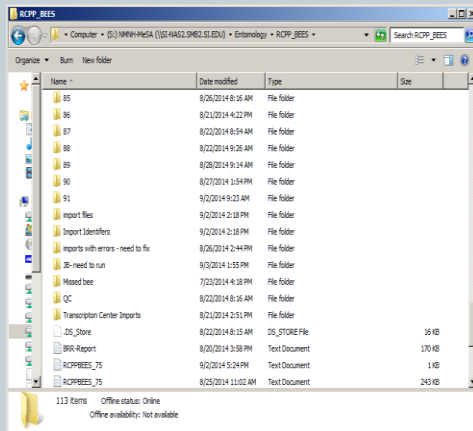
# Image and Post Processing



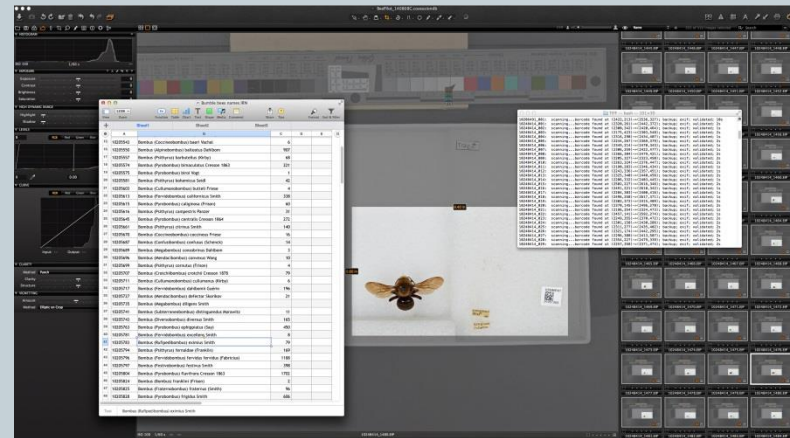
Capture Image



Crop (2820x1680)



Folders in MeSA



Rename files with barcode number, embed metadata & generate EMu friendly report



# Image and Post Processing





# Background: 2014 Record Estimate



TOTAL: 20,835,735 (-/+ 10%) records

- Pinned 18,813,906 largest problem
- Slides 1,392,676 methods exist
- Fluid 522,717 major problem
- Papered 106,436 done (Odonata)

# Potential Solution: RCPP for 3D Objects



## DPO Conditions:

- Medium size 3D specimens
- 1 image per specimen
- 1 setting fits all
- 8 weeks
- Use of Transcription Center

## Entomology Conditions:

- Have Taxonomy in EMu
- Not much variation in size
- Neither flat nor a perfect sphere
- Be attractive for crowd-sourcing
- Candidates: Bees and Wasps

# The Importance of Digitization



- Smithsonian initiative
- Unlocks data that is stuck in drawers
- Improved access for the general public, educators, and the scientific community
- Allows data to be shared virtually saving a trip to the collection
- Less specimen handling after initial digitization
- Gets a snap shot of specimen condition

# Transcription Center Set Up



- Prior to RCPP production, a help file, template and in-line text were created with help from the TC team
- At time of import, a note is added to skeleton records that “tags” images for upload into transcription center
- Administrators have access to data sets and publish sets as needed
- Digital volunteers transcribe specimen labels into a template
- Keep in touch with volunteers through email and social media

# Transcription Center Data Entry





ID	NAME	ONLINE	OWNER	IMAGE	OPERATIONS
6775	The Bumblebee Project - Set 1	Yes	NMNH - Department of Entomology		<a href="#">Edit Project</a> <a href="#">Manage Assets</a> <a href="#">Review Transcriptions</a> <a href="#">Export</a>
6780	The Bumblebee Project - Set 2	Yes	NMNH - Department of Entomology		<a href="#">Edit Project</a> <a href="#">Manage Assets</a> <a href="#">Review Transcriptions</a> <a href="#">Export</a>
6781	The Bumblebee Project - Set 3	Yes	NMNH - Department of Entomology		<a href="#">Edit Project</a> <a href="#">Manage Assets</a> <a href="#">Review Transcriptions</a> <a href="#">Export</a>
6782	The Bumblebee Project - Set 4	Yes	NMNH - Department of Entomology		<a href="#">Edit Project</a> <a href="#">Manage Assets</a> <a href="#">Review Transcriptions</a> <a href="#">Export</a>
6789	The Bumblebee Project - Set 5	Yes	NMNH - Department of Entomology		<a href="#">Edit Project</a> <a href="#">Manage Assets</a> <a href="#">Review Transcriptions</a> <a href="#">Export</a>
6790	The Bumblebee Project - Set 6	Yes	NMNH - Department of Entomology		<a href="#">Edit Project</a> <a href="#">Manage Assets</a> <a href="#">Review Transcriptions</a> <a href="#">Export</a>
6791	The Bumblebee Project - Set 7	Yes	NMNH - Department of Entomology		<a href="#">Edit Project</a> <a href="#">Manage Assets</a> <a href="#">Review Transcriptions</a> <a href="#">Export</a>
6792	The Bumblebee Project - Set 8	Yes	NMNH - Department of Entomology		<a href="#">Edit Project</a> <a href="#">Manage Assets</a> <a href="#">Review Transcriptions</a> <a href="#">Export</a>
6793	The Bumblebee Project - Set 9	Yes	NMNH - Department of Entomology		<a href="#">Edit Project</a> <a href="#">Manage Assets</a> <a href="#">Review Transcriptions</a> <a href="#">Export</a>
6794	The Bumblebee Project - Set 10	Yes	NMNH - Department of Entomology		<a href="#">Edit Project</a> <a href="#">Manage Assets</a> <a href="#">Review Transcriptions</a> <a href="#">Export</a>



# Results: We aren't finished yet!

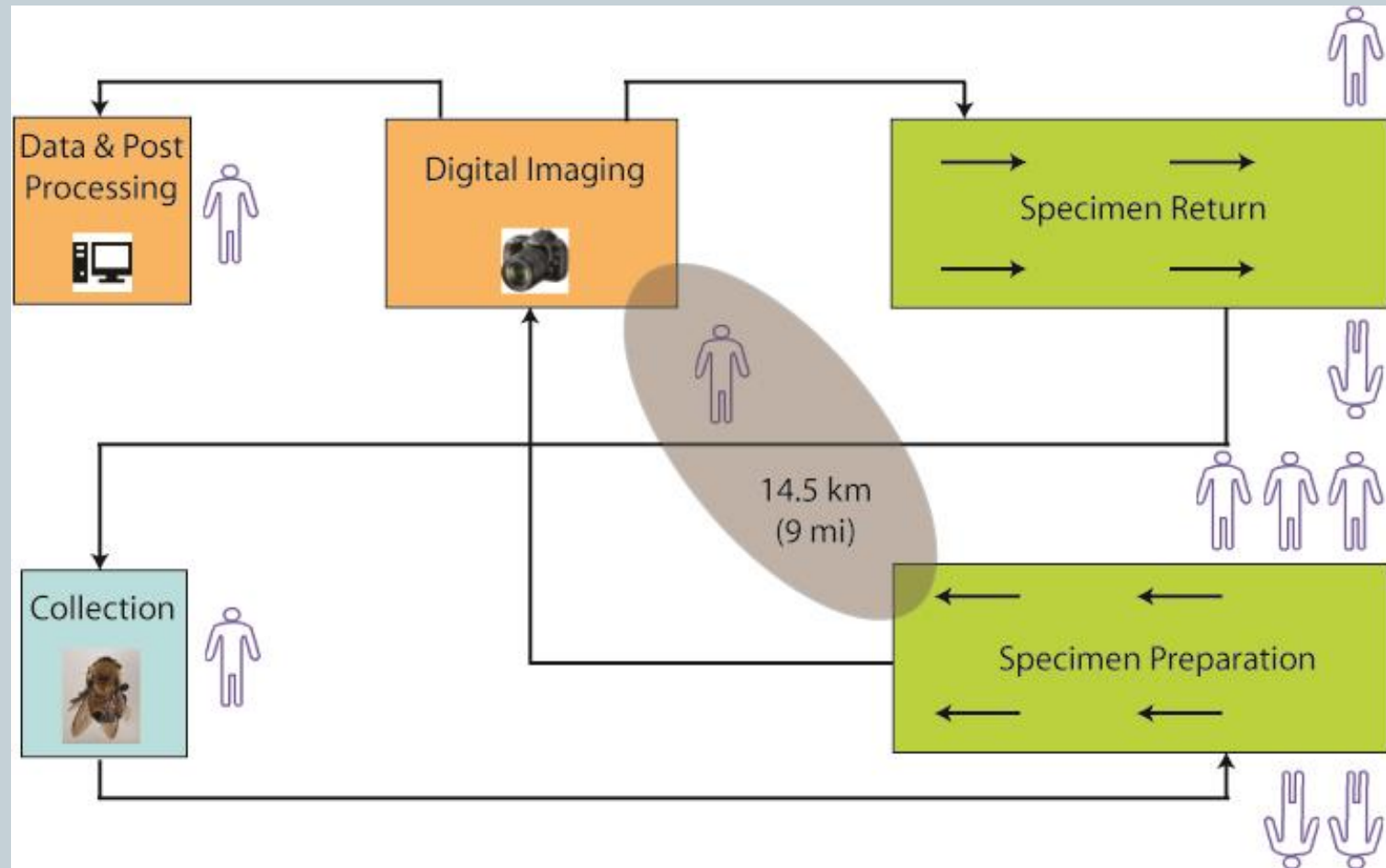


- Specimen handling worked better than anticipated: Reached goal in under 4 weeks
- Imaging never 100% utilized
- Optimized by adding staff to specimen processing
- Kept constant personnel in imaging (2) and data processing (1 including set up of transcription)
- Data Processing is not completed until after transcription
- Became the example of the Transcription Center and gained a lot of media attention

Personnel in Staging	Specimens Processed	Average per Person
4 x 9.5 days =228 hrs	6,732	29/hr
5 x 13 days =390 hrs	15,732	40/hr
6 x 10.5 days =378 hrs	13,675	36/hr
7 x 5 days=210 hrs	7,916	37/hr
	44,047 total	



# Bumble Bee Rapid Capture Workflow



# Traditional vs. RCPP



- One person manually enters data into EMu
- Average pace is 100 specimens per person per day

## ***The dirty little secret!!!!***

- 10 technicians digitizing 100 specimens a day would take 8 weeks to have full database records for 44,000 specimens, no image.
- 10 technicians in the rapid capture setting took 8 weeks to digitally capture the same amount of images which include labels, however the transcription rate is low and nearly one year later, the records are still skeletal.

# The Bumble Bee Team



## Co-managers:

Technical requirements: Ken Rahaim, OCIO-DPO

Logistics: Paul Kimberly, NMNH-ADS

Specimen & data handling: Patricia Gentili-Poole, NMNH-DoE

## Specimen Processing:

Erin Bilyeu

Laura Tancredi

Julie Hoskin

Amanda Lawrence

Catherine Winn

Hailey Hultquist

Ben Sullivan

Brian Harris

Sylvia Orli

Ingrid Lin

Jennifer Strotman

Ali Brooks



## Additional Support:

Rebecca Snyder, Informatics

Tom Hollowell, Informatics

Meghan Ferriter, Transcription Center

Desmond Smith, Entomology

## Imaging:

Eric Philcox

Brian

Leslie Cone

## Data Processing:

Cynthia Murphy

Jessica Bird

Digital Volunteers

## Funding:

OCIO-DPO

NMNH-ADS

Pixel Acuity

# What's next?



# The Conveyor Belt



# The Conveyor Belt



- Start date – Sept. 15, 2015
- One Herbarium sheet every 4-6 seconds
- Goal is 500,000 sheets in 6 months
- Currently 1.5 million sheets are digitized with 3.5 million to go
- Including all secondary costs it is approximately \$1.00 US per sheet (not including staff time)
- 4 technicians are currently preparing the specimens

# Fish Types



- 12,000 type specimens
- \$8.50 US per specimen



# Fish Types





# Office of the Registrar



- 500,000 documents
- \$1.00 US per document

VERTEBRATE ZOOLOGY

COMMON NAME <u>COOPER'S HAWK</u>		NC CATALOG NUMBER(006/003) <u>64375</u>
GENUS(071) <u>ACCIPITER COOPERII</u>		
SPECIES(070)		
SUBSPECIES(070)		
ORDER <u>FALCONIFORMES</u>	FAMILY <u>ACCIPITRIDAE</u>	
DATE COLLECTED DAY/MON/YEAR <u>11/19/83</u>		
COUNTRY (100) <u>US</u>		
STATE/PROV/INSE (102) <u>PA</u>		
COUNTY (103) <u>SNIDER COUNTY</u>		
SPECIFIC LOCALITY (W.P. REFERENCE)(104)		
MODIFIER OF SP., LOC.(106)		
ELEVATION (112)	DEPTH (114)	
OCEAN (107)		
SEA (108)		
SEA-WARDEN (109)		
LATITUDE-LONGITUDE (110) <u>41° 15' N 78° 15' W</u>		
COLLECTOR'S NAME (125) <u>LESLIE S. SWAMY</u>		COLLECTOR'S NUMBER (126)
SHIP-OR-TRAIL-STATION NUMBER(127)		
NUMBER OF SPECIMENS (130)		COLLECTION METHODS
DONOR'S NAME (180)		
ACCESSION NUMBER (156)		DATE CATALOGED IN NC (157) <u>10 FEB 1983</u>
OLD SHIP NUMBER, HORNS (200) <u>SNIDER COUNTY, PA</u>		
SEX (403) <input checked="" type="checkbox"/> MALE <input type="checkbox"/> FEMALE <input type="checkbox"/> ?		

A SKIN ONLY       F OTHER EXPLANATION \_\_\_\_\_       J SKIN, SKULL & SKELETON  
 B SKULL & SKELETON       K ALCOHOLIC, SKULL REMOVED  
 C ANATOMICAL       G SKULL ONLY       L POINTED HEAD  
 D ALCOHOLIC       H BODY SKELETON ONLY       M ENTIRE (EXHIBIT TYPE) POINT  
 E PARTIAL SKELETON       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 - Snider County, PA - No. 64375  
 Black Bird - Hawfinch No. 64376  
 Nuthatch - Nuthatch No. 64377  
 2 frames of photo  
 Pelt of Snake - Snake No. 64378

I would like to receive a letter acknowledging this contribution.

Name: Leslie S. Swamy

Address: 2302 Dresden St.

City: Hillsdale, N.J. State: N.J. Zip: 08042

# Office of the Registrar



- 500,000 documents
- \$1.00 US per document
- Estimated time to completion 1 year
- This is the wrong workflow for this project
- Outsourcing to document scanning company for approximately \$.20 US each

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.)

Haube - opposite Nashville - Lander County, Tennessee NE 64375  
Black Bird - ~~North~~ NE 64376  
Narrow-leaved - Mountain Island, North Carolina, Simpson Co. NO 63070  
2 frames of photo  
Pelt of ~~mountain~~ Pearly Wood, Simpson Co., North Carolina NE 64349

I would like to receive a letter acknowledging this contribution.

Name: Leslie S. Strang

Address: 2302 Dresden St.

City: Hillsport, Miss. State: Miss. Zip: 39001

# Digitarium's Small 3-D Object Conveyor



# Ortery's Multi-Camera/Multi-View

