



# Prototype Data Resource Registry for the GHG Community

NIST MML Office of Data and Informatics

Gretchen Greene – Group Lead Data Science

[gretchen.greene@nist.gov](mailto:gretchen.greene@nist.gov)

Collaboration with: Tamae Wong, Anna Karion, James Whetstone, Raymond Plante, Kim Tryka, Logan Mitchell, Conor Gately, Francesca Hopkins, Lorraine Sugar, Thomas Lauvaux, Kevin Gurney, Marcus Newrock



# Building Data Resource Registries

## Goal:

Facilitate Scholarly Discovery and Navigation to key data resources and information to support synthesis and validation of GHG science in a *federated* community

## Data Resources provide metadata, e.g. :

- Datasets – public, research, and private restricted
- Services – programmatic, custom, etc.
- Portals – Data Archives, Community, Research Project
- Organizations – Gov, Industry, Academic, Groups, Project
- Documentation - publication resources, key
- Tools/Software – Web Tools, Services, Applications

## Proven Model in operations:

- IVOA (International Virtual Observatory Alliance)
- NIST IMMR (Global Metrology network)
- Materials Genome Initiative (MGI)

**Focused Domain Expertise - reliability of the information – data quality, scholarship**



# Opportunity for FAIR Data Science

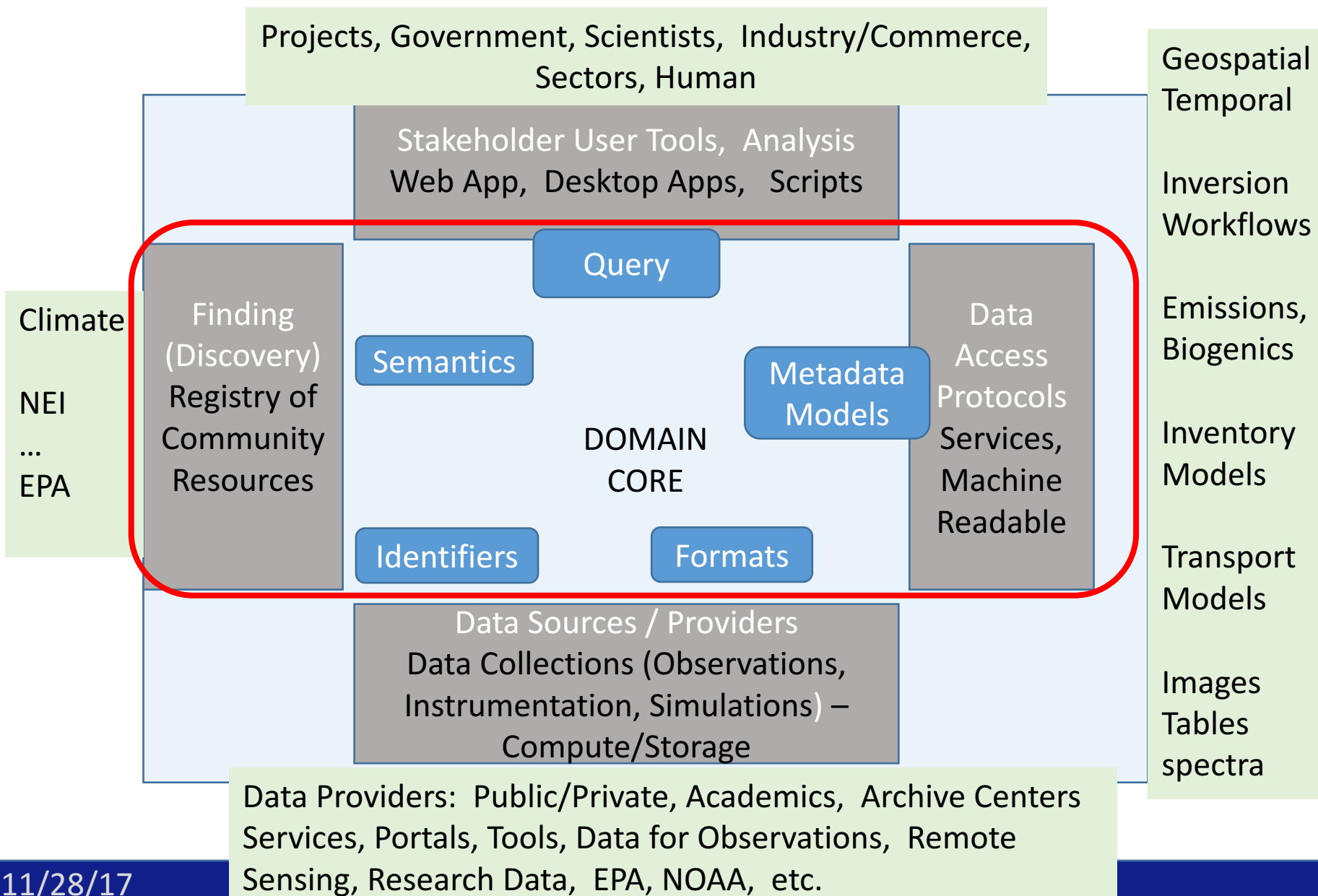
- FAIR – relatively new US legislative efforts to extend the *open data* initiatives to improve upon scientific investment and validation
  - **Findable** – Discoverable resources which represent the domain space information
  - **Accessible** – Services for data access (tools can access data) , protocols for machine use
  - **Interoperable** – agreed upon standards/protocols for describing (metadata) datasets
  - **Re-usable** – tools and analysis recognize the data formats, standards and deliver results to query

NOTE – FAIR does not preclude flexibility in use of methods or standards implementation

<https://www.nature.com/articles/sdata201618>

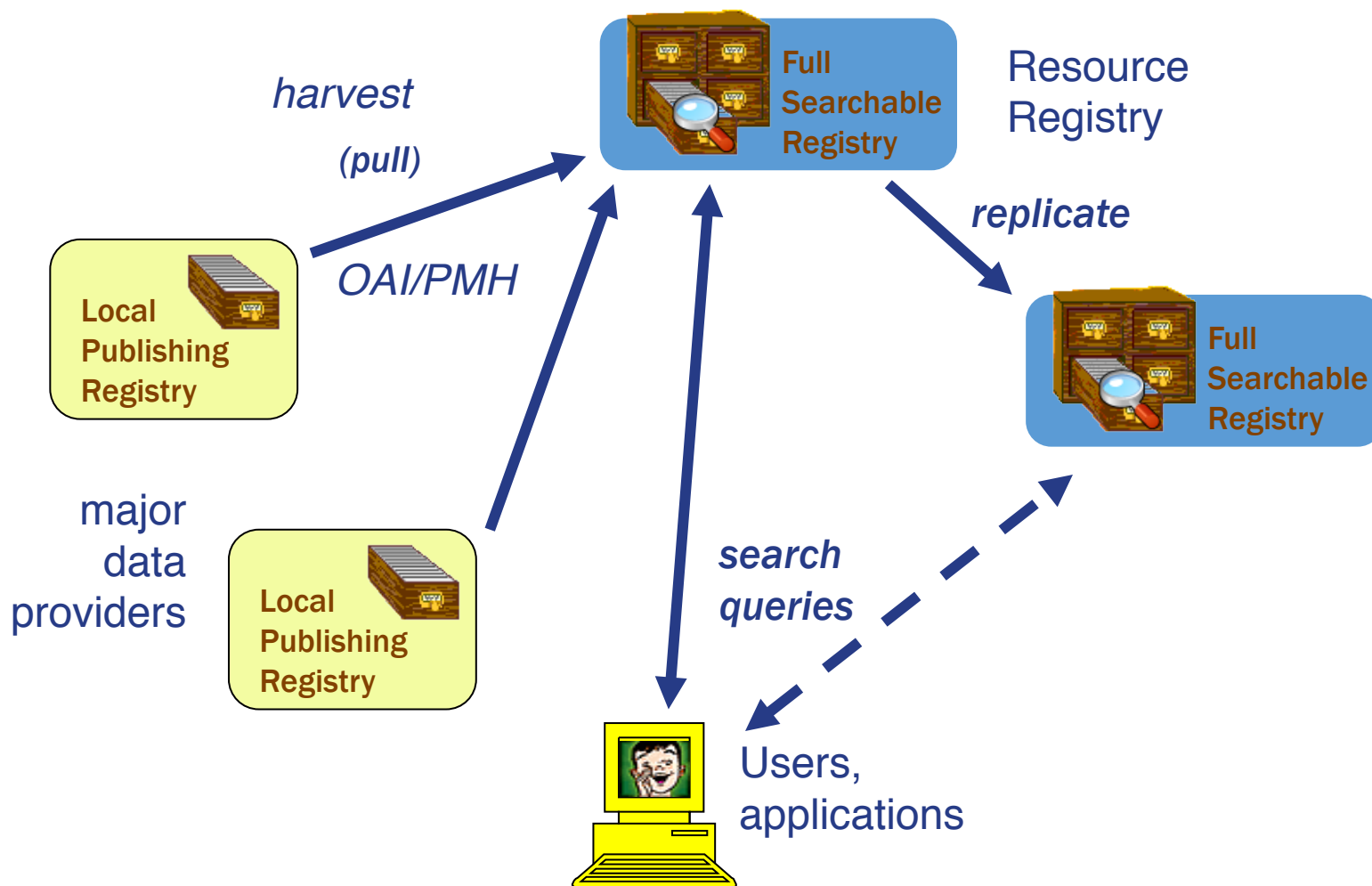


# GHG Registry Context -example FAIR Framework





# Federated Architecture





# Resource Registry in Practice

NAVO Directory

Find Astronomical Data Resources

Enter terms in the text box that describe the type of data you are looking for, or use the filters below to browse data by facet. Results will show catalogs and data collections that have these terms as part of their descriptions. Note that the search interface ignores slashes by default, so when searching for full or partial IDs you may want to put the search string in quotes (e.g. "vvo/mast.stsci.edu/bell").

All Fields binary stars

Examples: quasar, A01, binary stars, Chandra, GALEX, far ultraviolet

Search

Use the Data Discovery Tool to search and view data for a particular object or position.

Limit your search

Subject

Publisher

Resource Type

Content Level

Instrument

Facility

Validation Level

Rights

You searched for: binary stars

Start Over

Previous | 1 - 10 of 840 | Next

Sort by relevance | 10 per page

1. Speckle observations of binary stars (McAlister 1993)

ID: vvo/COB-MOAR/JAJ1001630

Title: Speckle observations of binary stars (McAlister 1993)

Short Name: JAJ1001630

Date: 13 Sep 2016 16:00:00

Publisher: COB

Description: Speckle interferometric observations from the Canada-France-Hawaii 3.6m and Cerro Tololo 4 m telescopes are reported for 1123 stars selected from the Yale Bright Star Catalogue in a continuing effort to detect new binaries among the bright stars. Thirty-two previously unresolved binaries have been detected, including companions to  $\lambda$  UMa and  $\delta$  S Mon. Measures of 107 previously resolved systems, many of which resulted from earlier speckle observations, are also presented. No evidence of duplicity within a specific (in Defaulter) window of detectability was found for 884 bright stars. These observations combined with two previously published surveys represent the inspection of 2088 stars, representing 23% of the members of the Bright Star Catalogue. Many of the systems discovered earlier have shown significant orbital motions, and we present preliminary orbital elements for these systems. (Published online 2016 September 13, 16:00:00. See the full-text version of this article at <https://doi.org/10.1088/1539-3633/1001630>.)

<http://vao.stsci.edu/keyword-search/>

Secure | <https://registry.nist.gov>

Home Services Login Help Contact

Greenhouse Gases Data Resource Registry

A Collaboration of GHG Research Organizations

SEARCH FOR RESOURCES ADD YOUR RESOURCE

Find GHG Resources

This system allows for the registration of resources, bridging the gap between existing resources and the end users. The Greenhouse Gases Data Resource Registry functions as a centrally located service, making the registered information available for research to the global community.

This is being developed at the National Institute of Standards and Technology (NIST) and is made available to collect information from the global community. Please do not enter any proprietary data into this system.

Home Page

Services

Search for resources

Add your resource

Login

Help

Public <https://registry.nist.gov/>

NIST

Home Services Login Help Contact

Materials Resource Registry

Part of the Materials Genome Initiative

SEARCH FOR RESOURCES ADD YOUR RESOURCE

Find Materials Data

This system allows for the registration of materials resources, bridging the gap between existing resources and the end users. The Materials Resource Registry functions as a centrally located service, making the registered information available for research to the materials community.

This is being developed at the National Institute of Standards and Technology and is made available to solicit comments from the Materials Science community. Please do not enter any proprietary data into this system.

Home Page

Services

Search for resources

Add your resource

Login

Help

Contact

<https://materials.registry.nist.gov/>

NIST

Home Services Login Help

Materials Resource Registry

Search for Resources

Enter keywords, or leave blank to retrieve all records

All Resources Organizations Data Collections Databases Services Informational Sites

Search criteria used (Clear all):

Type

TYPE

Organization (503)

Collection (246)

Dataset (32)

Service (4)

Software (127)

Web Site (22)

ORIGIN OF DATA

Peter Brimmer, Feng Gähler - Pofit

<https://www.pofit.net/ku/ku.php?id=101>

Subject keywords: molecular dynamics, empirical potentials, pair potentials, GCM and two-band GCM, ADI electrostatics (double dipole interactions), MEAM, Stillinger-Weber, Tersoff and modified Tersoff

"Pofit" is a free implementation of the force-matching algorithm to generate effective potentials from ab-initio data.

PRISMS-PF

Predictive Integrated Structural Materials Science (PRISMS) Center

<https://github.com/PrismsCenter/prisms-pf>

Subject keywords: phase field, microstructures, finite element analysis, microstructural evolution

"PRISMS-PF" is a powerful, massively parallel finite element code for conducting phase field and other related simulations of microstructures.

Bureau International des Poids et Mesures

Home Services Dashboard Logout Help Contact

International Metrology Resource Registry

SEARCH FOR RESOURCES ADD YOUR RESOURCE

Find Resources

This system allows for the registration of resources, bridging the gap between existing resources and the end users. The International Metrology Resource Registry functions as a centrally located service, making the registered information available for research to the global community.

This is being developed at the Bureau International des Poids et Mesures and is made available to solicit comments from the global community. Please do not enter any proprietary data into this system.

Home Page

Services

Search for resources

Add your resource

Dashboard

Logout

Contact

<http://imrr.bipm.org/>

Internal only



# Next Steps

- Collect Feedback on Registry System value
  - Maintain public NIST registry and work with external deployments where possible. Encourage collaboration and feedback.
  - Identify metadata, granularity and usefulness to the GHG groups. Features of use, e.g. programmatic access
- Continue to Collaborate and support GHG Community for Data Interoperability
  - Share best practices and standards for interoperability
- NIST is participating in open source advancement
- Facilitate development of a federated architecture of scientific information in support of the GHG community
  - Distributed and shared expertise across data science space