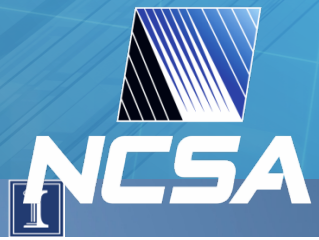
Broader partnership with CC IN2P3

Partnership meeting with IN2P3: May 2015

Athol Kemball
Dept. of Astronomy & NCSA
University of Illinois at Urbana-Champaign
Theme Lead: Physics & Astronomy (NCSA)

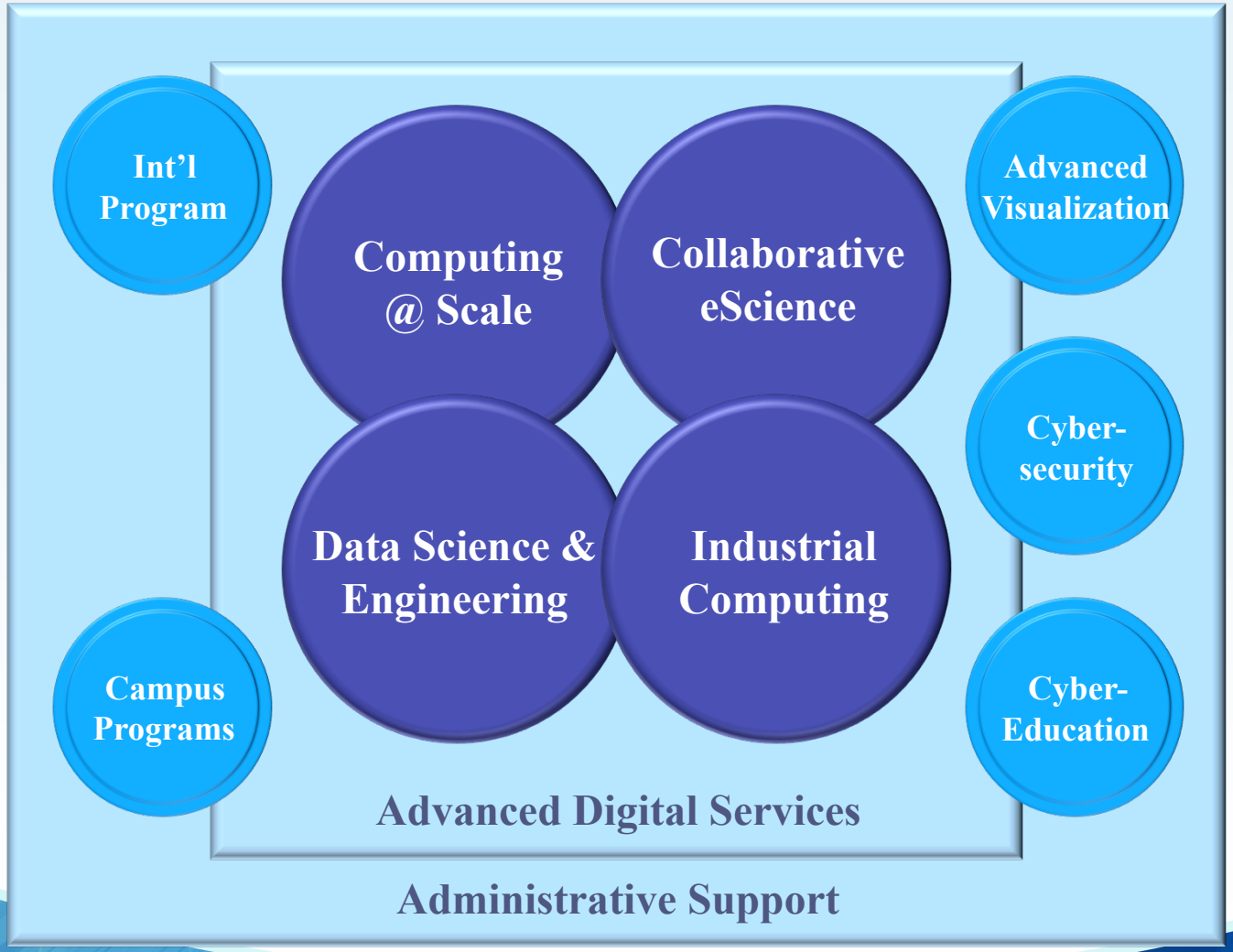


- NCSA: R&D unit of the U. of Illinois at Urbana-Champaign
  - One of original five supercomputing centers funded by the National Science Foundation (NSF) in 1985.
    - **Mission**: Provide state-of-the-art computing capabilities (hardware, software, HPC expertise) to scientists and engineers in the US
- Current profile:
  - Approximately 200 staff (160+ technical/professional staff), two facilities
  - Operating NSF's most powerful computing system: Blue Waters
  - Managing NSF's national cyberinfrastructure: XSEDE
  - Private Sector Program: over 2 dozen industrial partners

Source: Thom Dunning

- What does NCSA do ?
  - **Execute large projects** in compute- and data-intensive science (CDSE):
    - Build and operate large-scale national compute & data resources (e.g. Blue Waters).
    - Common national cyberinfrastructure (e.g. XSEDE).
    - Domain-specific CDSE project execution (e.g. DES, LSST).
  - **Leading-edge research and education** at national level in CDSE.
  - **Promote economic development** at state and federal level.
  - **Foster interdisciplinary CDSE research** within the University of Illinois in partnership with faculty, research scientists, & students.

NCSA Major Activities



Source: Thom Dunning

- A broader partnership between NCSA and CC IN2P3
  - **Many overlapping interests:**
    - Data-intensive science and technology
    - Computing at scale: science and technologies
    - Storage technologies
    - Facility operations
    - ….
    - Open to any proposals

The National
DATA SERVICE
Home
About
Projects
News
Get Involved
NDS CONSORTIUM WORKSHOP
Join us in Austin March 25-27 where we will focus on establishing the governance for the NDS Consortium, growing the pilot projects, and sharing the latest on funding efforts.
READ MORE ABOUT THE WORKSHOP →
NCSA

The National Data Service (NDS)

The National Data Service (NDS) is an emerging vision for how scientists and researchers across all disciplines can find, reuse, and publish data. It builds on the data archiving and sharing efforts already underway within specific communities and links them together with a common set of tools designed around the following capabilities:

Search

The NDS will allow users to easily search for data across disciplinary boundaries. As users hone in on data of interest, they can easily switch to discipline-specific tools.

Publish

The NDS will connect users to tools for building and sharing collections of data. It will help users find and deliver data to the best repository for data-publishing.

Link

The NDS will create robust connections between data and published articles. When researchers reference an article, they have ready access to the underlying data.

Reuse

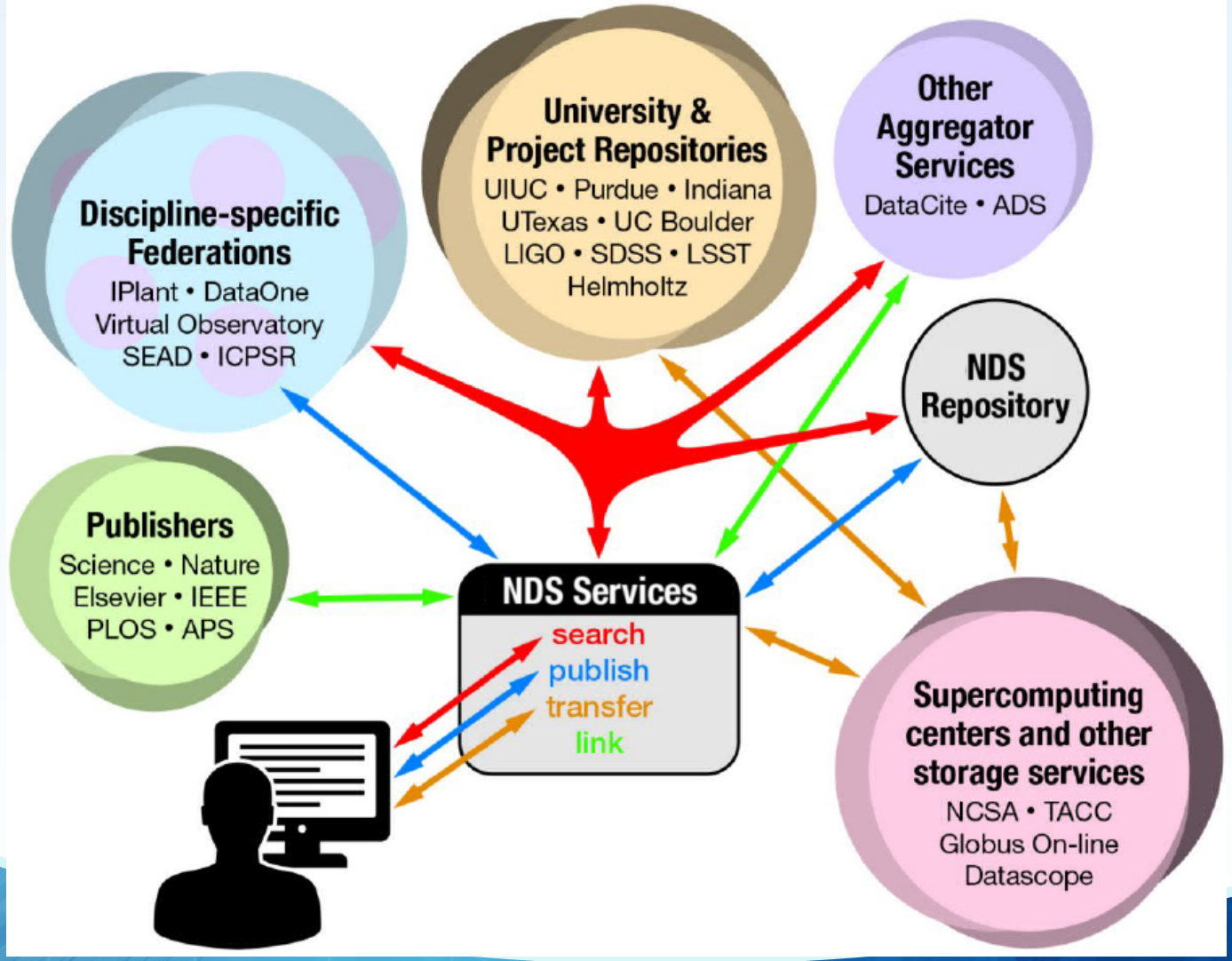
The NDS will not only provide access to data for download, it will provide tools for transferring data to processing platforms or allow analysis to be attached to the data.

Description of NDS goals and capabilities

- https://youtu.be/BPT1FNFAvnc

- Core NDS capabilities:
  - Searching for data
  - Moving data, between repositories and computing platforms
  - Sharing and publishing data
  - Creating, maintaining, and tracking links between data and literature.

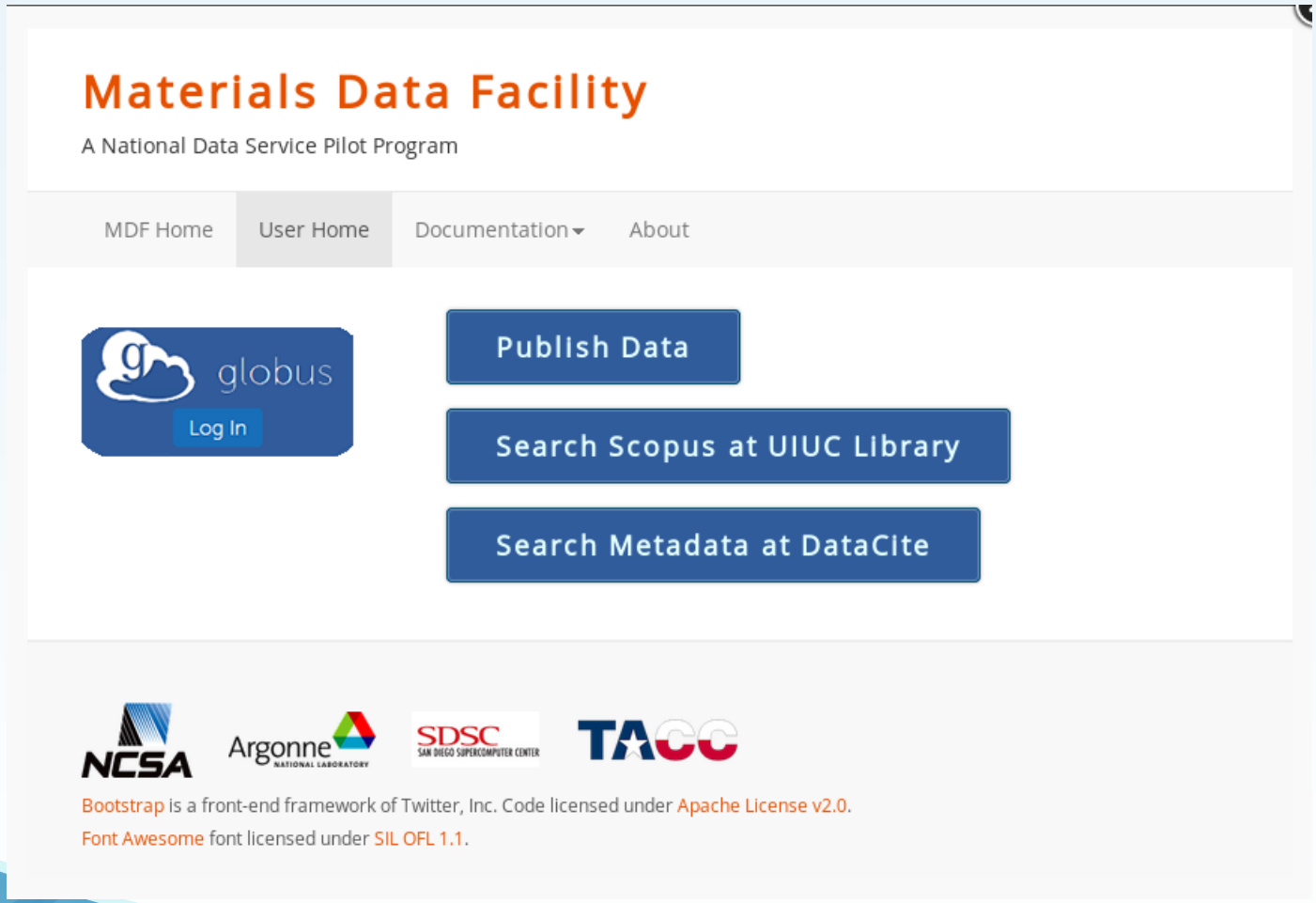
NDS architecture



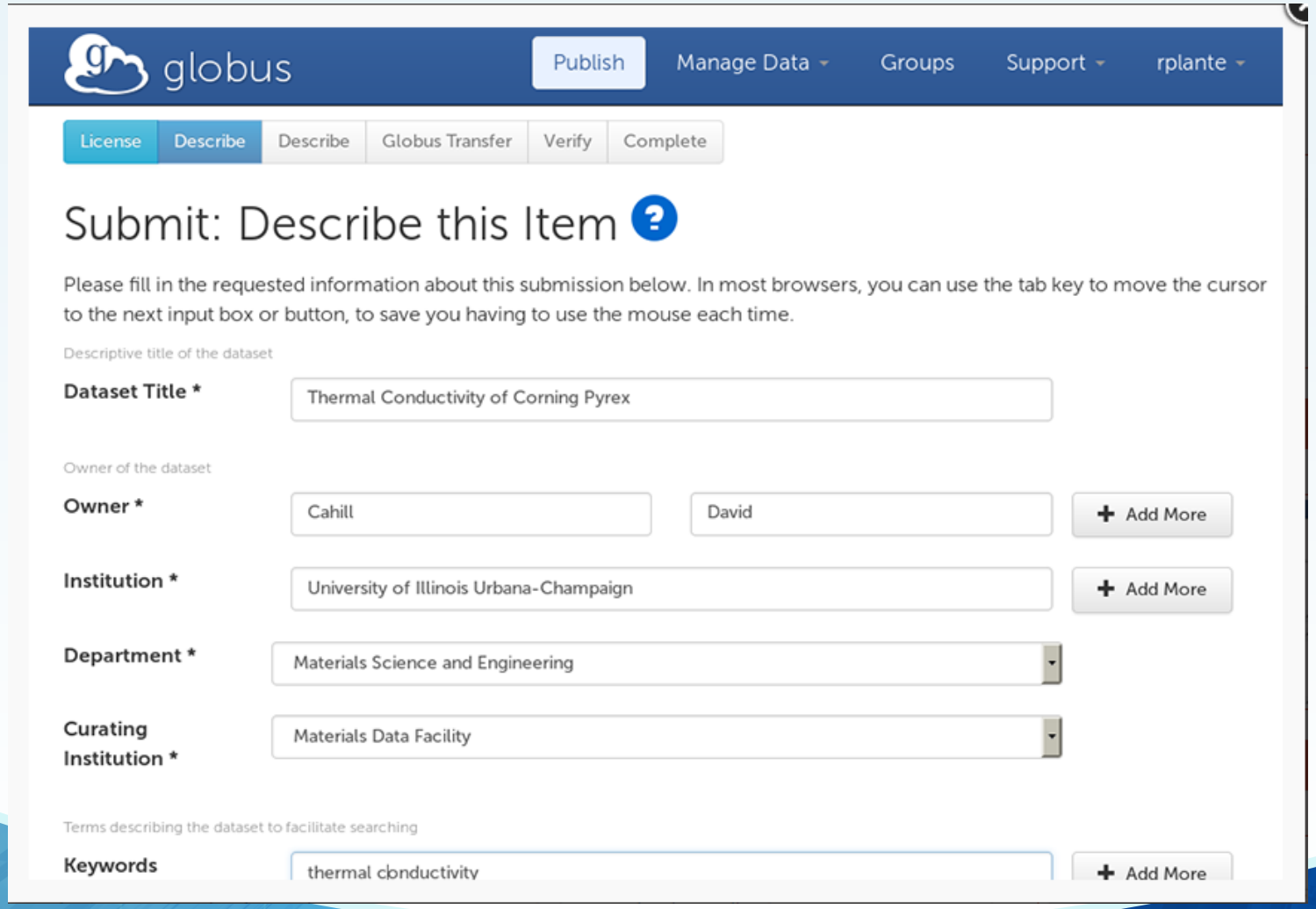
Pilot projects

- OLDRADA: Data-Literature Linking
  - *Linking data and literature together within an open, flexible standard*
- yt-Hub: Advanced Data Presentation and Analysis
  - *Visualization and analysis using yt.*
- Materials Data Facility: An NDS Publishing Protocol for Materials Science
  - *Generic NDS publishing portal for materials science*

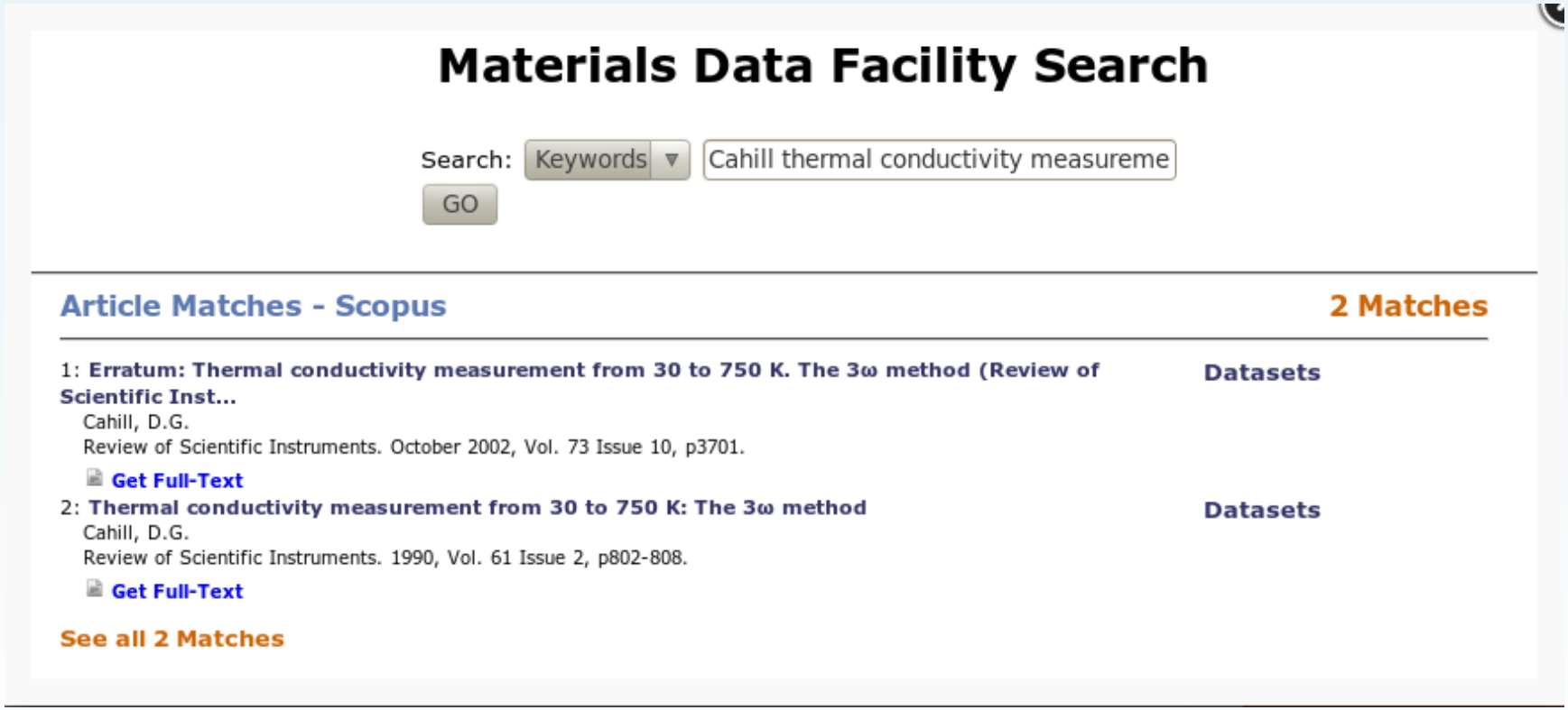
MDF: User home

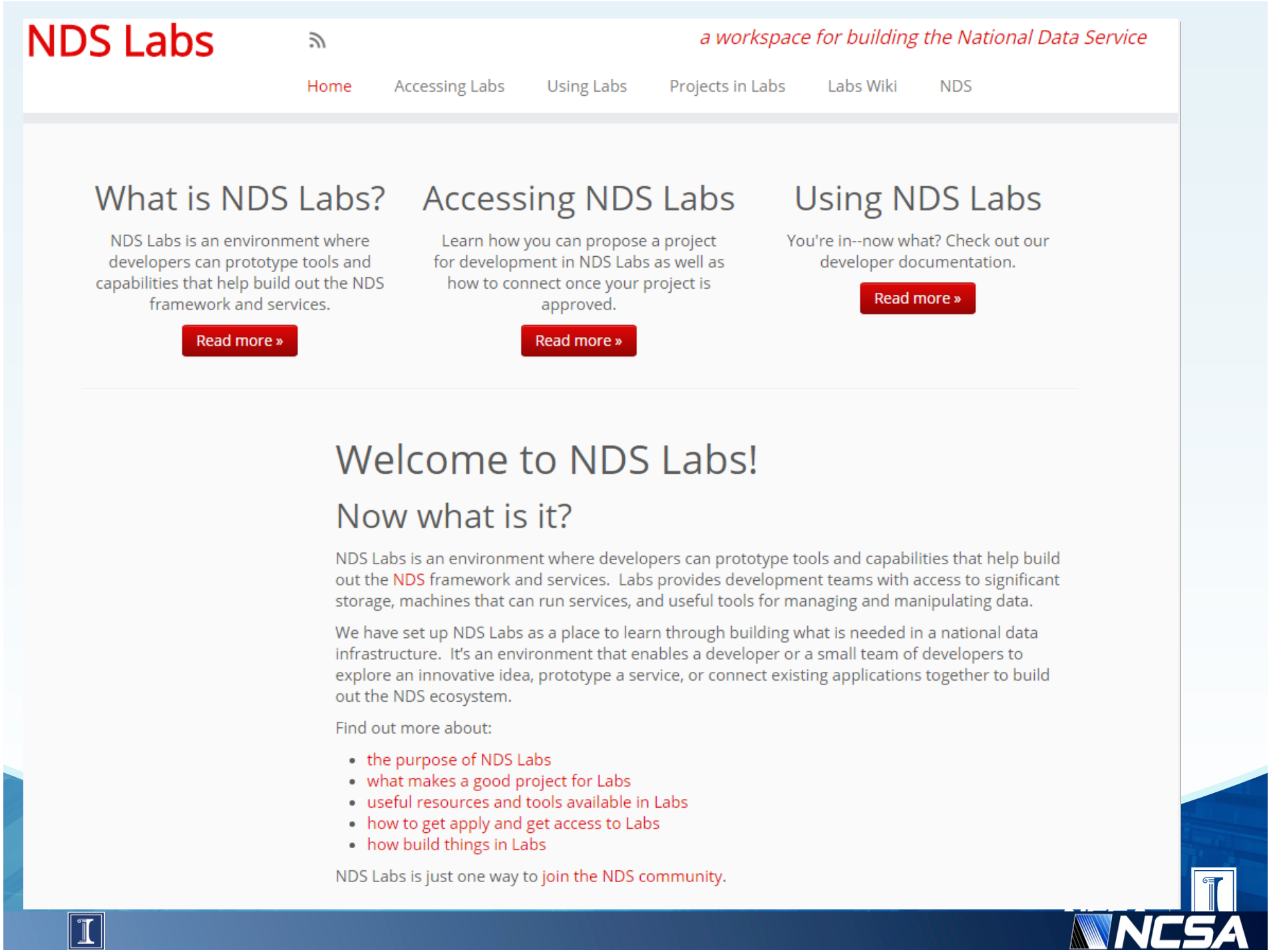


MDF: Publish

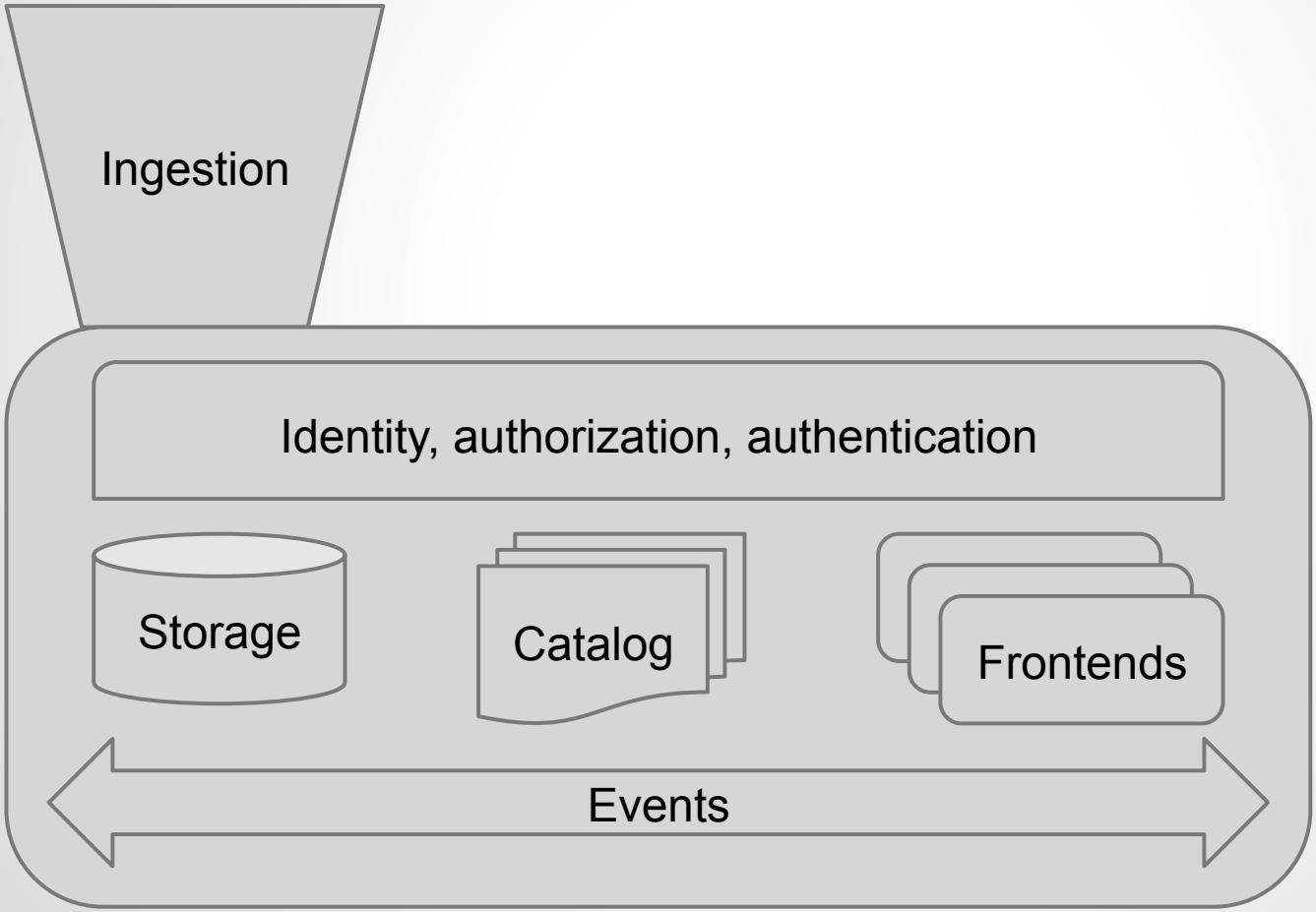

globus
Publish
Manage Data
Groups
Support
rplante
License
Describe
Describe
Globus Transfer
Verify
Complete
Submit: Describe this Item
Please fill in the requested information about this submission below. In most browsers, you can use the tab key to move the cursor to the next input box or button, to save you having to use the mouse each time.
Descriptive title of the dataset
Dataset Title *
Thermal Conductivity of Corning Pyrex
Owner of the dataset
Owner *
Cahill
David
Add More
Institution *
University of Illinois Urbana-Champaign
Add More
Department *
Materials Science and Engineering
Curating Institution *
Materials Data Facility
Terms describing the dataset to facilitate searching
Keywords
thermal conductivity
Add More

MDF: Search

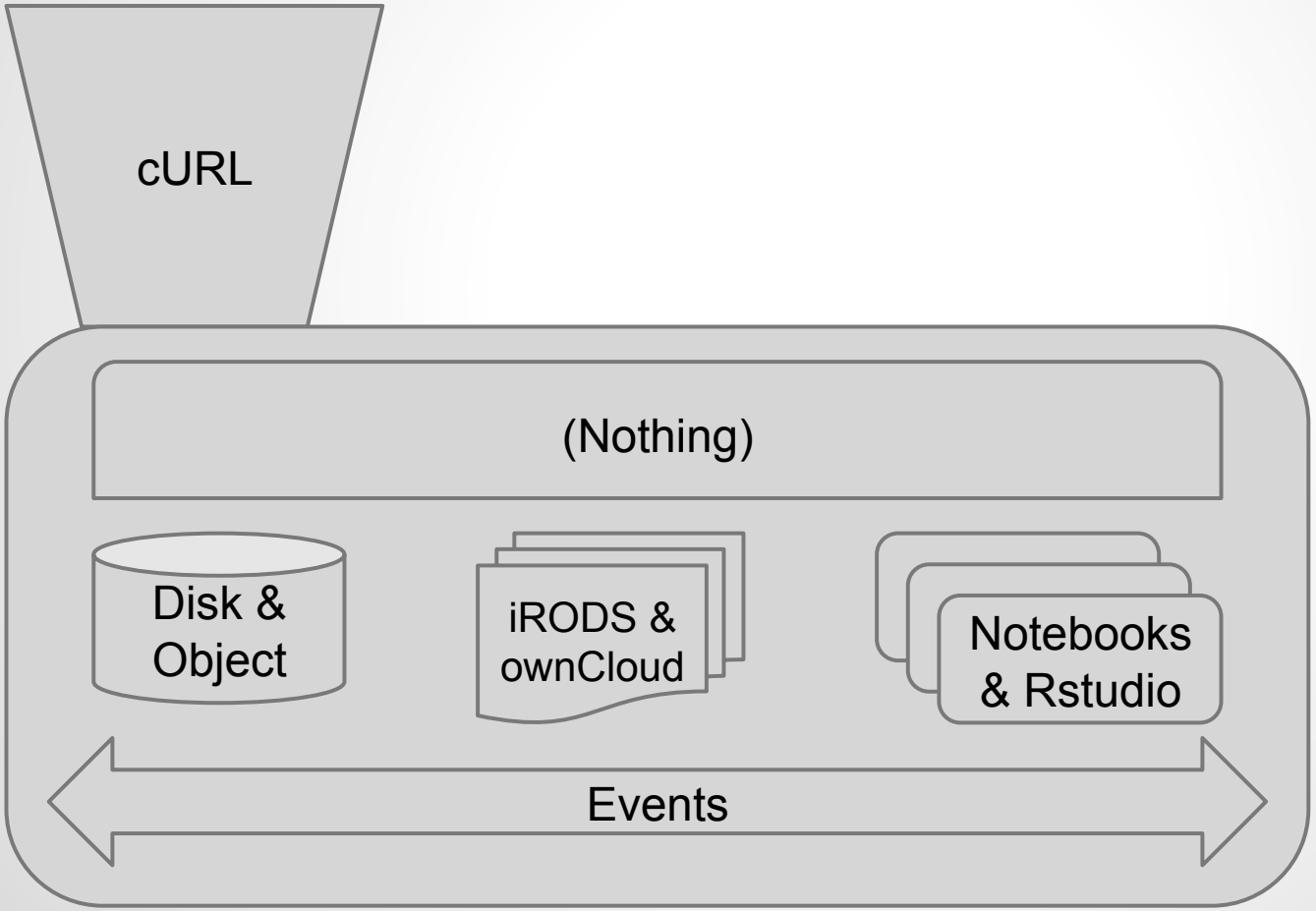

Materials Data Facility Search
Search:
Keywords
Cahill thermal conductivity measureme
GO
Article Matches - Scopus
2 Matches
1: Erratum: Thermal conductivity measurement from 30 to 750 K. The 3ω method (Review of Scientific Inst...
Cahill, D.G.
Review of Scientific Instruments. October 2002, Vol. 73 Issue 10, p3701.
Get Full-Text
Datasets
2: Thermal conductivity measurement from 30 to 750 K: The 3ω method
Cahill, D.G.
Review of Scientific Instruments. 1990, Vol. 61 Issue 2, p802-808.
Get Full-Text
Datasets
See all 2 Matches


NDS Labs
a workspace for building the National Data Service
Home
Accessing Labs
Using Labs
Projects in Labs
Labs Wiki
NDS
What is NDS Labs?
NDS Labs is an environment where developers can prototype tools and capabilities that help build out the NDS framework and services.
Read more »
Accessing NDS Labs
Learn how you can propose a project for development in NDS Labs as well as how to connect once your project is approved.
Read more »
Using NDS Labs
You're in--now what? Check out our developer documentation.
Read more »
Welcome to NDS Labs!
Now what is it?
NDS Labs is an environment where developers can prototype tools and capabilities that help build out the NDS framework and services. Labs provides development teams with access to significant storage, machines that can run services, and useful tools for managing and manipulating data.
We have set up NDS Labs as a place to learn through building what is needed in a national data infrastructure. It's an environment that enables a developer or a small team of developers to explore an innovative idea, prototype a service, or connect existing applications together to build out the NDS ecosystem.
Find out more about:
the purpose of NDS Labs
what makes a good project for Labs
useful resources and tools available in Labs
how to get apply and get access to Labs
how build things in Labs
NDS Labs is just one way to join the NDS community.
NCSA

NDS Labs: Experimental environment

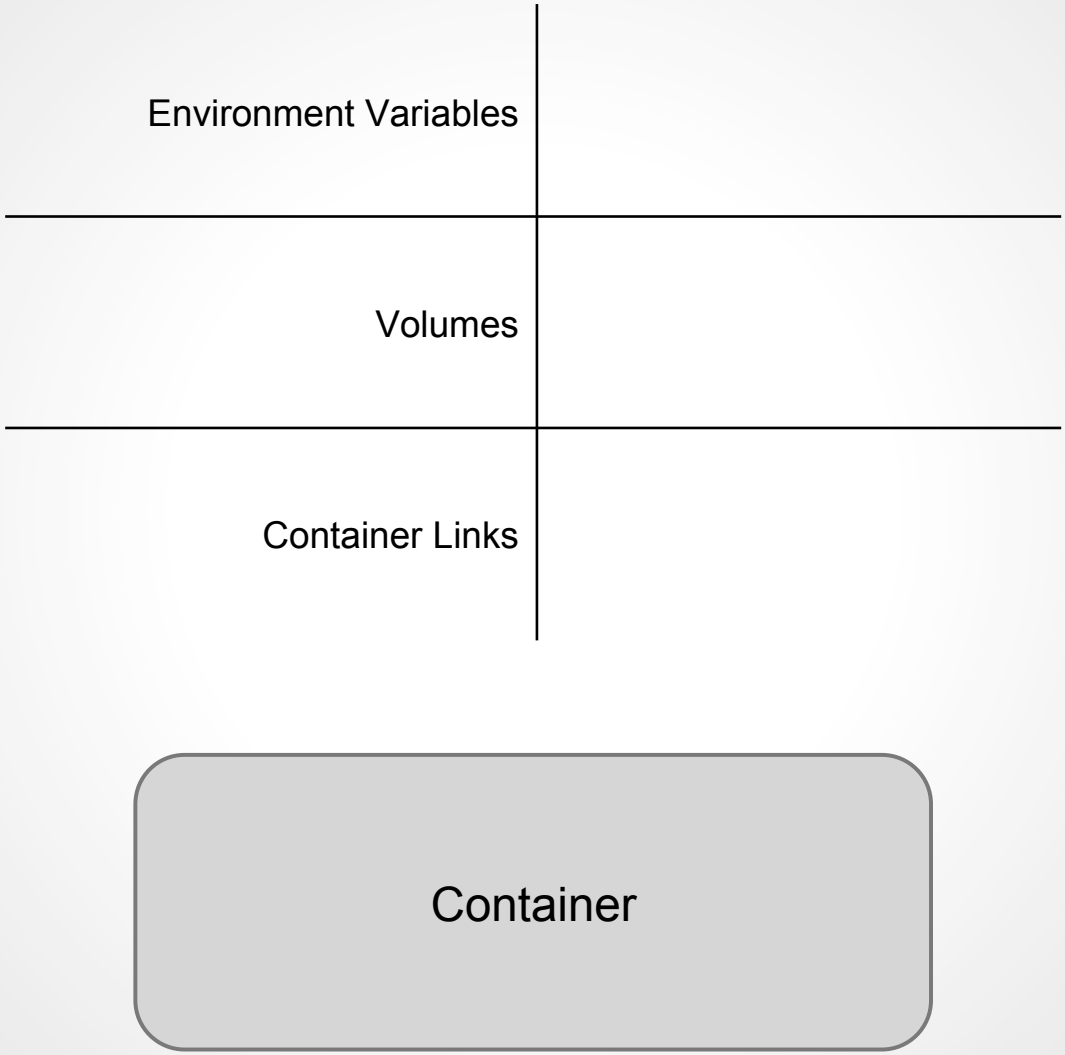


“Reference” Implementation



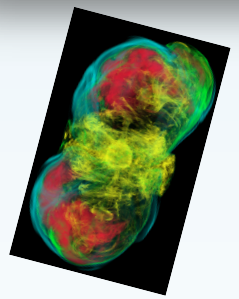
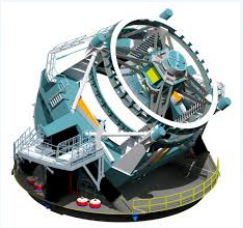
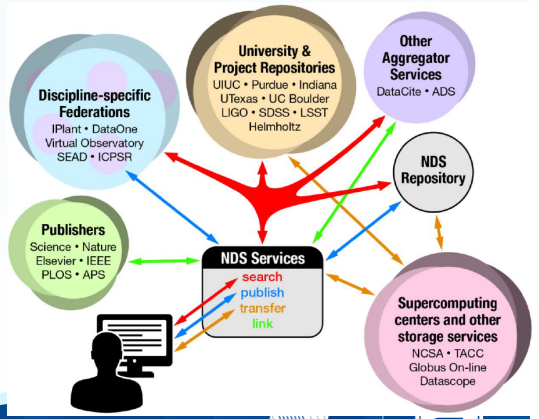
http://bitbucket.org/nds-org/
http://bitbucket.org/nds-org/nds-labs/
http://github.com/nds-org/

Container Recipes
Cluster Orchestration System
Container Orchestration System


Environment Variables
Volumes
Container Links
Container

Environment Variables	Dataset information, federation information
Volumes	Persistent storage volumes and ephemeral storage volumes
Container Links	Databases, transformers, extractors, services

Summary

- NCSA would like a broad partnership with CC IN2P3.

- We are open to different areas of collaboration.
  - Data-intensive computing (e.g. NDS).
  - Advanced computation
  - Facility operation.
  - Infrastructure technologies.