

*Connecting people and resources to  
accelerate discovery by empowering the  
science gateway community*



# What are science gateways and why should I care?

*Nancy Wilkins-Diehr, San Diego Supercomputer Center*

*wilkinsn@sdsc.edu*

*PI, Science Gateways Community Institute*



Award Number  
ACI-1547611

**science gateway** /sī' əns gāt' wā' / n.

1. an online community space for science and engineering research and education.
2. a Web-based resource for accessing data, software, computing services, and equipment specific to the needs of a science or engineering discipline.

Gateways can be many different things.

In a nutshell, they are web interfaces to remote resources.

We all use gateways every day.



# Gateways are changing the way science is conducted in so many ways



NATURE | LETTER  
日本語要約

## New deep-sea species of *Xenoturbella* and the position of Xenacoelomorpha

Greg W. Rouse, Nerida G. Wilson, Jose I. Carvajal & Robert C. Vignati

Affiliations | Contributions | Corresponding author

Nature 530, 94–97 (04 February 2016) | doi:10.1038/nature16544  
Received 19 September 2015 | Accepted 15 December 2015 | Published online 2016

**Acknowledgements** We thank the crew of the R/V *Westward* and the ROVs *Tiburón* and *Doc Ricketts* for their skill and patience in obtaining these "purple socks". We also thank S. Johnson for very helpful comments on the manuscript, S. Mirarab for discussions on methods and N. Holland for comments on the manuscript. This work was supported by the David and Lucile Packard Foundation, the Scripps Institution of Oceanography and the National Science Foundation Assembling the Genome of the *Xenoturbella* (DEB1036368 to G.W.R.).



## I-TASSER

Protein Structure & Function Predictions

(The server completed predictions for 381509 proteins submitted by 92441 users from 100 countries)  
(The template library was updated on 2018/02/21)



National Science Foundation  
4201 Wilson Boulevard  
Arlington, Virginia 22230

Contributor to  
BRAIN Initiative

NSF 14-044

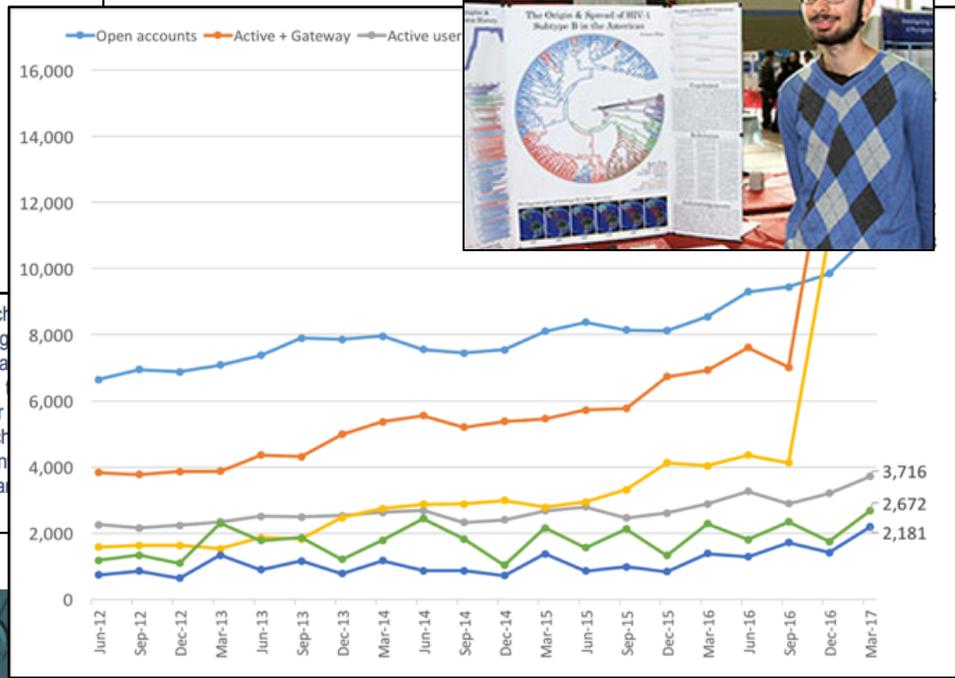
Dear Colleague Letter: BRAIN EAGERS to Enable Innovative Neurotechnologies to Reveal the Functional and Emergent Properties of Neural Circuits Underlying Behavior and Cognition

Date: March 7, 2014

This Dear Colleague Letter is aimed at identifying opportunities to leverage and synthesize technical and conceptual innovation across disciplines and scales to accelerate progress toward an integrated understanding of neural circuits in behavior and cognition, or more simply "catching circuits in a net". The neuroscience research community and specialists in other areas including, but not limited to, genetics, physiology, synthetic biology, engineering, physics, mathematics, statistics, behavior and cognition are encouraged to work across disciplines to develop new approaches and neurotechnologies focused at understanding the properties of circuits that underlie behavior and/or cognition in an organism. Projects that take advantage of existing DBI investments in informatics, computing and infrastructure, such as the [Neuroscience Gateway](#), in novel ways are also eligible.

**Budding Scientist Wins State Fair Prize Using CIPRES Science Gateway**

*10th Grader Creates Timeline, Map of How HIV Spread*



# There are thousands of gateways in many different fields

We are building a catalog so people can find them.

catalog.sciencegateways.org

## Science Gateways Catalog

[Log In](#)
[Sign Up](#)

Total Entries: 337

Sort by  
Newer ▾

**GATEWAY (329)**

SOFTWARE (8)

SGCI CLIENT (24)

USED IN CLASSROOM (17)

FORMAL (20)

Mathematics (8)

**PHYSICAL (153)**

Chemistry (38)

Physics (20)

Earth Sciences (84)

Space Science (23)

**LIFE (142)**

Biology (130)

**SOCIAL (65)**

Anthropology (4)



**Worldwide PDB (wwPDB)**

<https://www.wwpdb.org/>

Since 1971, the Protein Data Bank archive (PDB) has served as the single repository of information about the 3D structures of proteins, nucleic acids, and complex assemblies. The Worldwide PDB (wwPDB)...  
[More](#)

**Categories:** Physical,Life,Interdisciplinary,Chemistry,Biology,Bioinformatics,Biomedical Engineering

wwPDB
structural biology data
biocuration service
macromolecules

SHOW DETAILS



**Skyline**

<https://skyline.ms/project/home/begin.view>

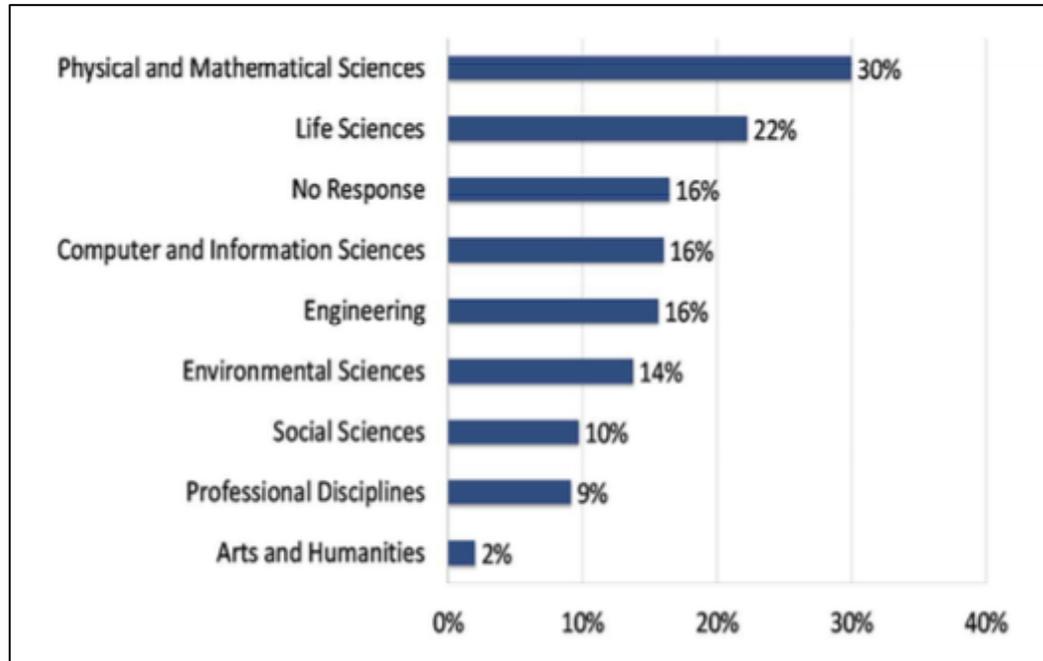
Skyline is a freely-available, open-source Windows client application for building Selected Reaction Monitoring (SRM) / Multiple Reaction Monitoring (MRM), Parallel Reaction Monitoring (PRM), DIA/SWAT...  
[More](#)

Skyline
Molecules
spectrometer data
Panorama

	Field of Science	Portal Homepage
Public Modeling Portal	Stellar Astronomy and Astrophysics	<a href="#">Visit Portal</a>
Public Database	Chemistry	<a href="#">Visit Portal</a>
Portal for inference of large phylogenetic	Systematic and Population Biology	<a href="#">Visit Portal</a>
Portal Anatomy	Neuroscience Biology	<a href="#">Visit Portal</a>
Portal Chemistry Grid (GridChem)	Chemistry	<a href="#">Visit Portal</a>
Portal Gateway	Geography and Regional Science	<a href="#">Visit Portal</a>
Natural Hazards Engineering Infrastructure	Engineering	<a href="#">Visit Portal</a>
	Advanced Scientific Computing	<a href="#">Visit Portal</a>
Portal Gateway	Genetics and Nucleic Acids	<a href="#">Visit Portal</a>
	Molecular Biosciences	<a href="#">Visit Portal</a>
Portal Server	Biological Sciences	<a href="#">Visit Portal</a>
Portal Automated Atomic Model Simulation	Physical Chemistry	<a href="#">Visit Portal</a>
Portal Gateway	Engineering Infrastructure Development	<a href="#">Visit Portal</a>
Portal Simulation Modeling of Hydrodynamic Systems with UltraScan	Biophysics	<a href="#">Visit Portal</a>
	Biochemistry and Molecular Structure and Function	<a href="#">Visit Portal</a>
Portal Gateway	Materials Research	<a href="#">Visit Portal</a>
	Geosciences	<a href="#">Visit Portal</a>
Portal Characterization Lab	Materials Research	<a href="#">Visit Portal</a>
Computational Nanotechnology and	Emerging Technologies Initiation	<a href="#">Visit Portal</a>
Earthquake Engineering Simulation	Earthquake Hazard Mitigation	<a href="#">Visit Portal</a>
Repository of Mathematical Formulae	Mathematical Sciences	<a href="#">Visit Portal</a>
Portal Geography	Earth Sciences	<a href="#">Visit Portal</a>
Portal Gateway	Chemistry	<a href="#">Visit Portal</a>
	Computer and Information Science and Engineering	<a href="#">Visit Portal</a>
Neuroscience Gateway	Neuroscience Biology	<a href="#">Visit Portal</a>
Rosetta Online Server that Includes	Biophysics	<a href="#">Visit Portal</a>



# 5000 survey respondents from a variety of disciplines indicate gateways are important to their work



Types of gateways	Percent
Data collections	75%
Data analysis tools, including visualization and mining	72%
Computational tools	72%
Tools for rapidly publishing and/or finding articles and data specific to my domain	69%
Educational tools	67%
Platforms for fostering group or community collaboration	63%
Simplified interfaces that eliminate the need to learn coding	62%
Citizen science and other public engagement resources	47%
Workflows that automate or capture tasks or processes	42%
Scientific instruments, such as telescopes, microscopes, or sensors	39%

# NSF is investing in gateways too

Press Release 16-088

## NSF commits \$35 million to improve scientific software

### Science Gateways Community Institute

The second award, led by the University of California, San Diego, establishes the **Science Gateways** Community Institute, a multi-institutional consortium that will increase the capabilities, number and sustainability of **science gateways**. Gateways are mobile or web-based applications that provide broad access to the nation's shared cyberinfrastructure to scientists and citizens alike.

"Gateways foster collaborations and the exchange of ideas among researchers and can democratize access, providing broad access to resources sometimes unavailable to those who are not at leading research institutions," said Nancy Wilkins-Diehr, associate director of the San Diego Supercomputer Center and principal investigator for the project. "Sharing expertise about basic infrastructure allows developers to concentrate on the novel, the challenging, and the cutting-edge development needed by their specific user community."

## SGCI launched in 2016

### Our Partners

This program launched in March 2017, and we're taking applications at any time. These partner organizations offer services or resources that are complementary to the work of the SGCI. Feel free to contact our partners directly, or contact us at [help@sciencegateways.org](mailto:help@sciencegateways.org) to request a referral.

Click on our partners' logos to learn more about the services they offer in partnership with SGCI.



OPEN HEALTH SYSTEMS LABORATORY



Extreme Science and Engineering  
Discovery Environment



**<insert public service  
announcement for another, related  
NSF software institute/>**

A nighttime cityscape featuring tall skyscrapers with illuminated windows. In the foreground, a multi-lane highway shows light trails from moving vehicles, including a yellow taxi. A road sign is visible on the left side of the highway.

# **NSF S12-S212 Conceptualization Geospatial Software Institute**

***Towards a National Geospatial Software Ecosystem  
PI Shaowen Wang, [shaowen@illinois.edu](mailto:shaowen@illinois.edu)***



# Addressing Biggest Scientific and Societal Challenges

- Climate change
- Disaster
- Emergency
- Food security
- Population growth
- Sustainability
- Urbanization
- Etc.

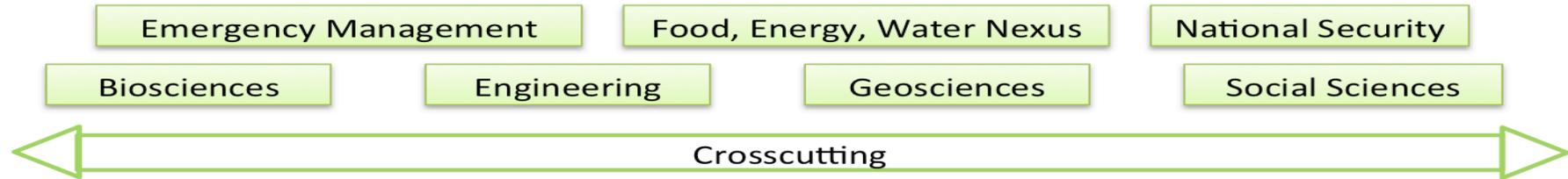
# Geospatial Software

- Software for transforming geospatial (geo & spatial) data into information, knowledge, and intelligence
- Fusion of rapidly changing multidisciplinary sciences and technologies
  - Gaps

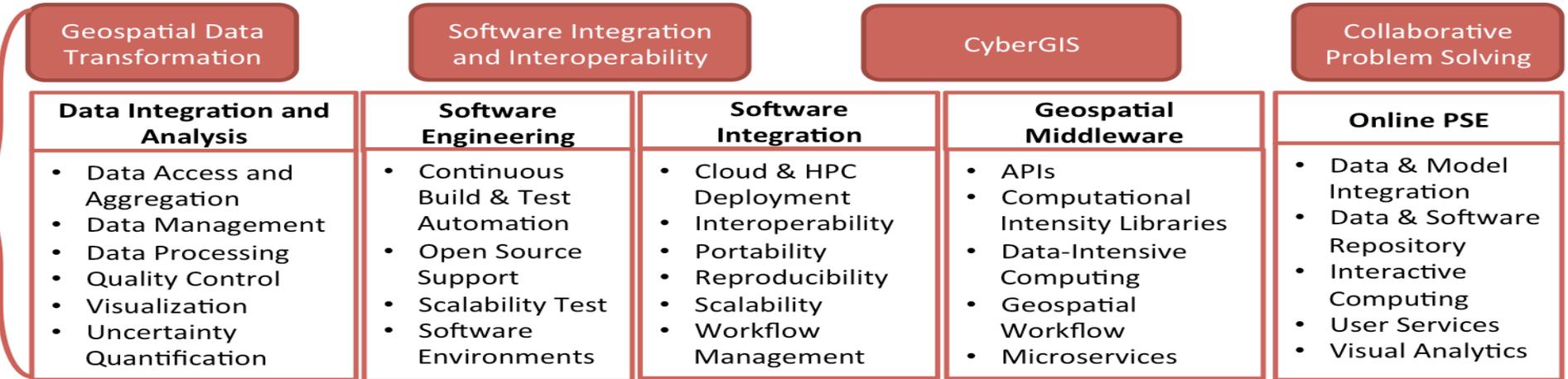


# Ecosystem

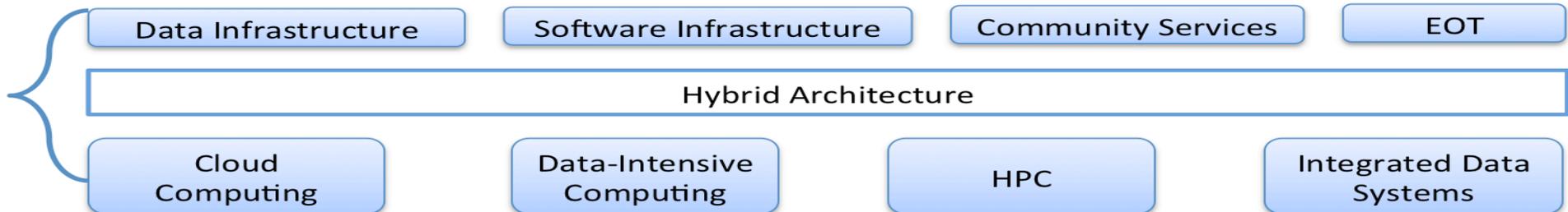
**Domain Communities and Sciences**



**GSI Capabilities and Services**



**Advanced CI Capabilities**

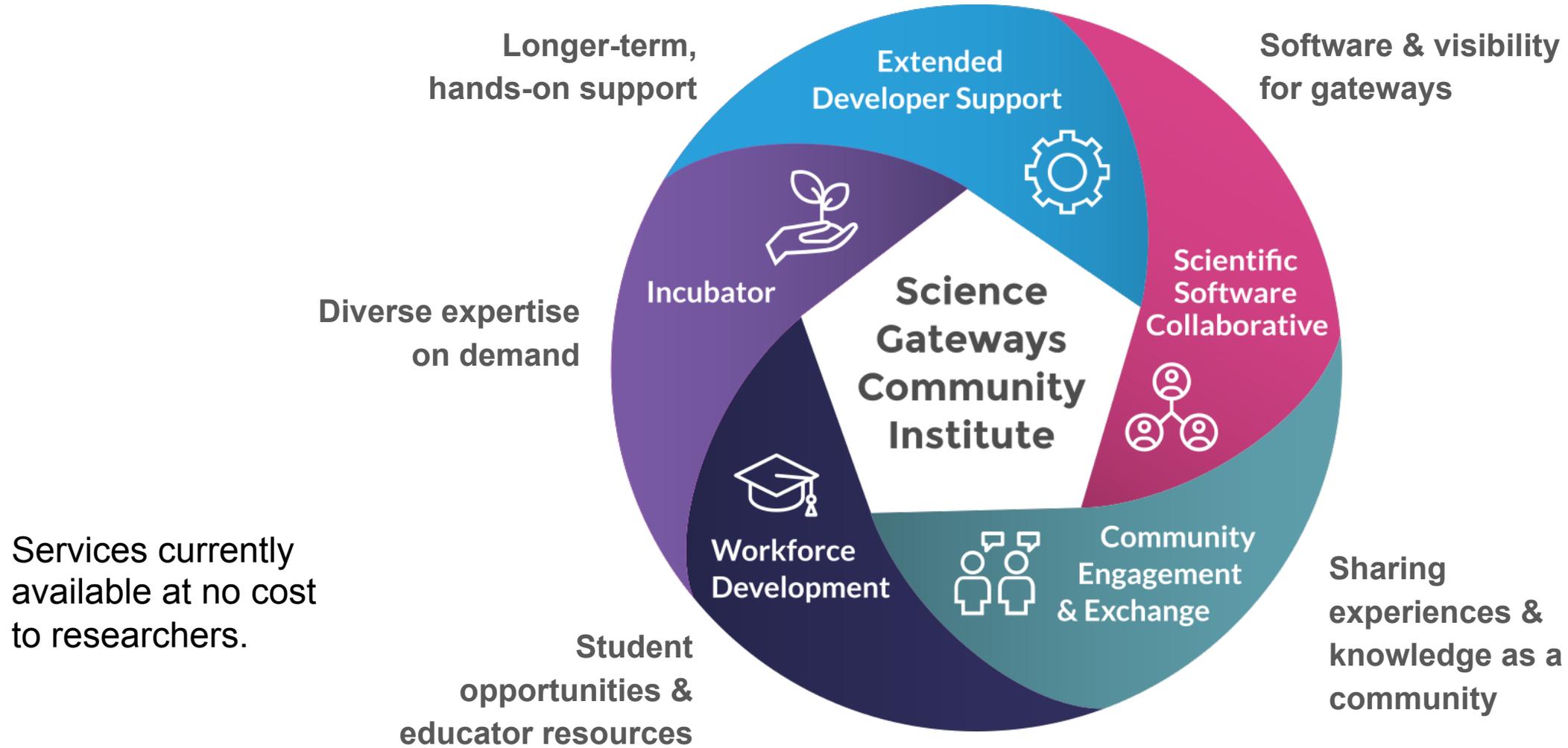


\* EOT – Education, Outreach, and Training; HPC – High Performance Computing; PSE – Problem-Solving Environment

**<end public service announcement  
for another, related NSF software  
institute/>**

# Science Gateways Community Institute

*Designed to help researchers build gateways more effectively*



# Some early clients

- Variety!
  - Size
    - Small, single PI projects
    - Large multimillion dollar awards
  - Type
    - Citizen science
    - Cloud or campus computing
    - Classroom/education
    - Instrument+computing
      - Nobel prize winning domain, commercial interest
  - Research area
    - Drug discovery, ecological science, cryo-EM, application parallelization, network monitoring, neuroscience, coastal flooding
  - Funding source

Chem Compute

DIBBs: Integrated Platform for Applied Network Data Analysis (PANDA)

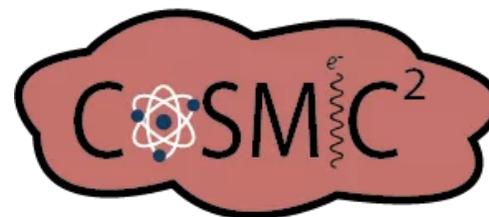


CitSci.org

CloudLaunch



QUBES



nSides



Coastal Emergency Risks Assessment

ADCIRC Coastal Circulation and Storm Surge Model + SWAN Wave Model



<https://sciencegateways.org/consulting/clients>



**Here are a few examples of client projects in the Earth sciences**

# CitSci.org

CitSci.org

Home About Projects Protocols Partners Maps Data Services Blog

Login Donate!

Helping you do great science

Create projects, build datasheets, add data, and view results in real-time.

Sign Up! >

● ○ ○ ○ ○ ○

432 projects	726,020 measurements	3,709 members	65,062 locations	1,549 protocols	Start a Project >
-----------------	-------------------------	------------------	---------------------	--------------------	-------------------

PI: Greg Newman, Colorado State University

Consultants: Tomislav Urban and Lucas Nopoulos,  
Texas Advanced Computing Center

- Background: CitSci.org is a platform on which anyone anywhere can create and enact citizen science projects.
- Objective: Advance gateway features and performance
  - Improve the long term maintenance and sustainability of the citsci.org code base

- Progress: CitSci.org base code added to GitHub
  - VM created at TACC with identical specs to CSU server
  - Currently working on optimization of DB code





# **Aquavit Science Gateway**

PI: Jack Smith, Marshall University

Consultant: Shawn Rice, Purdue University

- Background: Aquavit is a HUBzero-based collaboration portal to support multiple institutions in water quality monitoring and watershed modeling projects
- Objectives: Assist in the integration with the USGS/EPA/USDA Water Quality Portal (WQP)
  - Provide the ability for users to easily upload sensor data to EPA's STORET database via the EPA's Water Quality eXchange (WQX) protocol

- Progress:
  - Helped set up a development environment and workflow suitable for development on Aquavit's content management system (CMS).
  - Overhauled tutorial component
  - Provided CMS documentation and ongoing support/consultation for Aquavit developers.





# Coastal Emergency Risks Assessment (CERA)

PI: Jason Fleming, Seahorse Coastal Consulting

Consultant: John Gentle, Texas Advanced Computing Center

- Background: CERA is transformative knowledge resource and modeling platform for water research. Inclusive of support for multi-domain: data management, computation and modeling, advanced data visualizations and decision support tools.
- Objectives: Improve CERA web mapping capability
  - 50 scripts automate workflow, but creation of thousands of images induces huge load on the server during peak hurricane season

- Progress:
  - Extensive server configuration, including SVN, Python, virus protection installation
  - Apache Mapserver installation and testing, Postgresql installation
  - Django installation, link to Python, Postgresql, Apache, copy data to new server
  - Develop and test new post-processing workflow
  - Test completely upgraded gateway





# EcologyPlus Gateway

PI: Teresa Mourad, Ecological Society of America

Consultant: Shawn Rice, Purdue University

- Background: EcologyPlus is a mentoring platform that connects diverse college students and early career scientists with timely and relevant career opportunities and a community of peers and professionals in ecology and related careers across all sectors.
- Objectives: Provide a platform to match mentors and mentees; provide fresh dynamic content that matches career interests of mentees; provide tools (mentoring, resources, digital badges, LMS) for mentor-mentee career development support.

- Progress: Hub up and running; application submission and review process complete; ability to browse list of mentors and request mentorship; students can fill in a career plan and submit it for evaluation; content relevant to students retrieved from multiple outside sources and displayed; badging





# Hydroshare

PI: David Tarboton, Utah State University

Consultant: Martin Hunt, Purdue University

- Background: HydroShare is a hydrologic information system operated by CUAHSI. Users share and publish data and models in a variety of flexible formats, in order to make this information citable, shareable and discoverable for the advancement of hydrologic science. HydroShare includes a repository for data and models, and tools (web apps) that can act on content in HydroShare providing users with a gateway to high performance computing and computing in the cloud.
- Objectives: Multi-year consulting engagement that covers usability, cybersecurity and business planning consultations

- Progress: Simplifying access to files being worked on in Jupyter, design and implement a plan for managing containers, including the ability to run containerized code from within a Jupyter container and run containers on HPC resources.





# Ocean Observatories Initiative

PI: Ivan Rodero, Rutgers University

Consultant: Choonhan Youn, San Diego Supercomputer Center

- Background: OOI has established a network of interactive, globally distributed sensors with near real-time data access, enhancing capabilities to address critical issues such as climate change, ecosystem variability, ocean acidification, and carbon cycling. The OOI allows multiple scales of marine observations to be integrated into one observing system accessible to the scientific community.
- Objectives: OOINet uses uFrame to process raw data and present it in meaningful ways. A machine-to-machine (M2M) API provides programmatic access. OOI CI provides other data delivery methods such as a THREDDS server, a raw data archive, and an Alfresco server for cruise data, however push-based data delivery is not currently supported.

- Progress: Deployment of a proof-of-concept/prototype message broker system; implementation of software agents to automatically pull data from OOI via M2M API and insert it to the message broker system; implementation of a easy-to-use mechanism for the end use to be able to subscribe to OOI data through the message broker system



# SimCCS

PI: Yinzhi Wang, Indiana Geological and Water Survey

Consultant: Sudhakar Pamidighantam, Indiana University

- Background: SimCCS is a decision support tool for integrated assessment of carbon capture, utilization and storage technology developed by members of the U.S.-China Clean Energy Research Center (CERC), Advanced Coal Technology Consortium. It produces optimal solutions for integrated CCUS infrastructure deployment with constraints of emissions cap, emissions tax and time. SimCCS Gateway will make this tool available to all the academic and business partners within the consortium and eventually to other potential users as well.
- Objectives: Develop a community-based, open source SimCCS Gateway framework that integrates with high-performance computing platform for all the academic and industrial partners within the CERC as well as other potential users in a broader community.

- Progress: Integrate the desktop version of SimCCS with the gateway service; integrate new LANL releases into the SimCCS Gateway Client; build, deploy and test a Django-based SimCCS Gateway that is capable of handling the entire workflow



# Thank you

- We'd love to hear from you
- [www.sciencegateways.org](http://www.sciencegateways.org)