



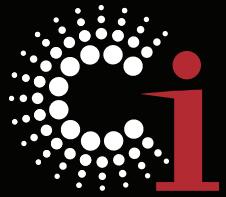
Hybrid Strategies for Research Data Management

Vas Vasiliadis, Computation Institute

vas@ci.uchicago.edu



computationinstitute.org

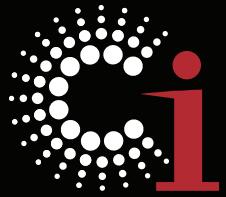


The Computation Institute

= UChicago + Argonne

= Cross-disciplinary nexus

= Home of the Research Cloud



Molecular biology

Cosmology

Metagenomics

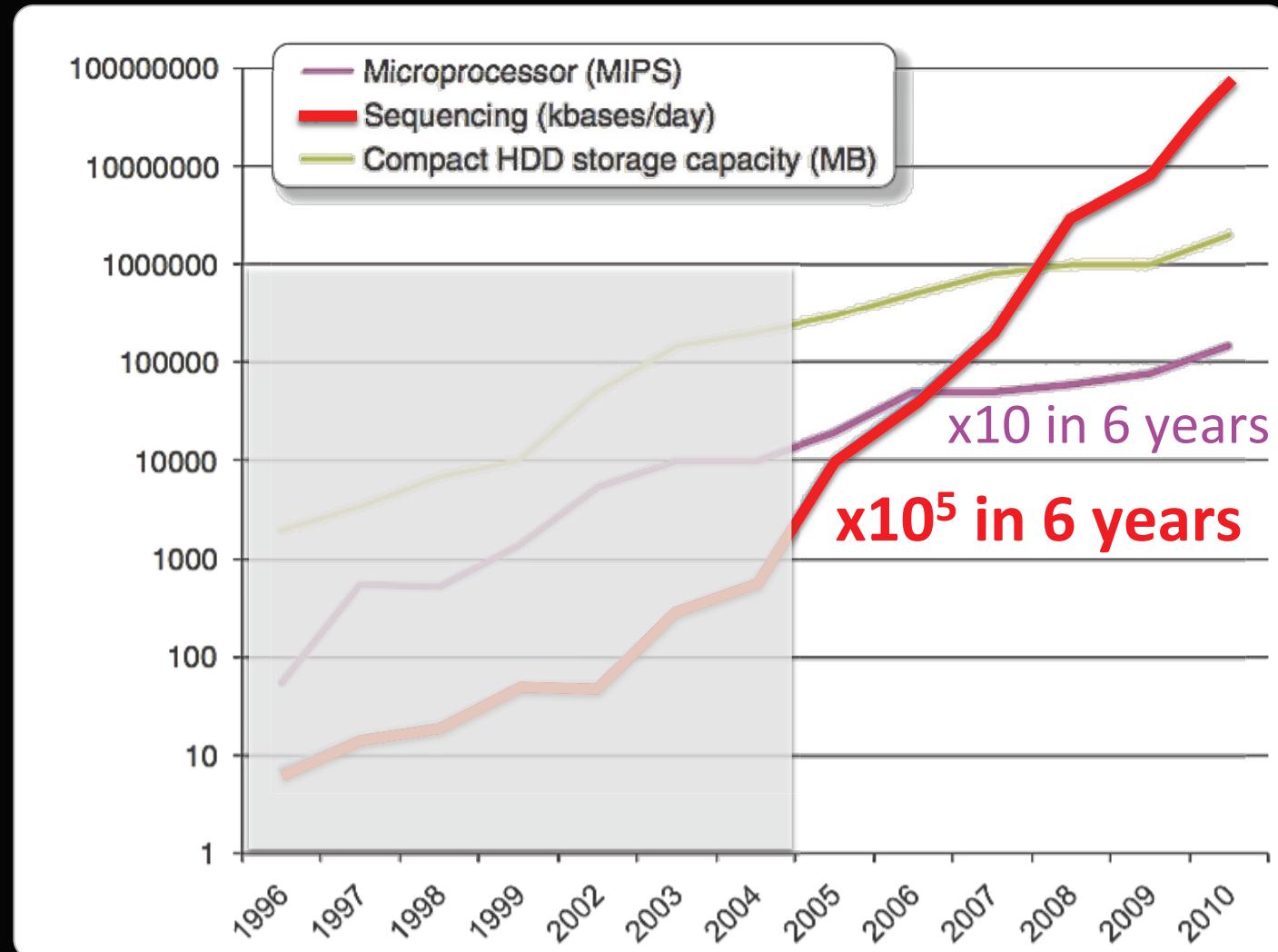
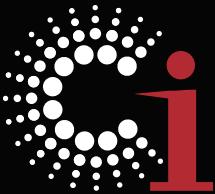
Economics

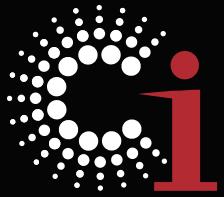
Genetics
High energy physics

Climate change

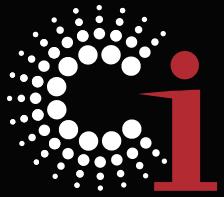
Linguistics

Visual arts

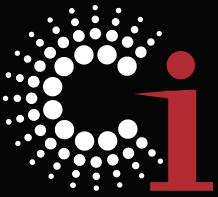




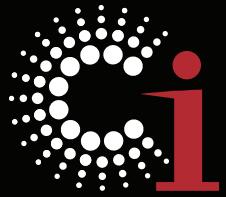
1 PB data in last experiment
Accessed by 800 scientists
worldwide



1.2 PB of climate data
Delivered to 23,000 users

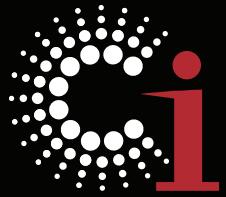


computationinstitute.org

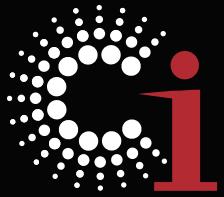


We have exceptional
infrastructure for the 1%

How can the 99%
manage this?

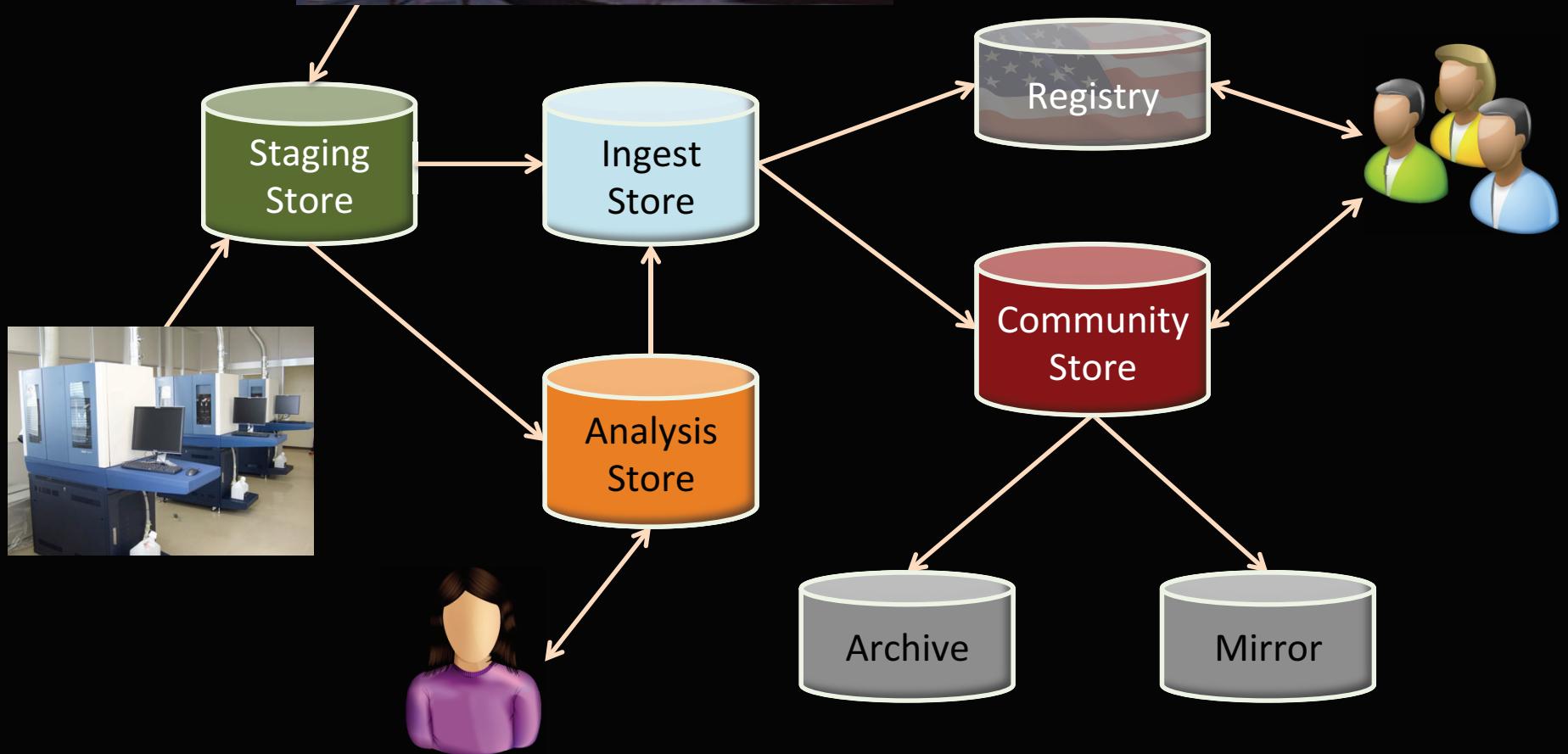


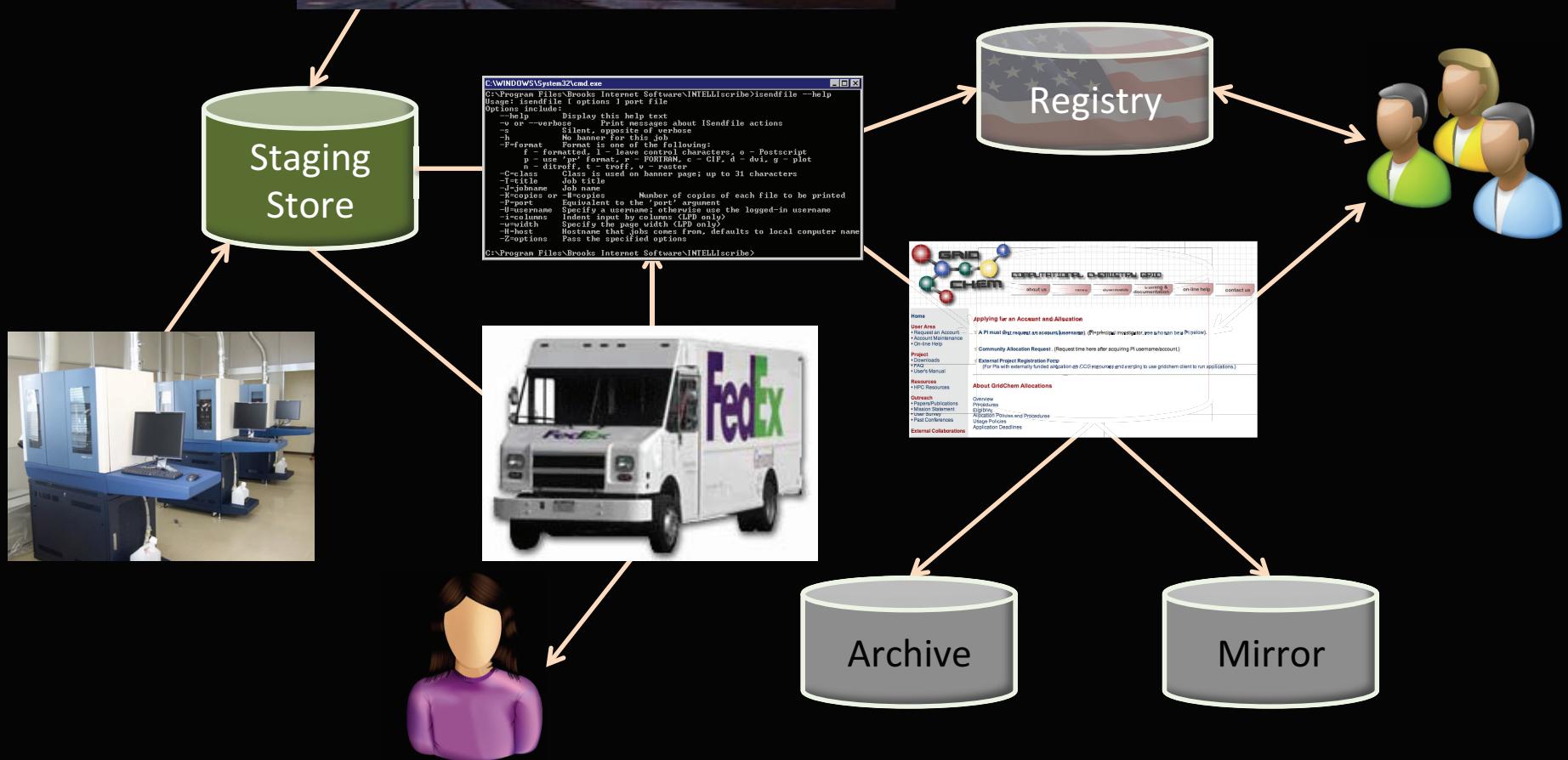
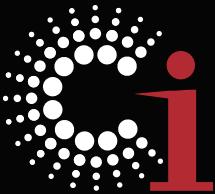
What would a “dropbox for science” look like?

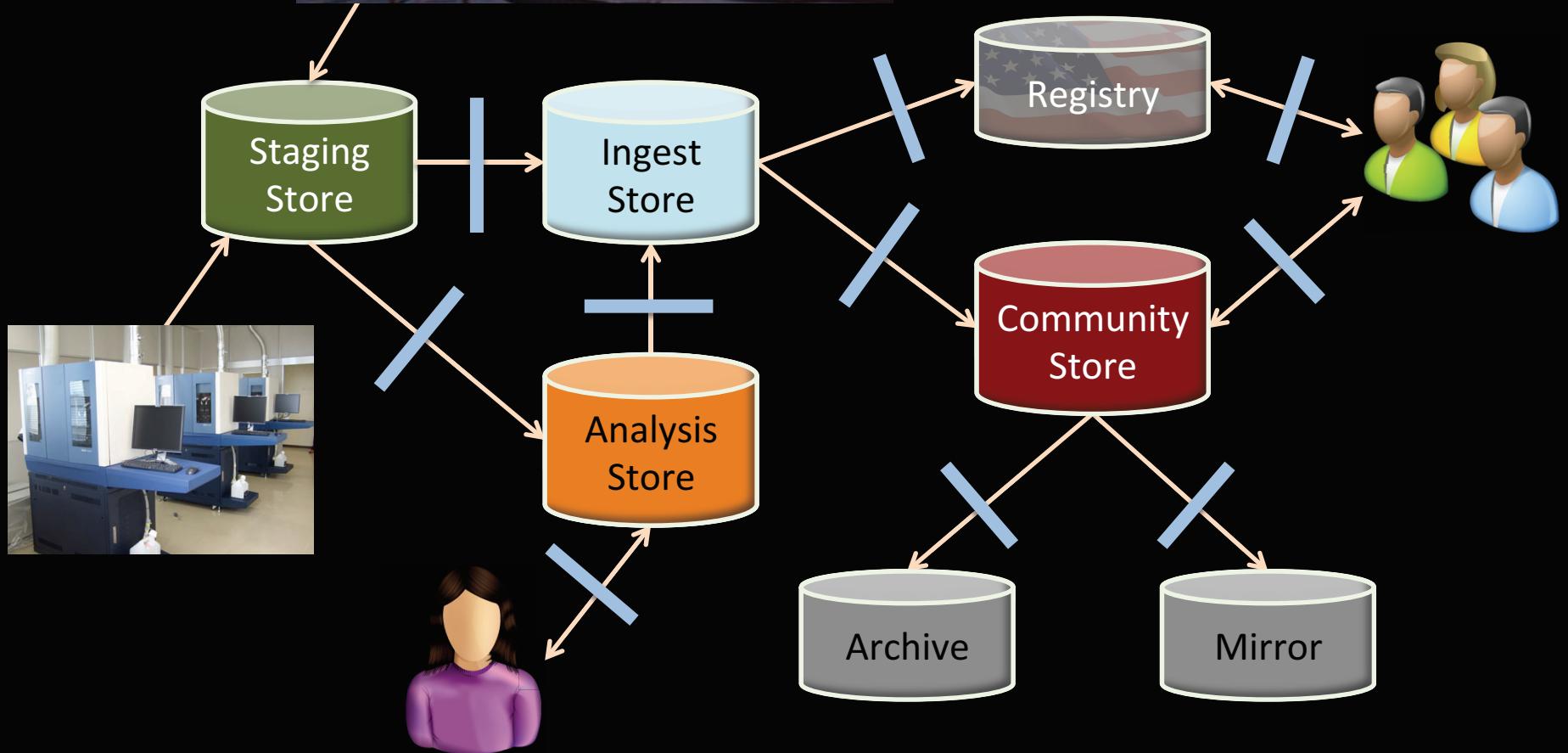
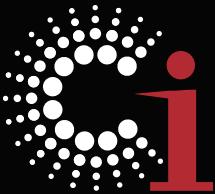


- Collect
- Move
- Replicate
- Share
- Analyze
- Catalog
- Publish
- Search
- Archive
- Backup

...among distributed research groups



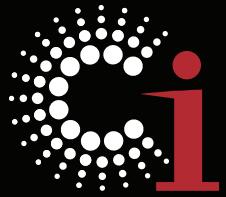




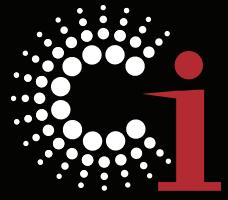


- Collect
- Move
- Replicate
- Share
- Analyze
- Catalog
- Publish
- Search
- Archive
- Backup

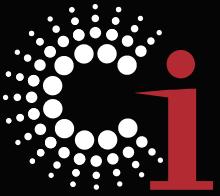
-as-a-Service



Security
Privacy
Reliability
Scalability
Control



A great user experience



www.globusonline.org

Research Data Management-as-a-Service



globus
online

Globus
Transfer

Globus
Storage

Globus
Collaborate

Globus
Catalog

Globus Integrate (Globus Nexus, Globus Connect)

SaaS

PaaS

Supported by

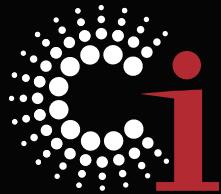


THE UNIVERSITY OF
CHICAGO

Argonne
NATIONAL LABORATORY

Mirror

computationinstitute.org



Communities using Globus

XSEDE

Extreme Science and Engineering
Discovery Environment

NERSC



APS
physics



Cal

NCSA

**Carnegie
Mellon
University**

KSU

NGS



 **THE UNIVERSITY
OF AUCKLAND**
NEW ZEALAND

Te Whare Wānanga o Tāmaki Makaurau


Information Sciences Institute




EMORY


ESnet
Energy Sciences Network


BERKELEY LAB
Lawrence Berkeley National Laboratory


**INDIANA
UNIVERSITY**




**Los Alamos
NATIONAL LABORATORY**
EST. 1943


CORNELL
U N I V E R S I T Y


Ole Miss


CERN

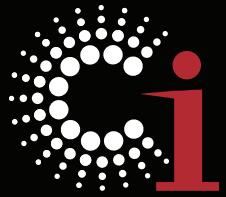

**THE UNIVERSITY OF
CHICAGO**



NEW YORK UNIVERSITY


JÜLICH
FORSCHUNGSZENTRUM


Argonne
NATIONAL LABORATORY
computationinstitute.org



What does it mean for us as IT resource managers?



installers → brokers





developers → integrators



InCommon®



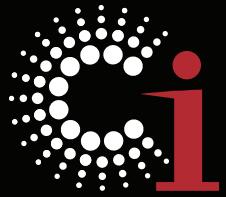
Shibboleth®



CI Logon

MyProxy
Credential Management Service

GSI-OpenSSH



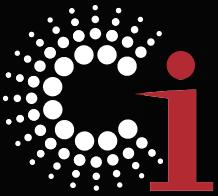
administrators → curators

(of the user experience)

Cloud? What cloud?

1 : 1 : 0

UX : Dev : Ops



myExperiment beta

myExperiment makes it really easy to find, use and share scientific workflows and other files, and to build communities.

All Search

Use myExperiment to...

- Find Workflows
- Find Files
- Share Your Workflows and Files
- Create and Find Packs of Items
- Create and Join Groups
- Find People and Make Friends
- Send Messages
- Get Feedback
- Tag and Rate things
- Write Reviews and Comments
- Build your Profile and

Explore

Find Workflows

About myExperiment
Join the Mailing List
Give us Feedback
For Developers
The myGrid Project
Taverna Workflow Workbench
The BioCatalogue Project
myExperiment Publications

Register or Login:

Username or Email:

Password:

Remember me:

Or use OpenID:
 (eg: name.myopenid.com)

Login

Forgot Password?

myExperiment has over 1250 users, 100 groups, 490 workflows, 130 files and 40 packs

The screenshot displays the nanoHUB.org website, which is an NNC project. The top navigation bar includes links for Home, My HUB, Resources, Members, Explore, About, and Support. A large banner on the right side promotes "ONLINE SIMULATION AND MORE FOR NANOTECHNOLOGY". On the left, a sidebar for "myExperiment" lists various features such as Find Workflows, Share Your Work, Create and Find Items, and Send Messages. The main content area features the "NANOHUB" logo and information about the "FUNDAMENTALS OF NANOELECTRONICS Online Course Spring 2012". To the right, four key features are highlighted: SIMULATE with over 160 tools for nanoelectronics, nanophotonics and more; RESEARCH & COLLABORATE via groups, question board and more; TEACH & LEARN with tool-powered curricula, courses, seminars and more; and SHARE & PUBLISH tools and research through our easy upload process. Below this, a "RESOURCES" section shows a search bar and a list of popular tags including nanoelectronics, course lecture, material science, Illinois, nano/bio, nanotransistors, research seminar, devices, nanophotonics, quantum transport, tutorial, transistors, molecular electronics, nano electro-mechanical systems, NEGF, carbon nanotubes, nanomedicine, education/outreach, UIUC, band structure, ABACUS, atomic force microscopy, quantum dots, MOSFET, nanowires, and More tags. A "FEATURED" section on the right lists items such as the MIT Atomic Scale Modeling Toolkit, Nanotechnology: Considerations for Facility Design, Greg Lush's contributions, and Topics For Introductory Materials Classes.

ONLINE SIMULATION AND MORE
FOR NANOTECHNOLOGY

Home My HUB Resources Members Explore About Support

myExperiment

Use myExperiment to:

- Find Workflows
- Find Files
- Share Your Work
- Files
- Create and Find Items
- Create and Join Groups
- Find People and Friends
- Send Messages
- Get Feedback
- Tag and Rate Tools
- Write Reviews and Comments
- Build your Profile

Learn More >

1 2 3 > || <

RESOURCES

Search

Popular Tags:

- nanoelectronics
- course lecture
- material science
- Illinois
- nano/bio
- nanotransistors
- research seminar
- devices
- nanophotonics
- quantum transport
- tutorial
- transistors
- molecular electronics
- nano electro-mechanical systems
- NEGF
- carbon nanotubes
- nanomedicine
- education/outreach
- UIUC
- band structure
- ABACUS
- atomic force microscopy
- quantum dots
- MOSFET
- nanowires
- More tags

FEATURED

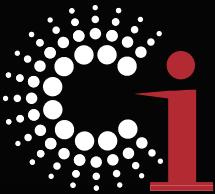
MIT Atomic Scale Modeling Toolkit : Tools for Atomic Scale Modeling - in Tools

Nanotechnology: Considerations for Facility Design - in Online Presentations

Greg Lush, University of Texas at El Paso - Contributions: 30

Topics For Introductory Materials Classes - in Topics

computationinstitute.org



login

MG-RAST

metagenomics analysis server



Browse Metagenomes

search for metagenomes



About



Register



Contact



Help



Upload *



News

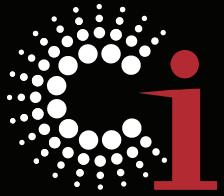
MG-RAST (the Metagenomics RAST) server is an automated analysis platform for metagenomes providing quantitative insights into microbial populations based on sequence data.

# of metagenomes	35,586
# base pairs	9.24 Tbp
# of sequences	85.21 billion
# of public metagenomes	7,167

The server provides web based upload, quality control, automated annotation and analysis for samples up to 10GBp. Comparison between large numbers of samples is enabled via pre-computed abundance profiles. MG-RAST was launched in 2007 and has over 5000 registered users and 35,586 data sets. The current server version is 3.1.2.

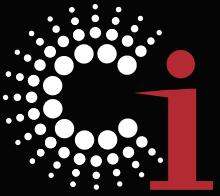
Updates

MG-RAST Version 3.1.2 released



Our vision for a 21st century cyberinfrastructure

To provide **more** capability for
more people at **substantially lower cost** by creatively
aggregating (“cloud”) and
federating (“grid”) resources in
a hybrid world



U.S. DEPARTMENT OF
ENERGY



THE UNIVERSITY OF
CHICAGO



CardioVascular Research Grid

Argonne
NATIONAL LABORATORY

computationinstitute.org