

Sharable Resources

CA Big Data

Biomedical Workshop

Oct. 9 – 10, 2015

Palm Desert, California



Sharable Resources

- What to share?
 - Data
 - Software
 - Services/APIs
 - Hardware
 - Cloud resources
 - Expertise

Data Sharing

- Benchmarks for Bioinformatics
- Globus endpoints for fast data sharing:
globus.org
- Welllderly ~2000 genomes from healthy elderly people (VCF file, summarized data)
- Issues to consider
 - Policy
 - Resources (cost, bandwidth, storage)

Shared Services

- MyVariant.info
- MyGene.info
- globus.org for fast data transfer/sharing services
- globus.org/genomics for large scale genomics on the cloud
- PDB RESTful services
- Others ... (-> Catalog)
- Issues
 - Highly scalable APIs, concurrency

Shared Hardware

- Cloud to host research services
 - Learning curve
 - Payoff (time, resources)
 - Educational grants from cloud providers
- Open Science Grid

Action Items

- Create API working group
 - Define best practices
 - White paper (Chunlei Wu)
- Big Sharing
 - White paper, what services are available (-> Catalog discussion group)
 - Biomedical informatics benchmarks (Ravi Madduri)
- Match Making/Use Cases
 - MyVariant.info – PDB Variant Mapping to Structure(C. Wu, P. Ping, P.Rose)