

A composite image of space exploration elements. In the foreground, a rocket with a NASA logo is launching. In the middle ground, a satellite with solar panels and a lunar lander are visible. In the background, a large, dark moon and a smaller, reddish planet (Mars) are seen against a starry sky. An American flag is partially visible in the lower center.

# GeneLab

Daniel C. Berrios MD MPH PhD

Terri G. Thompson PhD

# Flight Research: International Space Station

- Almost as soon as the International Space Station was habitable, researchers began using it to study the impact of microgravity and other space effects on several aspects of our daily lives.
- This unique scientific platform continues to enable researchers from all over the world to put their talents to work on innovative experiments that could not be done anywhere else.



# GeneLab Strategic Plan

“... there is a significant demand from the scientific community for NASA-funded research opportunities on the ISS that cannot be met using traditional management tactics based on single Principal Investigator (PI)-led investigations. Given the factors that limit the scientific output required to address the significant biological problems required for human exploration beyond low-earth orbit, *NASA will develop and implement a new multi-investigator approach based on high content bioinformatics analytics and with open science and data.*”

-GeneLab Strategic Plan, 2014

The image is a promotional graphic for the GeneLab project. At the top, it shows the International Space Station (ISS) in orbit against a black background with a bright sun or star on the right. The NASA logo is in the top right corner, and the text "National Aeronautics and Space Administration" is in the top left. The word "GeneLab" is written in large, white, sans-serif font in the center. Below it, the tagline "Expanding the Impact of Biological Research in Space" is written in a smaller white font. At the bottom, there is a horizontal strip of four small images: a person in a lab coat using a pipette, a glowing blue molecular structure, a person looking through a microscope, and another glowing blue molecular structure. The website "www.nasa.gov" is printed in small white text at the bottom left of the graphic.

National Aeronautics and Space Administration

NASA

# GeneLab

Expanding the Impact of Biological Research in Space

www.nasa.gov

# GeneLab Goals

## The GeneLab Goals

1. “Develop an integrated repository and bioinformatics data system for analysis and modeling “
2. “Enable the discovery and validation of molecular networks that are influenced by space conditions through ground-based and flight research using next-generation omics technologies”
3. “Engage the broadest possible community of researchers, industry, and the general public to foster innovation”
4. “Strengthen international partnerships by leveraging existing capabilities and data sharing “

-GeneLab Strategic Plan, 2014

# Motivations

- **Maximize ROI for ISS Utilization**: Open-access, systems-biology spaceflight experiments will provide foundational science that maximizes return-on-investment for rare and costly spaceflight opportunities and remove research “bottleneck”
- **Create a PI Multiplier Effect**: Open access greatly expands the community of researchers using ISS derived data for investigations – ISS research investments will yield numerous follow-up investigations and next generation hypothesis-driven research
- **Leverage NASA and External Partner Strengths**: Brings together NASA’s strengths in Space Biology and “big data” analysis with commercial, government and international partners through a scaled and iterative approach that capitalizes on existing databases, analytical tools and biotech capabilities
- **Maximize Utilization of Cutting Edge Bioanalytical Tools and Techniques**: Multiple omics datasets and integrated data system allows scientists to interrogate ISS derived samples using state-of-the-art high throughput genomics, proteomics, metabolomics and bioinformatics tools
- **Speed the Pathway to Translation**: Allows researchers to discover emergent properties in data to identify and understand pathways/macromolecules influenced by space stressors
- Directly responsive to 2011 Decadal Survey Recommendations and OSTP Open Data Initiative.
- Similar goals and efforts NASA ARMD and SMD to increase ROI and PI Multiplier effect and engage broad communities. ARMD/AvS/DashLink, SMD/SS/PDS, SMD/ES/EOS-EOSDIS & NEExchange.

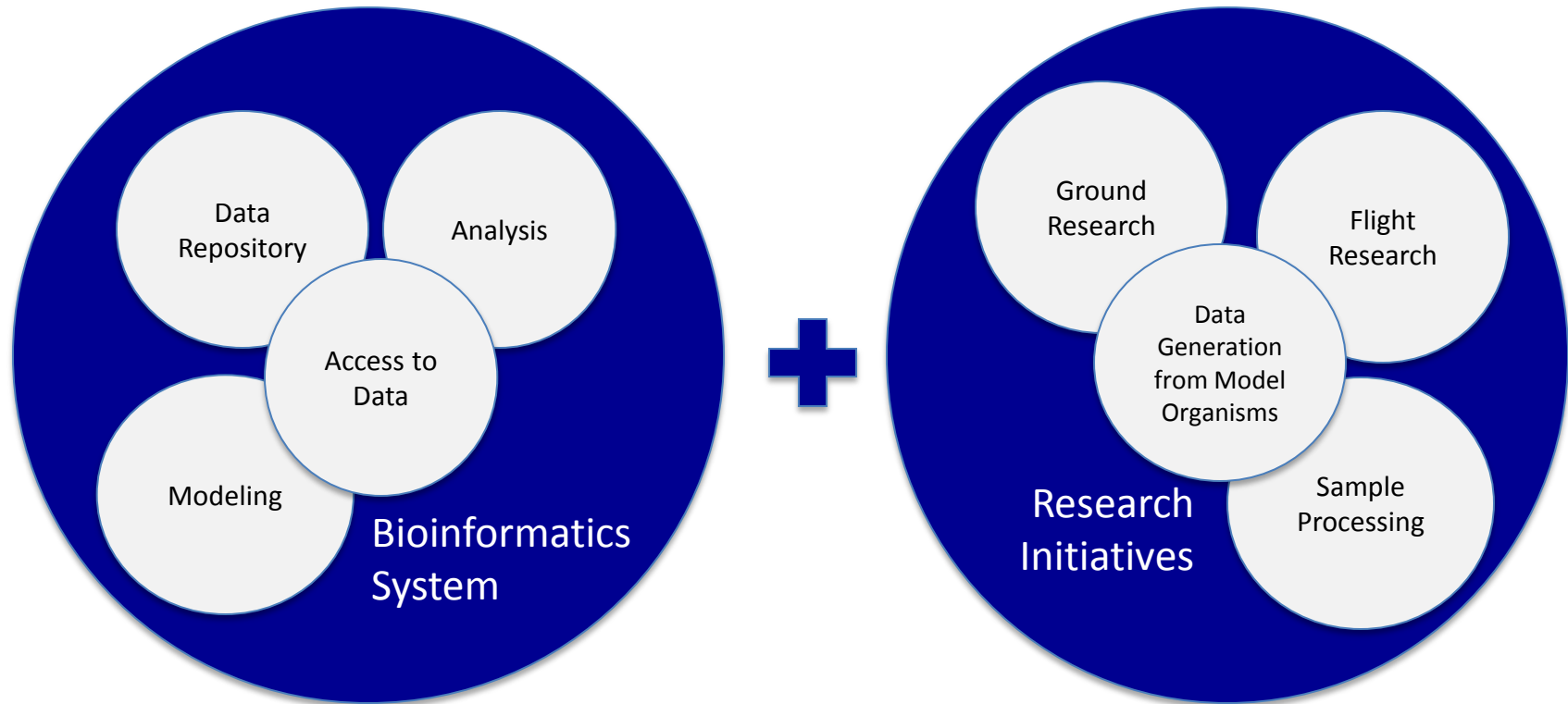
# ISS Research

- The unique conditions and workflows aboard the ISS are captured for a flight experiment and associated with the results.
- This information is ingested, stored, indexed and distributed from NASA's Life Science Data Archive.
- NASA GeneLab is expected to capture and distribute 'omics' data and experimental and process conditions most relevant to research community in their statistical and theoretical analysis of NASA's omics data.



NASA astronaut Barry "Butch" Wilmore setting up the Rodent Research-1 Hardware in the Microgravity Science Glovebox aboard the International Space Station

# Background: GeneLab Vision



**Vision: A centralized collaboration space for data deposition, retrieval, analysis and modeling to develop next generation science related to spaceflight**

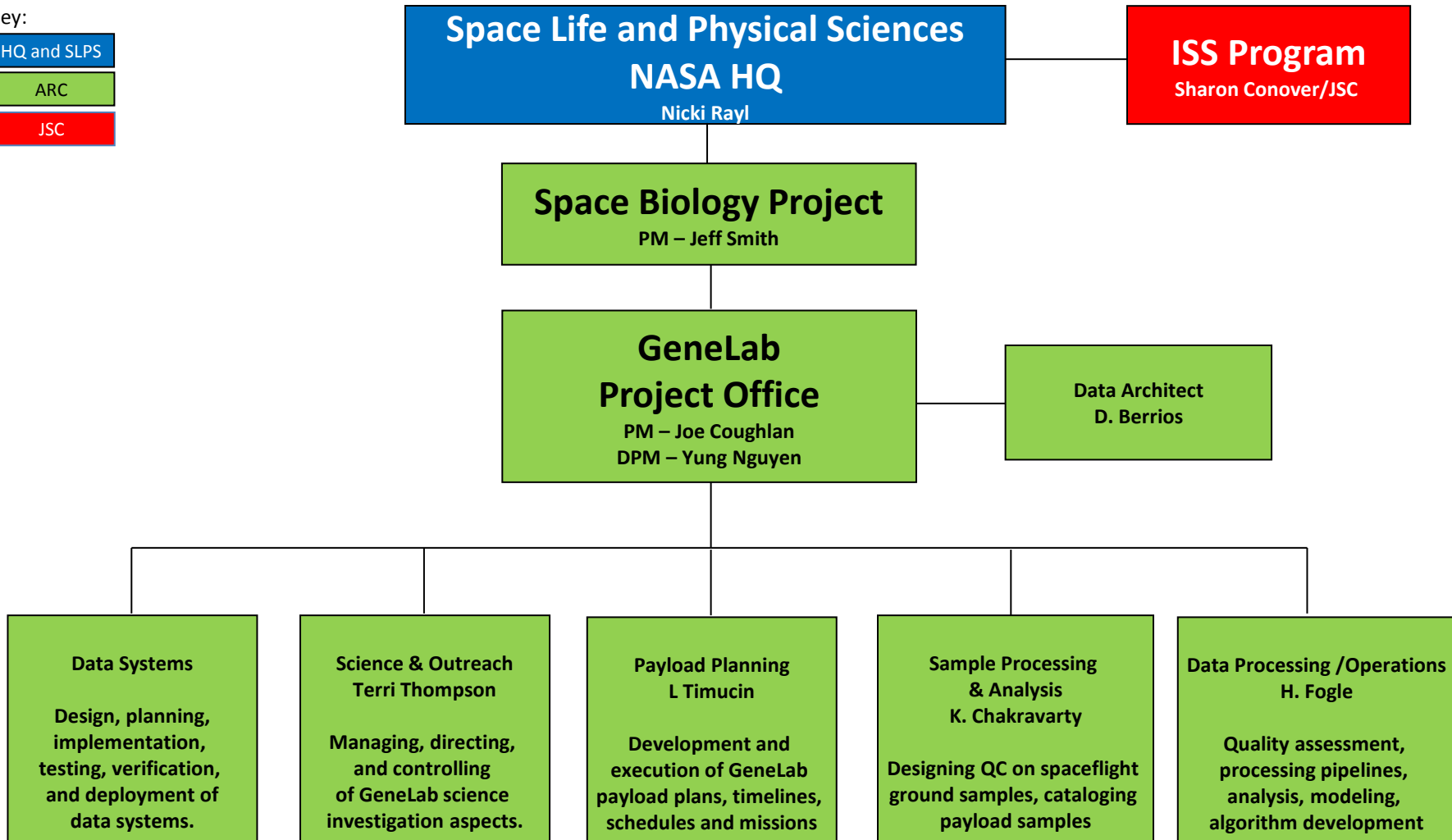
# GeneLab Organization

Key:

HQ, and SLPS

ARC

JSC





# GeneLab Users

- **Traditional Space Biology PI Community** – zoologists, botanists, microbiologists, cellular and molecular biologists
- **Non-Traditional Genomics and Systems Biology Communities**
- **NASA Human Research Program Omics Data** – GeneLab *potential* host for One Year Twins Study Data – Pilot Study for Human Omics Data. *Potential* repository for ground radiation studies using model organisms
- **CASIS** – Opportunity for data mining to identify targets for drug development, personalized medicine and systems biology of model organisms and humans yielding knowledge for terrestrial applications and acceleration of translation pathway
  - CASIS funded investigators encouraged to submit raw data to GeneLab
    - CASIS participated in 2013 GeneLab RFI
    - CASIS representation at September 2013 Omics Workshop
    - CASIS representation on continuing GeneLab Steering Committee and planning efforts

# Concept of Operations



## Experiment on ISS

Crew performs experimental protocol and harvests tissues.



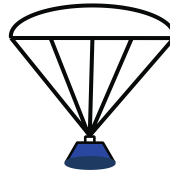
## Launch

Experiment is prepared and launched according to approved NRA.



## Return to Earth

Material sent back to earth for processing in investigators lab. Controls (ground and/or flight) processed at the same time.



## Process Samples

Extracted DNA, RNA and/or protein sent to validated omics center to generate sequence, transcript or protein expression data.

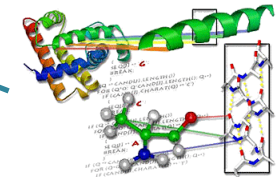


## Data Sharing

Data shared with larger scientific community. Results feedback to GeneLab and other databases accelerating scientific discovery by leveraging a bigger community.

## Data Collection & Hosting

Data returned to investigator or GeneLab for analysis. Raw data uploaded into GeneLab database for public viewing.



## Next Generation Research

Iterative research solicitations for experiments utilizing GeneLab data for ground validation and next generation flight research.



## Modeling and Validation

Wet lab validation and computational modeling.

# Mission Types

Mission	Type	Definition	Example
Dedicated	Reference Data	Mission is entirely dedicated to GeneLab objectives; the Science Definition Team (SDT) defines the experiment and requirements; SDT is selected through the NASA Research Announcement Process	Micro-16 (tissue TBD)
Collaborative	Sample Sharing	GeneLab obtains specimens/samples from the existing PI space flight and ground control experiment	Rodent Research (Mouse) Bioculture System Validation (Mouse cells)
	Augmentation	GeneLab provides supplemental funding to a PI experiment to increase the quantity and/or type of specimens to obtain dedicated sample; augmentation requires NASA SLPS experiment review approval process	BRIC-19 (plant), BRIC-20 (plant)  (BRIC=Biological Research in Canisters)

# Launch – Flight Experiment Selection

Flight experiment solicitation and selection follows the NASA SLPS peer review process.

- NASA HQ through SLPS Program
- Space Biology Project Office
- Follow Decadal Survey recommendations
- Science Definition Team Solicitation Plan

[http://www.nasa.gov/sites/default/files/files/NACRS\\_SLPSResearch\\_022414T.pdf](http://www.nasa.gov/sites/default/files/files/NACRS_SLPSResearch_022414T.pdf)

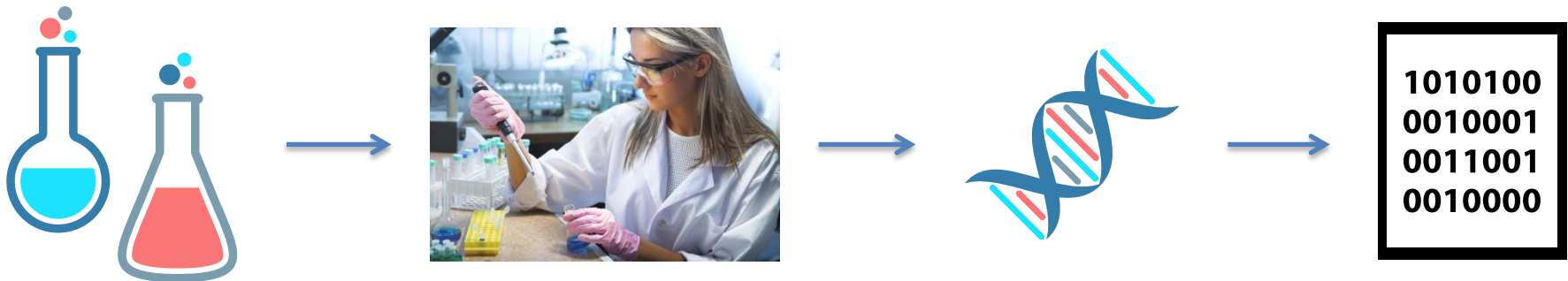
# Collaborative Datasets

Year	Payload		Mission Type
2015	BRIC-19	University of Wisconsin	Augmentation, plant
	BRIC-20	Ohio University	Augmentation, Plant
	RR-1*	NASA/CASIS	Sample Sharing, Rodent
	RR-2*	Loma Linda University; Florida State University	Sample Sharing, Rodent
	Bioculture* Validation	NASA	Sample Sharing, mouse cells
	Micro-9*		PI data only

\*Additional work proposed for 2015 current out of scope

# Process Samples

- Flight and/or ground samples received in Laboratory
- Researcher extracts DNA, RNA and/or Protein
- Extracted material sent to center for data generation
- Data is sent back to researcher from center



# Concept of Operations

## Overview of the flow of data through the system



### Experiment on ISS

- Crew performs experimental protocol and harvests tissues.

- **Data Collection & Hosting**
- Information sharing
- Reuse of data



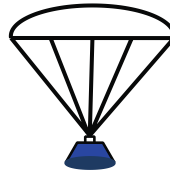
### Launch

Experiment is prepared and launched according to approved NRA.



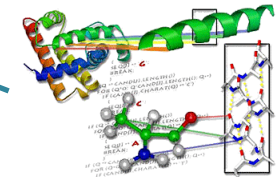
### Return to Earth

Material sent back to earth for processing in investigators lab. Controls (ground and/or flight) processed at the same time.



### Process Samples

Extracted DNA, RNA and/or protein sent to validated omics center to generate sequence, transcript or protein expression data.



### Data Sharing

Data shared with larger scientific community. Results feedback to GeneLab and other databases accelerating scientific discovery by leveraging a bigger community.



### Data Collection & Hosting

Data returned to investigator or GeneLab for analysis. Raw data uploaded into GeneLab database for public viewing.

### Next Generation Research

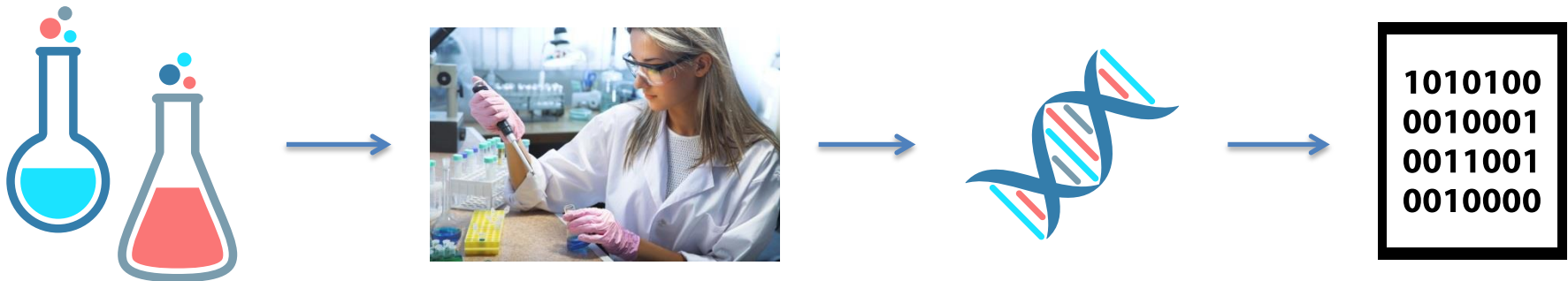
Iterative research solicitations for experiments utilizing GeneLab data for ground validation and next generation flight research.

### Modeling and Validation

Wet lab validation and computational modeling.

# Process Samples

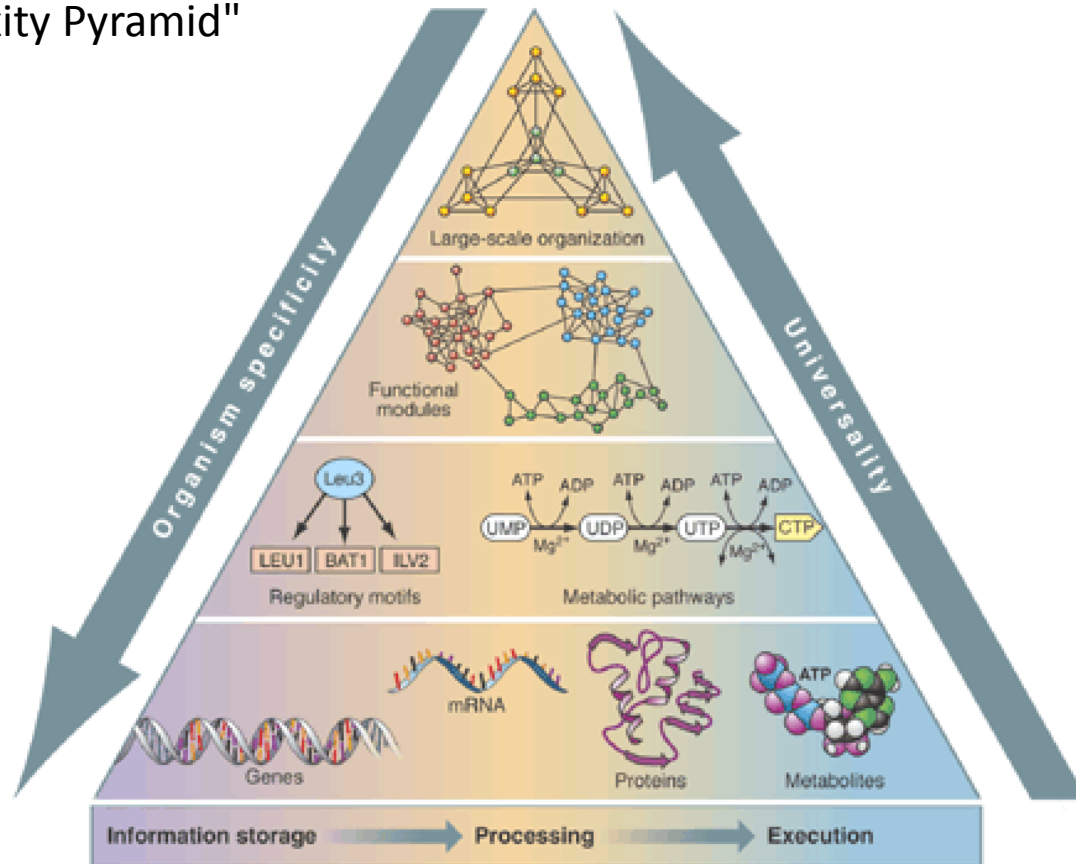
- Flight and/or ground samples received in Laboratory
- Researcher extracts DNA, RNA and/or Protein
- Extracted material sent to center for data generation
- Data is sent back to researcher from center





# Process Samples – Omics Data

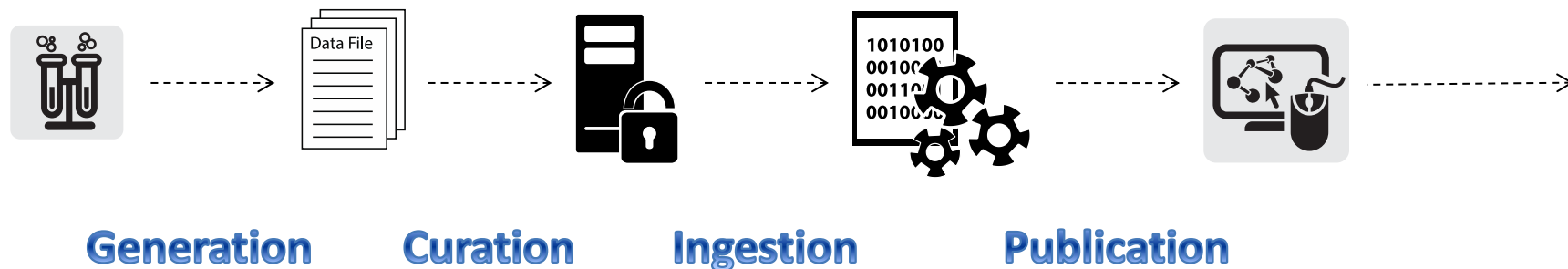
"Life's Complexity Pyramid"



(from Oltvai-Barabasi, Science, Oct 02)

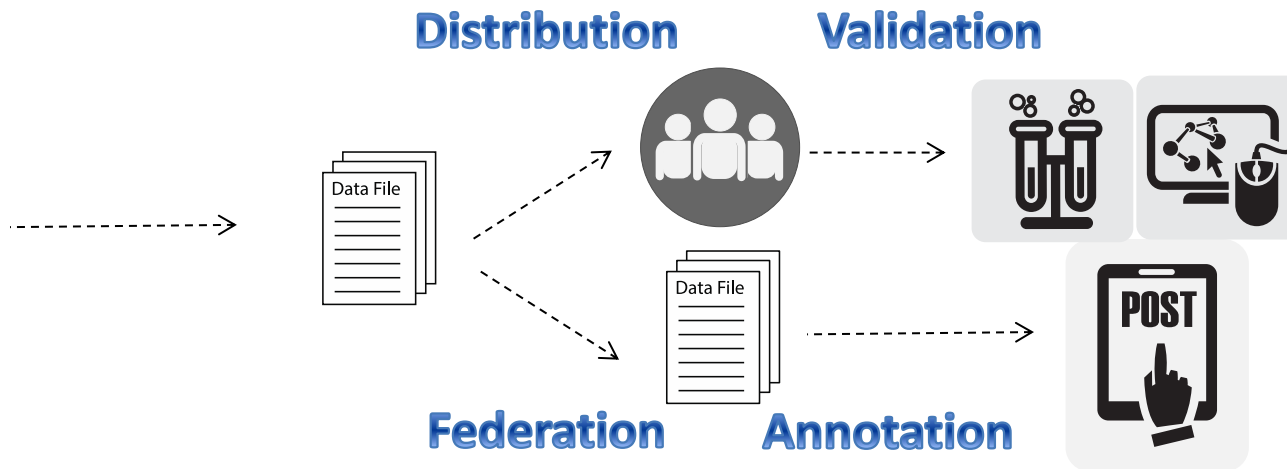
# Data Collection and Hosting

- Flight and/or ground samples are processed to generate Omic data
- Raw and metadata are assessed for completeness and quality by GeneLab personnel
- Completed data is uploaded into data system
- Datasets are published in the system and available to the public

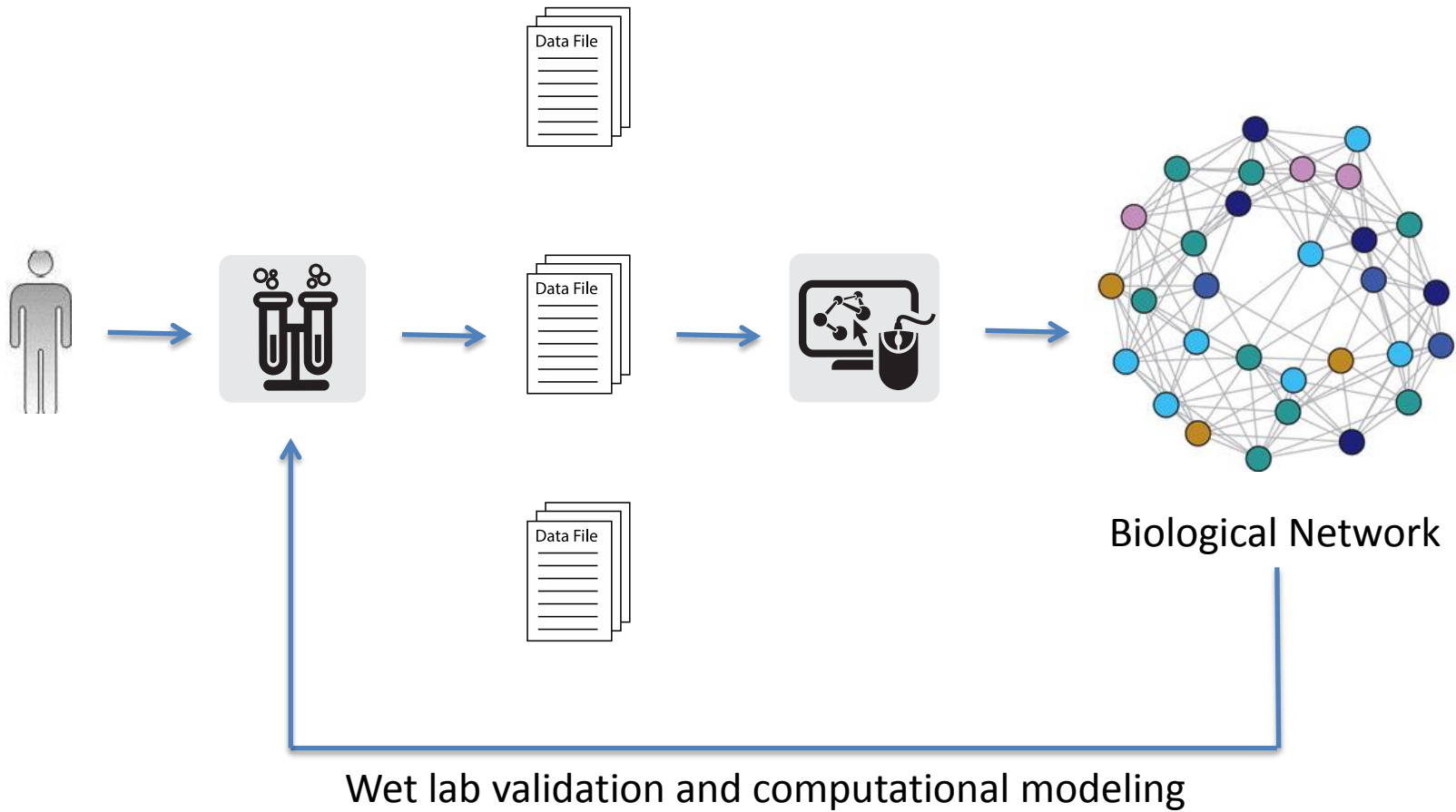


# Data Sharing

- Space biology research data hosted by GeneLab is freely available to all
- Researchers can download datasets for analysis, modeling, validation
- Researchers can publish data, methods, results and thoughts

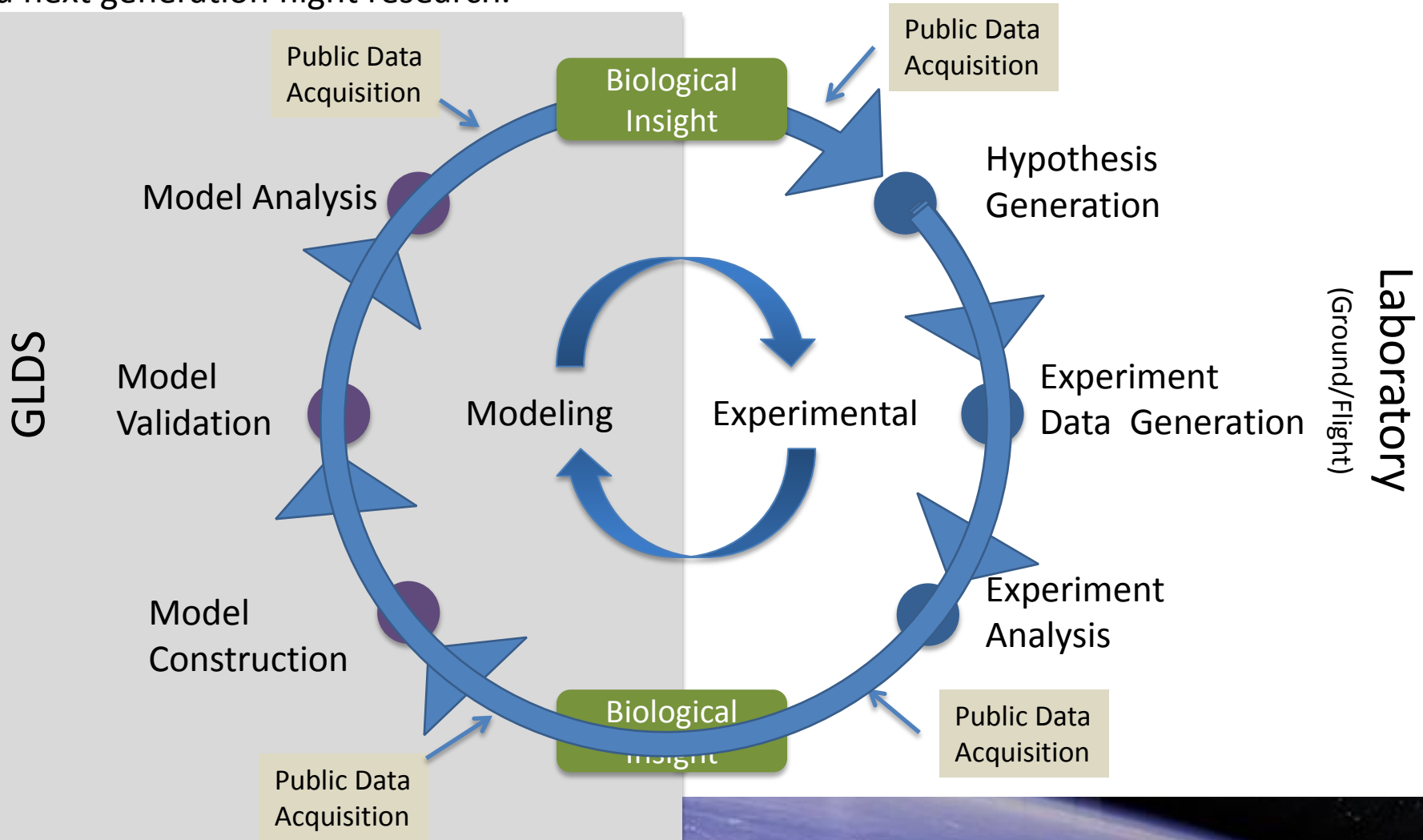


# Modeling and Validation



# Next Generation Research

Iterative research solicitations for experiments utilizing GeneLab data for ground validation and next generation flight research.



# New Grant Opportunities and Missions

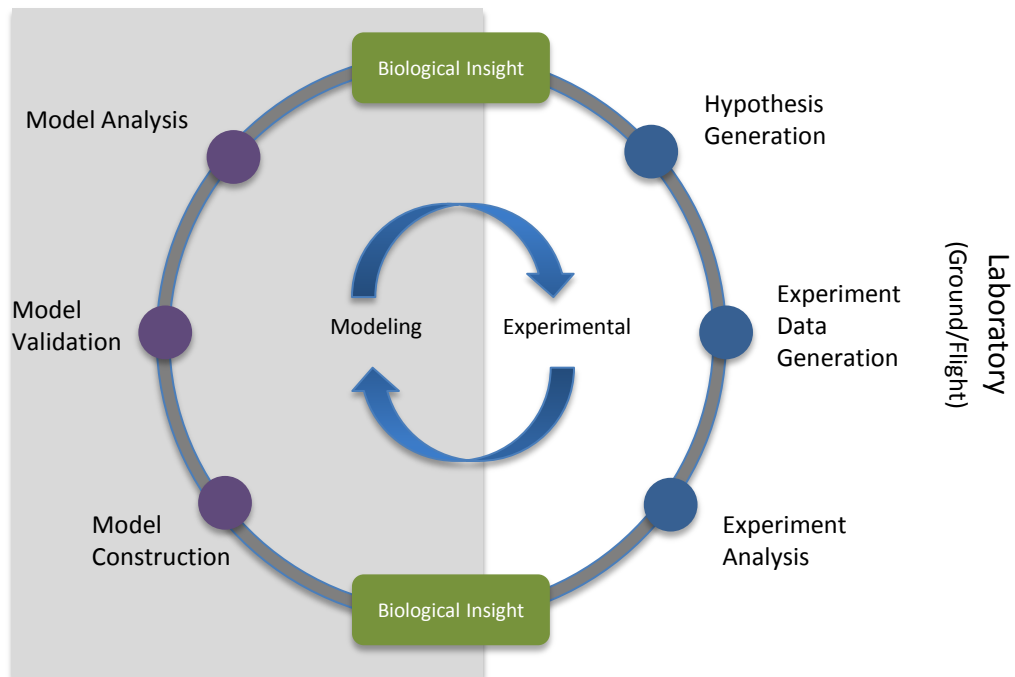
- Next generation of grant proposal
- New flight missions

## Launch

Experiment is prepared and launched according to approved NRA.



GLDS



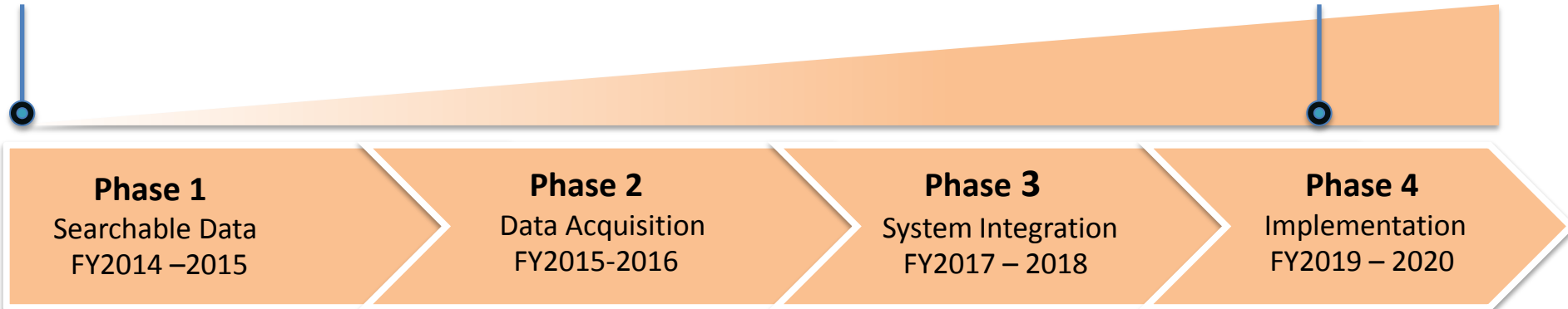
- Computational modeling
- New hypothesis

# GeneLab Phased Implementation

## Phased Implementation 2014-2020

Begin Implementation

Full Implementation



### IT Systems

- **System Requirements & Architecture**
- **Public Website**
- **Searchable Data Repository**
- Requirements level 1

### Science

- Omics Center Solicitation
- Protocol Development
- Data analysis validation
- **Initiate ground controls**
- **Collaborate with two manifested flight experiments**
- **SDT Solicitation for Dedicated Flight**

### IT Systems

- Link to Public Databases
- Beta Space Bioinformatics system

### Science

- Omics Center Selection
- Data analysis from initial ground studies
- Science Definition Teams Identified
- Outreach Program Plan

### IT Systems

- Integrated Platform across model organisms
- Build Community via collaborative science

### Science

- Continue ground controls and process enhancement
- Engage with Scientists external to NASA as part of Outreach Program
- Dedicated flight experiments

- Full science community engagement
- Development of analytical and Modeling tools
- Ongoing dedicated flight experiments
- Website and platform sustaining activities
- Continuous improvement

# Salient Questions / Science

- How should the current state of analytics for proteomics and metabolomics influence choice of global profiling design vs. more focused experiments?
- How to use resources wisely to generate longitudinal datasets and support experiment (organism/tissue/assay) diversity?



# Salient Questions / Data Systems

- What is the value of providing biocomputing capabilities in terms of increasing scientific throughput?
  - Access to high performance computing resources
  - Biocomputing tools repository
  - Is there analogy to collaborating teams of scientists in Earth Science

# Salient Questions / Science

- What are the data boundaries of scientific relevance vis-a-vis the environment of spaceflight experiments?
  - 60K+ ISS parameters
  - Selections by PIs may overlook parameters
  - Data archived approx. every 2 years, after which retrieval is much more difficult

# Salient Questions / Science

- How can the SDT maximize relevance of flight experiment data to ground-based experiments?
- How can we expand the user-base of the flight experiment data to include more researchers performing ground-based experimentation?

# Salient Questions / Data Systems

- How can GeneLab best support biological model development, analysis, validation? Best leverage distributed scientific knowledge?
  - Data Discovery/Federation
  - Data annotation/Knowledge Capture
  - Data visualization
  - Develop in-house capabilities
  - Integrate externally-developed capabilities
  - Direct users to external systems and tools

# Salient Questions / Data Systems

- How best to represent Omics metadata?
  - Myriad, rapidly evolving metadata standards
  - Newer Omics data types have no metadata standards yet
    - Data formats
    - metadata, protocols and analyses
  - How to capture the unique aspects of space biology experiments as metadata and related data?

# GeneLab Staff

Project Manager – [Joe Coughlan](#)

Deputy Project Manager – [Yung Nguyen](#)

Project Scientist – [Terri Thompson](#)

Outreach Lead – [Jon Rask](#)

Payload Lead – [Linda Timucin](#)

Data Systems Lead/Architect – [Daniel Berrios](#)

Lead Developer – [Chris Middour](#)

Database System/Web Engineer – [Jon Welch](#)

Bioinformatics Scientist – [Homer Fogle](#)

Ground Lab Science R&D Lead – [Kaushik Chakravarty](#)

Lab Manager – [Sam San-Huei Lan](#)

Lab Technician – [Rick Chen](#)

Project Analyst/Configuration Manager – [Nikita Gilkerson](#)

Project Coordinator – [Desireemoi Bridges](#)

# Backup Slides