

The AgBioData Research Coordinaton Network: Ensuring FAIR Agricultural Data through Community Based Standards

PAG 2022, Ag2PI Workshop

Feb 9, 2022

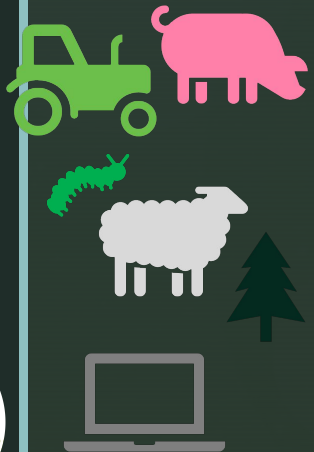
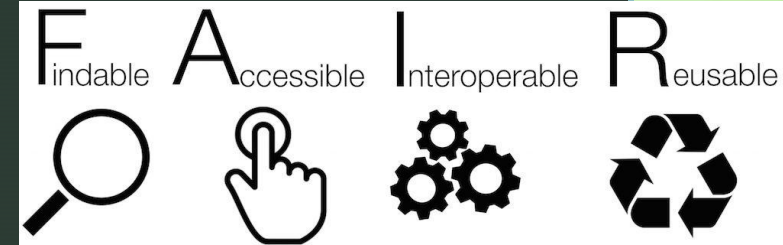


What is AgBioData (www.agbiodata.org)

- Est 2015 at PAG
- ~35 member GGB resources
- First 'official' meeting (Utah) April 2017
 - Harper, L., et al. (2017) [10.1093/database/bay088/5096675](https://doi.org/10.1093/database/bay088/5096675)
- NSF RCN- funded in Sept 2021

Aim1. Recommendations for and implementation of standards for FAIR data

Aim3. Create and test FAIR data management training materials and curriculum



Aim 2. Expand network to ensure inclusivity, diversity and equity

Aim 4. Develop a road map for GGB resource sustainability



Current working groups-1-3

- Data Federation
 - Development and implementation of, and training/education in, best practices to promote accessibility and interoperability.
- Data Collection Accuracy and Precision
 - Define current status of problems/solutions with data collection, accuracy and precision
 - Make recommendations on best practices and tools to improve accuracy of data collection for the future.
- Standards for Genetic Variation Data
 - Address the coordination and access of genetic variation data for agronomic species, plants, animals, fish, and insect vectors.
 - Prepare recommendations and best practices guidelines.



Current working groups-4-6

- Genome Assembly and Annotation Nomenclature
 - Create a unified nomenclature system for gene model names, annotations, and genome assemblies including guidelines on formatting files accompanying genome assembly announcements and publications.
- Genotype-Phenotype data
 - Improve the data collection and data sharing in order to help the identification of genotype-phenotype associations.
 - Define the data and metadata for phenotype-genotype studies using standardized terms (e. g. ontologies)
- Pan Genomes
 - Promote reproducible and well-documented analysis pipelines,
 - Identify roles that pan-genomes and their visualization and analyses play in biological research,
 - Review and perhaps improve ways of connecting phenotype data to pan-genome data



Current working groups-7-9

- Ontologies
 - Review/assess current ontology usage, identify gaps, opportunities and barriers for adoption. Develop recommendations, best practices and training materials for ontology use.
- Recruiting
 - Ensure equity, diversity and inclusivity within our working groups and across the Consortium by engaging a broadly representative population of stakeholders including: data generators, data consumers, genetics and genomic breeding database staff, data curators/managers, data publishers and funders.
- Education
 - Train researchers at all career points in FAIR data management. Develop and disseminate educational materials about FAIR data and GGB databases.



Future working groups

- Sustainability (Y3)
- Implementation (Y3)
- Emerging data types (community driven, Y2-Y3)



Synergies with Ag2PI community-making data FAIR

- Data formats/standards
- Metadata standards
- Data collection/curation best practices
- Ontologies



Participate: AgBioData (www.agbiodata.org)

- Interested in joining a working group ? Contact us
 - agbiodata@gmail.com
- Attend Community Networking Event March 15-17(2hrs/day)
 - Breakout rooms with WGs
- Attend monthly meetings/ webinars (1st Wed of the month)
- Join our mailing list (<https://www.agbiodata.org/user/register>)
- If you have a GGB resource – join the network
- Follow us on Twitter @AgBioData
- Join our Slack channel (<https://www.agbiodata.org/join-slack>)

