Welcome!

Midwest Big Data Hub & National Data Service
Data Science Tools & Methods Workshop

July 11-12, 2018
Big Data Hubs
Stimulate regional grassroots partnerships focused on Big Data

Midwest Hub Priority Areas

Supported by NSF awards 1550312, 1550320
What is the U.S. National Data Service?

National effort to bring together infrastructure supporting the publication, discovery, and reuse of data

→ From the Internet to the “Datanet”

1. Large-scale Data Service Interoperability
   - Distributed cloud and compute
   - Innovation in the gaps: services, software, integration

2. Incubator of Data Projects & Pilots
   - Quick start sandbox
   - Choose services based on features (not time to install)

3. Training Platform
Open Storage Network

Open/distributed storage network near computational and networking resources

- Share data with broader scientific community
- Support novel analysis on the data
- Leverage other data services and support

Supported by NSF award 1747490
Workshop Goals

1. Identify cross disciplinary data science tools and methods, as well as the application of data management and analytics tools, to address research challenges.
2. Expand knowledge and access to data tools, methods, and services.
3. Seed/foster cross-disciplinary collaborations around cyberinfrastructure R+D and domain applications.
Program Highlights

• Lunch – BOFS & Live Demos
• End of Day 1 – Themes emerging, project ideas
• End of Day 2 – Report out, commitments
BOF Questions

Self-assembling for future data infrastructure

• The most important data infrastructure gap on the horizon is…
• The most impactful pilot project for data access to spin up now is…
• We are poised to drive a Proof of Concept / pilot, looking for collaborators with...