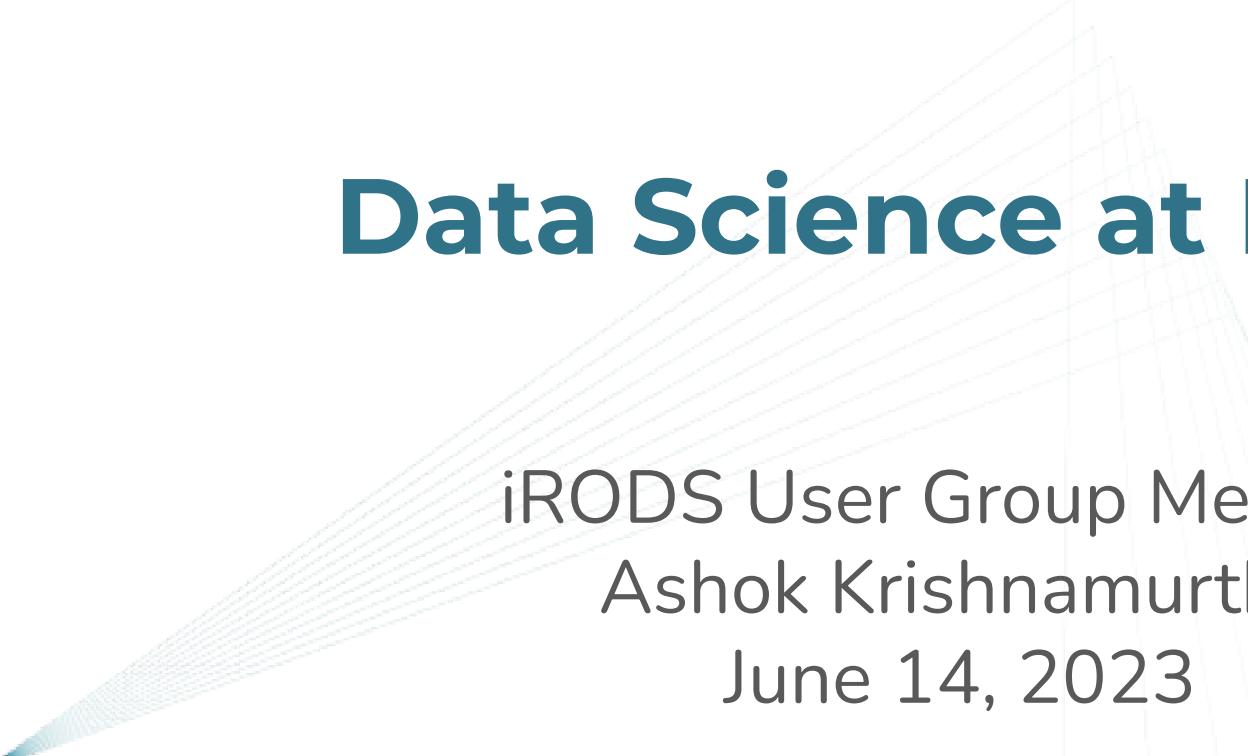


Data Science at RENCI



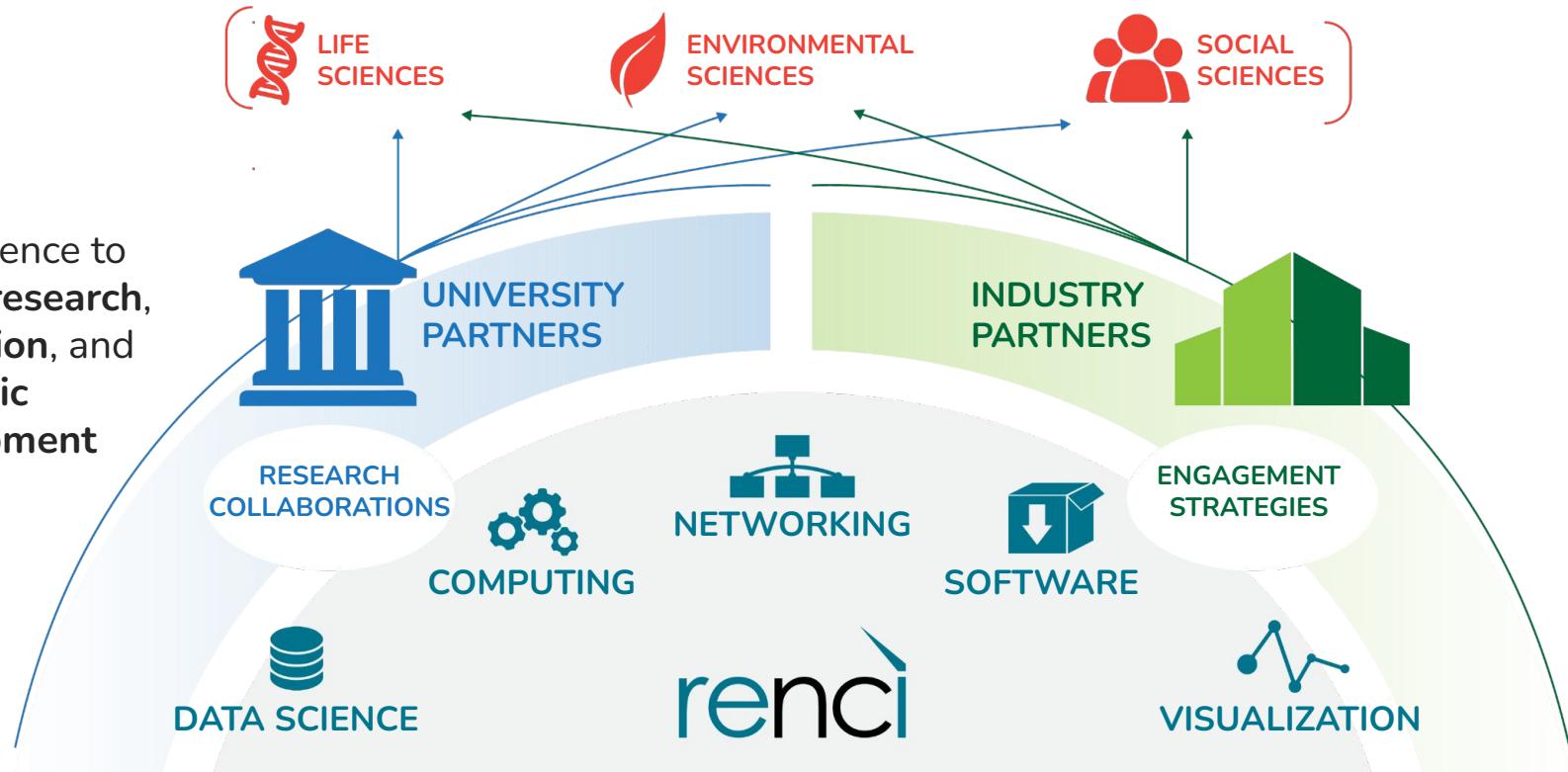
iRODS User Group Meeting

Ashok Krishnamurthy

June 14, 2023

RENCI at a Glance

Data science to enable research, innovation, and economic development



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

Duke
UNIVERSITY

NC STATE UNIVERSITY

RENCI at a Glance

RENCI develops and deploys **data science cyberinfrastructure** that helps researchers in academia, government, and business use data to drive discoveries, innovate, make informed decisions, and spur economic development.



Interdisciplinary
Collaboration



Community
Building

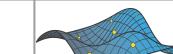


Innovative
Technology

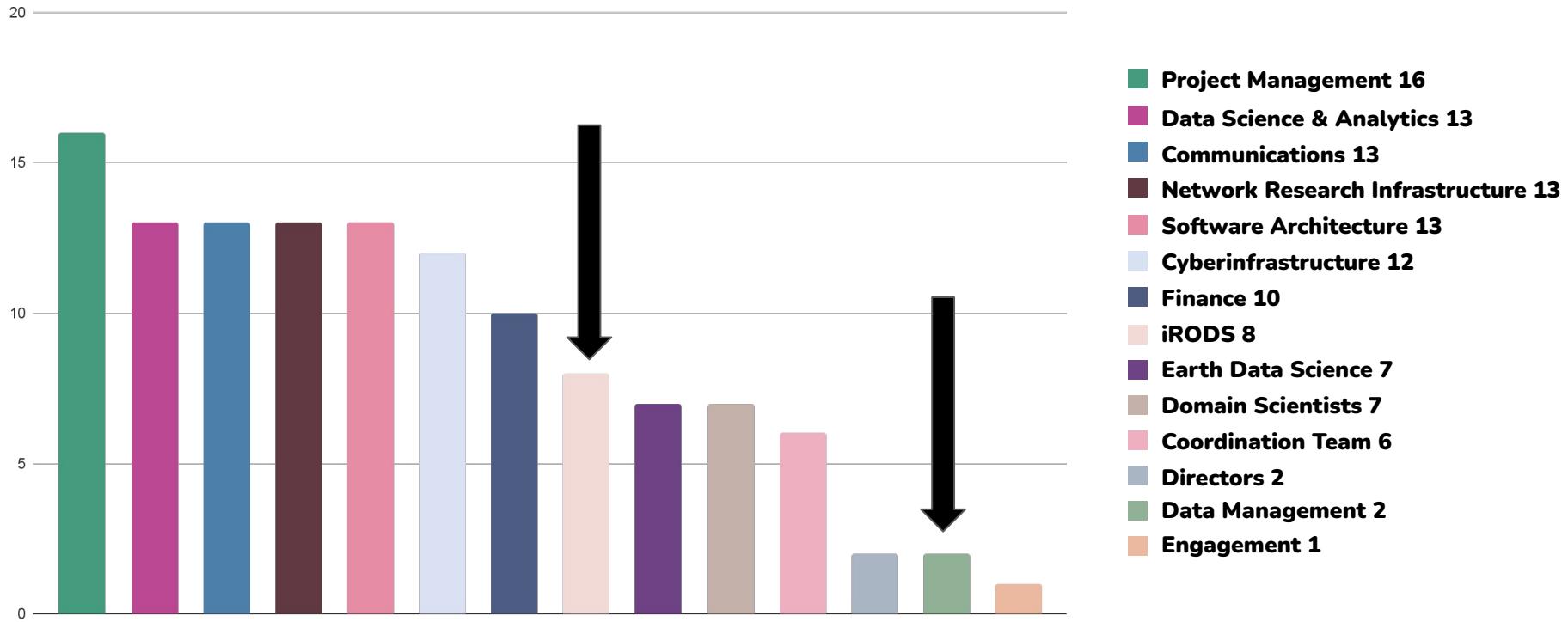


People-
centric

View our [strategic plan in full](#) to learn more about our foundational values.

Capabilities	Description	RENCI Projects
Clinical Data Linkage and Machine Learning	Integration of biomedical data and knowledge, including the open integration of knowledge from clinical sources. Creation of large knowledge graphs, and the use of algorithms applied to those graphs to infer new information relevant to researchers, such as mechanistic understanding of statistical associations.	 <small>*Not an officially approved NCATS logo</small> 
Open-source Data Management Software	Open-source data management software virtualizes data storage resources, so users can take control of their data, regardless of where and on what device the data is stored. RENCI's iRODS is used by research, commercial, and governmental organizations worldwide and released as a production-level distribution aimed at deployment in mission-critical environments.	 
Collaborative Cloud Platforms	Collaborative cloud-based platforms provide tools, applications, and workflows to enable researchers to find, access, share, store, cross-link, and compute on large scale data sets in secure workspaces to drive discovery and scientific advancement, leading to novel diagnostic tools, therapeutic options, and prevention strategies.	 
Earth Data Sciences	Hazard and risk assessment for coastal environments. Hurricane impacts and climate. Real-time prediction of storm surge, river flooding, and wave.	 
National/International Research Cyberinfrastructure	Design, development, deployment, operation and evaluation of federated multi-user national-scale research cyber-infrastructure facilities and testbeds, which combine advanced cloud, networking, data and other technologies.	   
Industry and Government Engagement	Identify and cultivate meaningful partnerships between academic, industry, and government institutions to accelerate data science.	 

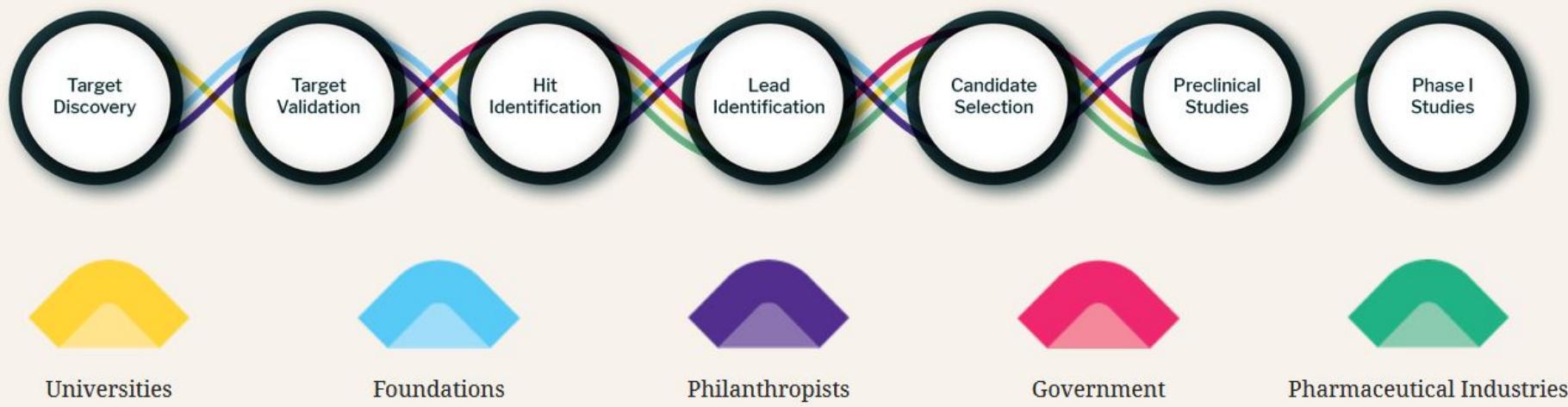
RENCI staff by division



Selected RENCI projects using iRODS

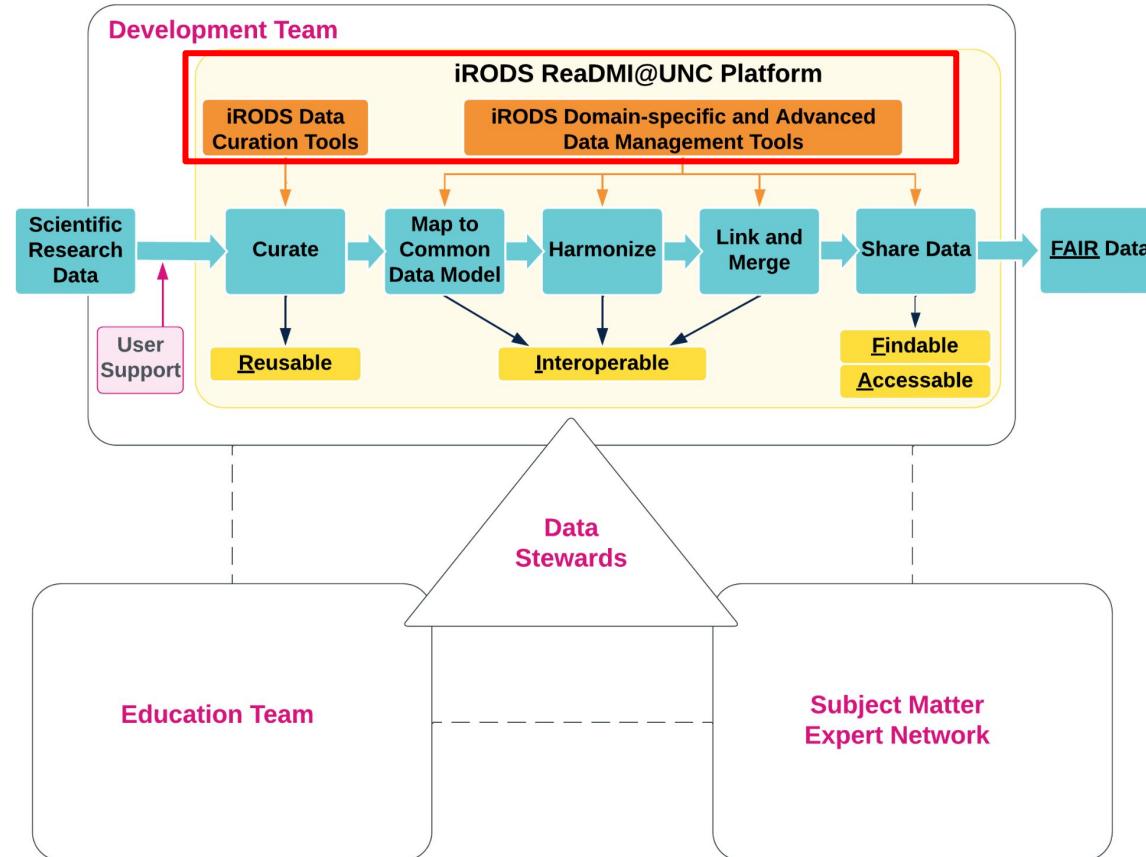
READDI Anti-viral Drug Discovery Center (READDI-AC)

Antiviral Innovation Engine



iRODS is used for data management in READDI-AC

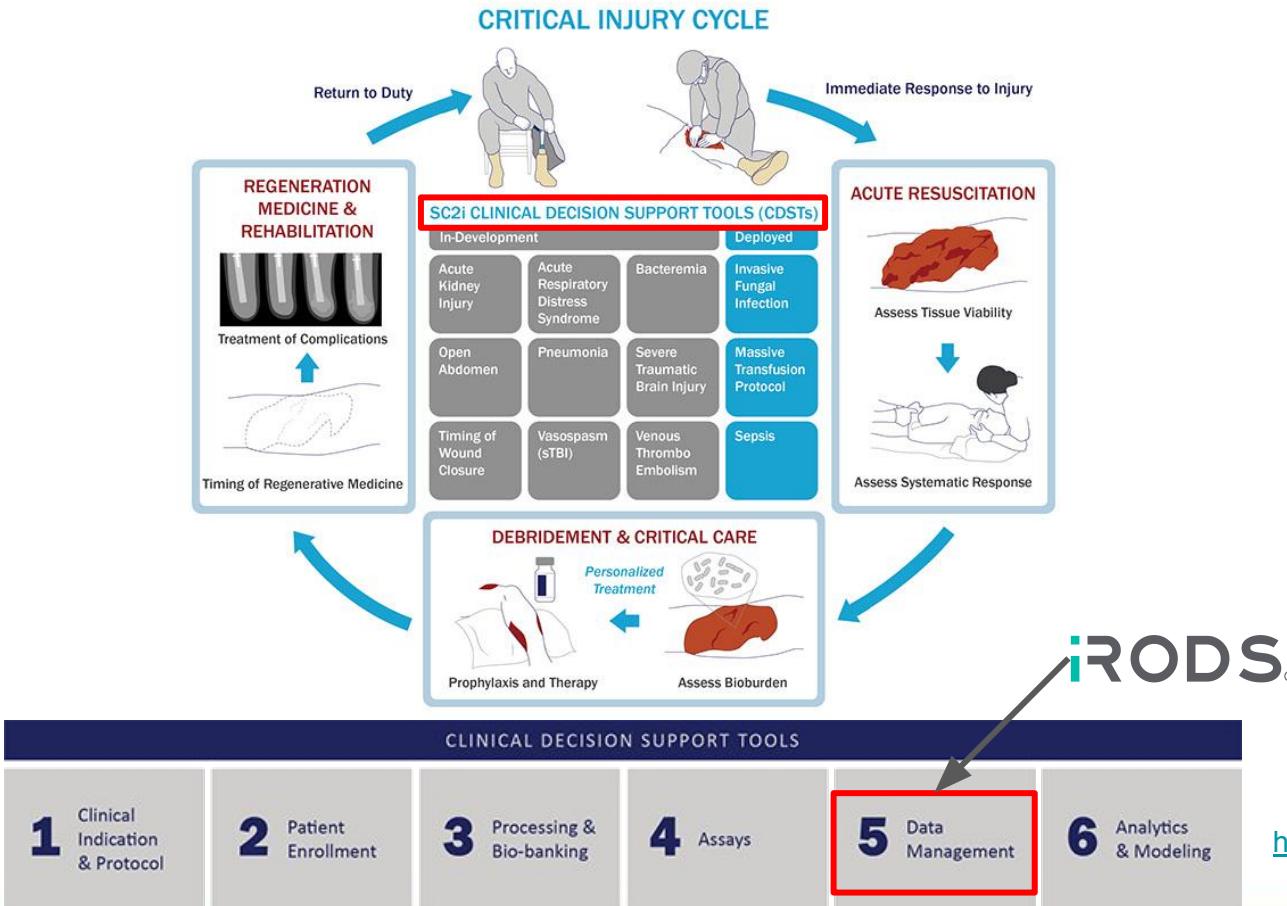
UNC Research Data Management Initiative (ReaDMI)



Partners:

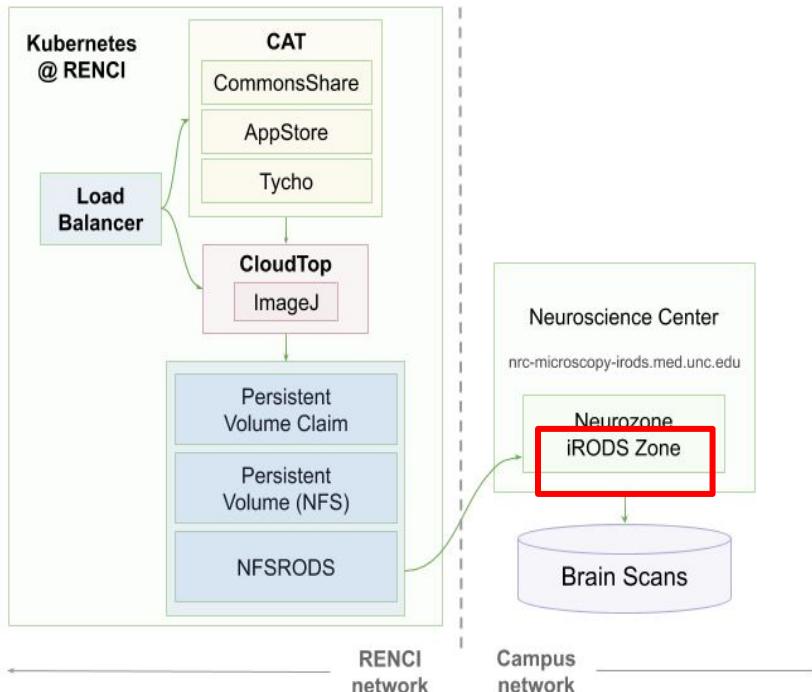
- SDSS
- RENCI
- Odum Institute
- NC A&T

Surgical Critical Care Initiative (SC2i)



iRODS as a Catalyst for Collaboration

BRAIN-I



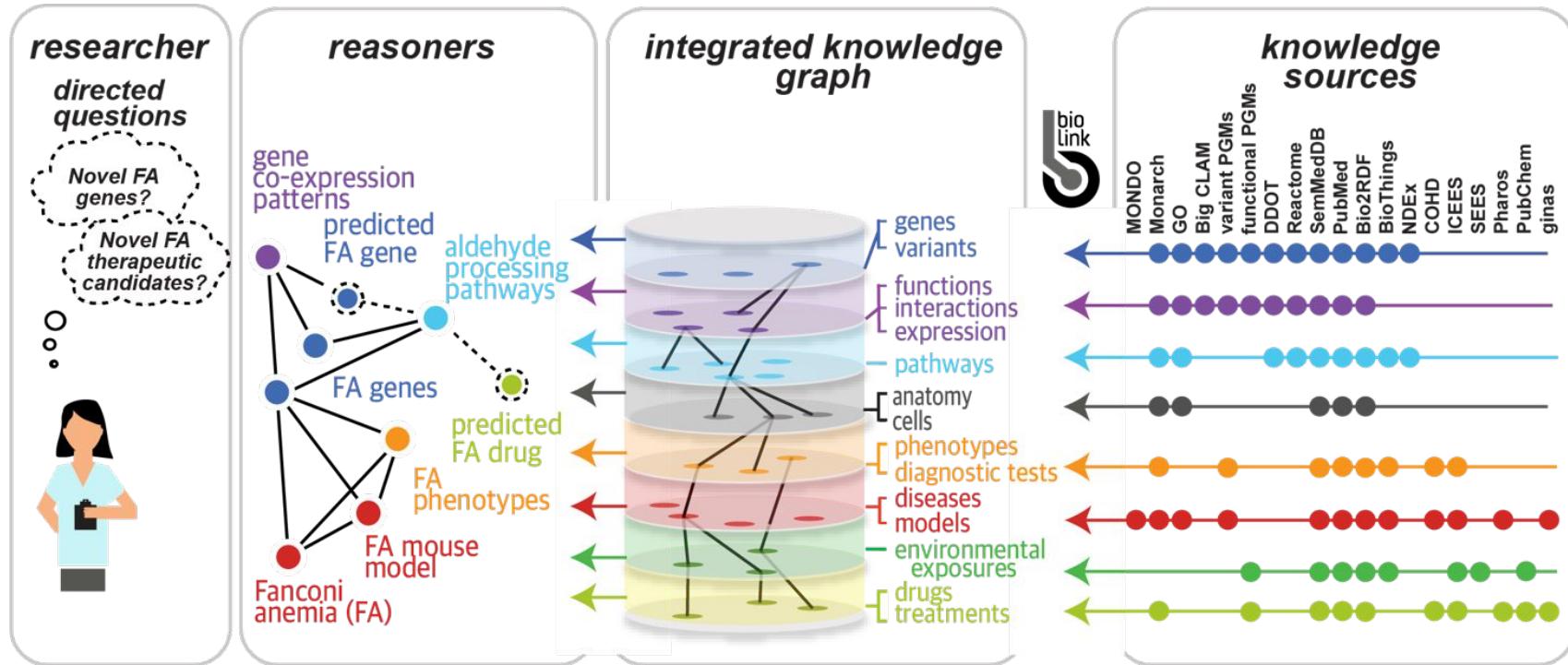
OMERO

An enterprise image data management platform that enables access, processing, sharing, analysis and where appropriate publication of scientific image data and metadata.

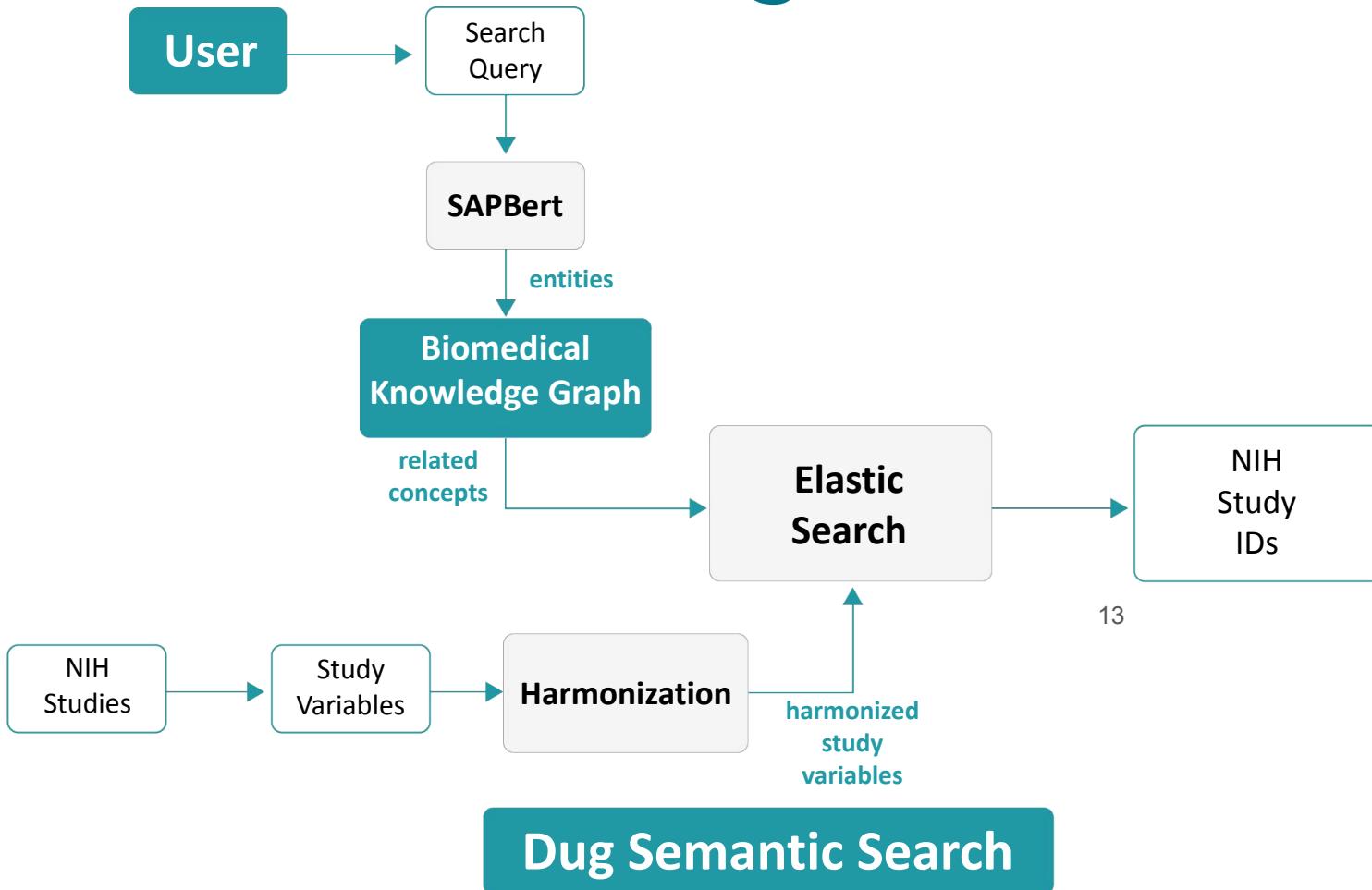
BRAIN-I has led to an iRODS Consortium project for imaging

A few other RENCI projects

Biomedical Data Translator: Data Integration



Dug



WHO?

WHAT?

WHERE?

SCIENCE!

WHY?



Genomics

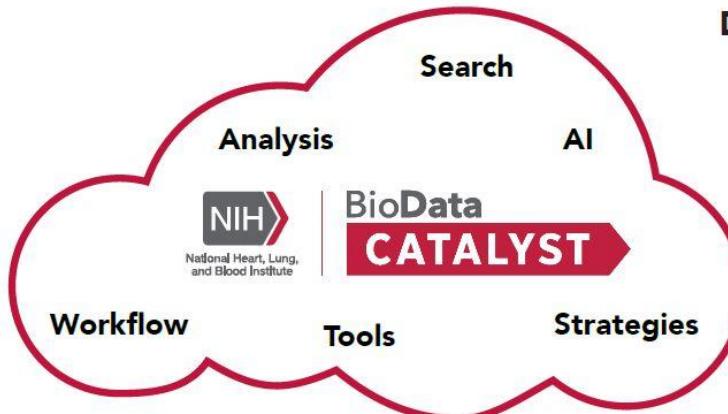


Clinical



Imagery

**DATA
HARMONIZATION**



- UNDERSTAND
- OPEN SCIENCE
- CROSS-LINK
- COLLABORATE
- SCALE
- SHARE
- INTEROPERATE

HOW?

Diagnostic Tools
Therapeutic Options



DISCOVERY

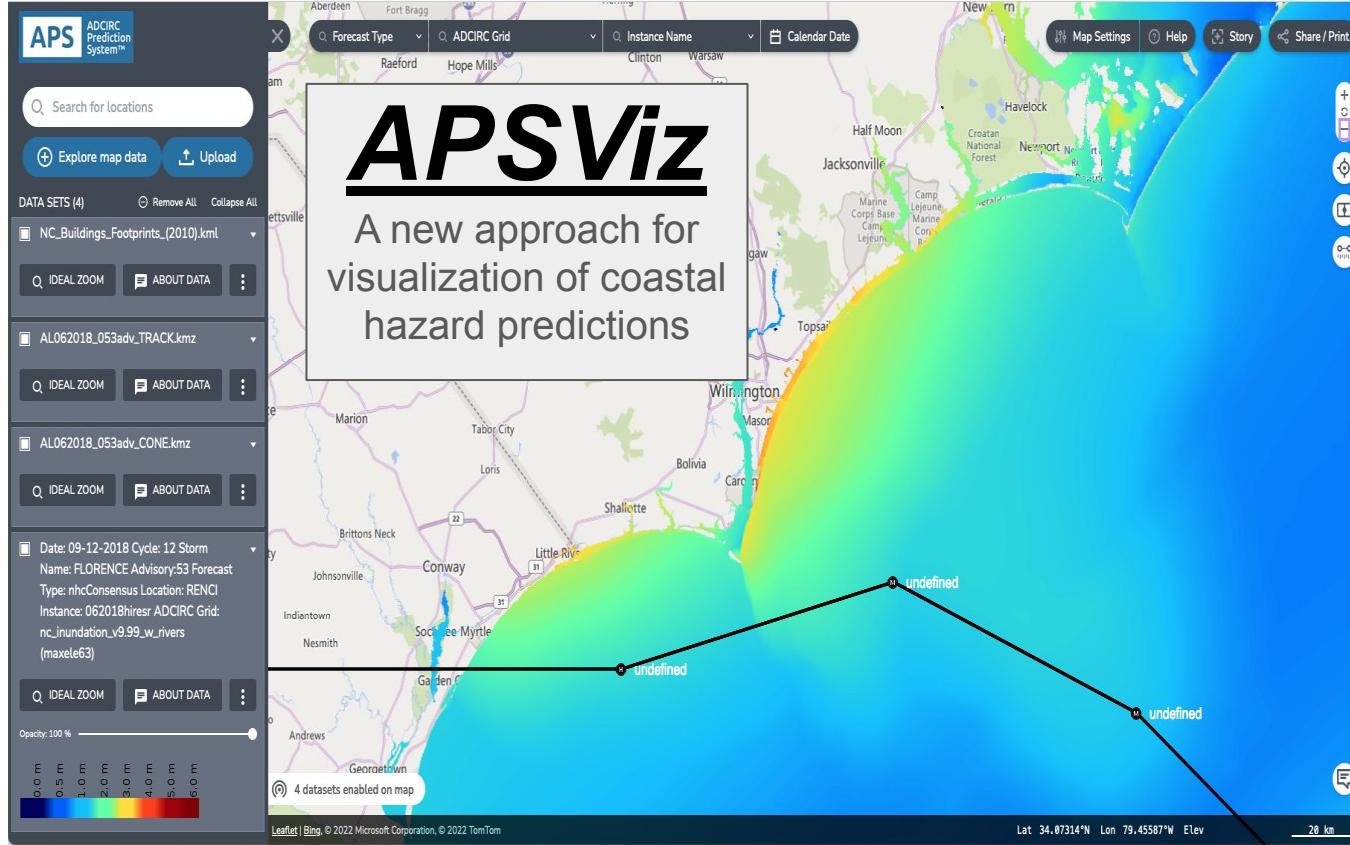


Prevention
Strategies



PATIENTS!

APSViz for Storm Surge Modeling





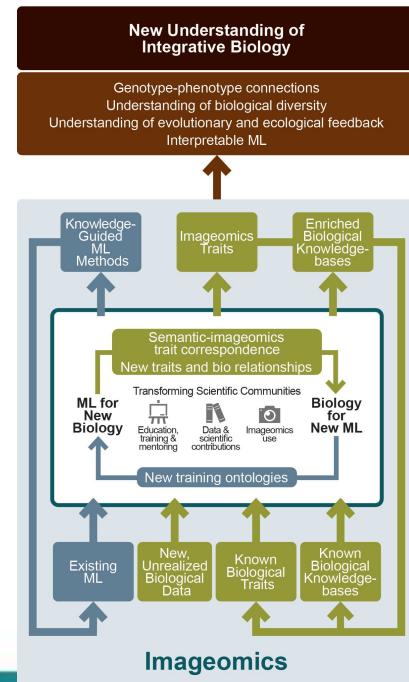
HDR Institute: Imageomics: A New Frontier of Biological Information Powered by Knowledge-Guided Machine Learning

Tanya Berger-Wolf (OSU), Henry Bart (Tulane), Anuj Karpatne (VT), Hilmar Lapp (Duke), Charles Stewart (RPI), Yasin Bakis (Tulane), James Balhoff (UNC), Bryan Carstens (OSU), Wei-Lun Chao (OSU), Wasila Dahdul (UCI), Nico Franz (ASU), Leanna House (VT), Paula Mabee (NEON), Murat Maga (UW), Dan Rubenstein (Princeton), Yu Su (OSU), Josef Uyeda (VT)

Imageomics : from images to biological traits through the power of biology-guided machine learning

Vision:

A new field of **Imageomics**:
Collaborative research, training,
and community-facing
environment for **extracting**
existing and new biological traits
from images of organisms, with
the necessary cyber-, information,
and model development
infrastructure.

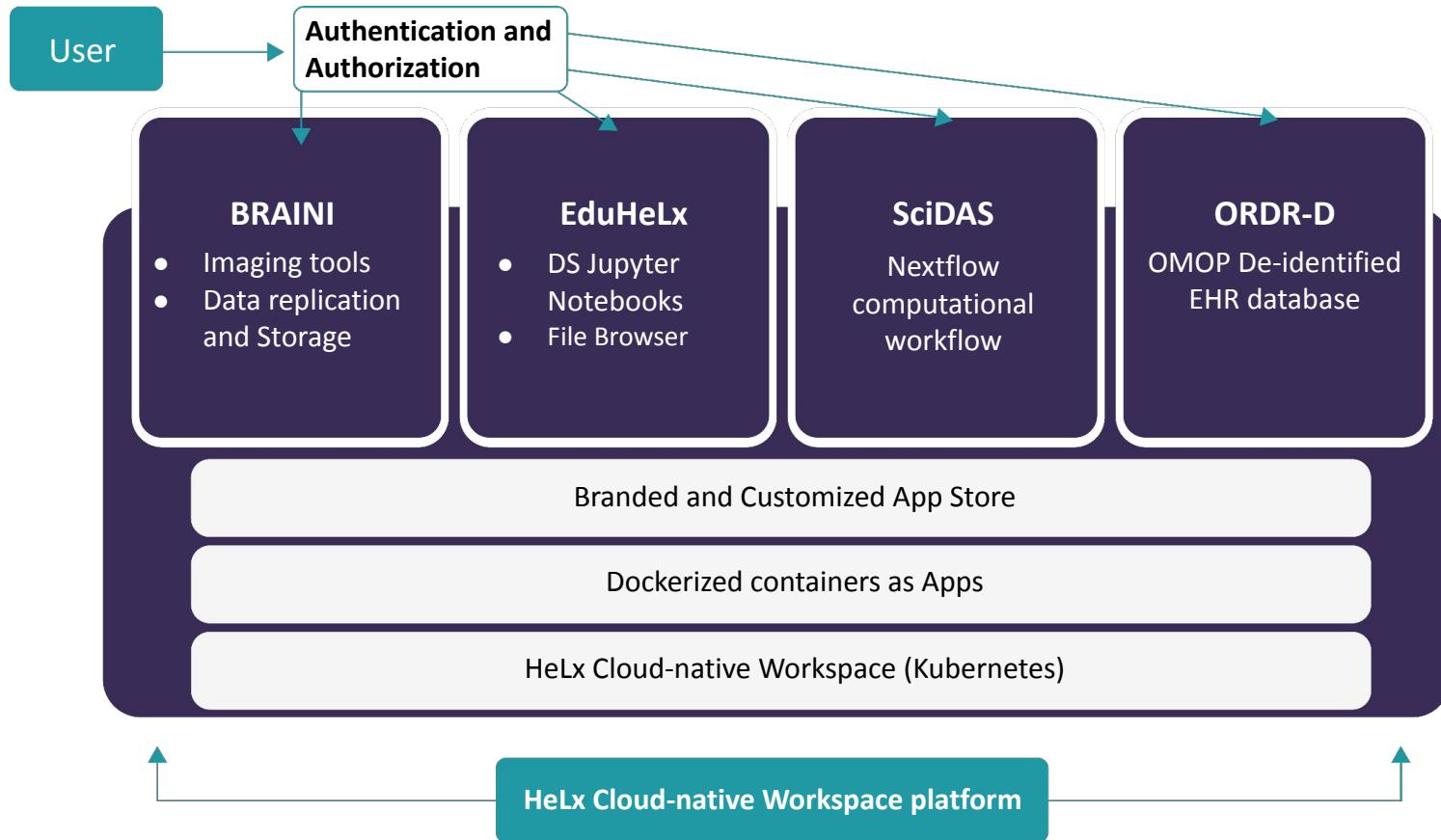


Images are the most abundant source for documenting life on the planet, and yet traits of organisms cannot be readily extracted from them. The analysis of traits, the integrated products of genes and environment, is critical for biologists to predict effects of environmental change or genetic manipulation and to understand the significance of patterns in the two billion year evolutionary history of life.

Contributions:

- Digital resources for the scientific community and practitioners in biology, data science, and ML
- Advancements in **Imageomics**-enabled biology
- Innovations in machine learning
- Advancements in interdisciplinary training and education
- Engagement of the broader public in the scientific process by leveraging images as the source of data to democratize science

HeLx



NSF FABRIC



30 sites across the U.S.
FAB: 4 continents

Open launch
planned for
October 2023

Source: <https://whatisfabric.net/>

RENCI Pilot STAR Program

Spring 2023 - Student Advancement at RENCI (STAR) Program

Pilot program launched Spring 2021

Students span from
High School to
Graduate

UNC and Non-UNC

Paid and non-paid
positions

Focus on diversity of
students and offered
projects

The majority of students are college
juniors/seniors and graduate students



The STARShip provides mutual benefits to
RENCI and students

Leverage sponsor funding, Work Study Program, and RENCI institutional commitments to fund student's work

The STAR Team provides:

- ★ Support for mentors and supervisors
- ★ Guidance on creating projects and mentoring the students
- ★ Support for students by creating a networking and engaging environment and outlining standardized requirements geared to educating students in presentation, science, and communication skills

The STAR Team is working on streamlining the process to launch a full robust program later this year

Thank you!

Questions?