Towards a unified MOD+GO resource
An Overview

# Current MODs model

From Valentina di Francesco

- ## User confusion for lack of homogeneity
  - User access interfaces
    - need different navigation skills and data access approaches for each resource
    - Semantic inconsistencies and different data structures for the same genomic entities
  - Analyses
    - human/model organism association for disease and phenotypes
    - functional annotation
    - homology representation

- ## Redundancy of operations at 6 resources
  - Data management systems for related data structures and types
  - System administration and IT support
  - Technical user support
  - Links to the same public resources which need updates and maintenance

- SGD
- WormBase
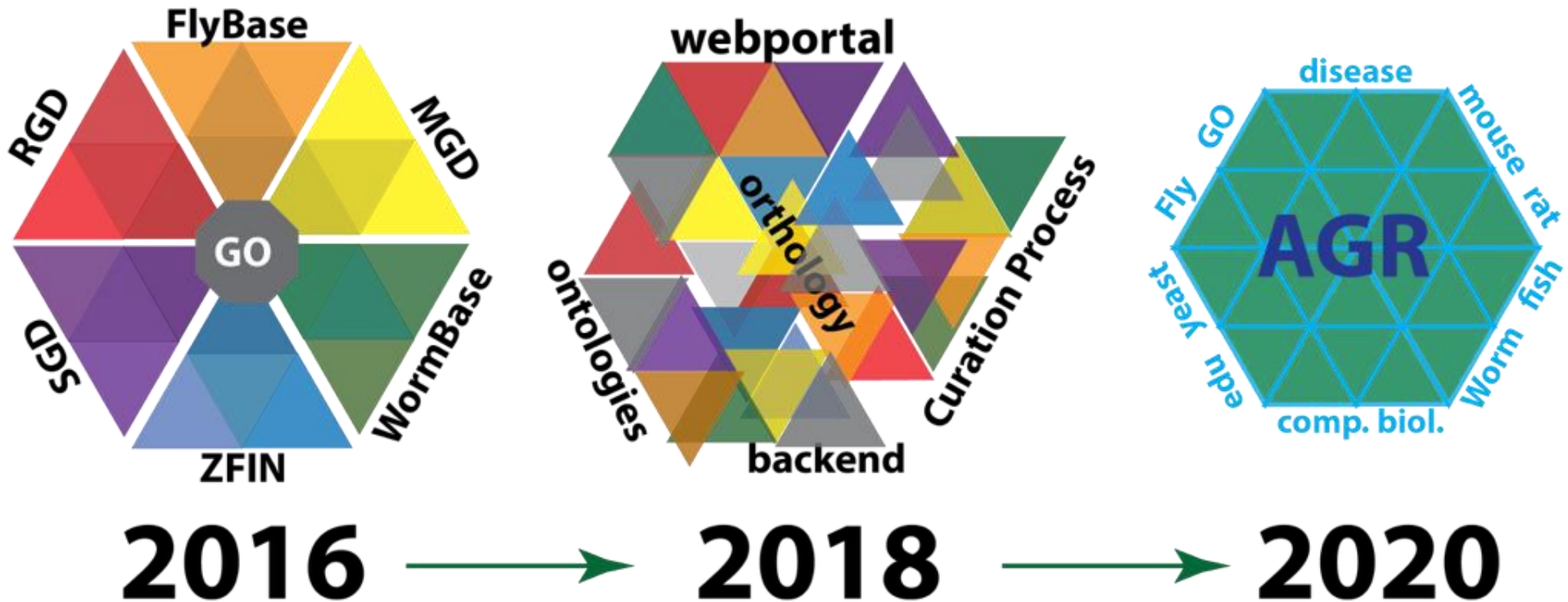- FlyBase
- Zfin
- MGD
- RGD
- GOC
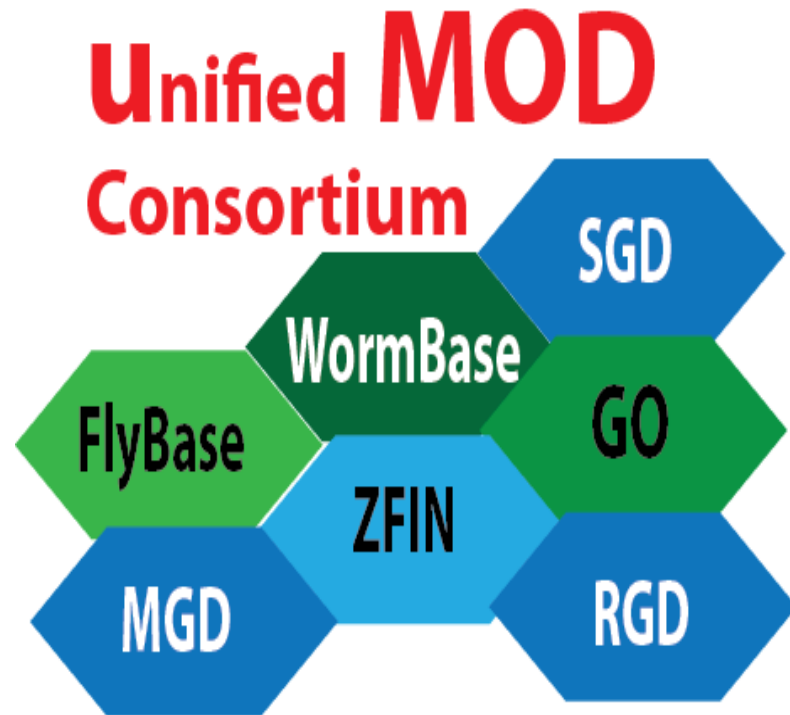
Alliance for Genome Resources:AGR

Alliance of Genome Resources

integrating ≥7 resources

2016 → **2018** → **2020**

GOAL: provide all researchers the quality of resource that represents the best of model organism databases
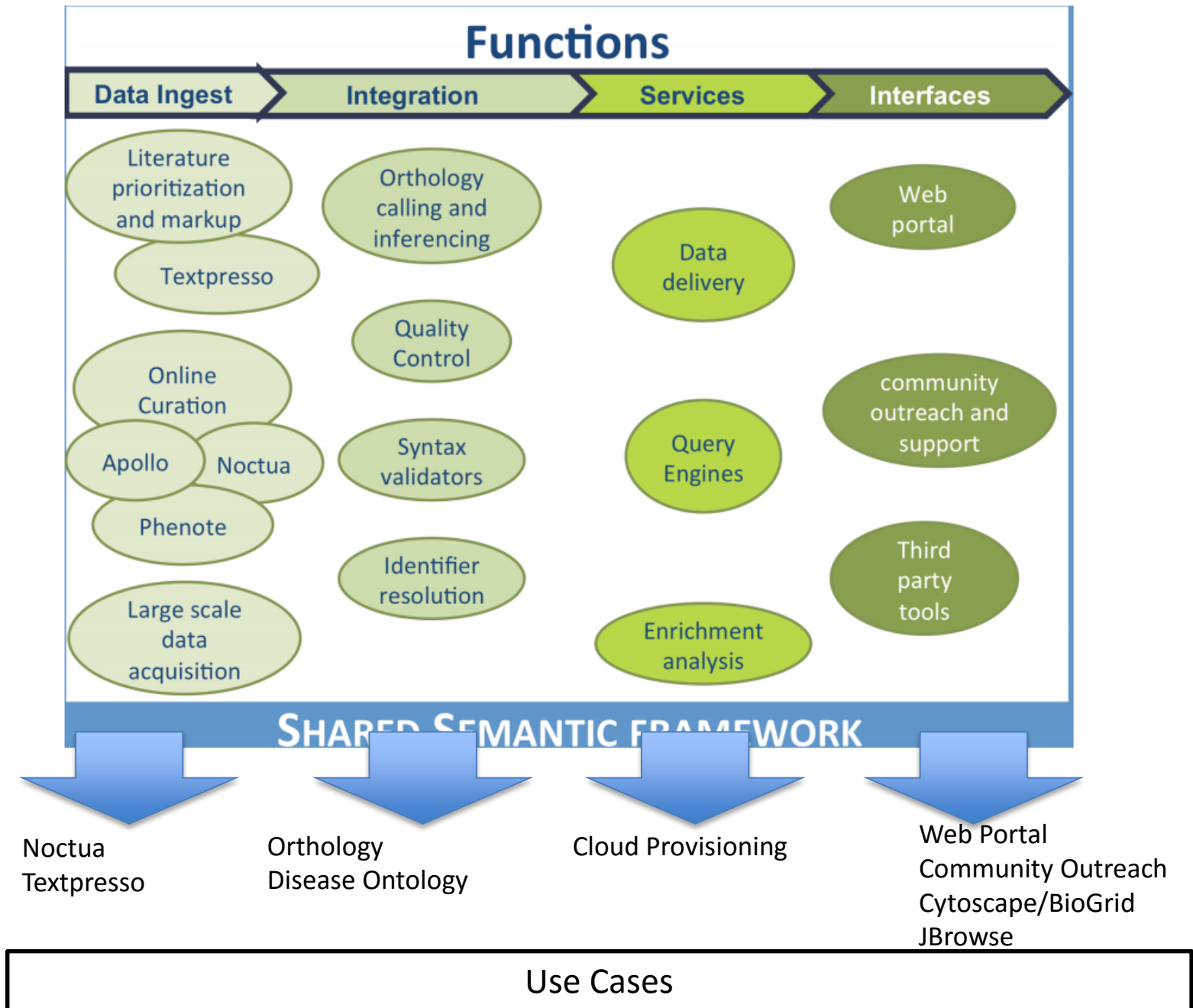
Future Alliance model

Unified
MODs
and GOC

Target user communities:
- Basic researchers
- Clinical researchers
- Computational biologists/Data scientists
- Educators/students

- ## Common web portal
  - ### User access interfaces
    - Unified search & display for common data elements
    - Preservation of elements/data types unique to a particular organism
  - ### Analyses
    - Common analysis tools for users
    - Preservation of analysis tools unique to a particular organism

- ## Streamlined operations
  - Common tool sets for curation, analysis, visualization
  - Common data management systems for related data structures & types
  - Common systems administration & IT support
  - Common technical user support

Relationship of proposed work groups to developing the "shared semantic framework"

Functions

Data Ingest → **Integration** → **Services** → **Interfaces**

- Literature prioritization and markup
- Textpresso
- Online Curation
- Apollo
- Noctua
- Phenote
- Large scale data acquisition
- Orthology calling and inferencing
- Quality Control
- Syntax validators
- Identifier resolution
- Data delivery
- Query Engines
- Enrichment analysis
- Web portal
- community outreach and support
- Third party tools

SHARED SEMANTIC FRAMEWORK

- Noctua
  Textpresso
- Orthology
  Disease Ontology
- Cloud Provisioning
- Web Portal
  Community Outreach
  Cytoscape/BioGrid
  JBrowse

Use Cases

Getting started

- Let's build on existing projects for building shared tools. Use these as examples of early AGR "wins".
  - *Launch formal AGR working groups around these projects.*
    - e.g., Noctua, Textpresso
- Let's build on existing agreements about common standards.
  - *Launch formal working groups around these if they don't already exist.*
    - e.g., AGR orthologys, DO disease ters annotations, JBrowse as AGR genome browser
- Stand up an ongoing working group to help assist with moving centralized infrastructure/tools into the cloud (Dev Ops group)
  - e.g., JBrowse and Cytoscape
- Start with a simple use case to launch AGR web portal development
  - The proposal on the table from the tech team is to use the existing Intermine infrastructure to drive the development of the initial portal (using the first use case from the Use case group).

Proposed Work Products and Working Groups:
Sept 2016-Dec 2016

AGR Work Products	Description	Working Group
AGR standard for orthologs	Orthologs to be used across AGR resources.	Orthology
Use Cases	Generation of feature requirements documentation for the AGR portal.	Use Case
Textpresso	Develop common literature triage and markup tool for all of AGR	No formal AGR working group yet although work is ongoing
Noctua	Development of common GO annotation curation tool for all of AGR	No formal AGR working group yet although work is ongoing
JBrowse	Implementation of centralized infrastructure for hosting Jbrowse	Need AGR working group
Interaction data and visualization	Identification of sources of interaction data for all AGR resources. Implementation of common Cytoscape viewer.	Need AGR working group
Disease Ontology	• OMIM terms get DO-id • All members migrate to using DO in their annotation processes • Harmonization of evidence codes • Identify/analyze human gene-disease sets and continue curation (RGD)	Phenotypes/Disease
Web portal	AGR web portal driven by use cases.	Need AGR working group
AGR User Support	Team to produce AGR educational and outreach materials and coordinate social media outreach	Need AGR working group
Cloud provision team	Team dedicated to moving work products into the AWS cloud environment.	Need AGR working group

Curation can be an internationally distributed activity

The AGR comprises distributed projects many of which have several performance sites and many telecommuters

Web-based curation tools are broadly used

Telecommuting biocurators are often embedded in other annotation groups rather than organism-specific research laboratories because
Biocurators now have more in common with each other than other researchers.
For example, WormBase curators work in ZFIN and FlyBase space.

Local biocuration centers would increase the user outreach, feedback and training in that country.

Noctua Common Annotation Tool

Noctua Training at EBI

September, 2016

AGR Portal Prototype

ALLIANCE of GENOME RESOURCES

All ▾ | Pax6 | 🔍

Category

🧬 Gene Ontology	105
🧬 Diseases	33
🧬 Genes	21

159 results for **"Pax6" ✖**

View As

[≣ List] [⊞ Table]

Page 1 of 4

< >

Sort By

Relevance ▾

Page Size

50 ▾

⬇ Download

Pax6

🧬 Genes

Name: paired box 6
Synonyms: *Pax6*
Source: RGD:3258
Species: *Rattus norvegicus*
Gene Type: gene

Pax6

🧬 Genes

Name: paired box 6
Synonyms: *Pax6*
Source: MGI:97490
Species: *Mus musculus*
Gene Type: protein coding gene

AGR 'in person' Meetings

- Leadership meeting December 6 in Denver
  - Iron out decision making processes and communication
  - Start planning "All hands meeting"
    - Likely next June

- AGR SAB and Community meeting March 7 in DC

AGR PI at this AFSP Meeting

- Judith Blake
  - PI for MGD and GOC

- Norbert Perrimon
  - PI for FlyBase

- Monte Westerfield
  - PI for ZFin