Methods, Protocols, and Analytical Tools for Specimen-Based Research in the Biological Sciences
2019 Digital Data in Biodiversity Research Conference

Discussion Session: Developing a Framework of Core Biodiversity Data Literacy Skills

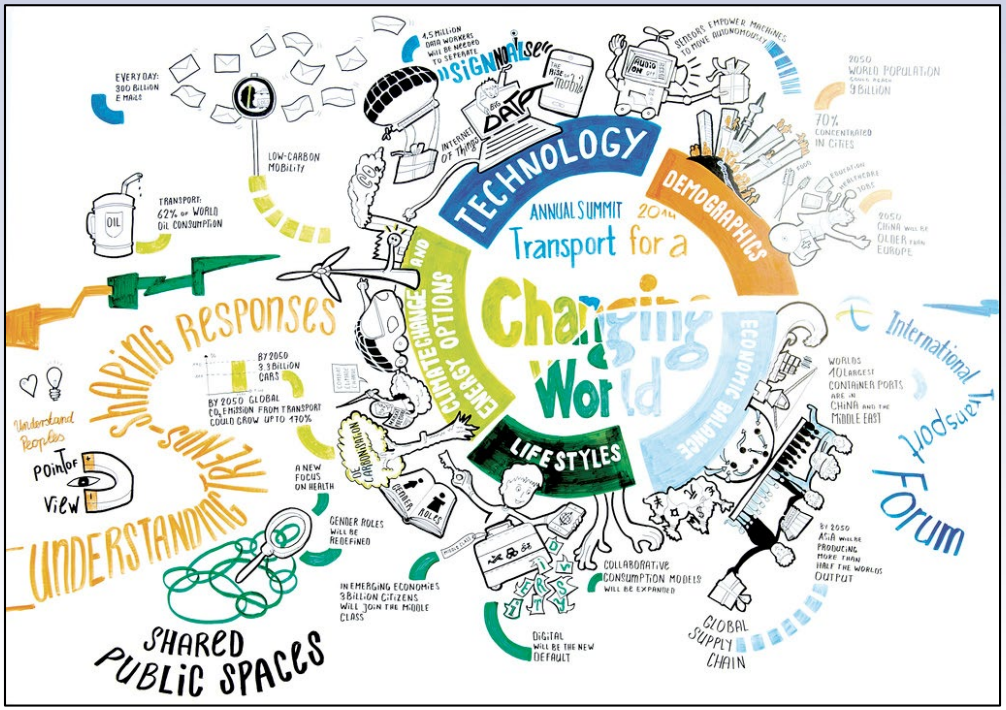
Yale, New Haven, CT
10 June 2019

Anna K. Monfils, Central Michigan University
Elizabeth R. Ellwood, La Brea Tar Pits & Museum
Lisa White, University of California Museum of Paleontology
Debra L. Linton, Central Michigan University
Natalie Douglas, Central Michigan University
Molly Phillips, University of Florida

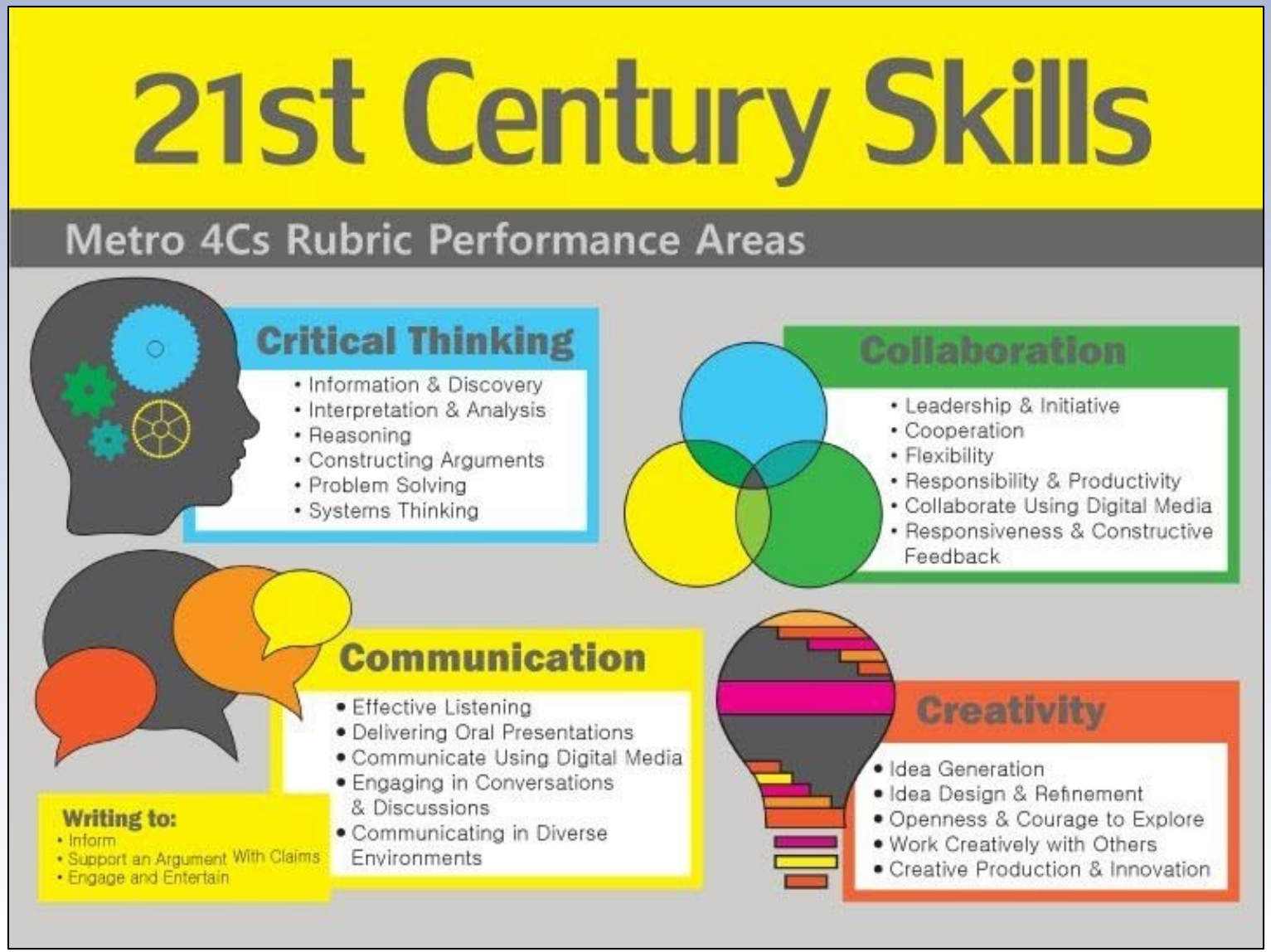
BLUE
Biodiversity Literacy in Undergraduate Education

Changing landscape for the next generation workforce

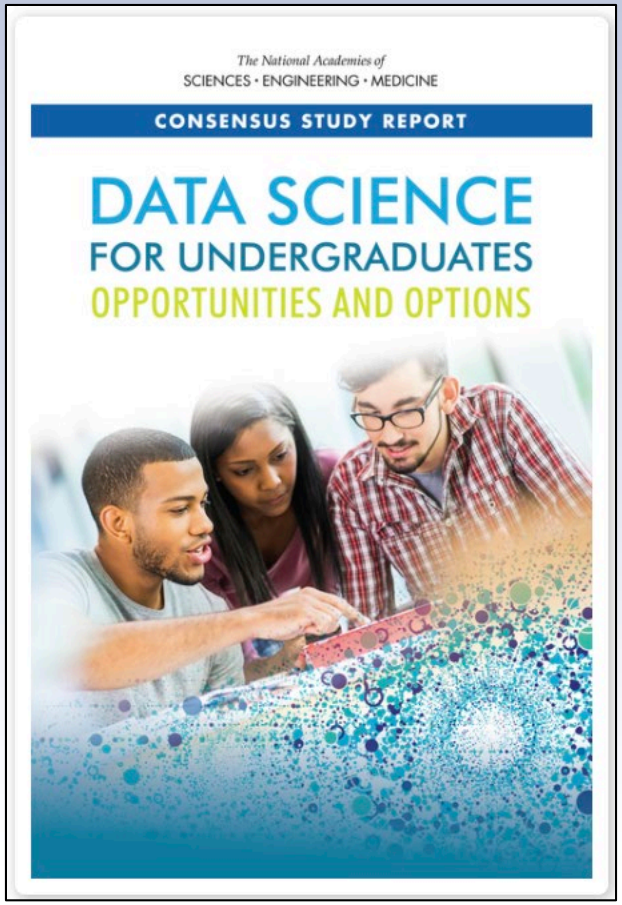
Increased technology, mobility, connectivity, diversity, resource needs, and explosion of data



Created by Changing World Animation


21st Century Skills
Metro 4Cs Rubric Performance Areas
Critical Thinking
• Information & Discovery
• Interpretation & Analysis
• Reasoning
• Constructing Arguments
• Problem Solving
• Systems Thinking
Collaboration
• Leadership & Initiative
• Cooperation
• Flexibility
• Responsibility & Productivity
• Collaborate Using Digital Media
• Responsiveness & Constructive Feedback
Communication
• Effective Listening
• Delivering Oral Presentations
• Communicate Using Digital Media
• Engaging in Conversations & Discussions
• Communicating in Diverse Environments
Writing to:
• Inform
• Support an Argument With Claims
• Engage and Entertain
Creativity
• Idea Generation
• Idea Design & Refinement
• Openness & Courage to Explore
• Work Creatively with Others
• Creative Production & Innovation

National Academies of Sciences, Engineering, and Medicine 2018. *Data Science for Undergraduates: Opportunities and Options*. Washington, DC: The National Academies Press.

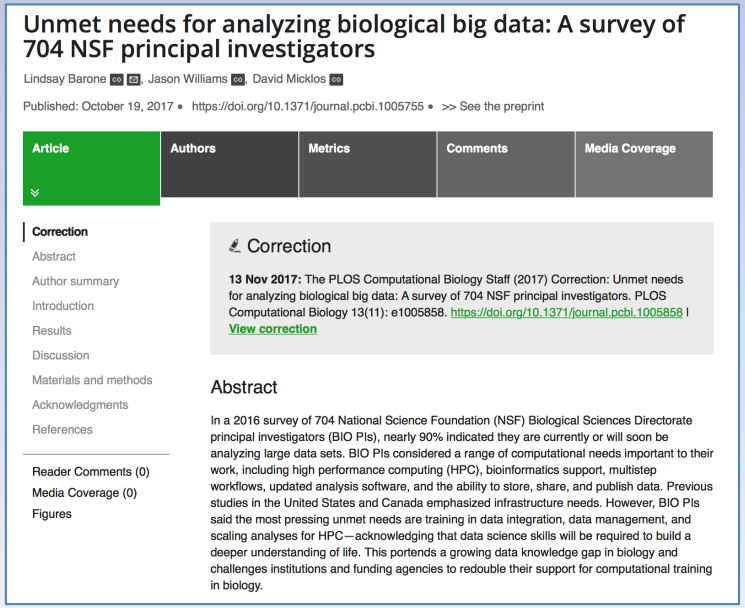


"…data science spans a broader array of activities that involve applying principles for data collection, storage, integration, analysis, inference, communication, and ethics.

In future decades, all undergraduates will benefit from a fundamental awareness of and competence in data science."

Barone L, Williams J, Micklos D. Unmet needs for analyzing biological big data: A survey of 704 NSF principal investigators. PLOS Computational Biology 2017, 13(10)

2016 survey of 704 National Science Foundation (NSF) Biological Sciences Directorate principal investigators;
90% are or will be analyzing big data sets



Unmet needs for analyzing biological big data: A survey of 704 NSF principal investigators

Lindsay Barone, Jason Williams, David Micklos

Published: October 19, 2017 • https://doi.org/10.1371/journal.pcbi.1005755 • >> See the preprint

Article | Authors | Metrics | Comments | Media Coverage

Correction
Abstract
Author summary
Introduction
Results
Discussion
Materials and methods
Acknowledgments
References
Reader Comments (0)
Media Coverage (0)
Figures

Correction

13 Nov 2017: The PLOS Computational Biology Staff (2017) Correction: Unmet needs for analyzing biological big data: A survey of 704 NSF principal investigators. PLOS Computational Biology 13(11): e1005858. https://doi.org/10.1371/journal.pcbi.1005858 | View correction

Abstract

In a 2016 survey of 704 National Science Foundation (NSF) Biological Sciences Directorate principal investigators (BIO PIs), nearly 90% indicated they are currently or will soon be analyzing large data sets. BIO PIs considered a range of computational needs important to their work, including high performance computing (HPC), bioinformatics support, multistep workflows, updated analysis software, and the ability to store, share, and publish data. Previous studies in the United States and Canada emphasized infrastructure needs. However, BIO PIs said the most pressing unmet needs are training in data integration, data management, and scaling analyses for HPC—acknowledging that data science skills will be required to build a deeper understanding of life. This portends a growing data knowledge gap in biology and challenges institutions and funding agencies to redouble their support for computational training in biology.

"…the most pressing unmet needs are training in data integration, data management, and scaling analyses for HPC—***acknowledging that data science skills will be required to build a deeper understanding of life***."

NSF 2026 Idea Machine

"…In biology, at least the last 10 years have been a transition into a "Big Data" science. While the technological obstacles of this transition were significant, the **highest barrier continues to be building computational and data skills within the science workforce**"

Reinventing Scientific Talent

https://nsf2026imgallery.skild.com/entries/reinventing-scientific-talent

Biodiversity Collections Network. 2019. Report: Extending U.S. Biodiversity Collections to Promote Research and Education. American Institute of Biological Sciences, Washington, DC

Report Release April 4th, 2019

Extending U.S. Biodiversity Collections to Promote Research and Education

3 of 8 *Next Steps* relate to education

- Develop **data access tools** designed to maximize the educational potential of collections and collections-based data.
- Champion broad scale adoption of **core biodiversity data literacy skills** and competencies in K-12 and undergraduate curricula.
- Support **enhanced training of emerging and established professionals for interdisciplinary work in biodiversity, data science, and informatics.**

BLUE

Biodiversity Literacy in Undergraduate Education

@BiodiversityEd

Biodiversityliteracy.org

BLUE-L@LISTS.UFL.EDU

http://tinyurl.com/bluecontact

BLUE Data Network Vision

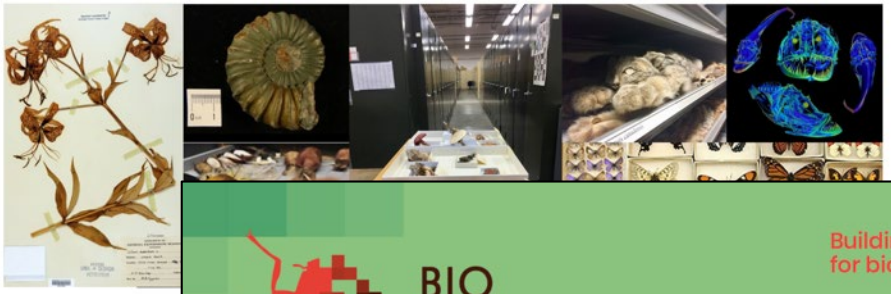
- Incorporate biodiversity science and data skills across the core undergraduate curriculum to improve undergraduate biology training and meet increasing workforce demands.

Defining Biodiversity Literacy and Core Competencies

5th Life Discovery – Doing Science Biology Education Conference

Microbiomes to Ecosystems:
Evolution and Biodiversity Across Scale, Space and Time

Digital Data in Biodiversity Research Conference



BIO DIVERSITY NEXT

Building a global infrastructure for biodiversity data. Together.

Better data - better science - better policies

ADVANCING THE INTRODUCTORY BIOLOGY EXPERIENCE

Home | Why Intro Bio? | Conference Information | Conference Registration

Nebraska › Advancing the Introductory Biology Experience

Advancing the Introductory Biology Experience

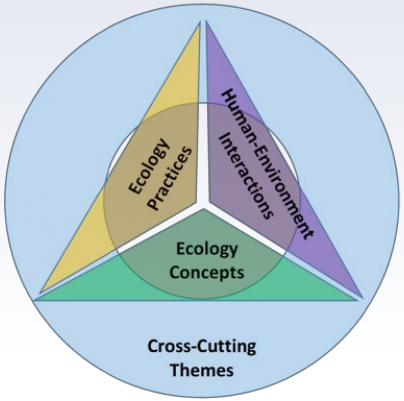
A conference to develop a unified vision for undergraduate introductory biology

Presented by the National Association of Biology Teachers

August 7-9, 2019

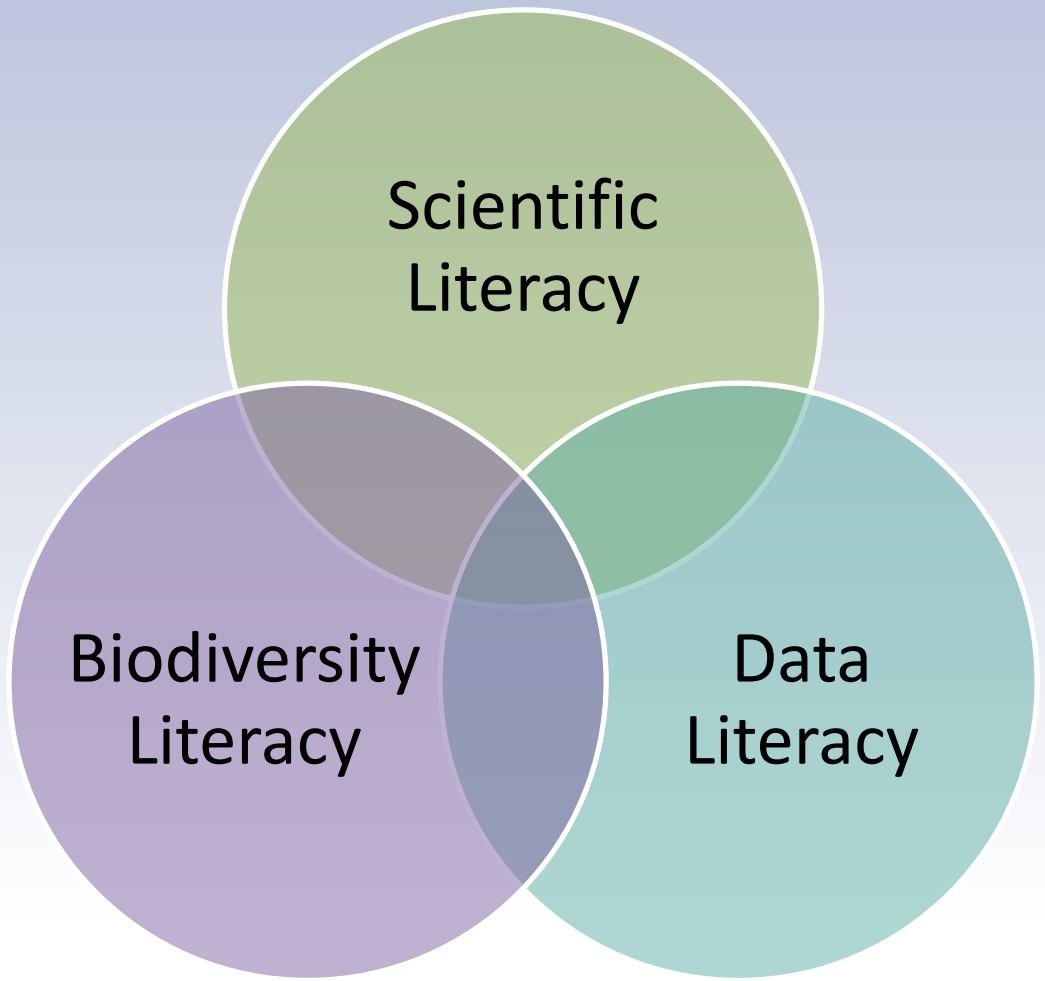
Howard Hughes Medical Institute, Chevy Chase, Maryland

4DEE



BLUE
Biodiversity Literacy in Undergraduate Education

Biodiversity Data Literacy



BLUE
Biodiversity Literacy in Undergraduate Education

Define Literacy: Having knowledge and skills

- *Science literacy*
  - *Science* is a way of knowing
  - knowledge and understanding of scientific concepts and processes
- *Data literacy*
  - reasoning with data for information and knowledge
- *Biodiversity literacy*
  - content and skills required to address questions relating to organismal biology

Biodiversity Data Literacy

Informed Citizen

- Understand biodiversity relevant science
- Communicate clearly about biodiversity science
- Make informed decisions about biodiversity issues

21st Century Workforce

- Discover sources of biodiversity data
- Evaluate the relative utility of specific biodiversity data resources
- Apply biodiversity data to discipline specific questions

Biodiversity Researcher

- Employ core data skills relevant to biodiversity data
- Collect, discover, manage, evaluate, and critically apply biodiversity data to complex evolutionary, ecological, geological, and conservation based questions.
- Integrate the process of science, core data skills, and biodiversity content when designing and executing biodiversity research

Biodiversity Data Literacy

Informed Citizen

- Understand biodiversity relevant science
- Communicate clearly about biodiversity science
- Make informed decisions about biodiversity issues

21st Century Workforce

- Discover sources of biodiversity data
- Evaluate the relative utility of specific biodiversity data resources
- Apply biodiversity data to discipline specific questions

Biodiversity Researcher

- Employ core data skills relevant to biodiversity data
- Collect, discover, manage, evaluate, and critically apply biodiversity data to complex evolutionary, ecological, geological, and conservation based questions.
- Integrate the process of science, core data skills, and biodiversity content when designing and executing biodiversity research

Core Biodiversity Data Literacy

The fundamental understanding of the process of science, core data skills, and biodiversity content needed to apply large aggregated biodiversity data sets to discipline specific questions.

See page 2 for comments

What is “unique” about biodiversity data and science?

“Biodiversity data are complex and cross-disciplinary in thinking and application. Working with biodiversity data involves integrating different data structure (e.g., integer, file, graph, tree, etc.) and type (e.g., nominal, interval, ratio, qualitative, etc.). The data have variable accuracy and precision, involve knowledge of multiple discipline-specific collection methods, bridge spatial scales, and span geological time. Biodiversity scientists must be able to apply data skills appropriately and diversely; recognize constraints and opportunities relative to data resources, tools and systems; and use higher order thinking and a growth mindset related to new data types, resources, applications and methodologies for analysis to produce novel solutions, meaningful content, and innovative science.”

See page 7 for comments

What core skills are the basis of biodiversity data literacy?

- What do **all** biology students need to be able to do relative to biodiversity data?
- What skills would we like to introduce or reinforce early in a students undergraduate experience?
- What skills and knowledge do biologist outside the field of biodiversity science need to be able to evaluate the utility of biodiversity data?

In your opinion, how important is it for undergraduates majoring in life sciences or closely related disciplines to be able to...

- Explain the role of data mining in hypothesis-driven processes within the life sciences?
- Identify appropriate sources for specific types of biodiversity data?
- Find and access various types of biodiversity data (data discovery)?
- Evaluate the utility of data to answer a scientific question?
- Evaluate the resolution and accuracy of data relative to collection methods?
- "Clean a dataset", i.e. remove data that is incorrect, incomplete, improperly formatted, or duplicated?
- Aggregate data from multiple sources?
- Evaluate the relative utility of data from multiple sources?
- Perform statistical tests commonly used with large datasets?
- Visualize biodiversity data using a variety of graphic formats (e.g., line graph, bar graph, distribution map, infographic)?
- Apply biodiversity data to a research question?
- Identify and use data standards when collecting data?
- Format data for long term use?
- Contribute data to open-data aggregators?

See page 5 for comments

Acknowledgements

- AIM-UP! Network Participants
  - J. Cook, S .Edwards, S. Ickert-Bond, & E. Lacey (NSF 0956129)
- iDigBio Education and Outreach Working Group
  - L. Page, L. Fortés, B. McFadden, G. Riccardi, & P. Soltis (NSF 1115210)
- QUBES
  - NSF 1346584, 1446269, 1446258, & 1446284
- CollectionsWeb
  - L.A. Prather, H. Bart, M. Blackwell, & J. Woolley (NSF 0639214)
- Biodiversity Collections Network (BCoN)
  - R. Gropp & A. Bentley, (NSF 144178)