



# 5th National Big Data Health Science Conference

Friday Feb 2, 2024 / 09:30AM - 10:15AM

“*Enhancing Institutional Data Science Capacity*”



Raphael D.  
Isokpehi, Ph.D.

# Enhancing Institutional Data Science Capacity

Raphael D. Isokpehi, PhD

*Training, Workforce Initiatives, and Community Engagement (TWICE)*

*NIH Office of Data Science Strategy (ODSS)*



Unlocking the Power of  
**BIG DATA**  
IN HEALTH  
Empowering Scientific and  
Healthcare Communities with  
Data Analytics



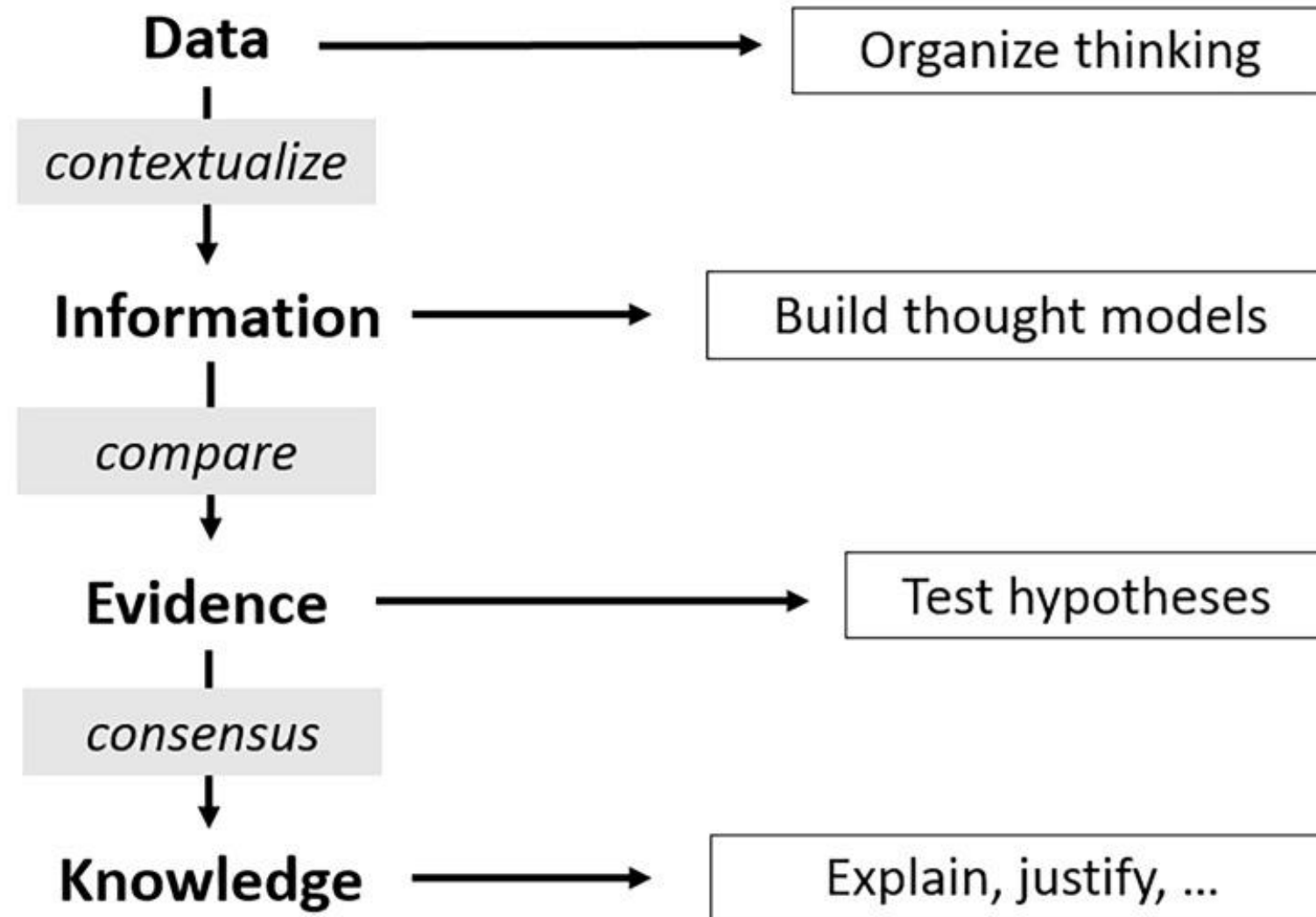
**5<sup>TH</sup> NATIONAL BIG DATA  
HEALTH SCIENCE CONFERENCE**  
February 2-3 2024 | Columbia, SC

Keynote Talk on February 2, 2024

# **Part 1: Data, Big Data & Analytics of Designations of U.S. Higher Education Institutions**

# Data: Quantitative and Qualitative

Data are quantitative (numeric) or qualitative (non-numeric) values of variables.





# Big Data: Voluminous Amounts of Data

Technological and other developments in the 21st Century have led to individual, organizational, and societal capacities to produce Big Data for diverse purposes.

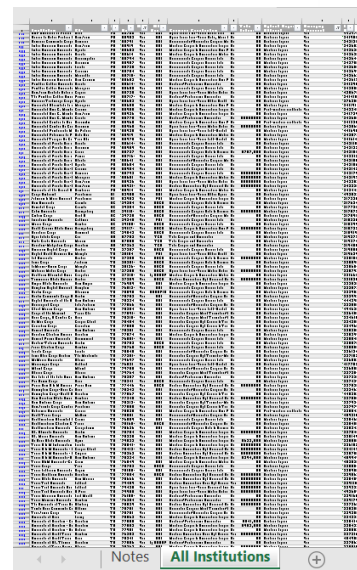


# Analytics of Dataset on Designations of Accredited U.S. Higher Education Institutions

**Data Source:**  
Office of Science  
U.S. Department of Energy

Link	Eligibility Dates Covered	Significant Changes from Prior Posting
<a href="#">Excel</a>	December 21, 2023 to Present	Updated total federal R&D expenditures using the National Science Foundation's FY 2022 Higher Education Research and Development Survey data.

## Dataset of U.S. Higher Education Institutions



Data Fields
Institution
City
State
Zip
Minority Serving Institution (MSI)?
MSI Designation
Carnegie Classification
R1?
Total Federal R&D Expenditures
Highest Degree Offered
Emerging Research Institution
IPEDS Unit ID

12 Data Fields  
1 Header Row  
6388 Institutions



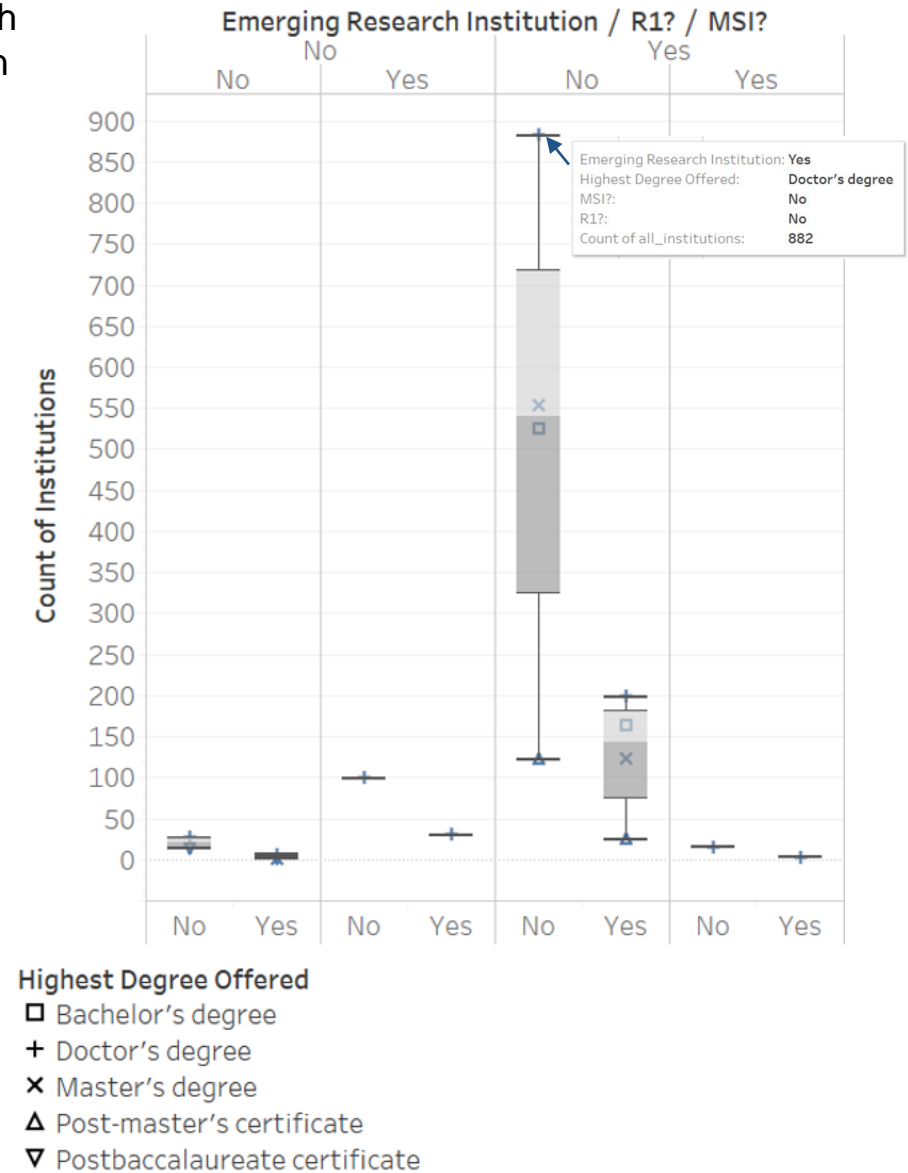
Carnegie Classification	
Associate's Colleges: High Career & Technical-High Nontraditional	85
Associate's Colleges: High Career & Technical-High Traditional	99
Associate's Colleges: High Career & Technical-Mixed Traditional/Nontraditional	113
Associate's Colleges: High Transfer-High Nontraditional	109
Associate's Colleges: High Transfer-High Traditional	106
Associate's Colleges: High Transfer-Mixed Traditional/Nontraditional	98
Associate's Colleges: Mixed Transfer/Career & Technical-High Nontraditional	112
Associate's Colleges: Mixed Transfer/Career & Technical-High Traditional	102
Associate's Colleges: Mixed Transfer/Career & Technical-Mixed Traditional/Nontraditional	95
Baccalaureate Colleges: Arts & Sciences Focus	215
Baccalaureate Colleges: Diverse Fields	299
Baccalaureate/Associate's Colleges: Associate's Dominant	97
Baccalaureate/Associate's Colleges: Mixed Baccalaureate/Associate's	90
Doctoral Universities: High Research Activity	133
Doctoral Universities: Very High Research Activity	147
Doctoral/Professional Universities	185
Master's Colleges & Universities: Larger Programs	316
Master's Colleges & Universities: Medium Programs	179
Master's Colleges & Universities: Small Programs	156
Not applicable, not in Carnegie universe (not accredited or nondegree-granting)	2,600
Special Focus Four-Year: Arts, Music & Design Schools	67
Special Focus Four-Year: Business & Management Schools	47
Special Focus Four-Year: Engineering and Other Technology-Related Schools	12
Special Focus Four-Year: Faith-Related Institutions	243
Special Focus Four-Year: Law Schools	30
Special Focus Four-Year: Medical Schools & Centers	36
Special Focus Four-Year: Other Health Professions Schools	224
Special Focus Four-Year: Other Special Focus Institutions	25
Special Focus Four-Year: Research Institution	22
Special Focus Two-Year: Arts & Design	28
Special Focus Two-Year: Health Professions	186
Special Focus Two-Year: Other Fields	49
Special Focus Two-Year: Technical Professions	48
Tribal Colleges and Universities	35
<b>Grand Total</b>	<b>6,388</b>

# Emerging Research Institutions (ERI) Designation

The term "emerging research institution" means an institution of higher education with an established undergraduate or graduate program that has less than \$50,000,000 in Federal research expenditures. H.R.4346 - Chips and Science Act

## Analytics of Dataset on Designations of Accredited U.S. Higher Education Institutions

Carnegie Classification	Highest Degree Offered / Emerging Research Institution											
	Associate's degree		Bachelor's degree		Doctor's degree		Master's degree		Post-master's certificate		Postbaccalaureate certificate	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Baccalaureate Colleges: Arts & Sciences Focus		96		23		84		10				2
Baccalaureate Colleges: Diverse Fields		88		24		176		10				1
Doctoral Universities: High Research Activity			15	118								
Doctoral Universities: Very High Research Activity			129	18								
Doctoral/Professional Universities				184				1				
Master's Colleges & Universities: Larger Programs				219		50		47				
Master's Colleges & Universities: Medium Programs				85		58		36				
Master's Colleges & Universities: Small Programs			1	61		77		17				
Special Focus Four-Year: Medical Schools & Centers			1	35								
Special Focus Four-Year: Other Health Professions Schools	2	60		102	1	46		13				
Special Focus Four-Year: Research Institution			16	6								
<b>Grand Total</b>	<b>2</b>	<b>244</b>	<b>162</b>	<b>875</b>	<b>1</b>	<b>491</b>	<b>134</b>	<b>3</b>				





# Part 2: NIH Strategic Plan for Data Science





# National Institutes of Health



Search NIH

[Virtual Tour](#) | [Staff Directory](#) | [En Español](#)

[Health Information](#)

[Grants & Funding](#)

[News & Events](#)

[Research & Training](#)

[Institutes at NIH](#)

[About NIH](#)

[Home](#) » [About NIH](#)

## WHO WE ARE

The National Institutes of Health (NIH), a part of the U.S. Department of Health and Human Services, is the nation's medical research agency — making important discoveries that improve health and save lives.



### The NIH Director

Monica M. Bertagnolli, M.D., is the 17<sup>th</sup> director of the National Institutes of Health, officially taking office on November 9, 2023.



### NIH Leadership

NIH leadership plays an active role in shaping the agency's research planning, activities, and outlook.



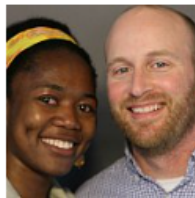
### Organization

The NIH is made up of 27 different components called Institutes and Centers.



### Staff Directory

The NIH Enterprise Directory (NED) is an electronic directory of people who work at the NIH.



### Voices of the NIH Community

NIH's project to capture an oral history of the research experience.



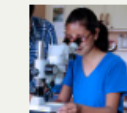
### History

For over a century, NIH scientists have paved the way for important discoveries that improve health and save lives.

## NIH at a Glance



Explore the Bethesda campus and how NIH turns discovery into health.



### Mission and Goals



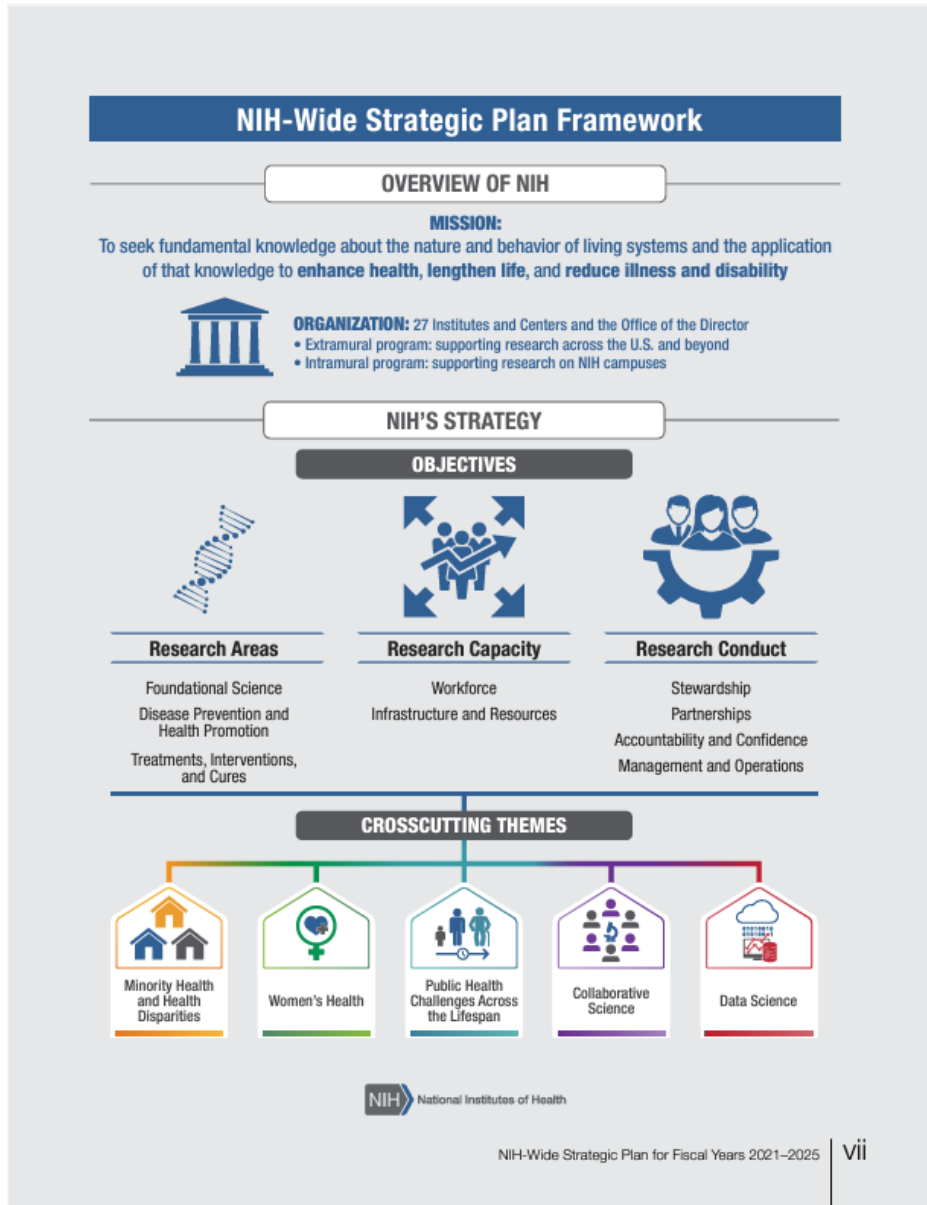
NIH Director: Monica M. Bertagnolli, M.D.



Headquarters:  
Bethesda, Maryland,  
USA

# NIH-Wide Strategic Plan for Fiscal Years 2021-2025

## Data Science is a crosscutting theme at NIH



NIH Office of the Director

Office of the Principal Deputy Director

Administration and Services

Communications

Legislative Policy and Analysis

Executive Secretariat

Management

Science Policy

Research, Funding and Coordination

Office of Data Science Strategy

The 27 NIH Institutes and Centers

# List of NIH Institutes and Centers (ICs)

## Reference for Acronyms of NIH Institutes and Centers

<b>NIH Institutes</b>
Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD)
National Cancer Institute (NCI)
National Eye Institute (NEI)
National Heart, Lung, and Blood Institute (NHLBI)
National Human Genome Research Institute (NHGRI)
National Institute of Allergy and Infectious Diseases (NIAID)
National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
National Institute of Biomedical Imaging and Bioengineering (NIBIB)
National Institute of Dental and Craniofacial Research (NIDCR)
National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
National Institute of Environmental Health Sciences (NIEHS)
National Institute of General Medical Sciences (NIGMS)
National Institute of Mental Health (NIMH)
National Institute of Neurological Disorders and Stroke (NINDS)
National Institute of Nursing Research (NINR)
National Institute on Aging (NIA)
National Institute on Alcohol Abuse and Alcoholism (NIAAA)
National Institute on Deafness and Other Communication Disorders (NIDCD)
National Institute on Drug Abuse (NIDA)
National Institute on Minority Health and Health Disparities (NIMHD)
National Library of Medicine (NLM)
<b>NIH Centers</b>
Center for Information Technology (CIT)
Center for Scientific Review (CSR)
Fogarty International Center (FIC)
National Center for Advancing Translational Sciences (NCATS)
National Center for Complementary and Integrative Health (NCCIH)
NIH Clinical Center (CC)

<https://www.nih.gov/institutes-nih/list-institutes-centers>

# Data Science is a Priority Initiative at NIH

## Big Data for Health

- NIH supports the generation and analysis of Big Data
  - All of Us Research Program
  - Cancer Moonshot
  - Common Fund Programs
  - National COVID Cohort Collaborative (N3C)
  - Projects funded by Institutes and Centers

Notice of Special Interest (NOSI): Harnessing Big Data to Halt HIV

Notice Number:

NOT-AI-21-054

- Big Data for Health is a motivation for the NIH Strategic Plan for Data Science

### Big Data for Health

Two signature NIH projects that aim to garner health insights from human data are the All of Us Research Program and the Cancer Moonshot. The All of Us Research Program aims to gather data over time from 1 million or more people living in the United States, with the ultimate goal of accelerating research and improving health. Scientists plan to use All of Us Research Program data to learn more about how individual differences in lifestyle, environment, and biological makeup can influence health and disease. Participants in the All of Us Research Program may be invited to use wearable sensors that will provide real-time measurements of their health and environmental exposures, significantly expanding this type of research. The Cancer Moonshot aims to accelerate cancer research to make more therapies available to more patients, while also improving our ability to prevent cancer and detect it at an early stage. Data-intensive strategies include mining past patient data to predict responses to standard treatments and future patient outcomes, developing a three-dimensional cancer atlas to view how human tumors change over time, and a Cancer Research Data Commons.

Source: NIH Strategic Plan for Data Science 2018-2022



# The NIH Office of Data Science Strategy (ODSS)

Dr. Susan Gregurick, Associate Director for Data Science and Director of ODSS

Dr. Belinda Seto, Deputy Director of ODSS

- ODSS leads implementation of the NIH Strategic Plan for Data Science through scientific, technical, and operational **collaboration** with the institutes, centers, and offices that comprise NIH.

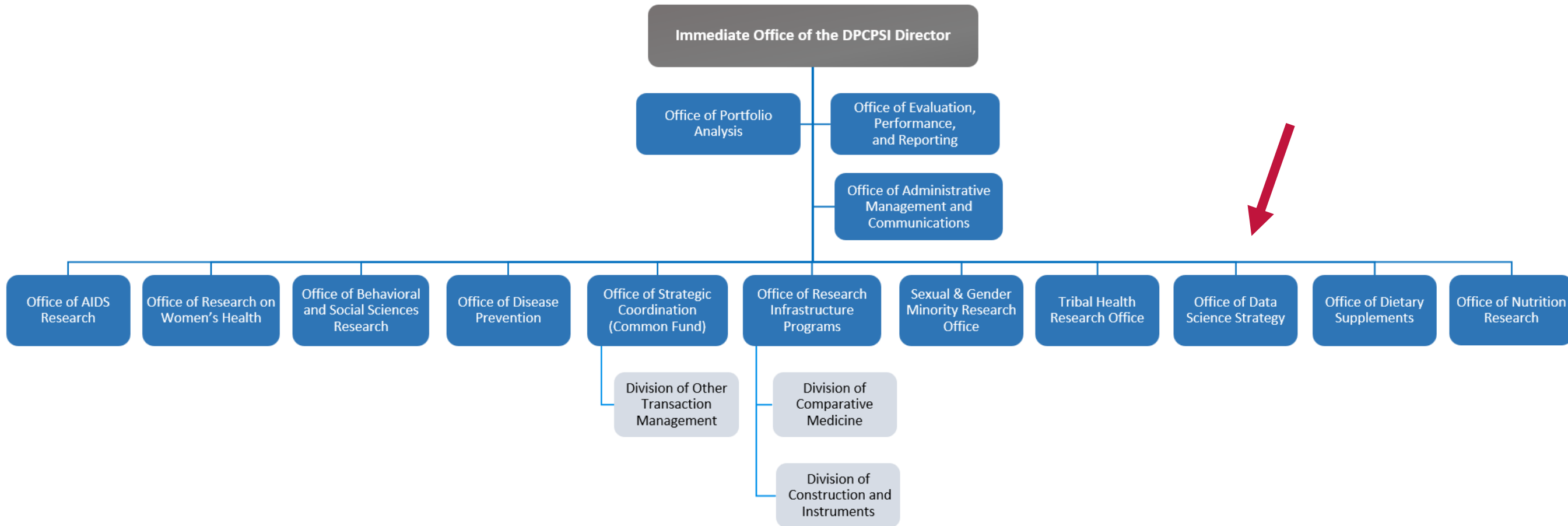
## Data Science Themes at ODSS

- Data Infrastructure
- FAIR (Findable, Accessible, Interoperable, Reusable) Data and Resources
- Tools & Analytics
- Training, Workforce Initiatives, and Community Engagement
- Clinical Informatics and FHIR ® (Fast Healthcare Interoperability Resources)
- Artificial Intelligence at NIH

<https://datascience.nih.gov/about/odss>

# Where is NIH Office of Data Science Strategy (ODSS)?

ODSS is within the NIH Office of the Director –  
*Division of Program Coordination, Planning and Strategic Initiatives (DPCPSI)*



# Training, Workforce Initiatives and Community Engagement (TWICE)

ODSS established TWICE to build a stronger and broader data science community for turning discoveries into health

	Within NIH	Extramural Community
<b>Training</b>	Recruit and support diverse data science trainees in the IRP	Support data science trainees from diverse backgrounds
<b>Workforce</b>	Facilitate recruitment and retention of diverse data science talents at NIH	Promote diversity of data science workforce in the biomedical research community
	Develop a pathway for early data scientists to join the NIH	Broaden the reach of data science among established investigators
<b>Community Engagement</b>	Enhance interconnectivity of data scientists of all levels at NIH	Enhance data science capacity, particularly in institutions serving underserved communities
	Provide training tools and resources to engage all communities in growing data science knowledge and skills	

## Five Overarching Goals

1. Improve Capabilities to Sustain the NIH Policy for Data Management and Sharing
2. Develop Programs to Enhance Human Derived Data for Research
3. Provide New Opportunities in Software, Computational Methods, and Artificial Intelligence
4. Support for a Federated Biomedical Research Data Infrastructure
5. **Strengthen a Broad Community in Data Science**
  - **Enhancing Institutional Data Science Capacity**



# NIH Strategic Plan for Data Science 2023-2028

## Request for Information (RFI) Deadline: March 15, 2024

Request for Information (RFI): Inviting Comments on the National Institutes of Health (NIH) Strategic Plan for Data Science 2023-2028

**Notice Number:**  
NOT-OD-24-037

### Key Dates

**Release Date:** December 15, 2023

**Response Date:** March 15, 2024

### Related Announcements

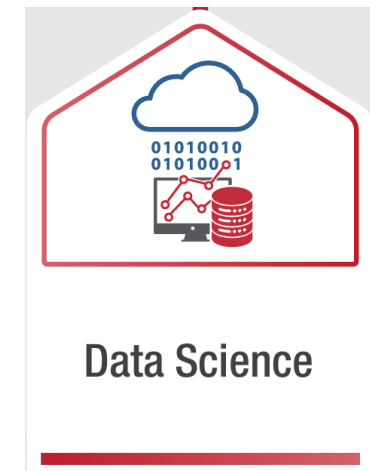
None

### Issued by

Office of Data Science Strategy (ODSS)

### Purpose

The purpose of this Request for Information (RFI) is to solicit public comments on the updated NIH Strategic Plan for Data Science, 2023-2028. The NIH is publishing this Notice to solicit input on topics under consideration for the strategic plan from its stakeholders, including members of the scientific community, academic institutions, the private sector, health professionals, professional societies, advocacy groups, and patient communities, as well as other interested members of the public.



# **Part 3: Enhancing Institutional Data Science Capacity: Programmatic Objectives**

# Enhancing Institutional Data Science Capacity is a Crosscutting Need

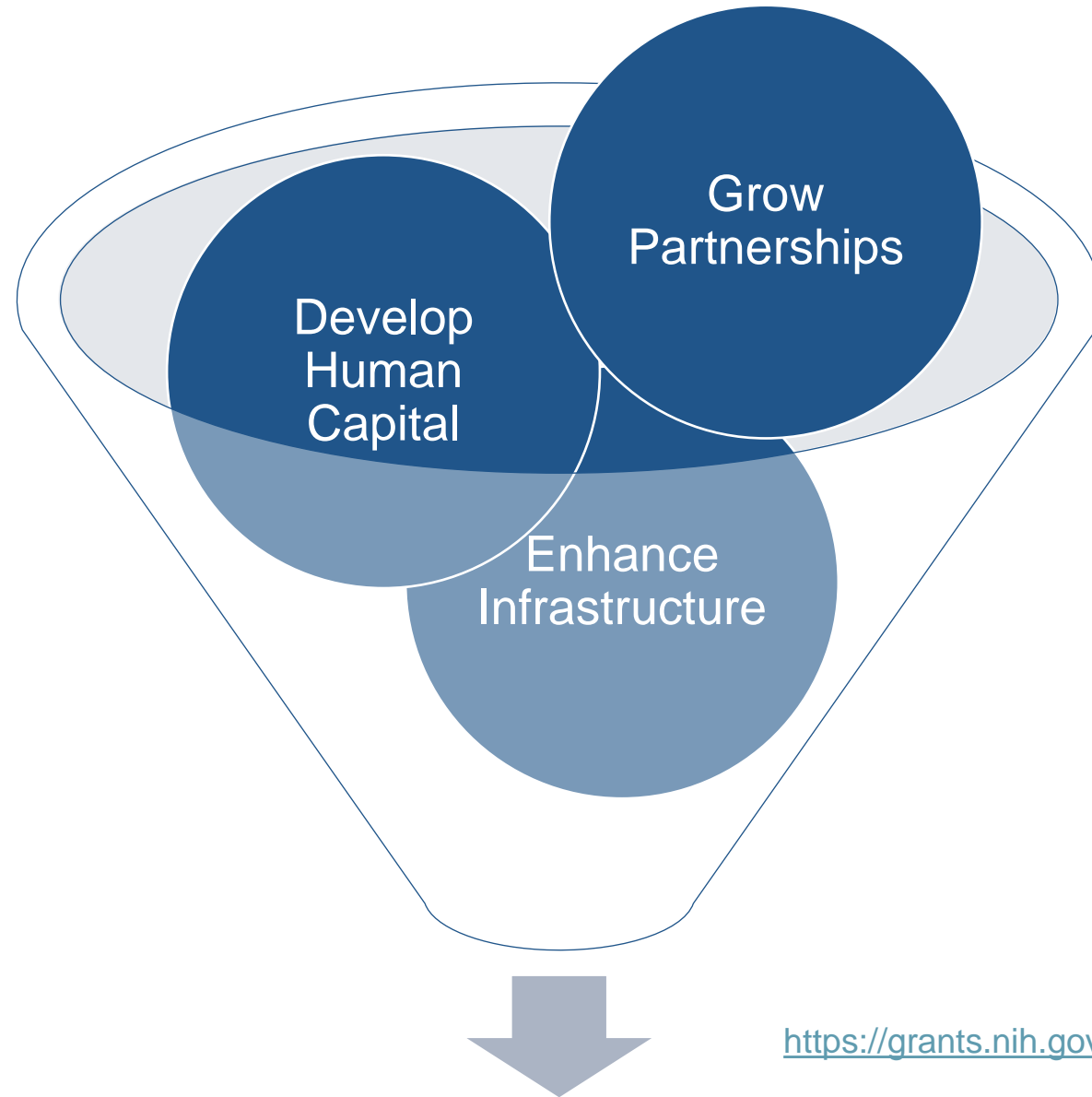
## Institution Designations/Classifications

To support applicants in assessing eligibility for Funding Opportunity Announcements, the Office of Science has compiled a list of institutions with the following specified: Minority Serving Institution designation, Carnegie classification, and Emerging Research Institution classification. Since these designations/classifications can change, files are dated to indicate the time period covered.

Link	Eligibility Dates Covered	Significant Changes from Prior Posting
<a href="#">Excel</a>	December 21, 2023 to Present	Updated total federal R&D expenditures using the National Science Foundation's FY 2022 Higher Education Research and Development Survey data.

- Total Institutions - **6388 records**
- Doctoral Universities: Very High Research Activity - **147 records**
- Emerging Research Institutions – **2608 records**
- Carnegie Classifications – **34**
- Highest Academic Degree Types - **9**

# Programmatic Objectives for Enhancing Institutional Data Science Capacity



<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-23-123.html>

Enhancing Institutional Data Science Capacity



# Data Science Topic Areas for Capacity Enhancement

Examples of data science areas include, but are not limited to:

- artificial intelligence;
- predictive analytics;
- machine learning;
- bioinformatics;
- cloud computing;
- computational science;
- software design and programming;
- supercomputing;
- statistics;
- clinical informatics;
- data visualization;
- modeling and simulation;
- data sharing and access;
- data management;
- data compression and standards;
- other data science topics

Crosscutting: Institutional awareness, knowledge and communication of:

- data ethics;
- risk management of cybersecurity.



# Data Science Community in Health Research

- **Data science literate**
  - Not intimidated by data science
  - Can read and understand reported outcomes resulting from data science approaches
  - Know where to find relevant resources
- **Data science savvy** – data science literate and
  - Will actively use data sciences approaches in research projects
  - Can initiate and/or participate in collaborations with data scientists
- **Data scientist**
  - Skilled in one or more data science areas
  - Can communicate what they learn and creatively display the information
  - Can formulate implications and implement follow up studies

# Programmatic Objective 1

## Objective 1: Grow Human Capital with Data Science Competencies

Promote data science knowledge, skills, abilities and behavior

### For Example:

- **Training courses and/or education events.**
- **Engagement of experts.**
- **Short-term mentored internship experiences.**
- **Training for members of Institutional Review Board (IRB).**



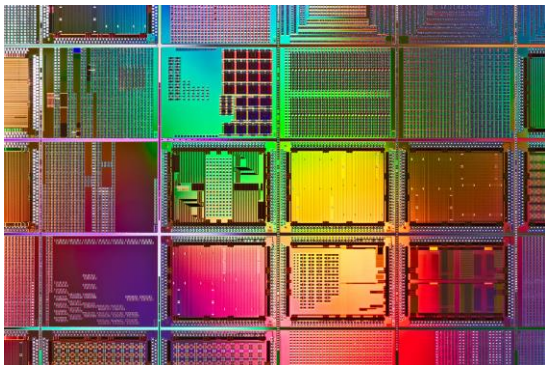
# Programmatic Objective 2

## Objective 2: Develop or Expand Institutional Infrastructure

Support data science research, training and education

### For Example:

- **Researchers' ability to conduct data science-relevant biomedical and health research.**
- **Researchers' access to controlled and registered datasets.**
- **Research with computational tools and datasets available in secure workspaces or workbenches of NIH cloud resources.**





# Programmatic Objective 3

## Objective 3: Build Data Science Partnerships

- Within institution;
- With other institutions;
- Across programs;
- With other organizations

### For Example:

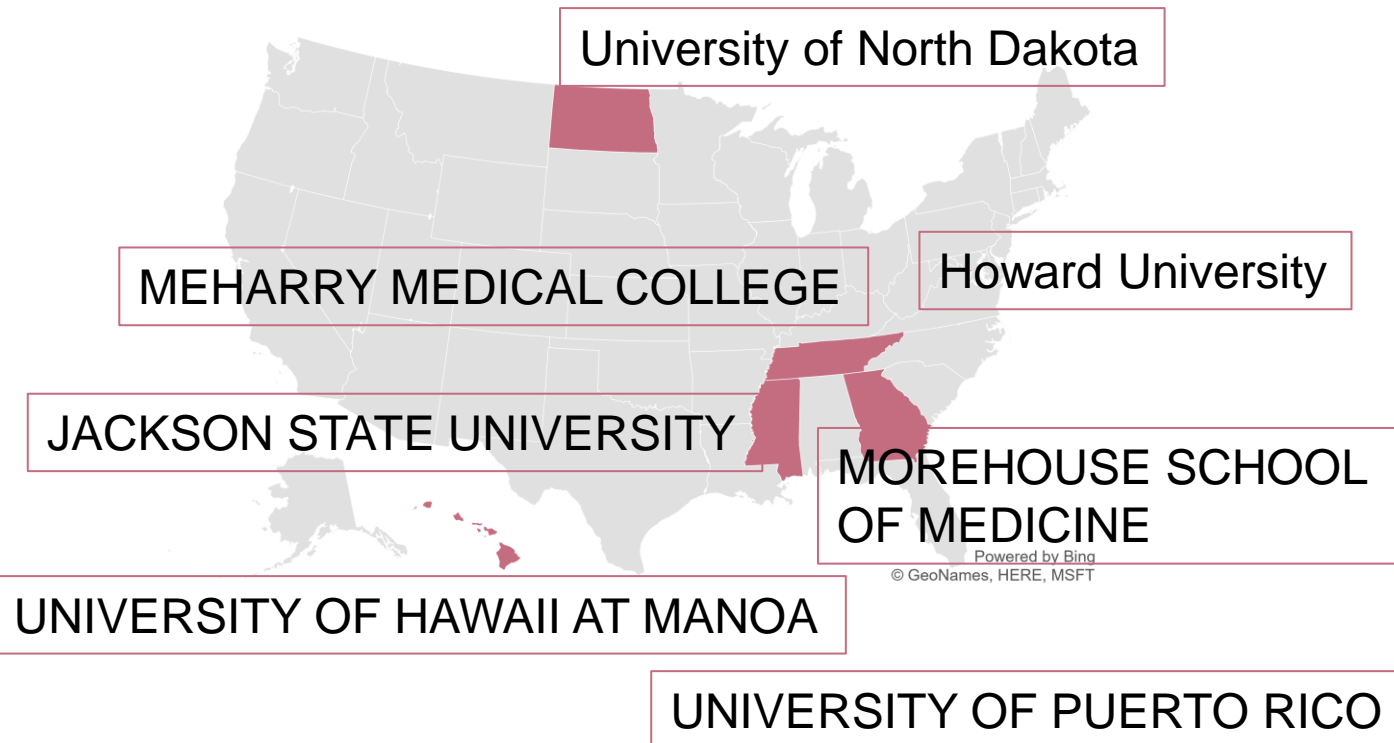
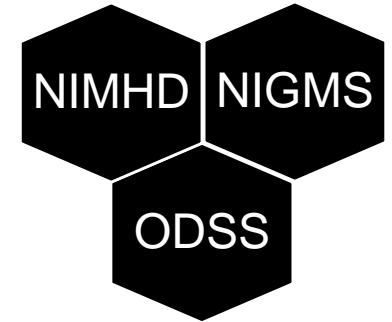
- **Collaboration between researchers from different disciplines conducting research with shared data science focus area.**
- **Learning communities to help participants to develop data science identity, equitable data practices, and sense of belonging to the data science community.**
- **Partnerships with academic and industry partners to provide opportunities for students and exposure to data science career pathways.**
- **Partnerships with other institutions or organizations that enhance data science knowledge and skills of the researchers and students in the institution.**



# **Part 4: Enhancing Institutional Data Science Capacity: ODSS Co-Funded Projects**

# Administrative Supplements to Enhance Data Science Capacity

Leverage existing NIH infrastructure-building programs to enhance data science capacity in institutions serving medically underserved communities and underrepresented students as described in the Notice of NIH's Interest in Diversity (NOT-OD-20-031).



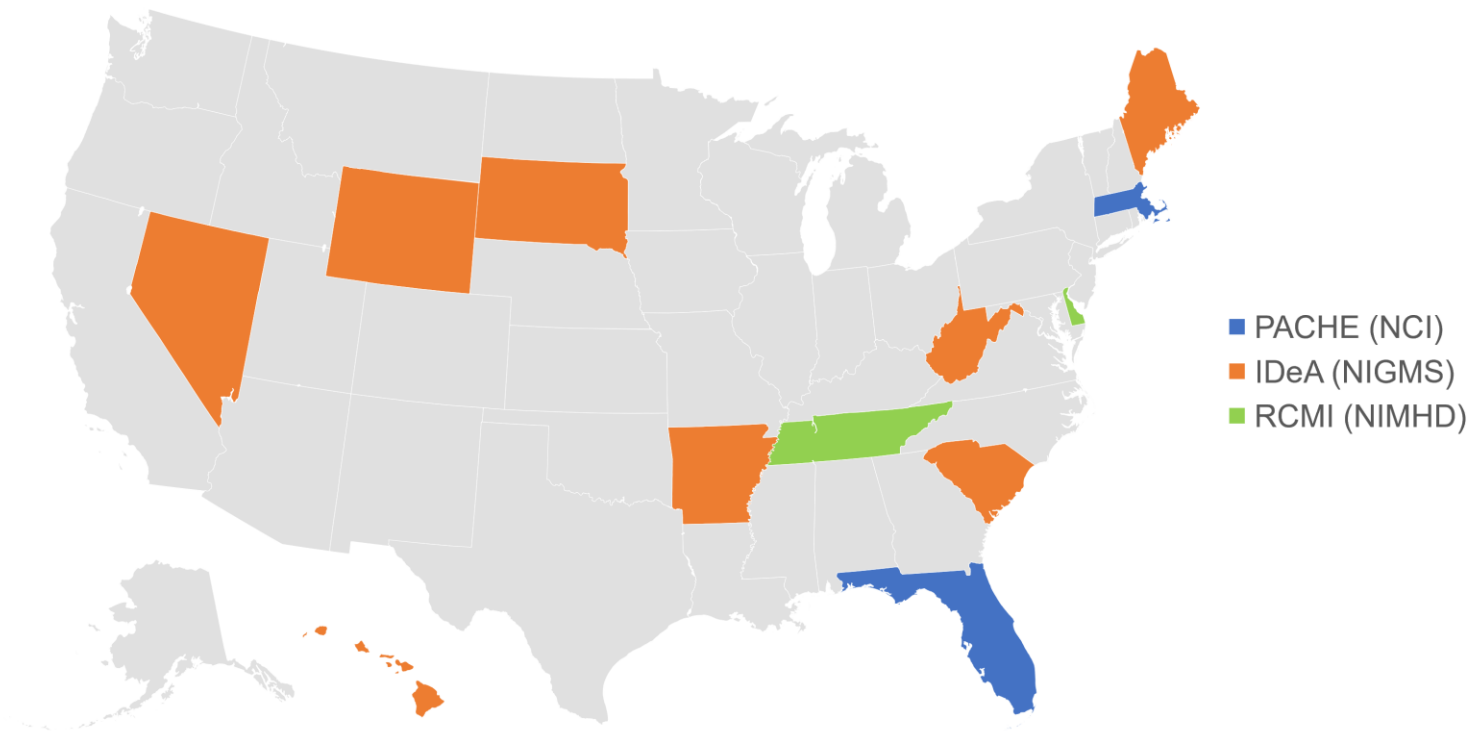
- Funded in FY22 and FY 23
- Implemented activities include:
  - Development and improvement of courses and curriculum
  - Providing training events
  - Building collaborations
  - Evaluation of impact of the activities and effectiveness of the overall projects
- Participants include undergraduate and graduate students as well as faculty investigators and community leaders

# NOT-OD-23-123 FY 2023 Awards

FY 2023: 19 Applications; 12 Awards

- **Partnerships to Advance Cancer Health Equity (PACHE, NCI)**
  - Florida A&M University, FL
  - University of Massachusetts Boston, MA
- **Institutional Development Award (IDeA, NIGMS)**
  - University of Arkansas for Medical Sciences, AR
  - University of Nevada Reno, NV
  - University of South Dakota, SD
  - University of South Carolina at Columbia, SC
  - University of Hawaii at Manoa, HI
  - University of Wyoming, WY
  - MaineHealth, ME
  - West Virginia University, WV
- **Research Centers in Minority Institutions (RCMI, NIMHD)**
  - Delaware State University, DE
  - Meharry Medical College, TN

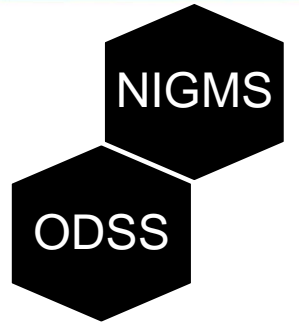
Locations of Awards for Administrative Supplements to Enhance Institutional Data Science Capacity (NOT-OD-23-123)



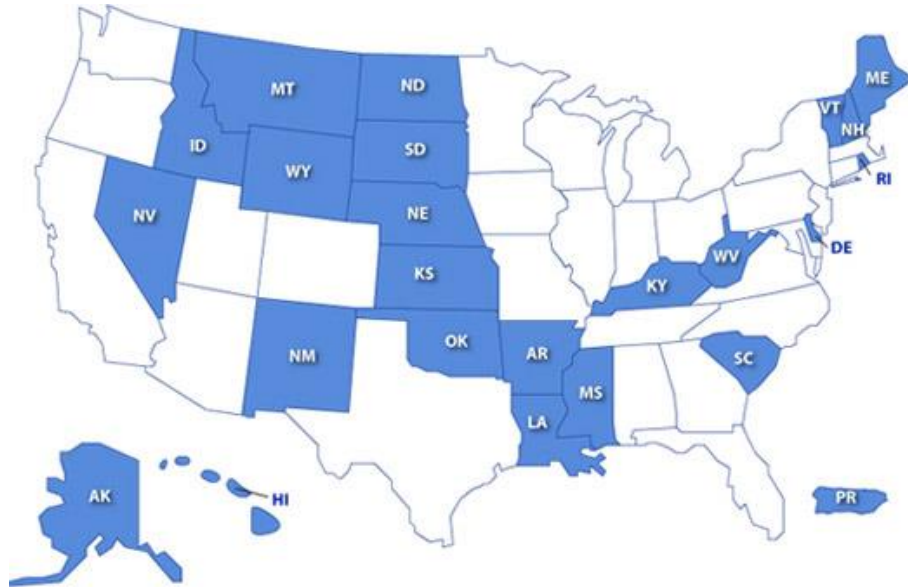
Powered by Bing  
© GeoNames, Microsoft, TomTom

# Administrative Supplements to Build Cloud-Based Learning Modules

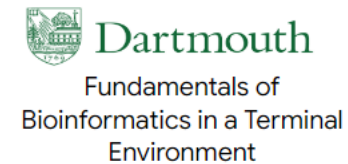
Provide funds to supplement NIGMS' Institutional Development Award (IDeA) Networks of Biomedical Research Excellence (INBRE) awards



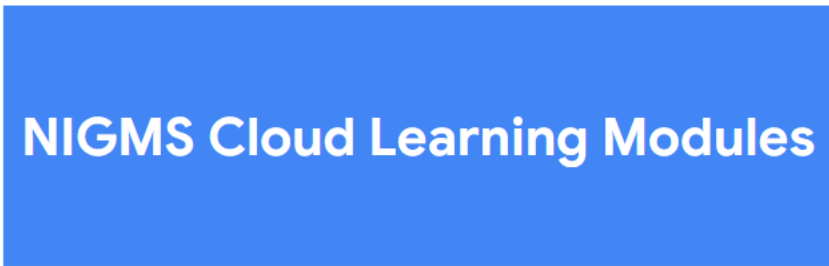
The NIGMS Sandbox is a collection of 12 biomedical research learning modules, built by NIGMS grantees, that run in the cloud



IDeA-Eligible States: Low NIH Funding



DNA Methylation Sequencing Data Analysis



Transcriptome Assembly, Refinement, and Analysis



Biomarker Discovery



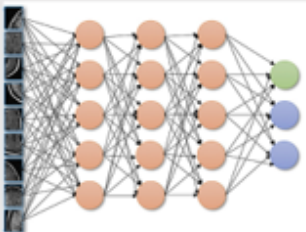
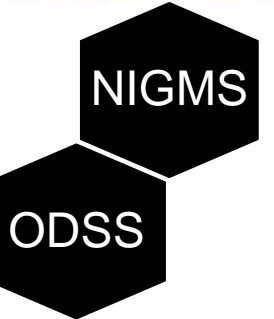
Consensus Pathway Analysis







# Administrative Supplements to Build Cloud-Based Learning Modules

## ODSS Co-funded Cloud-Based Learning Modules





### Biomedical Imaging Analysis using AI/ML Approaches

In this module you will learn how to generate a neural network, manipulate datasets, train a neural network on the dataset, apply the trained neural network to a new dataset, and quantify its performance.



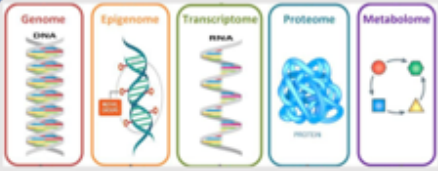

### Intro to Bioinformatics on the Command Line

In this module you will learn to use the Bash shell scripting language to work with common genomics file formats, create Conda environments, and troubleshoot command line errors.




### Introduction to Data Science for Biology

In this module you will learn how to create a simple decision tree using a structured dataset, evaluate model performance quantitatively, and understand why machine learning models require retraining from time to time.



### Integrating Multi-Omics Datasets

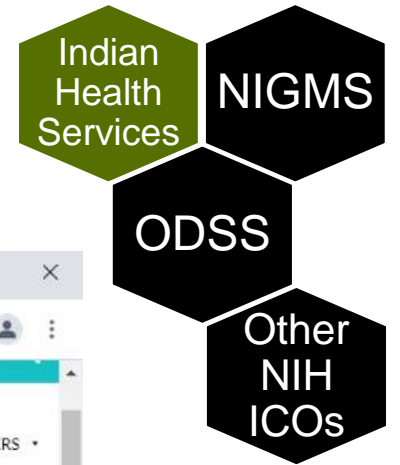
In this module you will learn how to analyze RNA-Seq, Epigenetics, and integrated multi-omics datasets in R with a Nextflow pipeline.



Course cards for Cloud-Based Learning Modules co-funded by NIH Office of Data Strategy (ODSS) and available at the NIGMS Sandbox (<https://github.com/NIGMS/NIGMS-Sandbox>)

# Co-funding Support for Native American Research Centers for Health

NARCH program supports opportunities for conducting research and career enrichment to meet health needs prioritized by American Indian/Alaska Native (AI/AN) tribes or tribally based organizations.



The screenshot shows a web browser window with the URL <https://www.npaihb.org/applied-biostatistics-and-data-science-for-american-indians-alaska-natives/>. The page header includes the NPaiHB logo and navigation links: NORTHWEST MEMBER TRIBES, EPICENTER, POLICY, PROGRAMS & PROJECTS, RESOURCES, EVENTS, CAREERS. The main heading is 'Applied Biostatistics and Data Science for American Indians & Alaska Natives'. The text describes a two-year training program for health professionals and students, including summer courses, seminars, and a capstone project. Contact information for Dr. Amy Laird is provided: [laird@ohsu.edu](mailto:laird@ohsu.edu). A photograph of several feathers is shown on the right side of the page. The footer contains the NPaiHB logo, address (2121 SW Broadway, Suite 300, Portland, OR, 97201), phone number (503-228-4185), and social media icons for Facebook, Twitter, YouTube, and Instagram.

<https://www.npaihb.org/applied-biostatistics-and-data-science-for-american-indians-alaska-natives/>

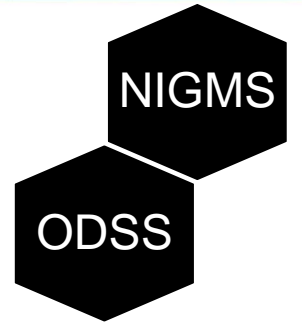


# Colorado Resource Center for Tribal Epidemiology Centers

ODSS Co-funded the Colorado Resource Center for Tribal Epidemiology Centers (TECs) to Provide Technical Assistance including Training on Data Science Topics



Apply | Visit | Give | Q



26  
FEB

### CIDA Short Course - Fundamentals of EHR Data 2024

A Biostatistics Short Course A six-week short course for researchers or students wanting to learn more about the foundations...

Monday, February 26, 2024 | 4:00 pm

Add to Calendar



27  
FEB

### CIDA Short Course - Fundamentals of Data Visualization

A Biostatistics Short Course A six-week short course for researchers or students wanting to learn more about foundations and...

Tuesday, February 27, 2024 | 5:00 pm

Add to Calendar



27  
FEB

### CIDA Short Course - Fundamentals of Statistical Literacy II

A Biostatistics Short Course A six-week short course offered in the spring for those wanting to follow-up to the companion...

Tuesday, February 27, 2024 | 6:00 pm

Add to Calendar



04  
MAR

### CIDA Short Course - Fundamentals of EHR Data 2024

A Biostatistics Short Course A six-week short course for researchers or students wanting to learn more about the foundations...

Monday, March 4, 2024 | 4:00 pm

Add to Calendar



# Educational Hub for Enhancing Diversity in Computational Genomics and Data Science

**Lead:** North Carolina A&T State University

- Workshops
- Course Integration Training
- Symposia
- Conferences
- Training
- Webinars
- Career Development

**Genomic Research and Data  
Science Center for Computation  
and Cloud-Computing**



**GRADS-4C**

Funded by NHGRI, NIMHD, All of Us Research Program and ODSS

<https://www.ncat.edu/news/2023/09/nih-grant-genomics-educational-hub.php>

<https://www.genome.gov/news/news-release/nih-awards-dollar-5-8m-to-create-genomic-data-science-educational-hub-for-early-career-researchers>

**Part 5: Enhancing Institutional  
Data Science Capacity:  
Resources, Funding Opportunities  
and Professional Opportunities**



Helping advance biomedical research by delivering access to industry-leading cloud providers.



The STRIDES Initiative aims to help NIH and its institutions accelerate biomedical research by reducing barriers in utilizing commercial cloud services. This initiative aims to harness the power of the cloud to accelerate biomedical discovery. NIH and NIH-funded researchers can take advantage of STRIDES benefits.

[Enroll Now](#)

### Gain access to

- Discounts on partner services
- Professional services consultations
- Access to training
- Potential collaborative engagements

>200  
Petabytes of  
Data

274M  
Compute  
Hours

>995  
NIH & NIH-funded  
Research  
Programs/  
Projects

\$41M  
Cost Savings

>4700  
People Trained

<https://datascience.nih.gov/strides>



ScHARe is a cloud-based platform for population science including social determinants of health (SDOH), and data sets designed to accelerate research in health disparities, health and healthcare delivery outcomes, and artificial intelligence (AI) bias mitigation strategies.

Get Involved    Acc

**ScHARe Think-a-Thons**



 **Think-a-Thon Schedule**

Think-a-Thons are held on the third Wednesday of each

Date	Time	Topic
January 17, 2024	2:00 – 4:30 p.m. ET	Preparing for a Data Sci Toward Goal 1

February 21, 2024	2:00 – 4:30 p.m. ET	<b>Data Science Projects 1 – Health Disparities and Individual SDOH</b> Toward Goal 2: Hands-on research collaboration session on individual SDOH: <ul style="list-style-type: none"> <li>• Explore the impact of individual SDOH on health outcomes</li> <li>• For researchers and students at all levels who want to collaborate on ScHARe to develop innovative and publishable research projects</li> </ul>	Registration coming
March 20, 2024	2:00 – 4:30 p.m. ET	<b>Data Science Projects 2 - Health Disparities and Structural SDOH</b> Toward Goal 2: Hands-on research collaboration session on structural SDOH: <ul style="list-style-type: none"> <li>• Assess the impact of structural SDOH on health outcomes</li> <li>• For researchers and students at all levels who want to collaborate on ScHARe to develop innovative and publishable research projects</li> </ul>	Registration coming
April 17, 2024	2:00 – 4:30 p.m. ET	<b>Data Science Projects 3 – Health Outcomes</b> Toward Goal 2: Hands-on research collaboration session on health care delivery and health outcomes: <ul style="list-style-type: none"> <li>• Investigate the influence of non-clinical factors on disparities in health care delivery and health outcomes</li> <li>• For researchers and students at all levels who want to collaborate on ScHARe to develop innovative and publishable research projects</li> </ul>	Registration coming

Individuals who need reasonable accommodation to participate should contact the ScHARe team at least five business days before the event.

**Past Think-a-Thons**

Date	Duration	Topic	Materials
November 15, 2023	2.5 hours	View video: <b>Preparing for AI 2: An Introduction to FAIR Data and AI-ready Datasets</b> Toward Goal 1: How to prepare an AI-ready dataset using gold standard data management principles, including: <ul style="list-style-type: none"> <li>• Making datasets findable, accessible, interoperable, and reusable (FAIR)</li> </ul>	View slides (PDF, 4 MB)

# Data Sharing and Reuse Monthly Seminar Series

The National Institutes of Health (NIH) Office of Data Science Strategy hosts a seminar series to highlight exemplars of data sharing and reuse on the second Friday of each month at noon ET. The seminar is open to the public and registration is required each month.



**Data Sharing and Reuse Seminar Series**

Learn more and register @ [bit.ly/NIHDataSeminars](https://bit.ly/NIHDataSeminars) | **#NIHData**

<https://datascience.nih.gov/nih-data-sharing-and-reuse-seminar-series>

# Enhancing Institutional Data Science Capacity

## Funding Opportunities

- Administrative Supplements to NIGMS Funded Awards for Building Cloud-Based Learning Modules
  - NOT-GM-24-006 **Application Due Date: February 15, 2024**
- Administrative Supplements to Enhance Institutional Data Science Capacity
  - NOT-OD-23-123 **Application Due Date: April 01, 2024**
- Broadening Opportunities for Computational Genomics and Data Science Education (UE5 Clinical Trial Not Allowed)
  - RFA-HG-23-002, **Application Due Date: June 10, 2024**



# Work at NIH: DATA Scholar Program

## Data and Technology Advancement National Service Scholar Program



- Recruit talents with **advanced data science expertise** to the NIH to use transformative approaches that lead to increased efficiency, innovative research, tool development and analytics.
- One to two years commitment.
- In addition to their own project, scholars participate in workgroups and collaborations, and contribute in many ways to the NIH
- <https://datascience.nih.gov/data-scholars-2023>

**Open for Applications: Early 2024**

**Contact:**

Dr. Bryan Kim

ds-workforce@nih.gov





# New Pilot: NIH DataPath Program

*Goal: To recruit and develop **early career** data science talents from diverse backgrounds to contribute their expertise to NIH operations while gaining valuable experience within the NIH environment.*

- Collaborate with the **U.S. Digital Corps Program**.
- Recruit **postbacs** and **post Masters** data talents to work at the NIH for 2 years.
- The program will:
  - Provide immediate data science capabilities for the NIH.
  - Foster a pipeline of skilled professionals who can contribute to the future of data-driven research and operations at NIH.



UNITED STATES  
**DIGITAL  
CORPS**

## Five Fellow Tracks:

- Cybersecurity.
- Data Science and Analytics.
- Design.
- Product Management.
- Software Engineering.

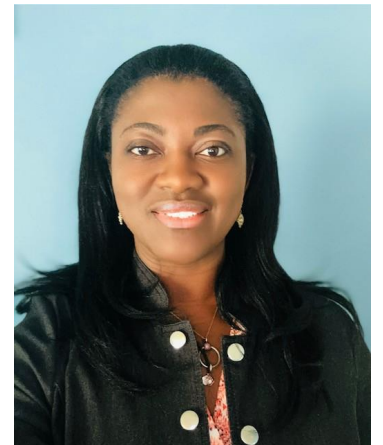
## Open for Applications:

Fall 2024

## Contact:

Evelyn Botchway

[ds-workforce@nih.gov](mailto:ds-workforce@nih.gov)



# Graduate Data Science Summer Program (GDSSP)



National Institutes of Health



- In collaboration with the NIH Office of Intramural Training and Education (OITE).
- Summer internship in the Intramural Research Program (IRP) at NIH.
- Must be current Master's student at the time of start of program.
- Must be able to commit 10 weeks in summer full-time.
- More information: <https://www.training.nih.gov/research-training/grads/summer-internship-program-sip/gdssp/>

# Enhancing Institutional Data Science Capacity

## Summary of Talk

- NIH supports the generation and analysis of Big Data for Health.
- Emerging Research Institution Designation.
- NIH Strategic Plan for Data Science 2023-2028 Request for Information - **Response Date: March 15, 2024.**
- Human Capital, Infrastructure and Partnerships are programmatic objectives for enhancing institutional data science capacity.
- Activities promote institutional awareness on data ethics and cybersecurity risk management are strongly encouraged.
- Funded projects to enhance institutional data science provide exemplars of activities that could be adapted.
- Resources and Upcoming Funding Opportunities.
- Data Science Workforce Initiatives at NIH: DATA Scholar, DataPath, and Summer Internship.

# ODSS Training, Workforce Initiatives, and Community Engagement (TWICE) Team and Contact



**Dr. Alison Lin**  
Lead



**Dr. Raphael Isokpehi**  
Program Director



**Dr. Bryan Kim**  
Program Director



**Evelyn Botchway**  
Program Analyst



**Nicholas Andrade**  
Training Specialist

- **Contact:** [ds-workforce@nih.gov](mailto:ds-workforce@nih.gov) ←
- **ODSS webpage:** <https://datascience.nih.gov/>
- **Data science funding opportunities:** <https://datascience.nih.gov/research-funding>
- **Repositories for sharing scientific data:** <https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/repositories-for-sharing-scientific-data>
- **Data science job opportunities at NIH (federal):** <https://datascience.nih.gov/jobs>