Using Digitized Herbarium Data in Research: Applications for Exploration, Taxonomy, and Ecology

Pam Soltis
University of Florida

FLORIDA MUSEUM

UF | UNIVERSITY of FLORIDA

iDigBio
Integrated Digitized Biocollections

NSF

Topics

- Intro to data
- Downloading data
- Cleaning data
- Georeferencing
  - GEOLocate
- Ecological Niche Modeling
  - Maxent
- ENM Analysis & Interpretation
- Uses of collections data
- Discussion/Q&A
- R-based methods

Workshop Leaders

- Pam Soltis
- Makenzie Mabry
- Shelly Gaynor
- Sarah Ellen Strickland
- Sydney Barfus
- Sebastian Fernandez
- Tyler Radtke
- Elizabeth White
- JT Miller
- Doug Soltis
- Contributions from many previous lab members/colleagues:
  - Blaine Marchant, Charlotte Germain-Aubrey, Andre Naranjo, Anthony Melton, Tal Kinser, Maria Kinser, Mike Belitz, Rhett Rautsaw, João Vidal, Jr.

FLORIDA MUSEUM™

UF | UNIVERSITY of FLORIDA

iDigBio
Integrated Digitized Biocollections

Logistics

- HiPerGator – UF's supercomputer
- GitHub repository
  - data
  - scripts
  - presentations
- Workshop evaluation survey
  - by email
- Pace/breaks


FLAS 22883
R. K. Godfrey Herbarium (FSU)
HERBARIUM
177581
FLORIDA STATE UNIVERSITY
The New York Botanical Garden
Woody Plants of U. S. A.
ACERACEAE Acer rubrum L. Female
Small tree, 8 m. tall, with two stems from base, 15 cm. in diameter.
16709-a April 22, 1979. Tree in full bloom, exclusively female (see no. 16708, a nearby male tree). Tips of calyx and corolla lobes light red as in no. 16708. Stigmas bright red.
16709-b June 4, 1979. Leaves dark green above, pale green below. Fruits pale tan.
Common in bottom of hollow, ungrazed woods of Quercus alba, Fraxinus americana, Populus grandidentata, Carya ovata.
WISCONSIN: Richland Co., 3 miles SE of Richland Center.
T 10 N; R 1 E; SE¼SW¼ sec. 25. alt. 250 m.
Coll.: M. Nee
HERBARIUM
UNIVERSITY OF FLORIDA
GAINESVILLE
135638

Herbaria Worldwide

NYBG STEERE HERBARIUM

Home Collections Discover Index Herbariorum Virtual Herbarium Loans Digitization

Index Herbariorum

3,400 herbaria

400,000,000 specimens!

Natural History Collections

1-2 billion specimens
in the US
3-4 billion specimens
worldwide

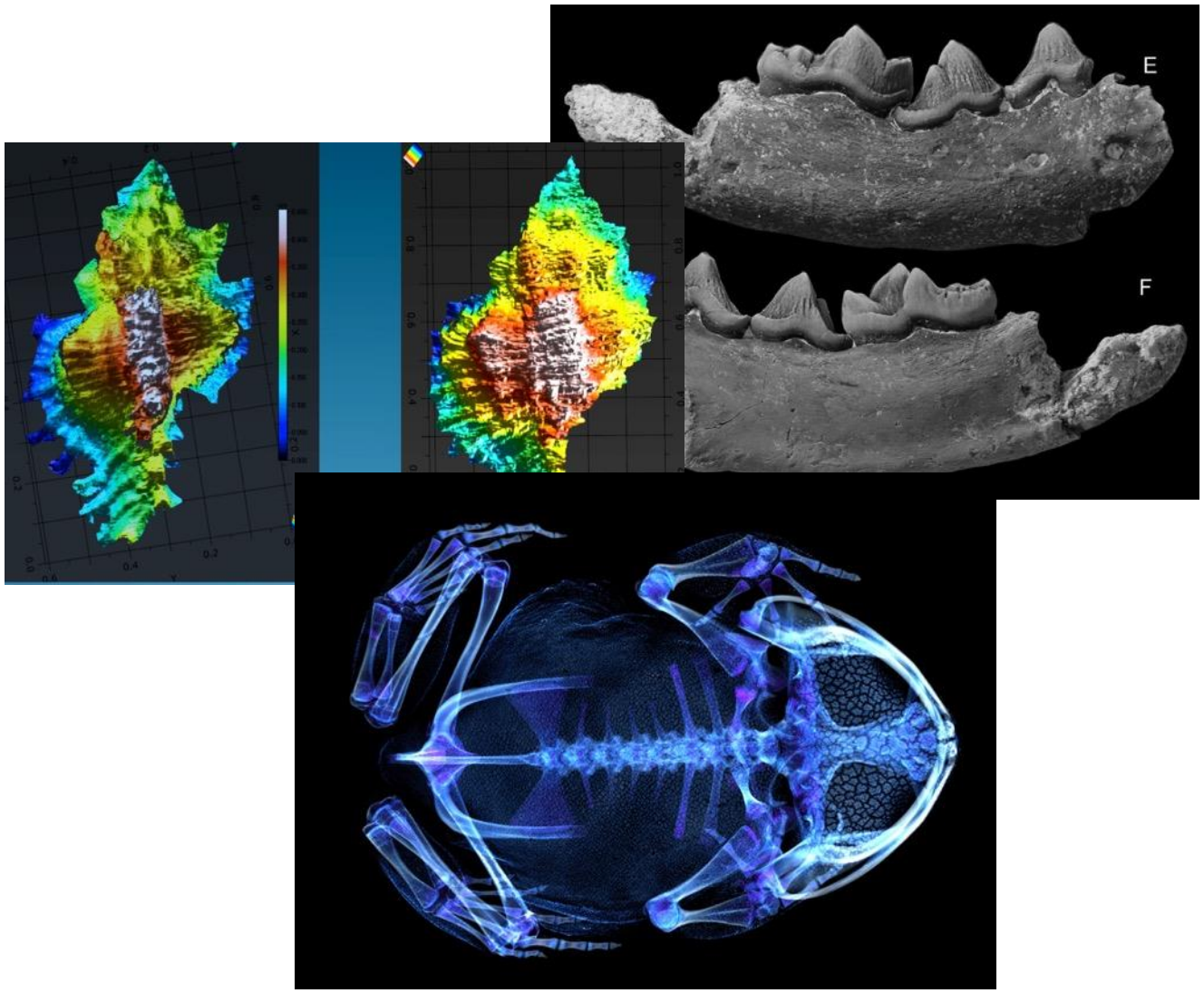
~1600 natural history collections in the US, thousands in the world

Smithsonian

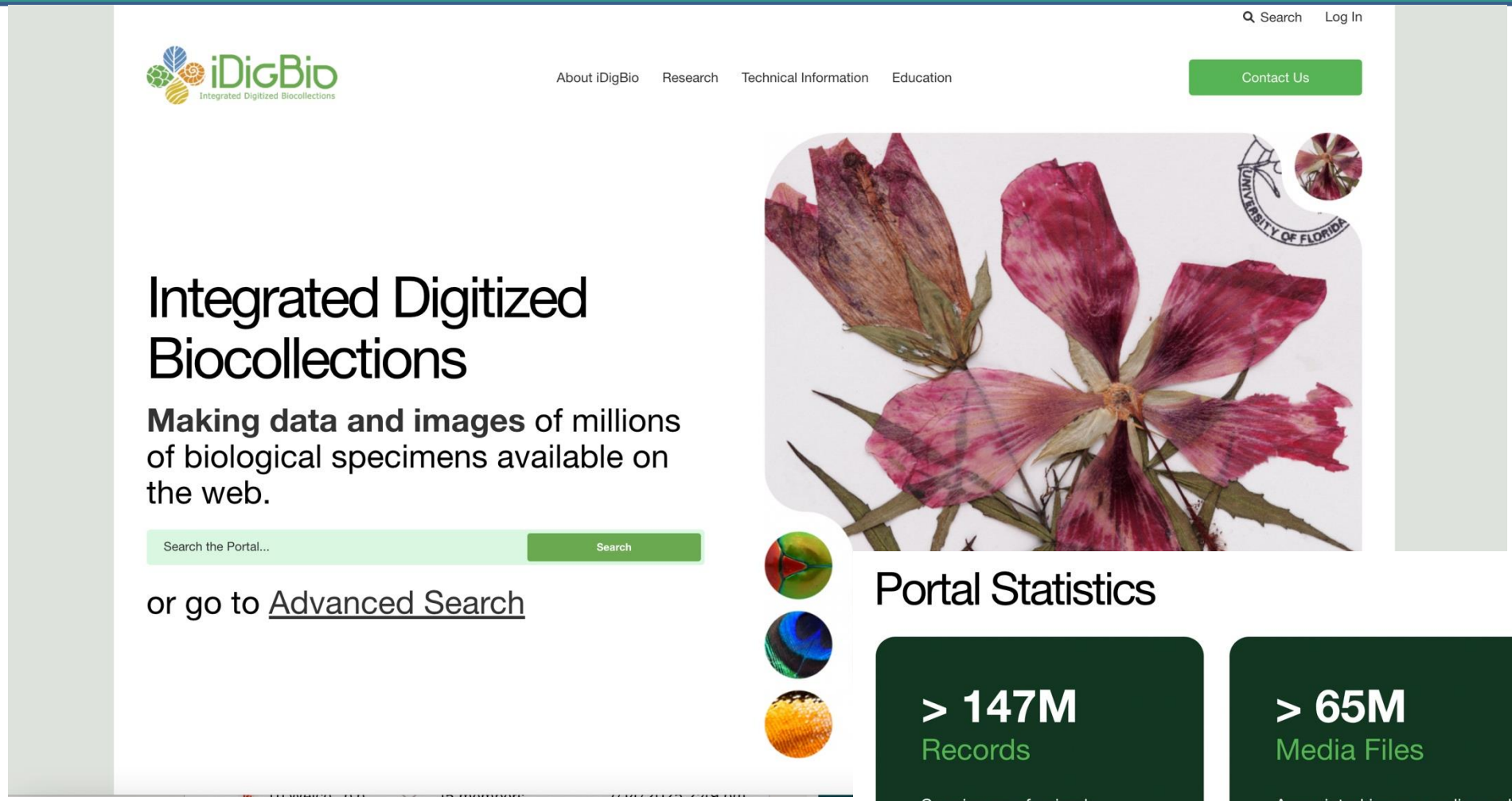
Digitization: Data and Images

- Scientific name –
  including authority
- Date
- Collector
- Location – state, county,
  specific site,
  GPS coordinates
- Associated species
- Notes

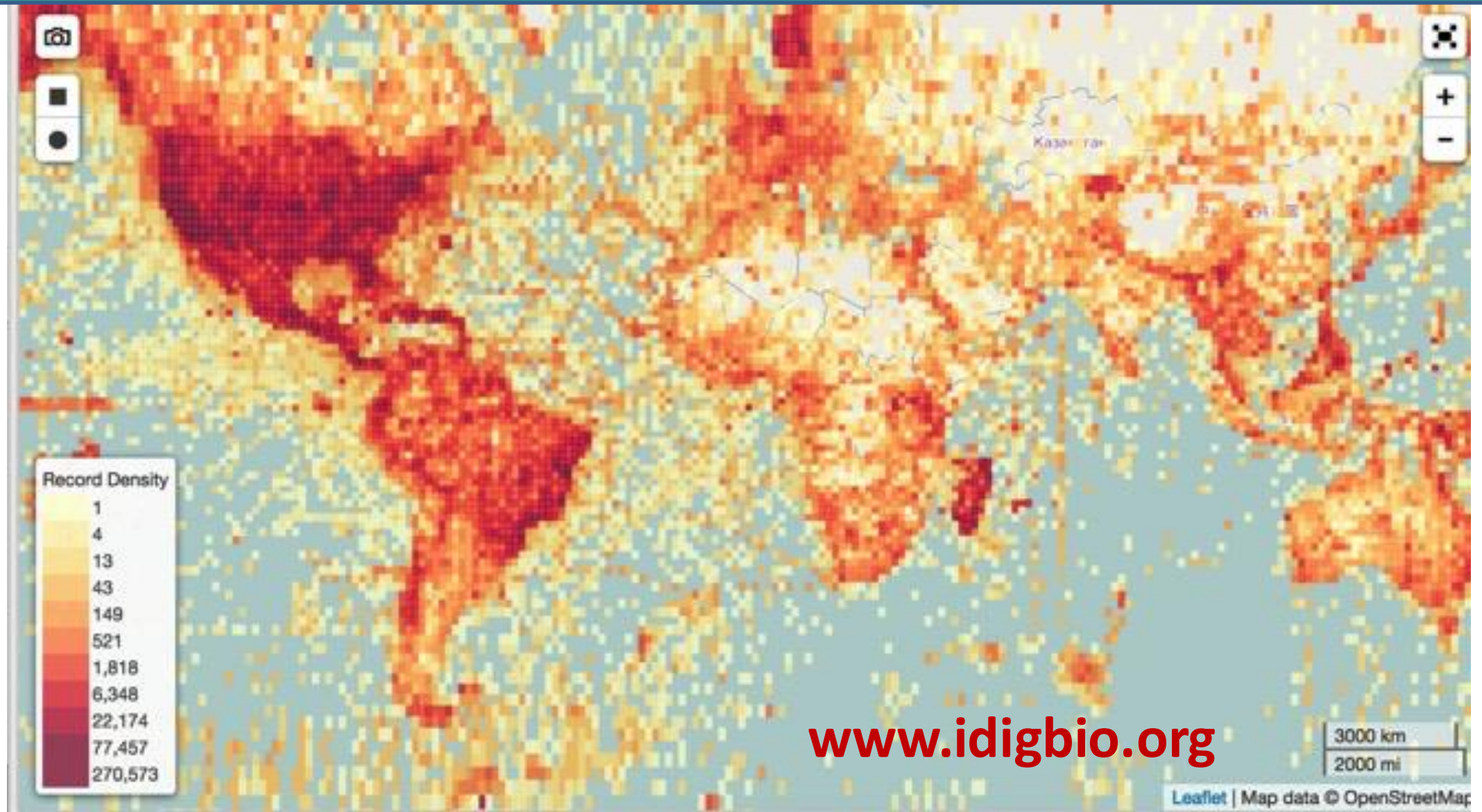
Digitization: Data and Images



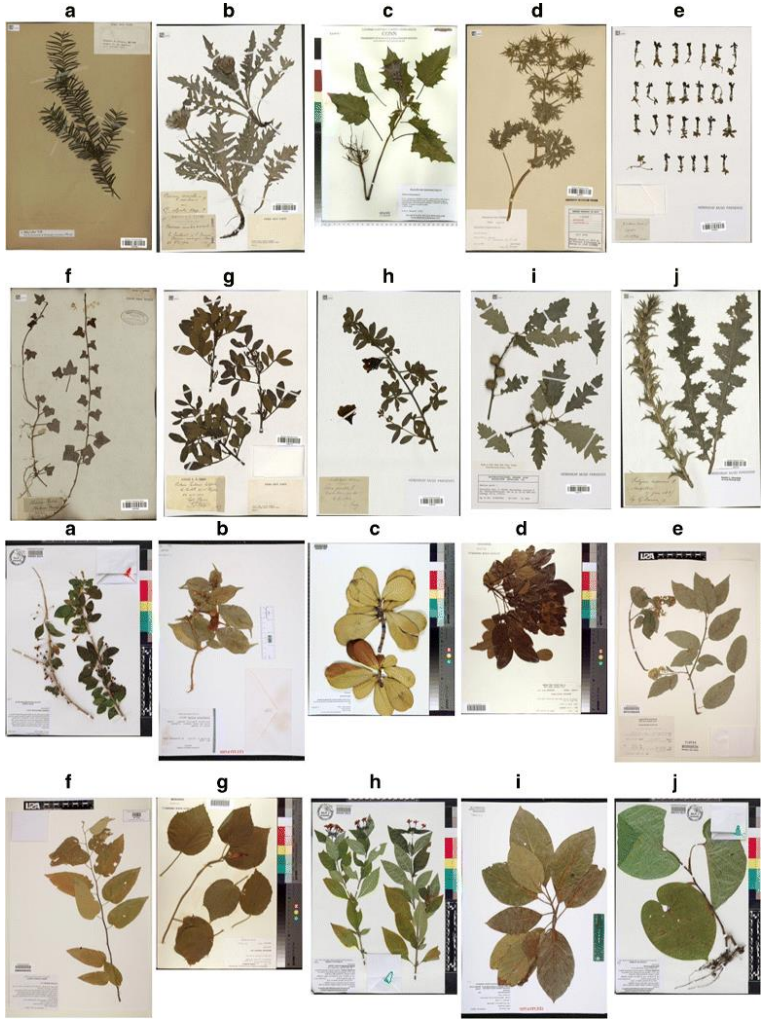
iDigBio: A Source for Specimen Records



Specimen Occurrences in iDigBio



Machine Learning & Biodiversity Research



Carranza-Rojas et al. 2017

Pearson et al. 2020

Other Data Aggregators

‘Big Data’ Research in Biodiversity Science

- Monitoring shifts in biodiversity
- Tracking invasive species
- Ecological Niche Modeling, climate change
- Past movements and climate change
- Tracking phenological shifts
- Landscape genetics
- Integration of ENM with phylogeny
- Community phylogenetics/assembly
- Biogeography

Overview of Day's Activities

9:00	Welcome and Overview of the Workshop – Pam
9:10	Darwin Core and Data Fields – Pam
9:15	Data Downloading Overview – Pam
9:20	***Activity***: Manual Data Downloads from iDigBio Portal – Pam
9:30	Intro to HiPerGator – Makenzie
9:35	Data Downloads – Shelly
9:45	***Activity***: R-based Data Downloads – Shelly
10:00	**Break**
10:30	Data Cleaning – Shelly
10:45	***Activity***: Data Cleaning
11:00	Georeferencing – Sarah Ellen
11:30	Environmental Variables & Data Exploration – Sydney
11:45	***Activity:*** Environmental Variables & Data Exploration – Sydney

Overview of Day's Activities

12:15	**Lunch**
1:15	Defining Accessible Area + Variable Selection – Makenzie
1:25	***Activity***: Defining Accessible Area + Variable Selection – Makenzie
1:45	Ecological Niche Models (ENMs) – Sebastian
2:00	***Activity***: ENMs – Sebastian
2:30	Null Models – Tyler
3:00	**Break**
3:30	Post ENM analyses + Future Predictions – Elizabeth
3:45	***Activity***: Post ENM analyses + Future Predictions – Elizabeth
4:15	Occupancy Models – JT
4:35	Closing with Examples – Doug
4:50	Final Q & A
5:00	**End**