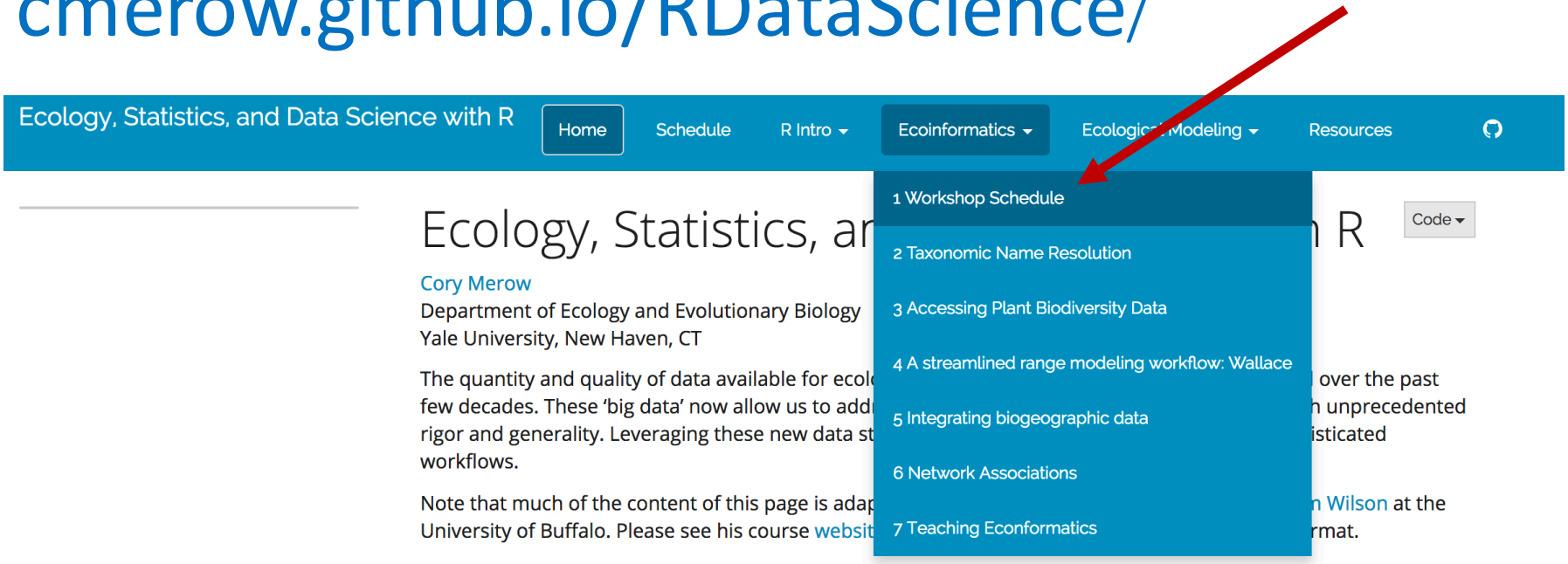
Integrating and Cleaning Biodiversity Data

A bunch of ecoinformaticians (see next)

9 Jan 2017 IBS Meeting

Install some packages while I'm talking…

cmerow.github.io/RDataScience/



Matt Aiello-Lammens

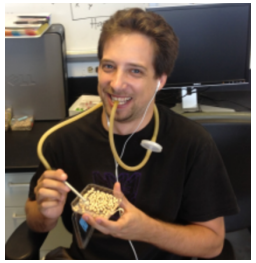
Rob Anderson

Brad Boyle

Brian Enquist

Jamie Kass

Drew Kerkhoff



Brian Maitner

Brian McGill

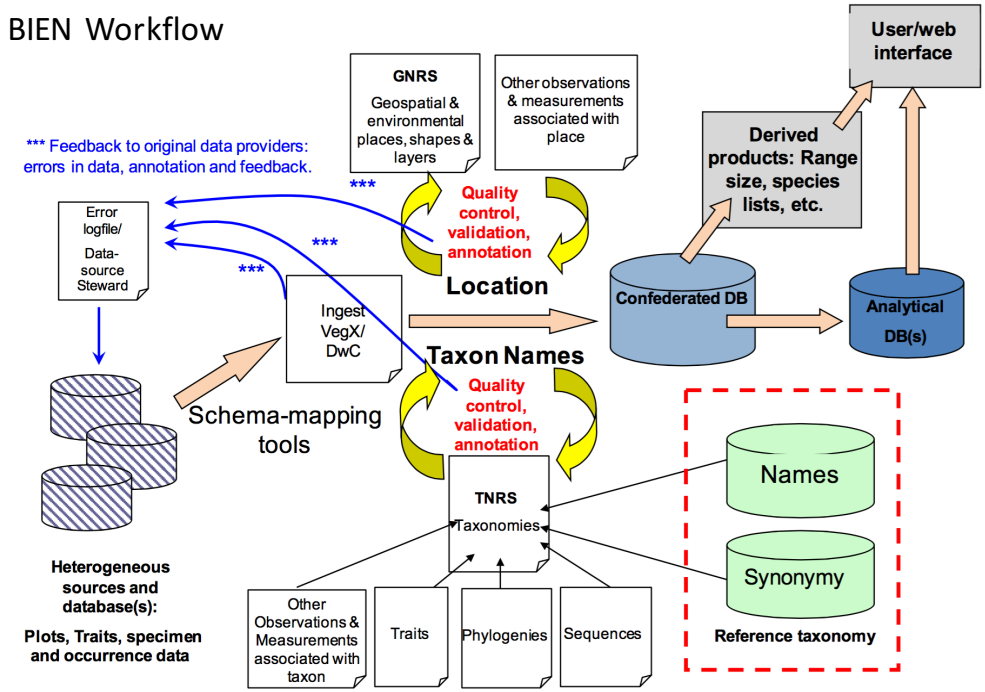
Cory Merow

Naia Morueta-Holme

Jens Christian Svenning

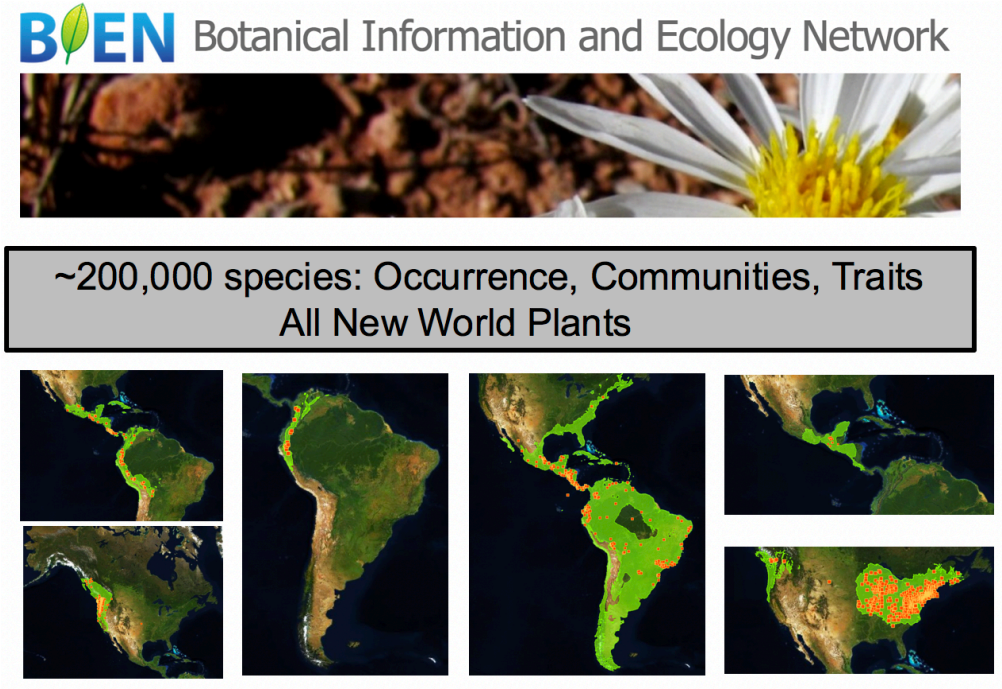
What is informatics

- Informatics – getting info from data
  - **Computation**



What is informatics

- Informatics – getting info from data
  - Computation
  - **Large data sets**

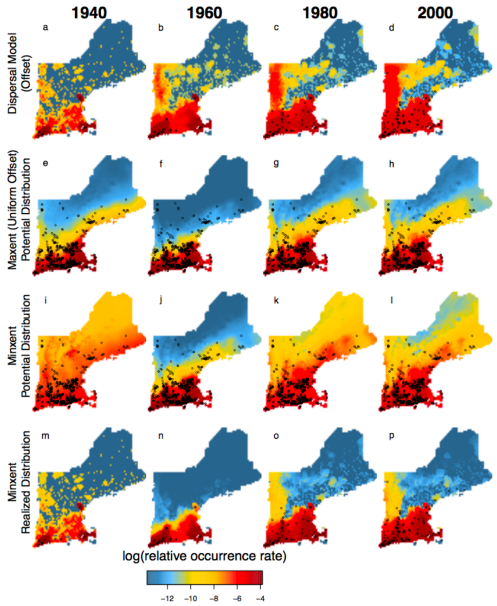


What is informatics

- Informatics – getting info from data
  - Computation
  - Large data sets
  - **Statistics**

New Models - Minxent

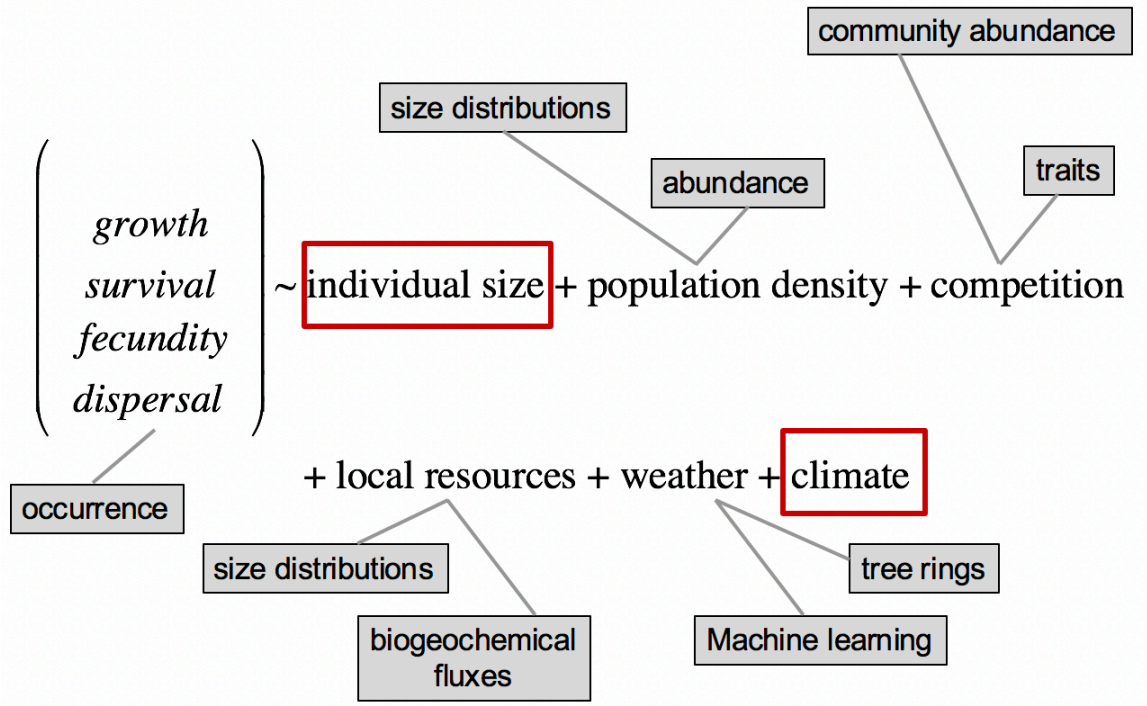
- Sampling Bias
- Native range info
  - Realized vs
  - Potential Distributions
- Expert maps
- Phylogenetic constraints



Merow et al. GEB 2016

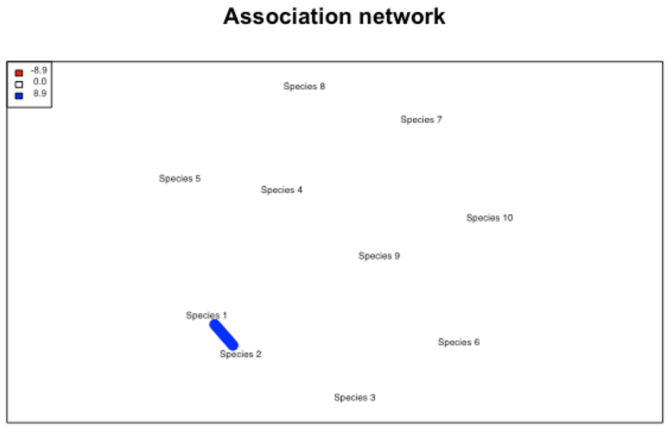
A case for informatics

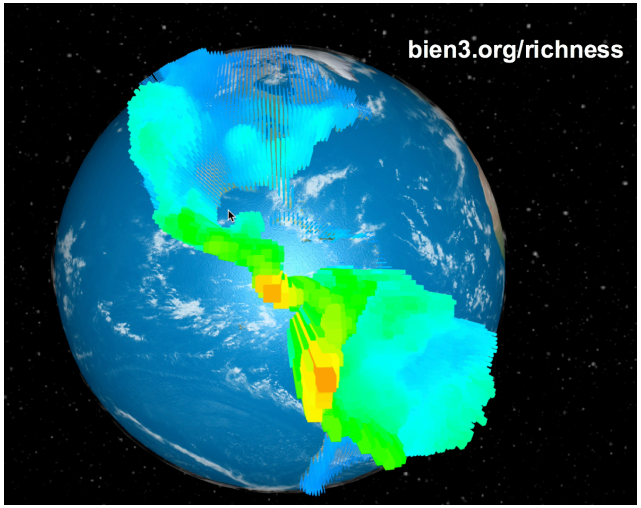
- **Combine data sets**

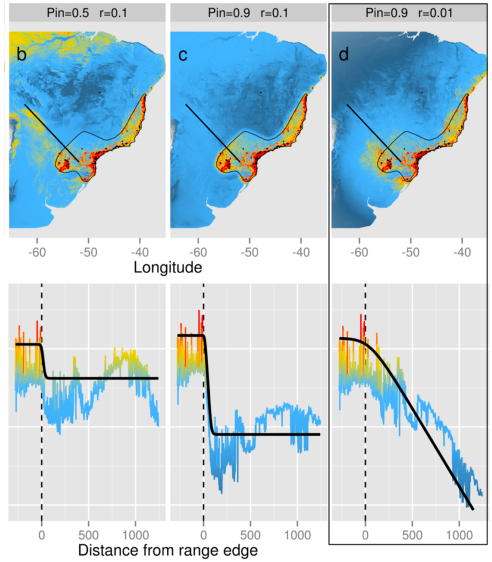


A case for informatics

- Combine data sets
- **Explore data sets**

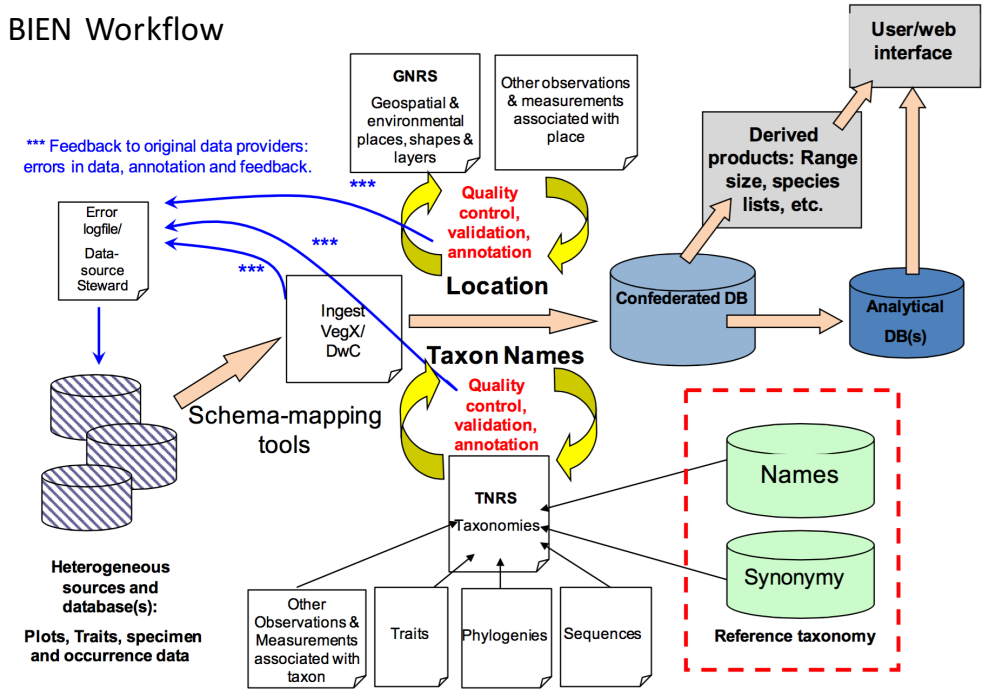






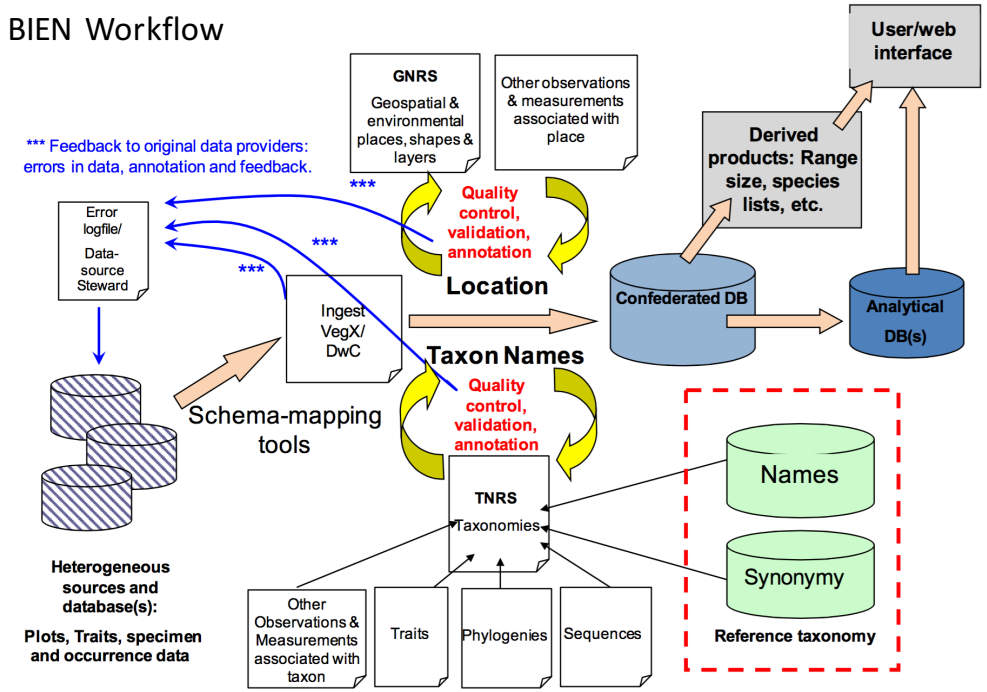
A case for informatics

- Combine data sets
- Explore data sets
- **Get our data organized**



A case for informatics

- Combine data sets
- Explore data sets
- **Get our data organized**

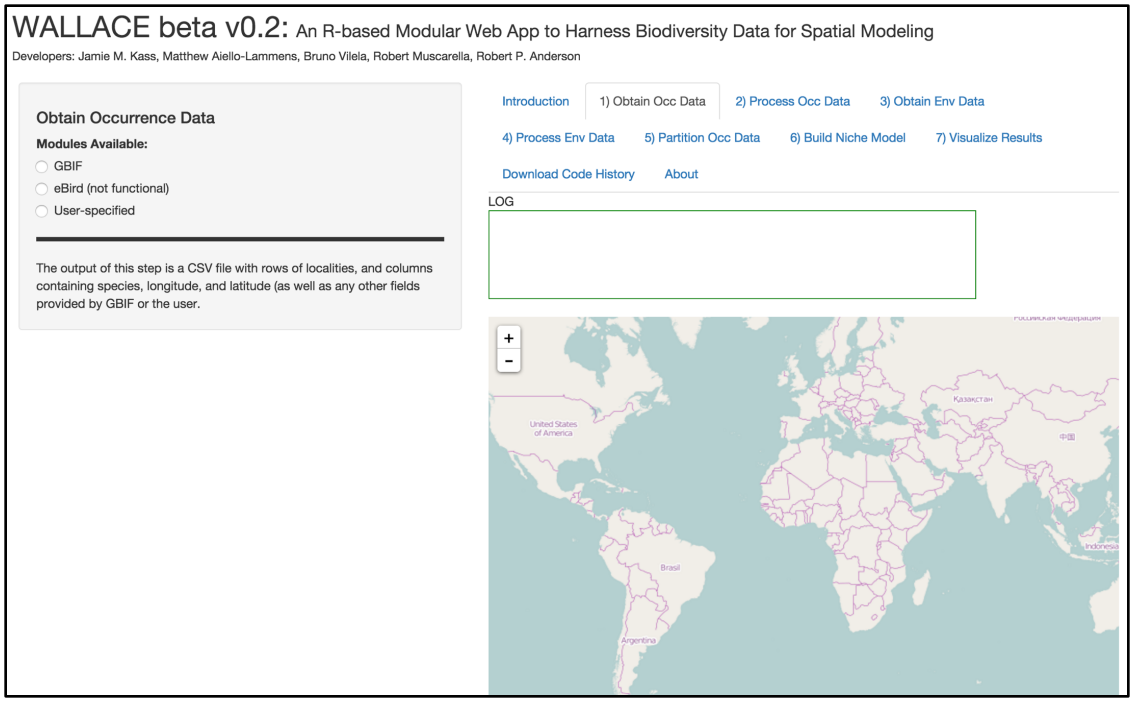


`RBien` or `BRI`?

BIEN_occurrence_species()
BIEN_ranges_load_species()
BIEN_plot_dataset()

A case for informatics

- Combine data sets
- Explore data sets
- **Get our data organized**



`RBien` or `BRI`?

BIEN_occurrence_species()
BIEN_ranges_load_species()
BIEN_plot_dataset()

Applied biodiversity informatics

Anderson (2012)

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Issue: *Blavatnik Awards for Young Scientists*

Harnessing the world's biodiversity data: promise and peril in ecological niche modeling of species distributions

Robert P. Anderson[1,2,3,4]

Agenda: Making data and modeling ready to address critical environmental issues of the 21st century

Agenda: applied biodiversity informatics

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Issue: *Blavatnik Awards for Young Scientists*

Harnessing the world's biodiversity data: promise and peril in ecological niche modeling of species distributions

Robert P. Anderson[1,2,3,4]

1. High-quality data, ready to be accessed when the particular problem presents itself

2. Software that achieves an appropriate balance between automation and supervision

3. Scientists capable of building, applying, and appraising high-quality models

THANK YOU

National Science Foundation
WHERE DISCOVERIES BEGIN

NSF DEB-1119915 and DBI-1650241

GC Digital Initiatives
THE GRADUATE CENTER

EBBE NIELSEN CHALLENGE
GBIF
bringing biodiversity information to life

http://www.andersonlab.ccny.cuny.edu/

The City College of New York

American Museum of Natural History

cmerow.github.io