MIDWEST BIG DATA HUB

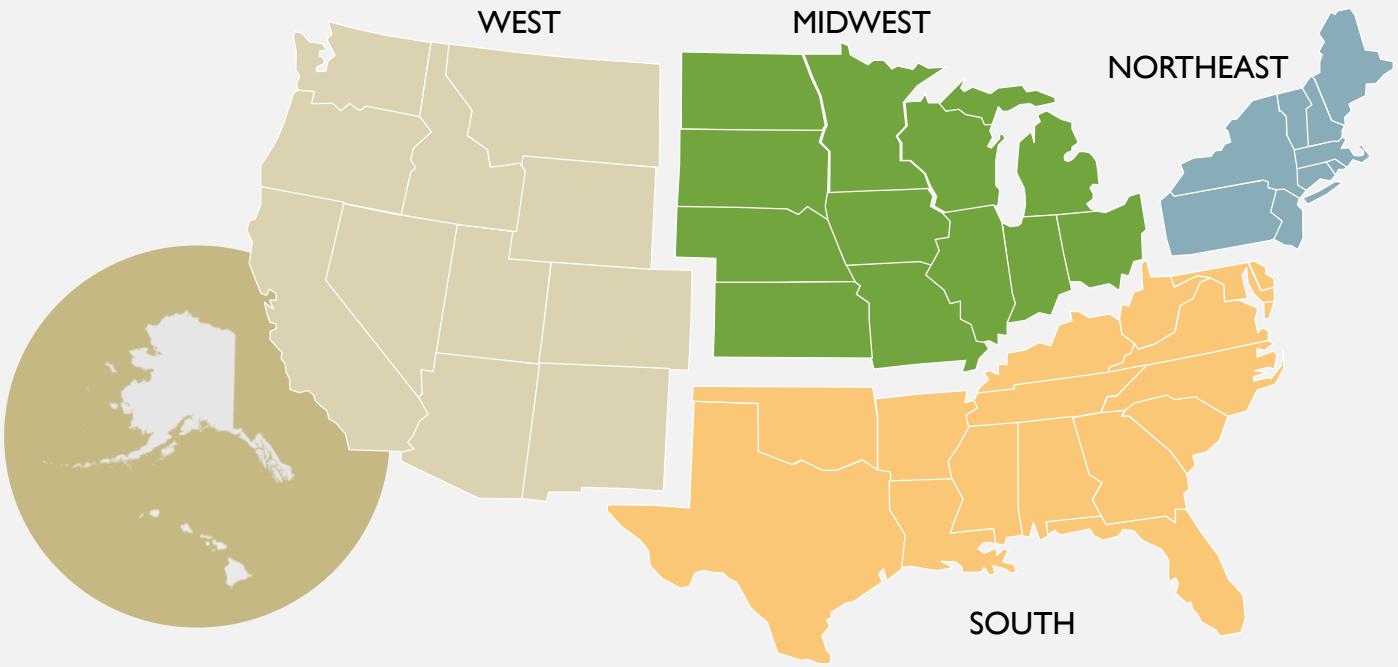
Accelerating the Big Data Innovation Ecosystem

OVERVIEW

- MBDH as a developing organization
- Thematic and Priority areas for the Midwest Hub
- Cross-cutting areas of interest
- Examples of projects and partnerships
- Opportunities and barriers for EU-US collaborations

MIDWEST BIG DATA HUB

REGIONAL BIG DATA INNOVATION HUBS



MIDWEST BIG DATA HUB

MBDH STRUCTURE & GOVERNANCE

- SEEDCorn grant (Seidel, $1.5M) to support launch with initial consortia
  - Indiana University (Plale)
  - Iowa State University (Nusser)
  - University of Michigan (Athey)
  - University of North Dakota (McGimpsey)

- Steering Committee – monthly meetings
  - PIs form initial Exec Committee
  - Thematic community leads (8)
  - Cross-cutting groups (2)
  - At-Large (5 elected)
  - At-large (2 by appointment)

  - Sub-committees (Sustainability; Governance)

- By-laws

MIDWEST BIG DATA HUB

THEMATIC COMMUNITIES IN THE MIDWEST HUB

Urban Science & Smart Cities

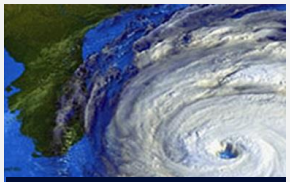
Business Analytics

Materials & Manufacturing

Agriculture

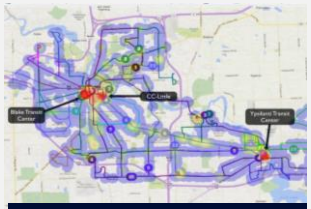
Health & Biomedicine

Neurosciences

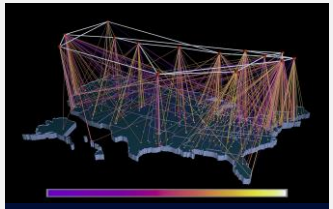


Community Resilience

F-E-W



Transportation



Network Sciences

MIDWEST BIG DATA HUB

CROSS-CUTTING "HORIZONTALS'

- Data, Tools, Services, and Methods
- Sensors, Robotics, Unmanned Systems
- Data science education, workforce development
- Data governance, sharing and use policies
  - Privacy, Security, and Ethics

MIDWEST BIG DATA HUB

"SPOKE" & PLANNING AWARDS FROM NSF

7 Awards - $3.3M

- Digital Agriculture - Unmanned Aircraft Systems, Plant Sciences and Education (UND; IA State; GPN; UN-L)
- Collaborative: Advanced Computational Neuroscience Network (ACNN) (UM; OSU; CWRU; UI)
- Collaborative: Integrative Materials Design (IMaD) - Leverage, Innovate, and Disseminate (UC; UIUC; UM; UW-M; NWU)
- Big Data Innovations for Bridge Health (UN-O; UN-L)
- Cyberinfrastructure to Enhance Data Quality and Support Reproducible Results in Sensor Originated Big Data (Purdue)
- Networked Resilience of Communities Facing Natural and Social Emergencies (Illinois)
- EAGER: Incorporating Citizen Science into Real-Time Sensor-Based Estimates of Traffic-Related Air Pollution Exposure (OSU)

MIDWEST BIG DATA HUB

FOSTERING THE BIG DATA COMMUNITY

- Industry engagement
  - Digital Agriculture All-Hands meeting (Ames, IA)
  - Data Science for Food, Energy, and Water workshop @ KDD 2016 (San Francisco, CA)
  - Big Data for Health and Medicine (Omaha, NE)
  - Big Data & Business Analytics Symposium (Detroit, MI) - Annual Event led by Wayne State University
- Developing expertise and building capacity in Data Science
  - Midwest Big Data Summer School (Ames, IA)
  - Data Quality in an Era of Big Data (Bloomington, IN)
  - Midwest Workshop on Big Neuroscience Data, Tools, Protocols & Services (Ann Arbor, MI)
  - Interdisciplinary Workshops of Big Data in Healthcare Outcome and Workflow for Early Career Researchers (Milwaukee, WI)
  - Food and Data: Interoperability through the food pipeline (Champaign)
- Hackathons, Training and Travel Support
  - Univ. of North Dakota BD Summit and Hackathon (Grand Forks, ND)
  - Students to Software Carpentry (Kansas City, MO)
  - Health Data Matters & LiveStories (Cleveland, OH)
  - Midwest Big Data Hackathon (Iowa City, IA)

MIDWEST BIG DATA HUB

EXAMPLES OF PROJECTS AND PARTNERSHIPS

- Microsoft Azure Cloud Services Partnership (all hubs)
  - providing $ 3M in credits across Hubs + technical support + training
  - Projects underway at Indiana Univ. (2) and Univ. of Minnesota (1)
- Open Commons Consortium-Adv. Center for CyberGIS
  - Demonstration project with the ROGER CyberGIS resource on opening ~200TB of NOAA environmental data
- Internet^2
  - Best practices for distributed teams sharing large data
- National Data Service
  - MBDH community access to hardware/software test-bed environments
  - Inclusion in the NDS Labs Workbench catalogue
- International Food Security at Illinois
  - Machine Learning: Farm-to-Table
- MBDH Data Science Meet-up Groups (Omaha; Fargo)

MIDWEST BIG DATA HUB

WHAT IS THE NATIONAL DATA SERVICE?

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

→ From the Internet to the "Datanet"

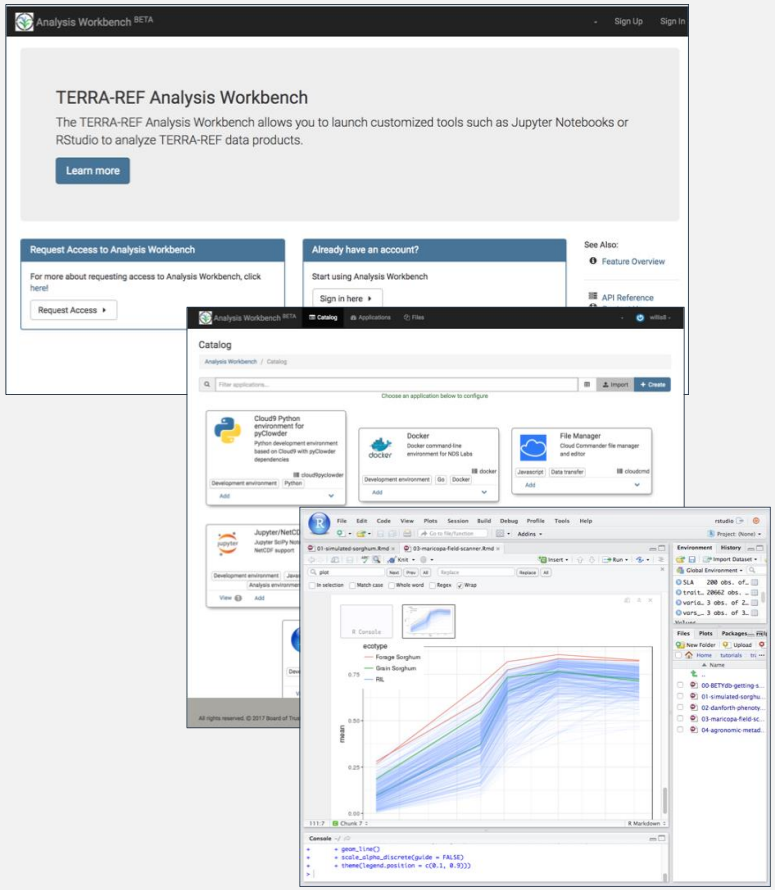
1. Large-scale Data Service Interoperability
   - Distributed cloud and compute
   - Innovation in the gaps:
     services, software, integration
2. Incubator of Data Projects & Pilots
   - Quick start sandbox
   - Choose services based on features
     (not time to install)

MIDWEST BIG DATA HUB

NDS LABS WORKBENCH (BETA)

- NDSC initiative started in January 2016
- Community-driven platform to share, discover, evaluate, develop, and test research data management and analysis tools
- Open platform -- community members recommend and/or contribute tools

USE CASE: TERRA-REF ANALYSIS WORKBENCH



http://www.terraref.ndslabs.org

- Initially used for workshop tutorial to provide consistent environments, scaling to support ~50 participants.
- Customized Labs Workbench instance hosted as part of TERRA-REF infrastructure at NCSA.
- Custom analysis and development environments with direct access to TERRA-REF data. Used by collaborators and alpha users.

MIDWEST BIG DATA HUB

MODELS AND OPPORTUNITIES

- CODATA (1966)
  - Committee on Data, International Council for Science (ICSU)
- NSF and Research Councils UK (RCUK) MOU (2014)
  - Directorate for Biological Sciences (NSF/BIO) and the Biotechnology and Biological Sciences Research Council (BBSRC) pilot
    - "reciprocal acceptance of peer review through unsolicited mechanisms …to help reduce …barriers to working internationally."
- Joint-Laboratory on Extreme Scale Computing (2009)
  - International collaboration on software challenges found in extreme scale high-performance computers
    - Partners: University of Illinois, INRIA, Argonne Nat'l Lab, Barcelona Supercomputing Center, Jülich Supercomputing Centre, and RIKEN AICS

THANK YOU

Melissa Cragin

info@midwestbigdatahub.org

#bdhubs

@MWBigDataHub

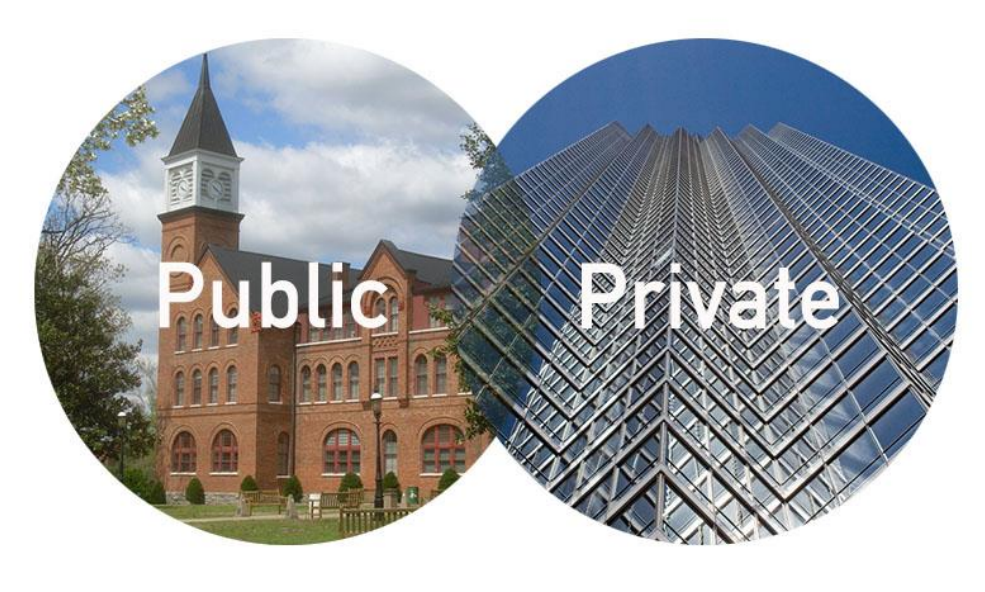
Start-up award from the National Science Foundation (NSF #1550320).

SAVE THE DATE!

- MBDH: Big Data and City Mobility in the Midwest
  - June 22-23 – Ann Arbor
- Sept. 30 – MBDH event : Transportation Data Challenge
  Lincoln, NE
  - Nebraska Transportation Center & Mid-America Transportation Center
- All Hands meeting – Oct. 2-3 – Omaha, NE
  - Partnering with Kiewit Corp, UN-L, UN-O, Creighton Univ., and local MeetUp groups
  - Oct. 4 – Bridge Health workshop
- MBDH Partners' Big Data Summit (Minneapolis, Oct. 26-27, 2017)

MIDWEST BIG DATA HUB

build and strengthen partnerships
across industry, academia, nonprofit and
civic groups, government



to address societal challenges,
spur economic development,
and foster a national big data ecosystem

MIDWEST BIG DATA HUB