Enabling Research via Shared Data, Resources, and Community
We gratefully acknowledge that funding for this workshop was provided by:

• U.S. Institute for Museum and Library Services (IMLS) LG-70-18-0202

• Clarivate Analytics, Inc., Web of Science Group
Program overview

• Workshop purpose and goals (Patty Mabry)
• The CADRE project (Val Pentchev)
• WoS data (Tim Otto)
• Keynote (Katy Borner)
• CADRE fellow presentations (Sam Hansen, Mike Park, Chao Min)
• Q&A and conclusion (Patty Mabry)
What is CADRE?

We responded to an IMLS solicitation to build a National Digital Platform.

Our central tenet: 
Academic libraries are challenged to provide sustainable, affordable, and standardized data and text mining services for licensed, big data sets, as well as open and non-consumptive data sets that are too large or unwieldy to work within existing research library environments, or that have no commercially viable data mining interface.

CADRE is the solution: a shared resource for researchers to solve the cost, silo, and expertise problems associated with hosting big research datasets.
Partners

University of Iowa Libraries  
University of Michigan Libraries  
Michigan State University Libraries  
University of Minnesota Libraries  
Ohio State University Libraries  
Penn State University Libraries  
Purdue University Libraries  
Rutgers University Libraries  

Health Partners  
Pervasive Technology Institute  
Midwest Big Data Hub  
South Big Data Hub  
West Big Data Hub  
Microsoft Research  
Web of Science Group  

This project was made possible in part by the Institute of Museum and Library Services LG-70-18-0202.
SHARE BIGDATA-GATEWAY FOR RESEARCH LIBRARIES (SBD-GATEWAY)

FEDERATED SINGLE SIGN ON

WEB QUERY INTERFACE

CLOUD COMPUTE GATEWAY

Machine Learning
Analytics Tool
Analytics Tool
Spark

Data Processing and Querying

Cloud Storage

Indexed Storage

Cloud Storage

Indexed Storage

Granular Dataset Permissions

AUTHORIZATION

RESEARCH ASSET COMMONS

saved and Shared Results, Tools and Datasets
Visualizations
Annotations
Data Derivatives
Metadata

ACCESS BY PROGRAMMING AND VISUALIZATION TOOLS

APPLICATION PROGRAMMING INTERFACE (API)

Text Analytics
Analytics Tool
Analytics Tool

Database Analytics
Analytics Tool
Analytics Tool

APPLICATION PROGRAMMING INTERFACE (API)

Cloud and Local Compute Resources

Cloud and Local Compute Resources

Jetstream
Commercial Cloud Services

CLOUD STORAGE

Indexed Storage

Relational Databases

Graph Databases

Future Databases

Raw Format Data Lake Storage

Web of Science

MSS Academic Group

USPTO Data

Future Databases

LOCAL AND INSTITUTIONAL STORAGE

Future Data Sources

Future Datasets
Workshop purpose and goals

• Main purpose of today’s workshop is to introduce CADRE to the Science of Science research community

• Desired workshop outcomes: participation in tutorial, ISSI researchers using CADRE, feedback on CADRE, more libraries joining CADRE.
Workshop purpose and goals

Our intended CADRE users are:

• Academic Libraries
• Library Science Researchers
• Science of Science Researchers, other researchers
• Others – e.g., University Administrators
Workshop purpose and goals: A CADRE Community of Practice

Researchers can help by:

• SHARING: Using the CADRE platform for research and sharing, data cleaning procedures/algorithms, analyses, results, pipelines and workflows.

• STANDARDS – eventually we hope the community will self organize and develop standards for sharing data and other research assets.

• FEEDBACK – Please give us your feedback on what you like about CADRE, any features or data you wish we would offer, your desire for future datasets. Also communicate with your institutions librarian about these.
VAL’s SLIDES
CADRE Fellows

Xiaoran Yan
CADRE related events

- **Apr. 2019**
  - 2019 CADRE meeting
  - CADRE Fellowship open
  - 1st Fellows announced
  - ISSI workshop & tutorial

- **Sep. 2019**

- **May. 2020**
  - 2020 CADRE meeting
  - BTAA Library Conference 2020
  - 2020 CADRE hack-a-thon
CADRE Fellowship program

• Gain access to the big bibliometric data sets
• Receive data and technical support for your project
• Join the CADRE community on Slack channels, GitHub repositories and other platforms
• Have early access to free cloud computing resources
• Receive travel scholarships
Utilizing Data Citation for Aggregating, Contextualizing, and Engaging with Research Data in STEM Education Research

Researchers: Michael Witt, Loran Carleton Parker, Ann Bessenbacher
Affiliation: Purdue University
MCAP: Mapping Collaborations and Partnerships in SDG Research

Researchers: Jane Payumo, Devin Higgins, Scout Calvert, Guangming He
Affiliation: Michigan State University
The global network of air links and scientific collaboration – a quasi-experimental analysis

Researchers: Katy Börner, Adam Ploszaj, Lisel Record, Bruce Herr II
Affiliation: Indiana University Bloomington and University of Warsaw
Measuring and Modeling the Dynamics of Science Using the CADRE Platform

Researchers: Russell Funk, Michael Park, Thomas Gebhart, Britta Glennon, Julia Lane, Raviv Murciano-Goroff, Matthew Ross, Jina Lee, Erin Leahey

Affiliation: University of Minnesota, University of Pennsylvania, New York University, Boston University, University of Arizona
Questions?
Understanding citation impact of scientific publications through ego-centered citation networks

Researchers: Yi Bu, Chao Min, Ying Ding
Affiliation: Indiana University Bloomington and Nanjing University
Questions?
Comparative analysis of legacy and emerging journals in mathematical biology

Researchers: Marisa Conte, Samuel Hansen, Scott Martin, Santiago Schnell
Affiliation: University of Michigan and University of Michigan Medical School
Systematic over-time study of the similarities and differences in research across mathematics and the sciences

Researcher: Samuel Hansen
Affiliation: University of Michigan
Questions?
Contact Us

https://cadre.iu.edu

cadre@iu.edu

@CADRE_Project