# Catalyzing an NDS ecosystem

lan Foster
Director, Computation Institute
Argonne National Laboratory
& The University of Chicago
@ianfoster
ianfoster.org



We aim to reduce sources of friction that currently impede data-driven science ...



## Discipline-specific **Federations**

IPlant • DataOne Virtual Observatory SEAD • ICPSR

## **University & Project Repositories**

UIUC • Purdue • Indiana UTexas • UC Boulder LIGO • SDSS • LSST Helmholtz

## **Other** Aggregator Services

DataCite • ADS

**FoNDS** Repository

### **Publishers**

Science • Nature Elsevier • IEEE PLOS • APS

**FoNDS Services** 

search publish transfer link

2101010100

Supercomputing centers and other storage services

> NCSA • TACC Globus On-line Datascope

1011010101010 NationalDo 0101001010

# Let's make a few simple things easy

## For example:

- "Dr. U shares data with colleagues"
- "Project V organizes data across partners"
- "Institution W sets up an institutional repository"
- "Dr. X searches across a dozen databases"
- "Agency Y creates a secure data commons"
- "Publisher Z links papers to research repositories"



# Three elements of an NDS ecosystem

Provide basic services for data sharing, publication, discovery, reuse; authentication, authorization, profile, group management; cross-domain search

Vanilla services

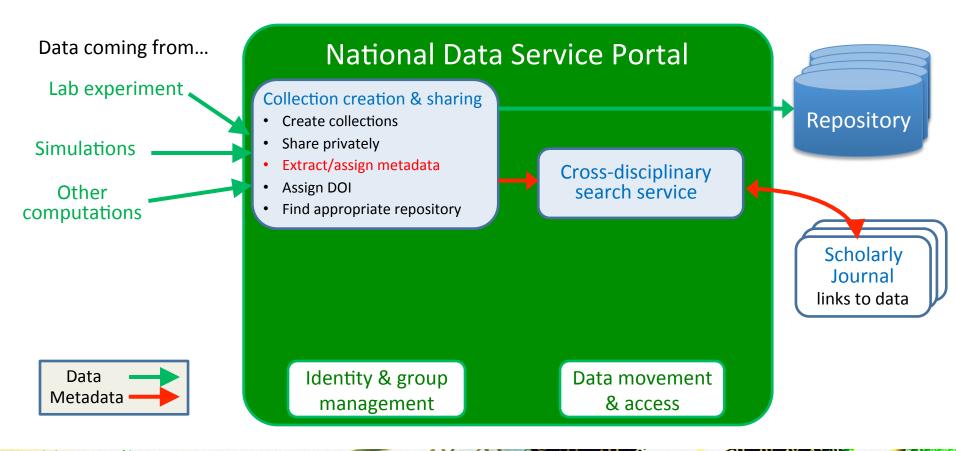
Pluggable framework

Identify interfaces and conventions to allow disciplinary extensions, replacements, and contributions to NDS infrastructure and services

### Foundational infrastructure

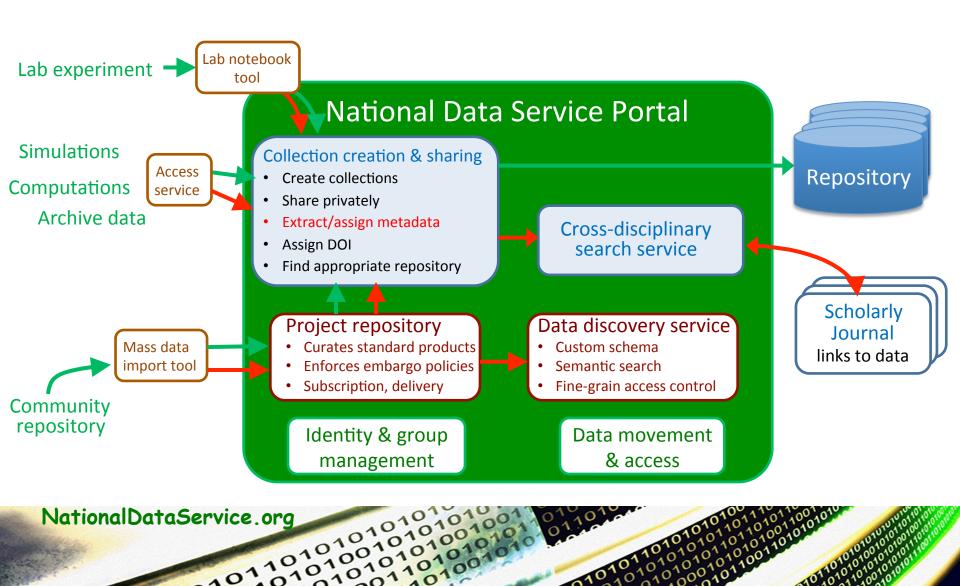
Deploy initial infrastructure to allow communities to jump-start disciplinary services: Scratch storage, repository storage, cloud computing services for hosting portals and services

# A generic data solution based on rough consensus and running code



National Data Service.org

# Organized as an extensible framework



# We want your help



Use cases
Infrastructure
Communities
Services