National Aeronautics and Space Administration

GeneLab Analysis Working Group kick-off meeting

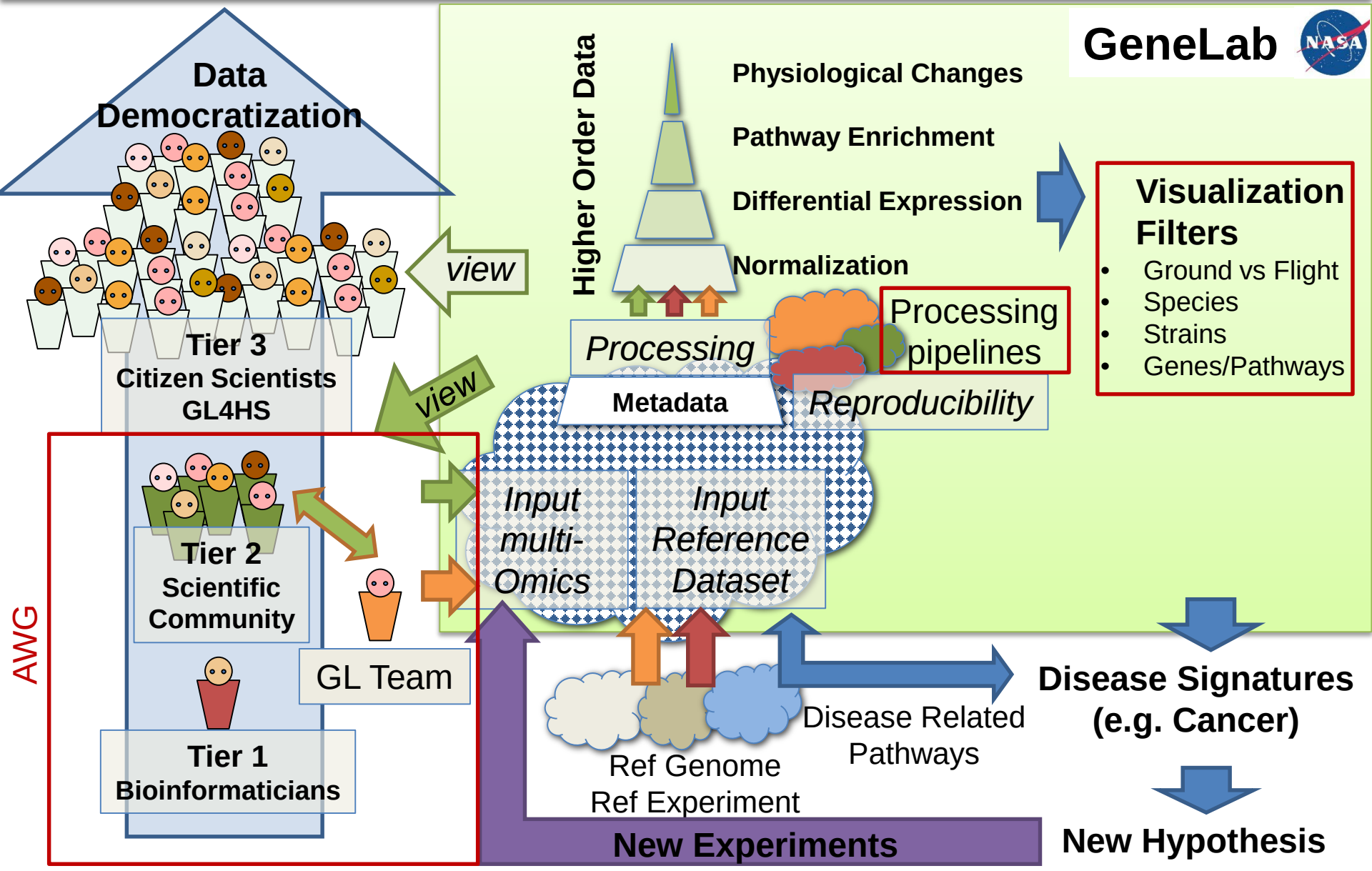
Sylvain Costes, Ph.D.
Project Manager

Agenda

- Goals to achieve for GeneLab AWG
  - GL vision
  - Review of GeneLab AWG charter
- Timeline and milestones for 2018
- Logistics
  - Monthly Meeting
  - Workshop
  - Internship
  - ASGSR
- Introduction of team leads and goals of each group
- Introduction of all members
- Q/A

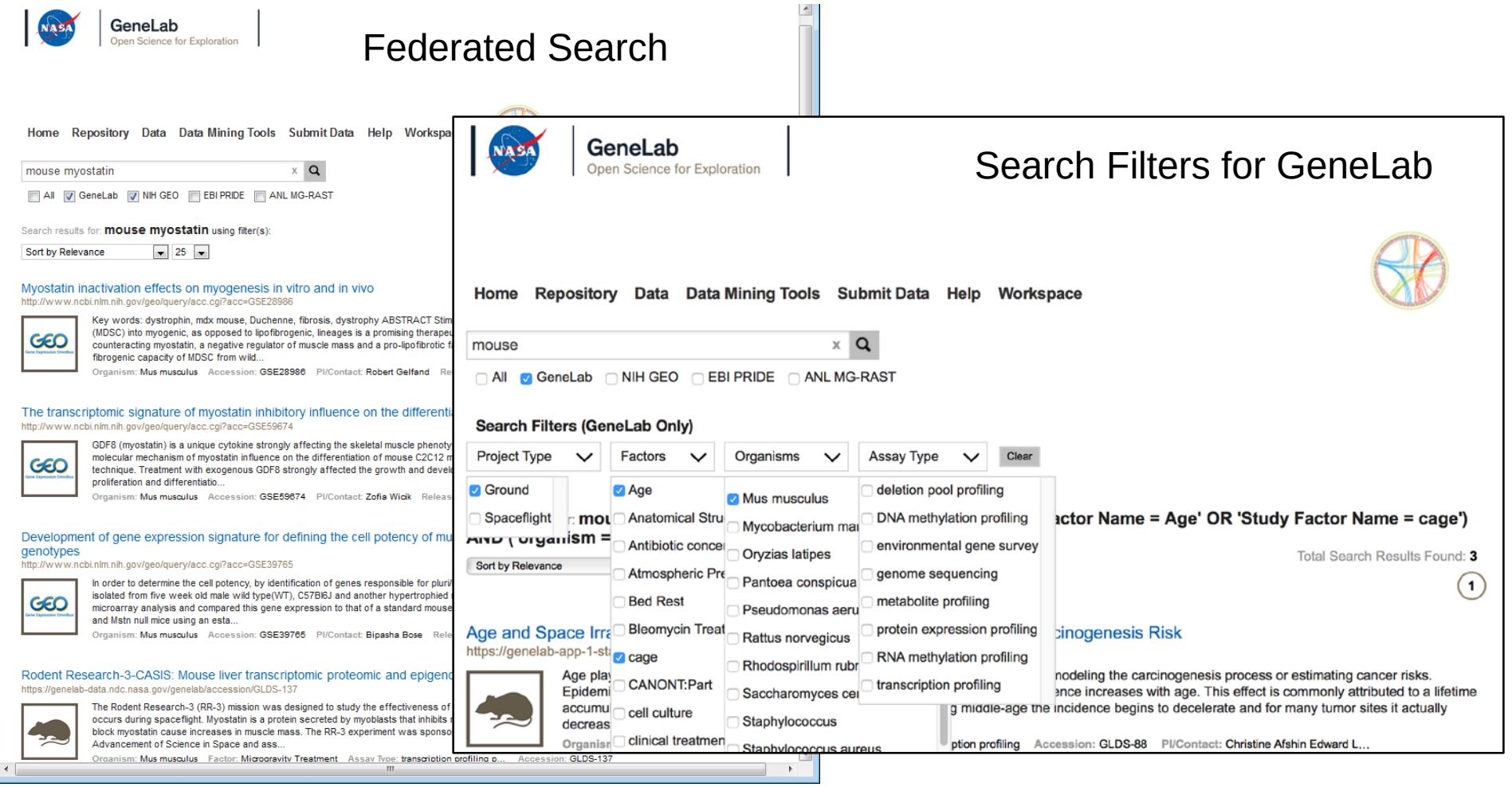
Three-tier Client Strategy to Democratize Data

GeneLab
Open Science for Exploration



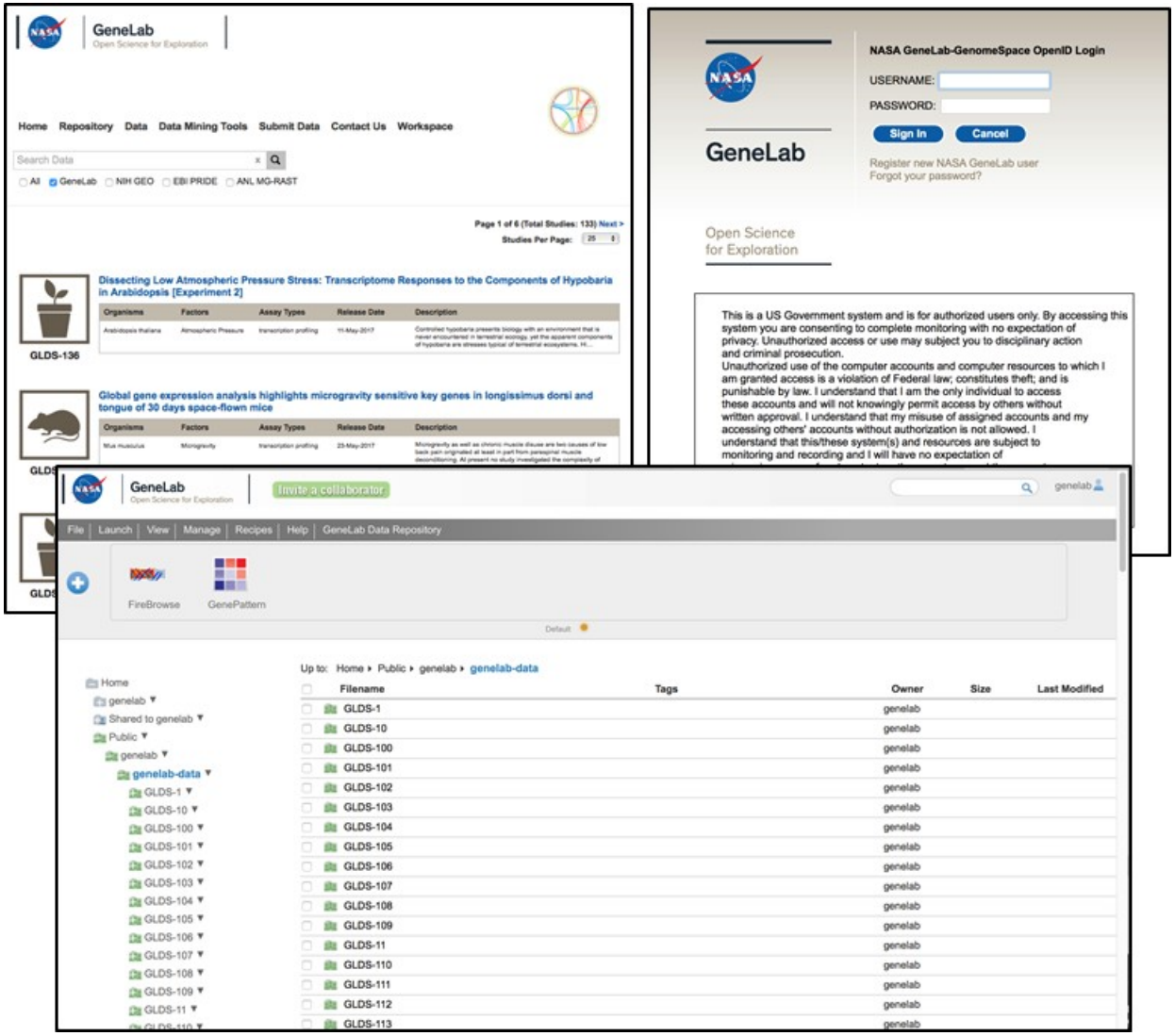
GLDS Phase 2 (Release 2.0) Google-like Search, Federated Search

Data federation/integration with heterogeneous bioinformatics external databases (GEO, PRIDE, MG-RAST)



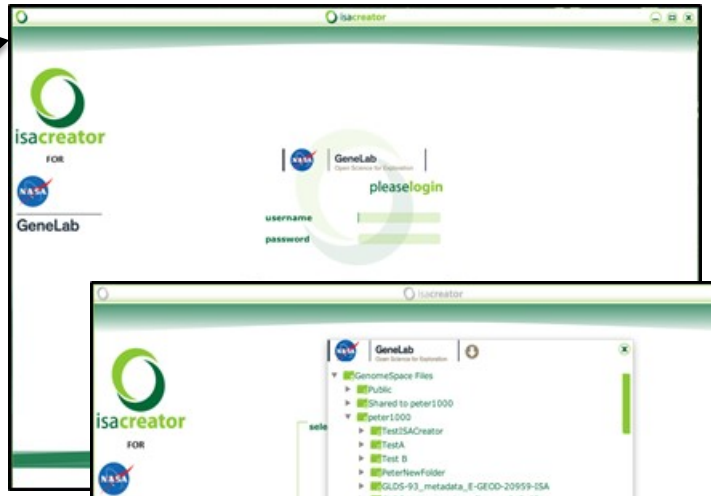
GLDS Phase 2 (Release 2.0) Customized NASA Collaborative Workspace

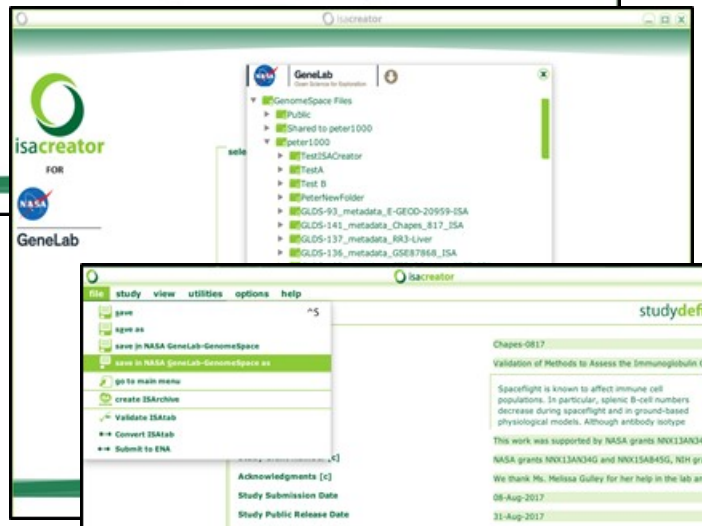
User Account Mgmt., Access Controls (e.g., Private, Shared, Public Folders)

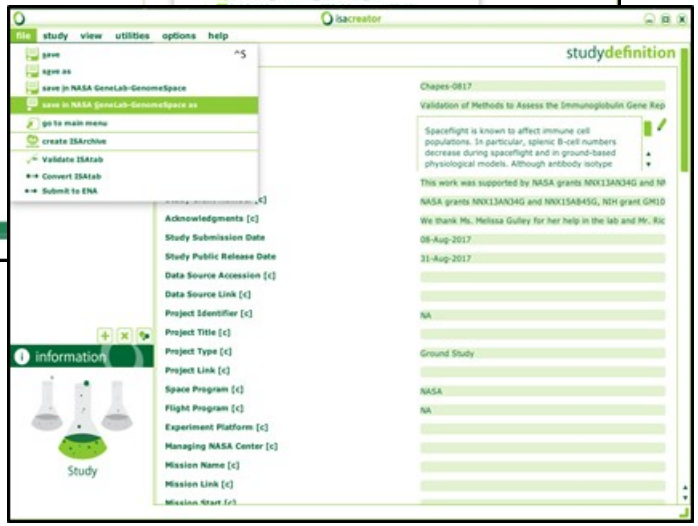


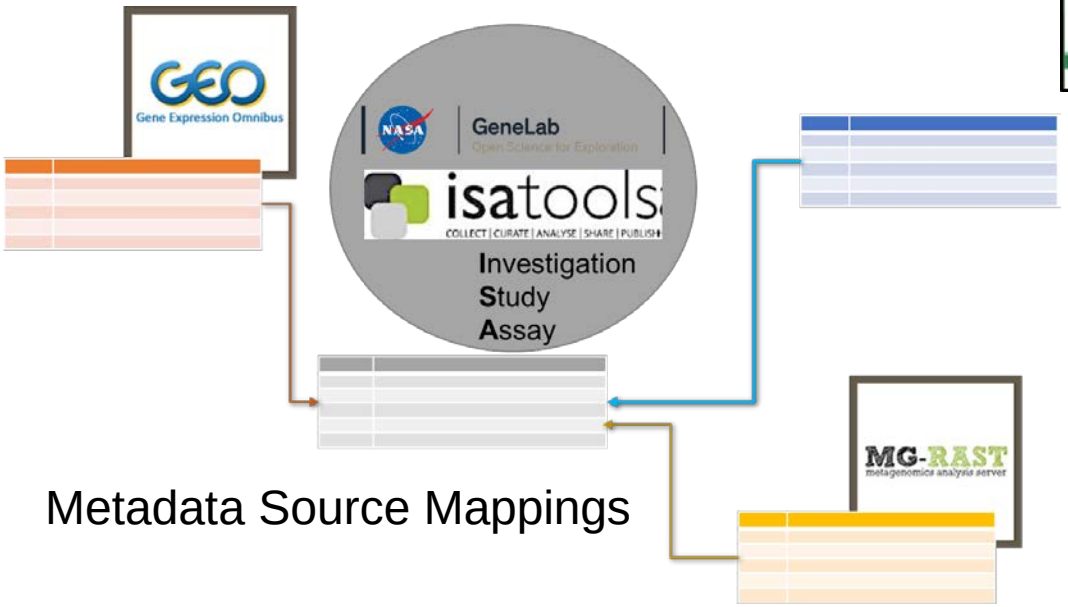
GLDS Phase 2 (Release 2.0) Metadata Curation via ISACreator Tool

GeneLab-GenomeSpace Integration with ISACreator for Streamlining Data Processing Operations









Metadata Source Mappings

Engaging the Scientific Community

GeneLab **Analysis Working Groups (AWG)** will be tasked with analyzing all data across the GLDS with relevance to a specific domain to generate higher-order data.

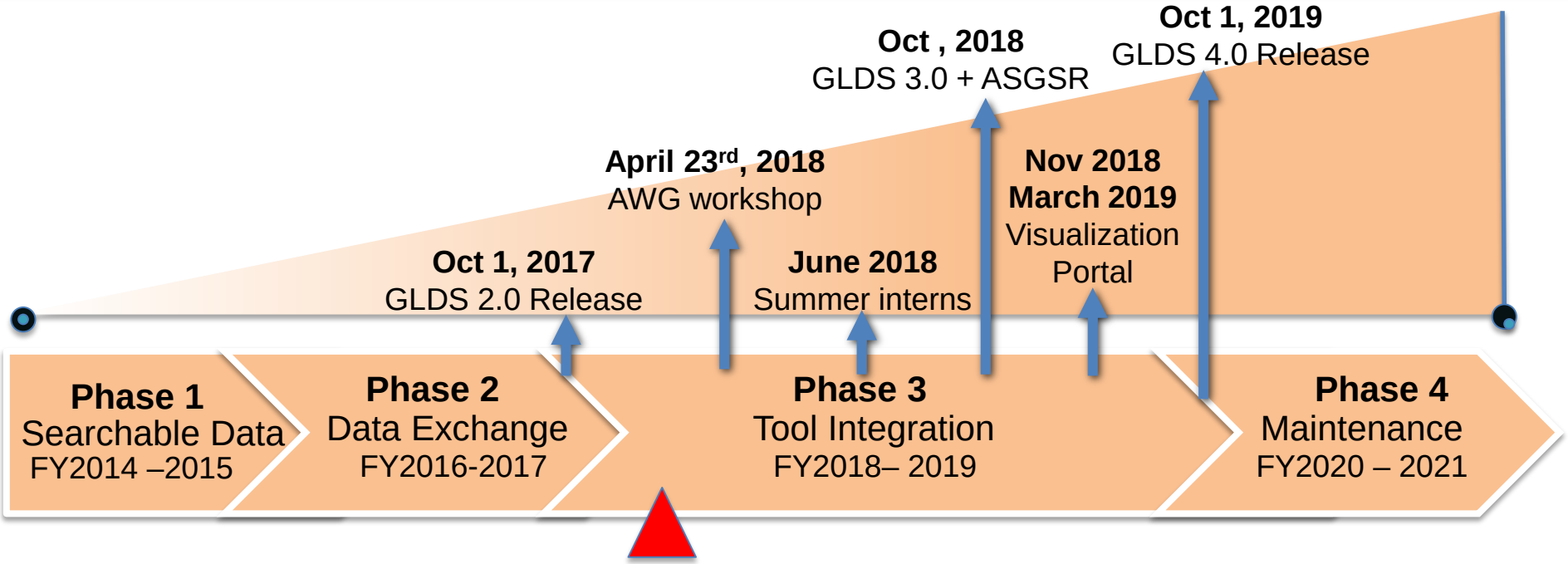
Goals:

1. Peer-reviewed publications describing AWG's comprehensive analysis.
2. Consensus data analysis pipelines relevant to AWG domains to be used on the GLDS will help domains harmonize their data.
   a) Summer interns will process all data based on AWG recommendation
   b) Processed "higher-order" data relevant to domains will be posted on the GLDS.
   c) Strategies needed to link metadata to processed data will be put in place for the visualization portal deployment
3. Critiques of the GLDS to be used for improving its utility; test driving passed along to scientific community via the AWG
   a) Access to galaxy toolshed and Jupyterlab notebook within GeneLab provided with CPU and RAM AWS resources
   b) Integration of GenomeSpace workspace with processing tools
   c) GLDS 2.0 search query needs to be improved – What should we do different?

AWGs emphasis:

1. Mammalian
2. Invertebrate
3. Plant
4. Microbial
5. Multi-omics systems biology challenge

Phased Implementation



Data System
- ✓ Public Website
- ✓ Searchable Data Repository
- ✓ Top Level Requirements
- ✓ New Data and Legacy Data

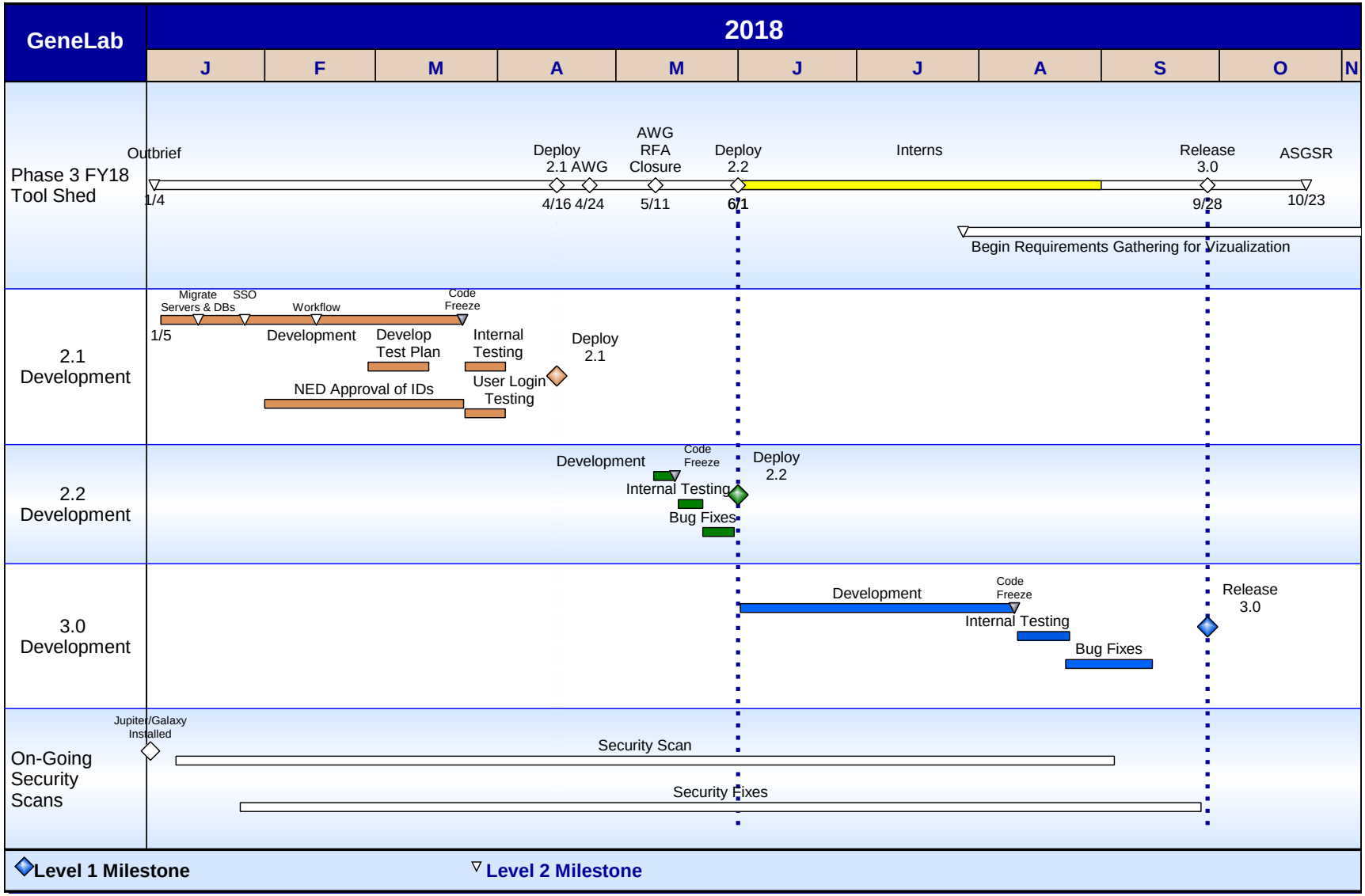
Data System
- ✓ Link to Public Databases via Data Federation
- ✓ Integrated Search (e.g., data mashup)

Data System
- Integrated Platform across model organisms
- Build Community via AWG
- Provide access to biocomputational tools for omics analysis
- Provide collaboration framework and tools

Open Source Maintenance
- User community becomes primary provider of new tools/knowledge
- Maintain integrity of data, and data system

GeneLab
Open Science for Exploration

Detailed Timeline for 2018



AWG Charter

- Please add…

Logistics

- Monthly Meetings
  - Logistics details please…
  - Mention need for members to get their credential sorted out for Galaxy account
  - Mention need for members to get acquainted with the data
  - Make sure all data are available within the toolshed
  - Communication – mention ResearchGate page

AWG Workshop

GeneLab will host a two-day in-person workshop in April to finalize the establishment of processing pipelines. In this workshop, each group will demonstrate the pipeline that will be used to analyze the datasets in the GeneLab Data Repository.

When: April 23-24, 2014

Where: Orlando, FL

Workshop agenda- with times are expected arrival

GeneLab will cover travel and accommodations cost for members to attend this two day conference. Spaces are limited and available on first come first served basis. **An email will be sent out after this meeting for confirmation of attendance.**

Student Internship Program

GeneLab will host students for 10 weeks during the summer of 2018 at NASA Ames Research Center. Students will be working on processing the data using the pipelines established by the AWG.

We encourage AWG members to recommend their students to apply. GeneLab cannot guarantee acceptance due to AWG participation but will preference students with GeneLab data processing experience. Applications will be managed by the NASA OSSI system, which has basic eligibility requirements.

When the posting is public, we will notify all AWG members.

Introduction to each AWG

- Please add slides for each group
  - Lead, support people, theme
  - List of confirmed members
  - Sam, can you take care of my group please, Thx!
  - Make sure final slide summarizes the key points (AWG workshop, internship, whatever logistics needed so it stays on while we have the Q/A)