Cold Spring Harbor Laboratory

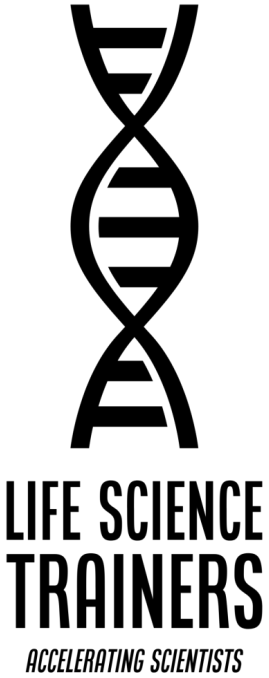
CURES and Data

Good data training in the classroom

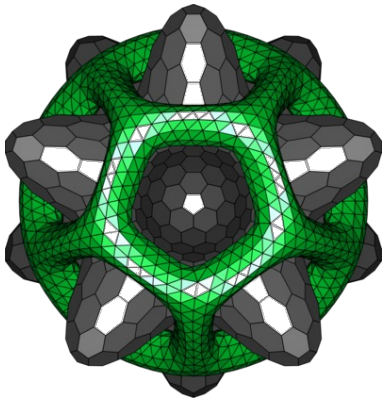
Jason Williams

Cold Spring Harbor Laboratory, DNA Learning Center

@JasonWilliamsNY

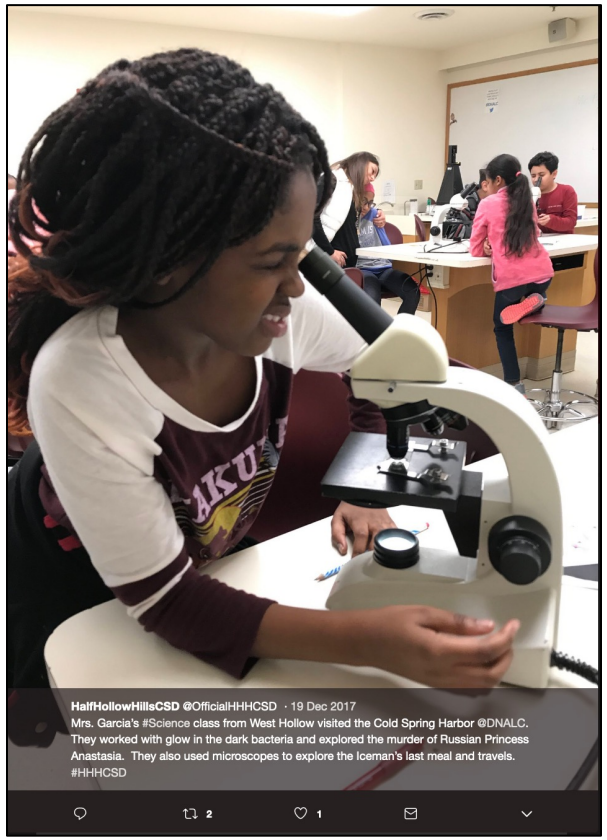


CyVerse



Journal of
Open Source Education

I went to school for…



I need to know this for my research…

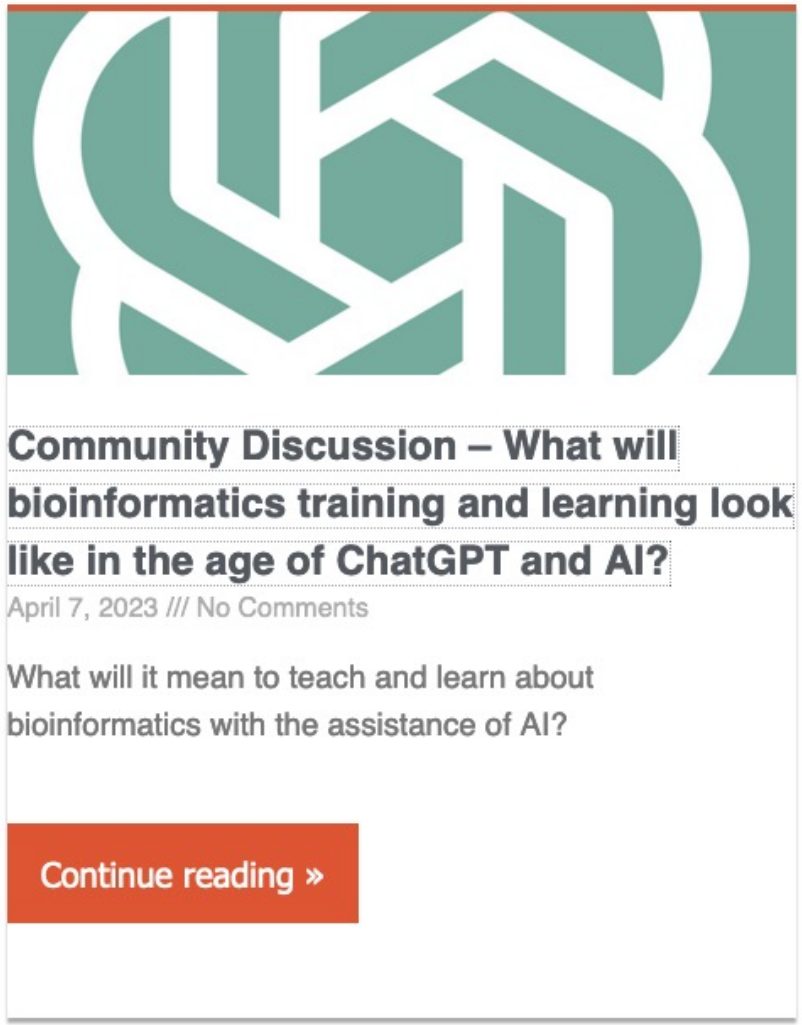
```
import urllib2
eutils = 'http://www.ncbi.nlm.nih.gov/entrez/eutils/'
efetch = 'efetch.fcgi?'
s = eutils + efetch

targets = ['J04243','M60064']
idString = 'id=' + ','.join(targets)
s += idString + '&db=nucleotide&rettype=fasta'
fileObject = urllib2.urlopen(s)
data = fileObject.read().strip()

entries = data.split('\n\n')
title, sequence = entries[0].split('\n', 1)
print title.split(' ',1)[0]

# prints:
# >gi|154102|gb|J04243.1|STYHEMAPRF
```

I don’t know what the future looks like



Community Discussion – What will bioinformatics training and learning look like in the age of ChatGPT and AI?

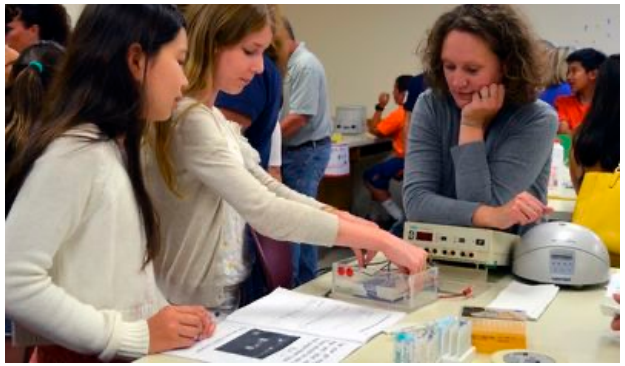
April 7, 2023 /// No Comments

What will it mean to teach and learn about bioinformatics with the assistance of AI?

Continue reading »

April post/Meeting info

CSHL DNA Learning Center



Hands on education in molecular biology/bioinformatics for secondary students (Grade 6-12); secondary and undergraduate faculty training; websites/multimedia

NIBLSE
Bioinformatics in Life Sciences Education

PLOS ONE

RESEARCH ARTICLE

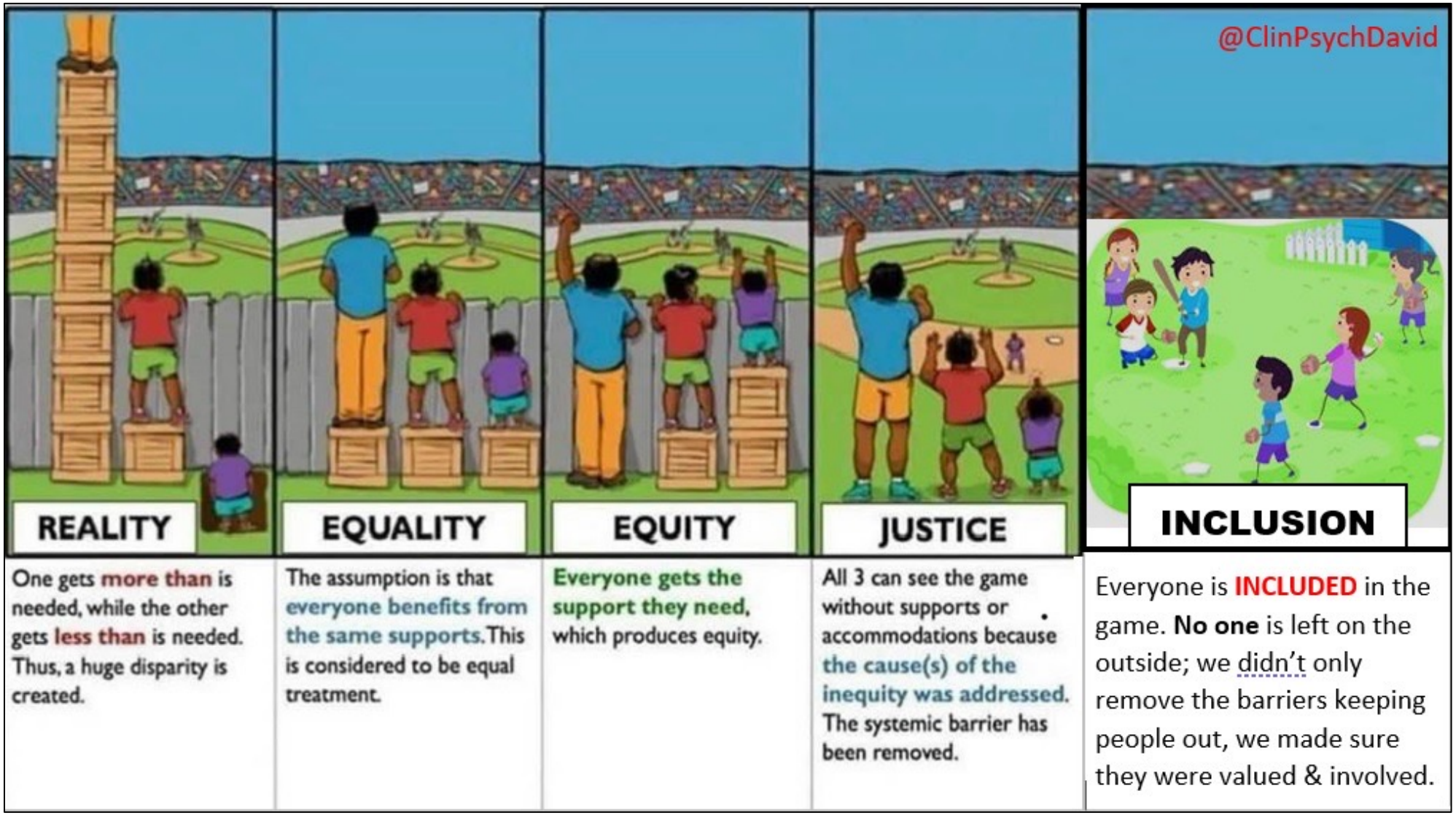
Barriers to integration of bioinformatics into undergraduate life sciences education: A national study of US life sciences faculty uncover significant barriers to integrating bioinformatics into undergraduate instruction

95%

of respondents indicate that bioinformatics should be integrated into the life science curriculum;

32% of faculty report achieving this

These gaps multiply and perpetuate



Credit: Saskatoon Health Service; Revised version - David Murphy

Undergraduate curriculum

PLOS ONE

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Bioinformatics core competencies for undergraduate life sciences education

Melissa A. Wilson Sayres, Charles Hauser, Michael Sierk, Srebrenka Robic, Anne G. Rosenwald, Todd M. Smith, Eric W. Triplett, Jason J. Williams, Elizabeth Dinsdale, William R. Morgan, James M. Burnette III, Samuel S. Donovan, Jennifer C. Drew, [...], Mark A. Pauley [view all]

Published: June 5, 2018 • https://doi.org/10.1371/journal.pone.0196878

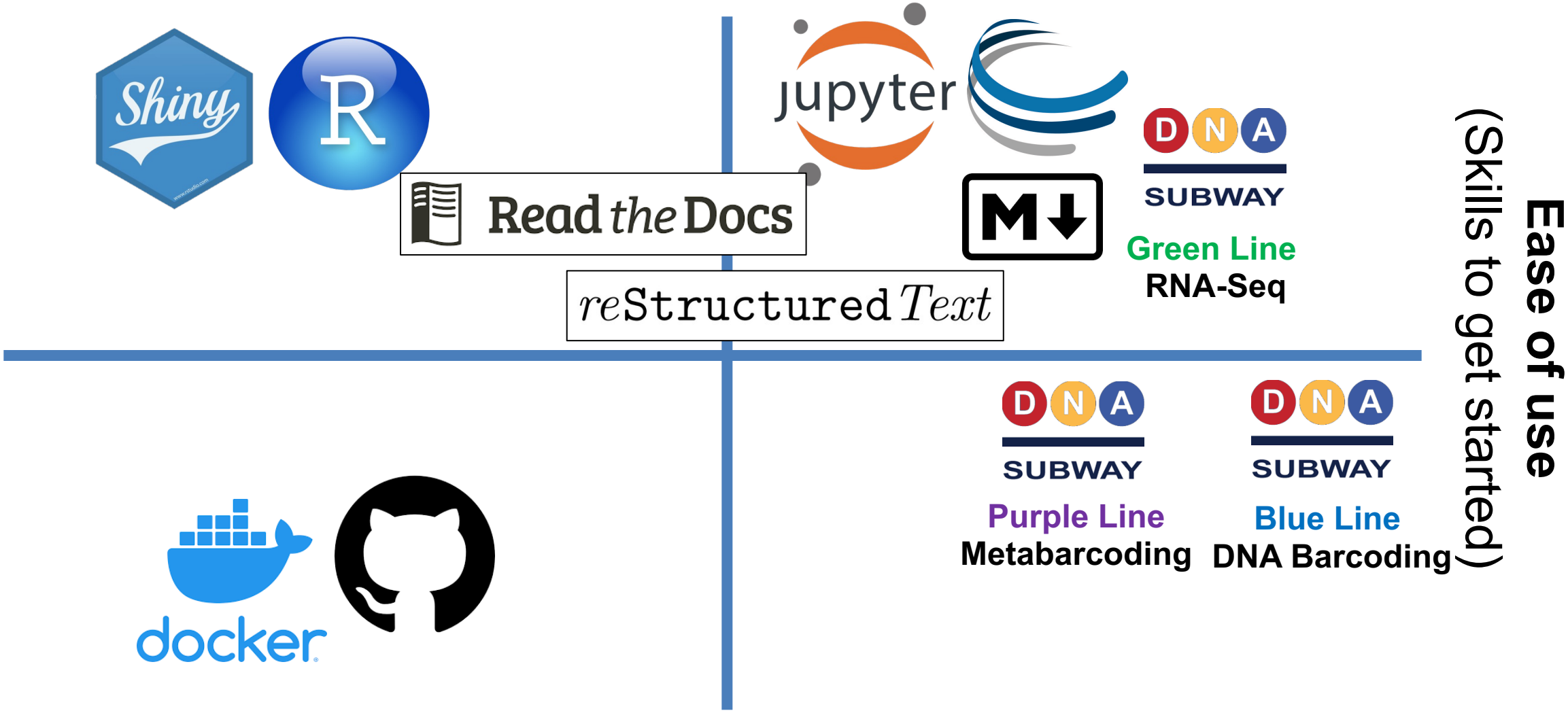
See the preprint

C1. Role of Bioinformatics | C2. Computational Concepts | C3. Statistical Concepts | C4. Bioinformatics Tools

C5. Data Retrieval | C6. Model | C7. Scripting | C8. Data Types | C9. Implications

POWER

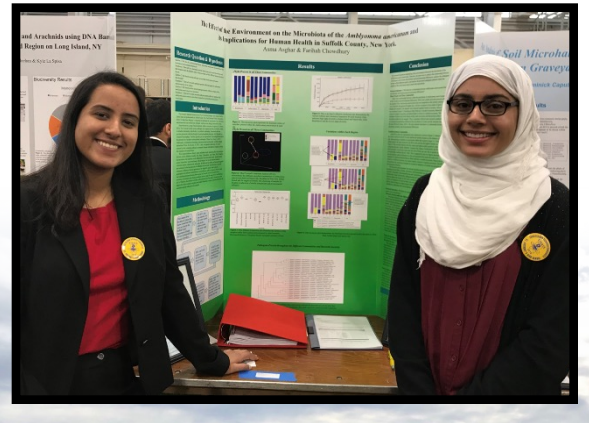
(number of uses/applications)



Ease of use

(Skills to get started)

DNA Barcoding Programs



NYC URBAN BARCODE PROJECT

URBAN BARCODE RESEARCH PROGRAM

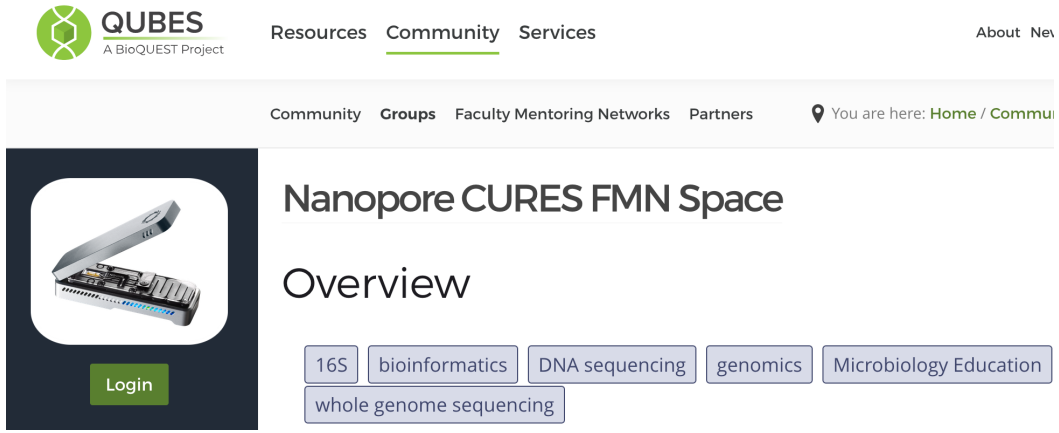
BARCODE LONG ISLAND

NIH NIGMS

SEPA SCIENCE EDUCATION PARTNERSHIP AWARD

SUPPORTED BY THE NATIONAL INSTITUTES OF HEALTH

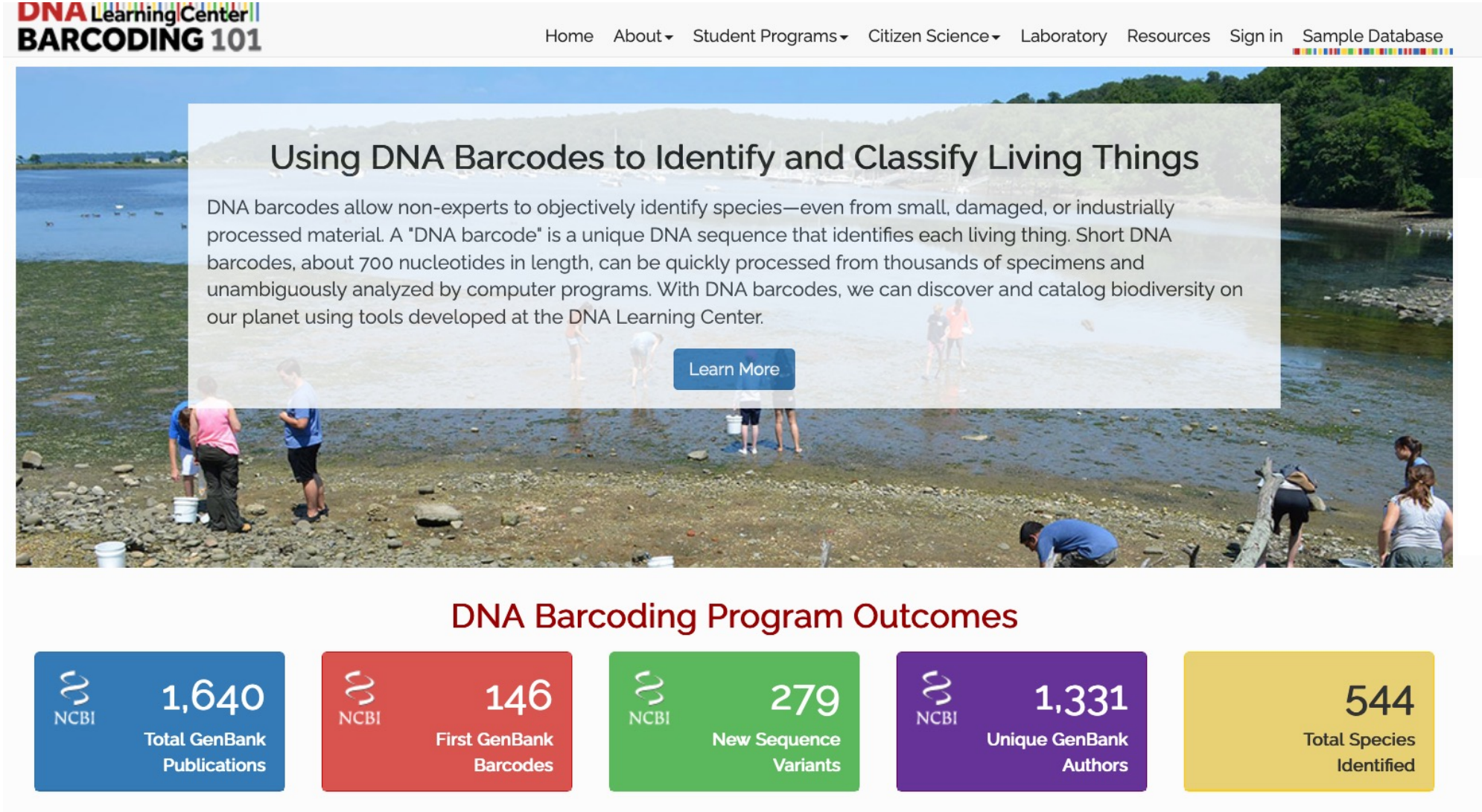
Nanopore Network



Developing Foundations for Nanopore DNA Sequencing Course-based Undergraduate Research Experiences at Minority-Serving Institutions

- Pilot (2-years)
- Simplify lab and bioinformatics protocols
- Support faculty needs and understand barriers to use

DUE# 2216349



DNA Barcoding Program Outcomes

1,640	Total GenBank Publications
146	First GenBank Barcodes
279	New Sequence Variants
1,331	Unique GenBank Authors
544	Total Species Identified

Banbury working group (May 2022)

"The illiterate of the 21st century will not be those who cannot read and write, but those who cannot learn, unlearn, and relearn"

– A. Toffler

Thanks!