Introduction to Science Gateways

Amit Majumdar

Data Enabled Scientific Computing Division

San Diego Supercomputer Center

SDSC Summer Institute 2019

Administrative and Technical tasks (barriers?) to use HPC

- **Write allocation proposals (proposals (peer-reviewed) for supercomputer time every year**
- **Understand HPC machines, policies, complex OS/software**
- **Install and benchmark complex applications on HPC resources**
- **Different machines have different schedulers**
- **Understand and manage remote authentication**
- **Figure out data transfer, file systems, storage**

What is a Science Gateway

Science gateways allow science & engineering communities to access shared data, software,

Catalyzes and democratizes science research for scientists and students

disciplines. Read more on Wikipedia.

From wiki: Science gateways provide access to advanced resources for science and engineering researchers, educators, and students. Through streamlined, online, user-friendly interfaces, gateways combine a variety of cyberinfrastructure (CI) components in support of a community-specific set of tools, applications, and data collections.:[1] In general, these specialized, shared resources are integrated as a Web portal, mobile app, or a suite of applications.[2] Through science gateways, broad communities of researchers can access diverse resources which can save both time and money for themselves and their institutions

Some common features

- **Easy web-based user interface**
  - Upload input files, data
  - Download output results
  - Some provide post processing/viz
  - In some cases programmatic access
- **Software/scientific applications already installed optimally on computing resources at the backend**
- **HPC resources - available via XSEDE (Extreme Science and Engineering Discovery Environment) or other**
- **(in most cases) Gateway team writes annual allocation proposal**

https://www.xsede.org/ecosystem/science-gateways

XSEDE About For Users Ecosystem Community Engagement News User Portal

Below is a complete list of current science gateways, to see a detailed project description please click on the name of the science gateway.

To update the information contained in this table, please contact help@xsede.org. To register your gateway, please complete the Gateway Registration Form.

Title	Field of Science	Portal Homepage
3D-Quantitative Phenotyping Gateway	Biological Sciences	Visit Portal
Asteroseismic Modeling Portal	Stellar Astronomy and Astrophysics	Visit Portal
Chem Compute	Chemistry	Visit Portal
CIPRES Portal for inference of large phylogenetic trees	Systematic and Population Biology	Visit Portal
Computational Anatomy	Neuroscience Biology	Visit Portal
Computational Chemistry Grid (GridChem)	Chemistry	Visit Portal
CyberGIS Gateway	Geography and Regional Science	Visit Portal
DesignSafe: Natural Hazards Engineering Research Infrastructure	Engineering	Visit Portal
Diagrid	Advanced Scientific Computing	Visit Portal
dREG gateway	Genetics and Nucleic Acids	Visit Portal
ENIGMA Bipolar BrainAge Analysis Upload Portal	Neuroscience Biology	Visit Portal
Galaxy	Molecular Biosciences	Visit Portal
GenePattern Server	Biological Sciences	Visit Portal
Globus Online	Engineering Infrastructure Development	Visit Portal
High-Resolution Modeling of Hydrodynamic Experiments with UltraScan	Biophysics	Visit Portal
I-TASSER	Biochemistry and Molecular Structure and Function	Visit Portal
Metaproteomics Gateway	Biochemistry and Molecular Structure and Function	Visit Portal
MP-Complete	Materials Research	Visit Portal
MyGeoHUB	Geosciences	Visit Portal
Nanoparticle Characterization Lab	Materials Research	Visit Portal
Network for Computational Nanotechnology and nanoHUB	Emerging Technologies Initiation	Visit Portal
NIST Digital Repository of Mathematical Formulae	Mathematical Sciences	Visit Portal
OpenTopography	Earth Sciences	Visit Portal
ParamChem Gateway	Chemistry	Visit Portal
PGA	Computer and Information Science and Engineering	Visit Portal
PICKSC Science Gateway	Physics	Visit Portal
Proteogenomics Gateway	Biochemistry and Molecular Structure and Function	Visit Portal
Providing a Neuroscience Gateway	Neuroscience Biology	Visit Portal
ROSIE, The Rosetta Online Server that Includes Everyone	Biophysics	Visit Portal
Science Gateways Platform as a Service (SciGaP)	Computer Systems Architecture	Visit Portal
SimCCS Gateway	Geosciences	Visit Portal
SimVascular	Fluid, Particulate, and Hydraulic Systems	Visit Portal
The Earth System Grid	Global Atmospheric Research	Visit Portal
The iPlant Collaborative Agave API	Integrative Biology and Neuroscience	Visit Portal
UCI Complex Social Science Gateway	Anthropology	Visit Portal
Unidata: Data Proximate Services in the Cloud	Atmospheric Sciences	Visit Portal
VLab - Virtual Laboratory for Earth and Planetary Materials	Materials Research	Visit Portal
WaterHUB - Platform for water education, research, data access, partnership and collaboration	Earth Sciences	Visit Portal

SDSC
SGW

Nancy Wilkins-Diehr, SDSC – Founding PI - http://sciencegateways.org/
Michael Zentner, SDSC current PI (Nancy retired)
Other institutions: Elizabeth City State in North Caronila, Indiana University, University of Notre Dame, Purdue University, the Texas Advanced Computing Center at the University of Texas, Austin, and the University of Michigan at Ann Arbor

https://sciencegateways.org/home

SGCI Science Gateways Community Institute

HOME ABOUT CONSULTING SERVICES RESOURCES

The SGCI serves the science gateway community.

We provide NSF-funded, online and in-person resources and services. Our goal is to facilitate—at little or no cost—the sharing of experiences, technologies, and practices of those working with science gateways. Read success stories about how we've supported gateway development and our community of gateway creators, users, and students.

Join the Gateways Community Forum! We've created a discussion forum open to all members of the gateway community. We welcome questions, event listings, and other community-relevant topics.

Register Now for Gateways 2019!

What is a Gateway?

Science gateways allow science & engineering communities to access shared data, software, computing services, instruments, educational materials, and other resources specific to their disciplines. Read more on Wikipedia.

Are You...

New to Gateways?

Building a Gateway?

Interested in Working with Us?

SGCI Consultation Gives CitSci a

Featured

- Science Node Article: Need help with your science gateway? SGCI Focus Week teaches best practices for building, operating, and sustaining science gateways
- Bootcamp Five Inspires a Shift—This Collaborative Workshop Is Now Called Gateway Focus Week
- A Hello and Farewell from SGCI's New Director
- Science Node Article: Join the Gateways 2019 conference

https://sciencegateways.org/about

We connect people and resources to accelerate discovery by empowering the science gateway community.

The Science Gateways Community Institute (SGCI) was funded by the NSF in 2016 to provide services, resources, community support, and education for creating and sustaining science gateways.

Science gateways are online interfaces that give researchers, educators, and students easy access to specialized, shared resources that are specific to a science or engineering discipline. For example, they may connect to or between instruments (such as telescopes or sensors), data collections, specialized software, or high-performance computing.

While the use of gateways can improve the productivity of researchers significantly, **the process of developing, operating, and sustaining a gateway can prove challenging and time consuming.** Our services and resources aim to fill that gap. The SGCI is organized into five service areas.

Consulting Services

- **Incubator:** Learn best practices from our consultants or Bootcamp
- **Extended Developer Support:** Get direct, custom development help

Resources and Learning Opportunities

- **Scientific Software Collaborative:** Find gateways or software components (or promote your own)
- **Community Engagement and Exchange:** Engage with and learn from the gateways community
- **Workforce Development:** Build your career as a student or young professional

Background

The Science Gateways Community Institute (SGCI) was founded with the **mission of providing resources, expertise, community support, and education to the creators of gateways serving science and engineering research and education.** Through these channels, we hope to speed the development and application of more robust, less expensive, and more sustainable gateways.

The SGCI is one of two Scientific Software Innovation Institutes funded by the National Science Foundation (NSF) in August 2016. We were motivated to create this institute in **recognition of the growing need for a "community center" that would bring together gateway creators** scattered across multiple domain areas. By sharing expertise about technologies and strategies, developers may concentrate on the unique, cutting-edge development needed by their specific user communities.

Our **partnership of seven universities** represents more than a decade of collaboration in service of science gateways. Our selection of resources and services is based on more than five years of qualitative and survey research to better understand how gateways succeed, how creators could benefit from supplemental services and resources, and where gateways might develop in the future.

ECSU
ELIZABETH CITY STATE UNIVERSITY
FOUNDED 1891

INDIANA UNIVERSITY

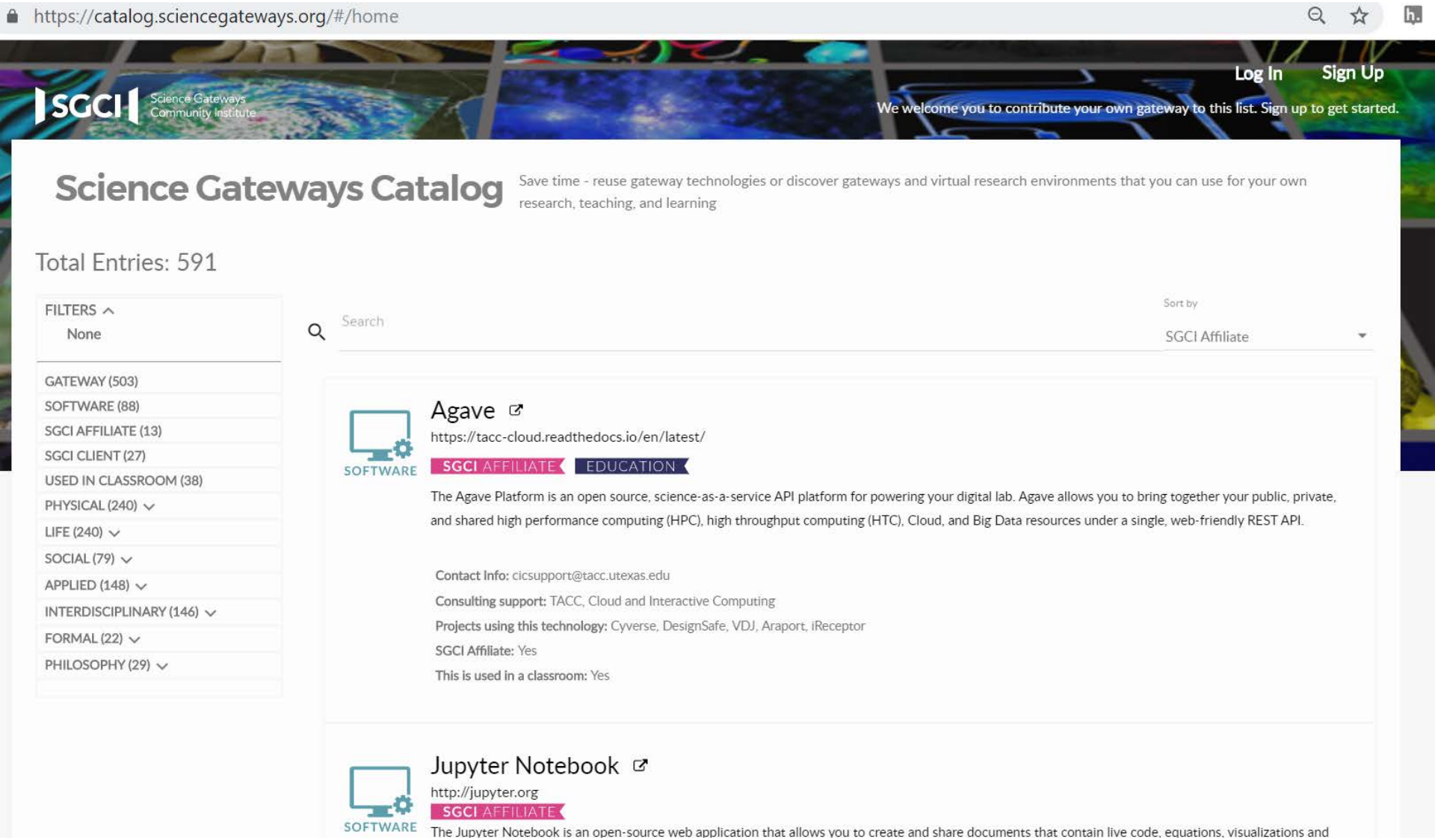
PURDUE
UNIVERSITY

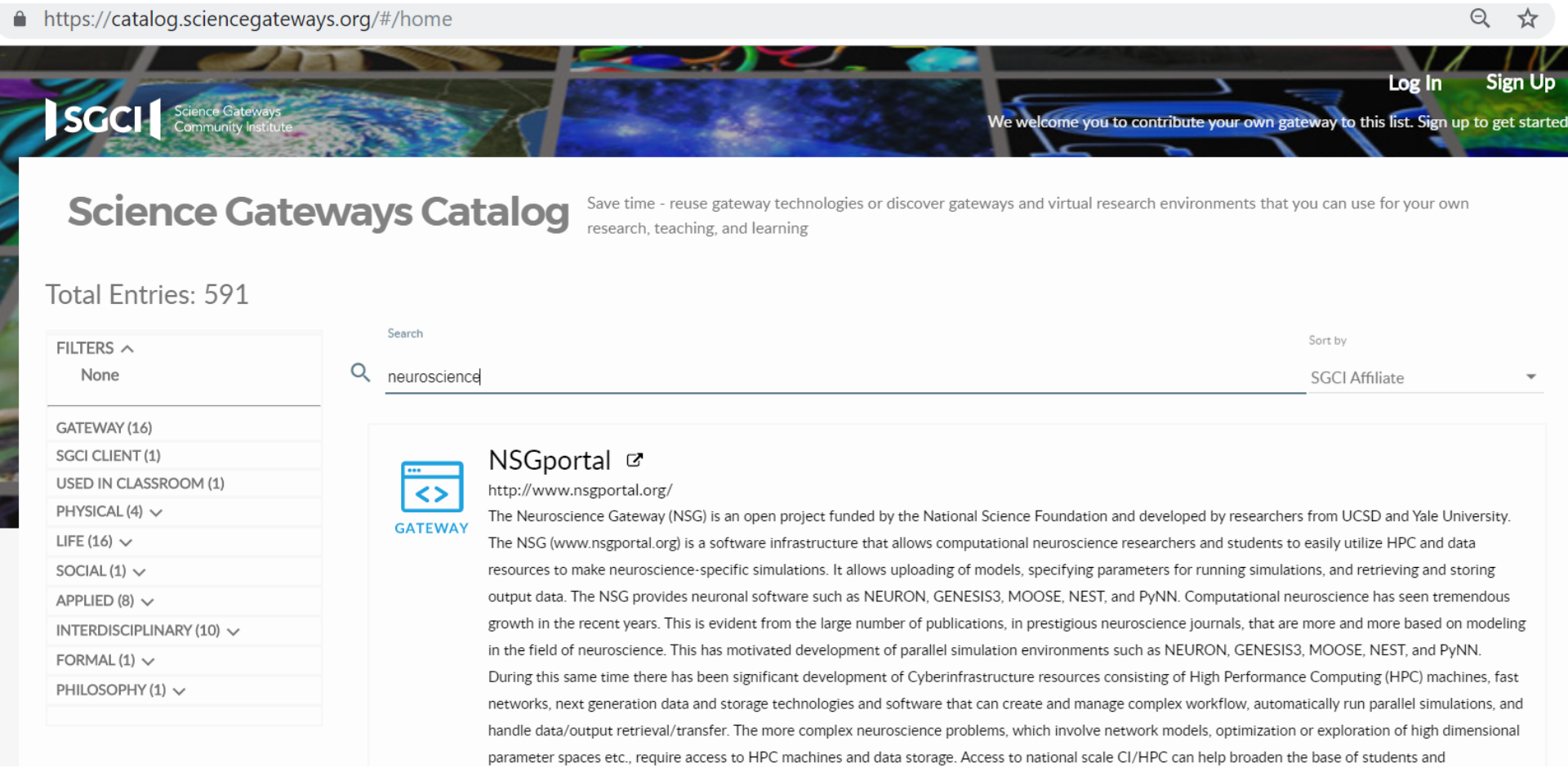
UC San Diego

UNIVERSITY OF
MICHIGAN

UNIVERSITY OF
NOTRE DAME

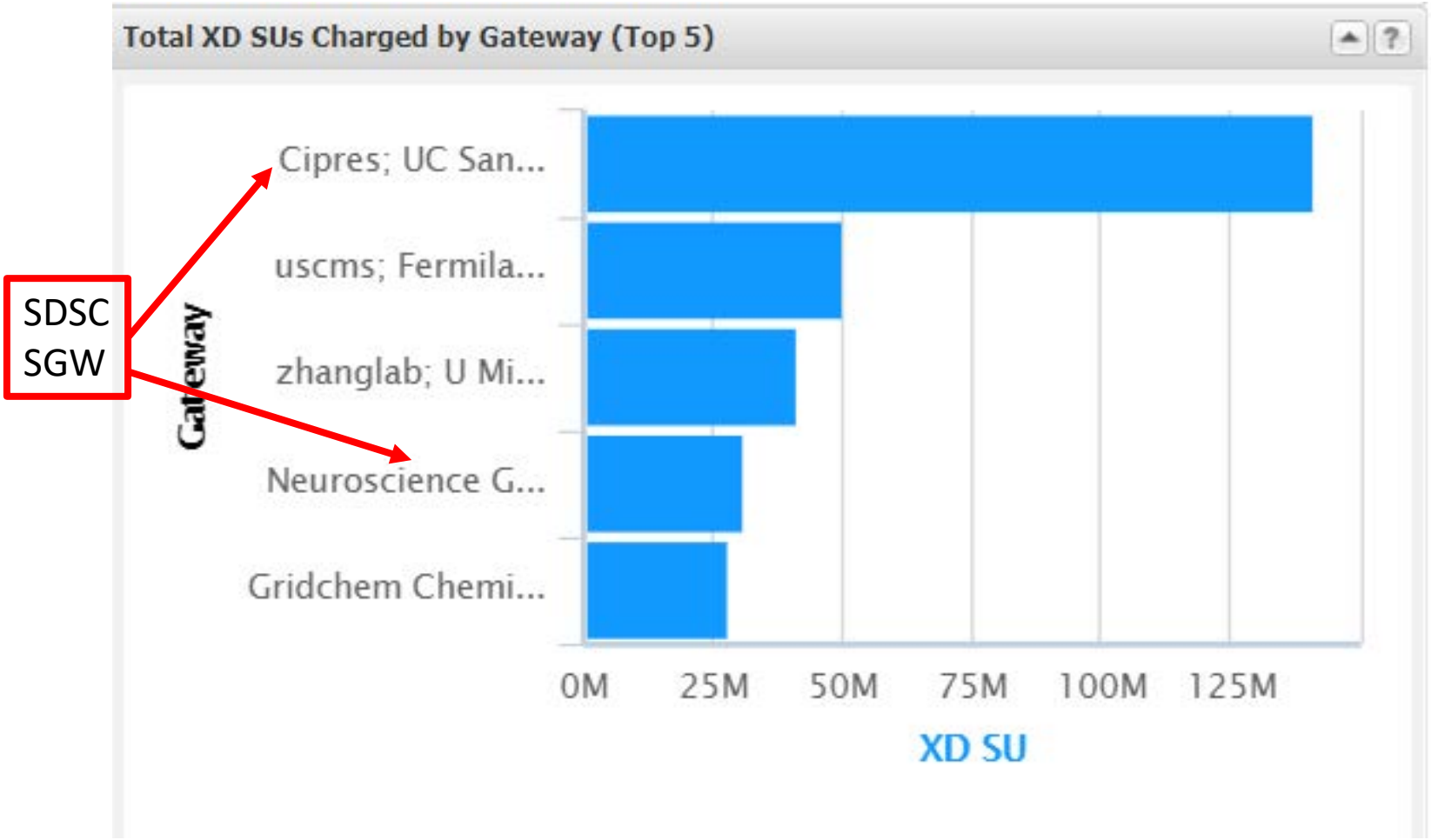
TEXAS
The University of Texas at Austin


https://catalog.sciencegateways.org/#/home
SGCI Science Gateways Community Institute
Log In
Sign Up
We welcome you to contribute your own gateway to this list. Sign up to get started.
Science Gateways Catalog
Save time - reuse gateway technologies or discover gateways and virtual research environments that you can use for your own research, teaching, and learning
Total Entries: 591
FILTERS
None
GATEWAY (503)
SOFTWARE (88)
SGCI AFFILIATE (13)
SGCI CLIENT (27)
USED IN CLASSROOM (38)
PHYSICAL (240)
LIFE (240)
SOCIAL (79)
APPLIED (148)
INTERDISCIPLINARY (146)
FORMAL (22)
PHILOSOPHY (29)
Search
Sort by
SGCI Affiliate
SOFTWARE
Agave
https://tacc-cloud.readthedocs.io/en/latest/
SGCI AFFILIATE
EDUCATION
The Agave Platform is an open source, science-as-a-service API platform for powering your digital lab. Agave allows you to bring together your public, private, and shared high performance computing (HPC), high throughput computing (HTC), Cloud, and Big Data resources under a single, web-friendly REST API.
Contact Info: cicsupport@tacc.utexas.edu
Consulting support: TACC, Cloud and Interactive Computing
Projects using this technology: Cyverse, DesignSafe, VDJ, Araport, iReceptor
SGCI Affiliate: Yes
This is used in a classroom: Yes
SOFTWARE
Jupyter Notebook
http://jupyter.org
SGCI AFFILIATE
The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and

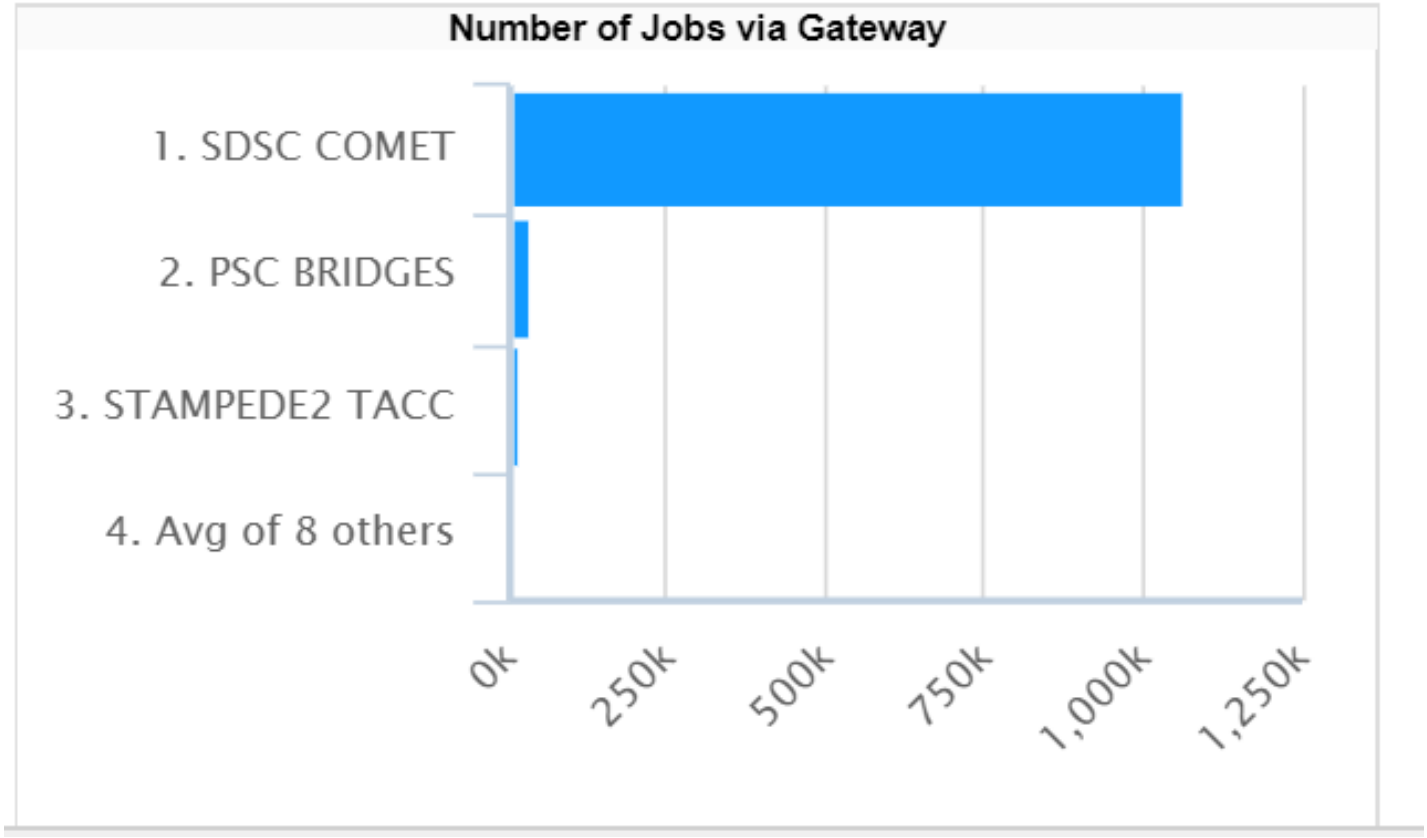

https://catalog.sciencegateways.org/#/home
SGCI
Science Gateways Community Institute
Log In
Sign Up
We welcome you to contribute your own gateway to this list. Sign up to get started
Science Gateways Catalog
Save time - reuse gateway technologies or discover gateways and virtual research environments that you can use for your own research, teaching, and learning
Total Entries: 591
FILTERS
None
GATEWAY (16)
SGCI CLIENT (1)
USED IN CLASSROOM (1)
PHYSICAL (4)
LIFE (16)
SOCIAL (1)
APPLIED (8)
INTERDISCIPLINARY (10)
FORMAL (1)
PHILOSOPHY (1)
Search
neuroscience
Sort by
SGCI Affiliate
GATEWAY
NSGportal
http://www.nsgportal.org/
The Neuroscience Gateway (NSG) is an open project funded by the National Science Foundation and developed by researchers from UCSD and Yale University. The NSG (www.nsgportal.org) is a software infrastructure that allows computational neuroscience researchers and students to easily utilize HPC and data resources to make neuroscience-specific simulations. It allows uploading of models, specifying parameters for running simulations, and retrieving and storing output data. The NSG provides neuronal software such as NEURON, GENESIS3, MOOSE, NEST, and PyNN. Computational neuroscience has seen tremendous growth in the recent years. This is evident from the large number of publications, in prestigious neuroscience journals, that are more and more based on modeling in the field of neuroscience. This has motivated development of parallel simulation environments such as NEURON, GENESIS3, MOOSE, NEST, and PyNN. During this same time there has been significant development of Cyberinfrastructure resources consisting of High Performance Computing (HPC) machines, fast networks, next generation data and storage technologies and software that can create and manage complex workflow, automatically run parallel simulations, and handle data/output retrieval/transfer. The more complex neuroscience problems, which involve network models, optimization or exploration of high dimensional parameter spaces etc., require access to HPC machines and data storage. Access to national scale CI/HPC can help broaden the base of students and

SDSC leads in developing science gateways

(8/18-7/19)



SDSC leads in serving science gateways
(8/18 – 7/19)



Gateways for many different fields

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

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

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & V

Archive › Volume 530 › Issue 7588 › Letters › Article

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. Vrijenhoek

Affiliations | Contributions | Corresponding author

Nature 530, 94–97 (04 February 2016) | doi:10.1038/nature16545

Received 19 September 2015 | Accepted 15 December 2015 | Published online 03 February 2016

Acknowledgements We thank the crew of the R/V *Western Flyer* and pilots of the ROVs *Tiburon* and *Doc Ricketts* for their skill and patience during hunts for these 'purple socks'. We also thank S. Johnson for verifying bivalve sequences obtained from *Xenoturbella* and L. Lundsten for hunting through many video files for imagery. We acknowledge the Cyberinfrastructure for Phylogenetic Research (CIPRES) Science Gateway[25] for computing resources, and thank M. Miller for additional resources, S. Mirarab for discussions on species tree methods and N. Holland for comments on the manuscript. This work was supported by the David and Lucile Packard Foundation via the Monterey Bay Aquarium Research Institute, Scripps Institution of Oceanography and the National Science Foundation Assembling the Tree of Life program (DEB1036368 to G.W.R.).

MENU

nature microbiology

Letter | OPEN | Published: 11 April 2016

A new view of the tree of life

Laura A. Hug, Brett J. Baker, Karthik Anantharaman, Christopher T. Brown, Alexand
J. Castelle, Cristina N. Butterfield, Alex W. Hernsdorf, Yuki Amano, Kotaro Ise, Yohey
Dudek, David A. Relman, Kari M. Finstad, Ronald Amundson, Brian C. Thomas & Jill

Nature Microbiology **1**, Article number: 16048 (2016) | Download Citation

National Science Foundation
4201 Wilson Boulevard
Arlington, Virginia 22230

NSF 14-044

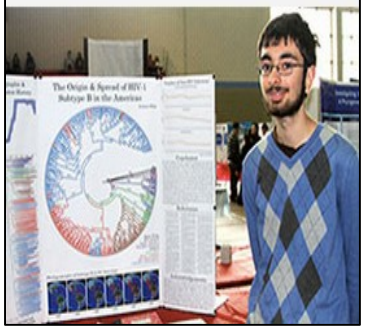
Dear Colleague Letter: BRAIN EAGERs to Enable Innovative Neurotechnologies to Reveal the Functional and Emergent Properties of Neural Circuits Underlying Behavior and Cognition

Date: March 7, 2014

This Dear Colleague Letter is aimed at identifying opportunities to leverage and synthesize technological 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 action". 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 neurotechnology focused at understanding the properties of circuits that underlie behavior and/or cognition in any organism. Projects that take advantage of existing DBI investments in informatics, computing and other 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



Neuroscience Gateway

NSG Team:
Amit Majumdar, Subha Sivagnanam, Kenneth Yoshimoto (SDSC)
Ted Carnevale (Yale school of medicine)

HPC Challenges for Computational and Data Processing Neuroscience

- **Modeling and data (EEG, fMRI etc.) projects start "small" and many are *forced* to stay "small"**
- **Rapid growth in development of complex neuronal network models, parameter sweep estimations, data processing etc. require HPC**
- **Not all neuroscientists, in the world, have access to large scale HPC**
- **Barriers of entry to HPC**
  - **Write peer-reviewed proposals for computer time**
  - **Understand HPC machines, policies, complex OS/software**
  - **Install and benchmark complex tools on HPC resources**
  - **Figuring out data transfer, management, storage issues**

➢ Neuroscience Gateway (NSG)

- Science gateway for neuroscience modeling and data processing – NSF, NIH funded
- https://www.nsgportal.org
- Free and open (non-commercial users), since 2013
- Variety of modeling and data analysis tools
- Calculations performed on XSEDE compute resources
- XSEDE allocation of 10,000,000+ SUs on various machines

The Neuroscience Gateway (NSG)

The NSG provides simple and secure access through portal and programmatic services, to run neuroscience related software and tools on HPC resources
http://www.nsgportal.org

1. Developed using CIPRES SDK framework, customized for neuroscience research.
2. Easy user interface – providing easy model/input data upload, running of codes
3. Provide neuronal simulation and data processing tools – widely used by neuroscientists
4. Ability to easily get to the results, download results

NSG – Portal and Programmatic Access

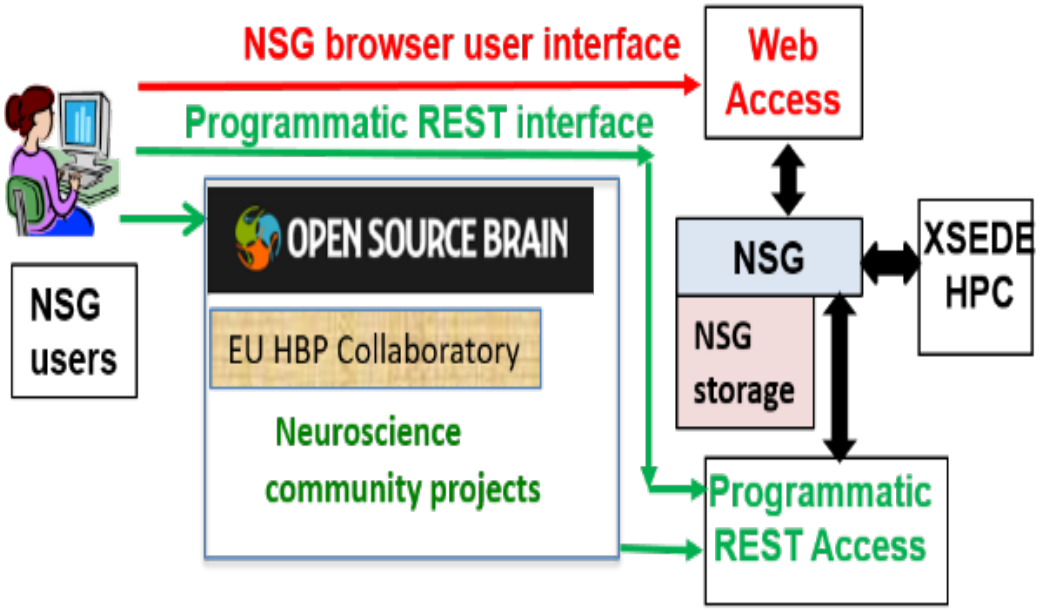
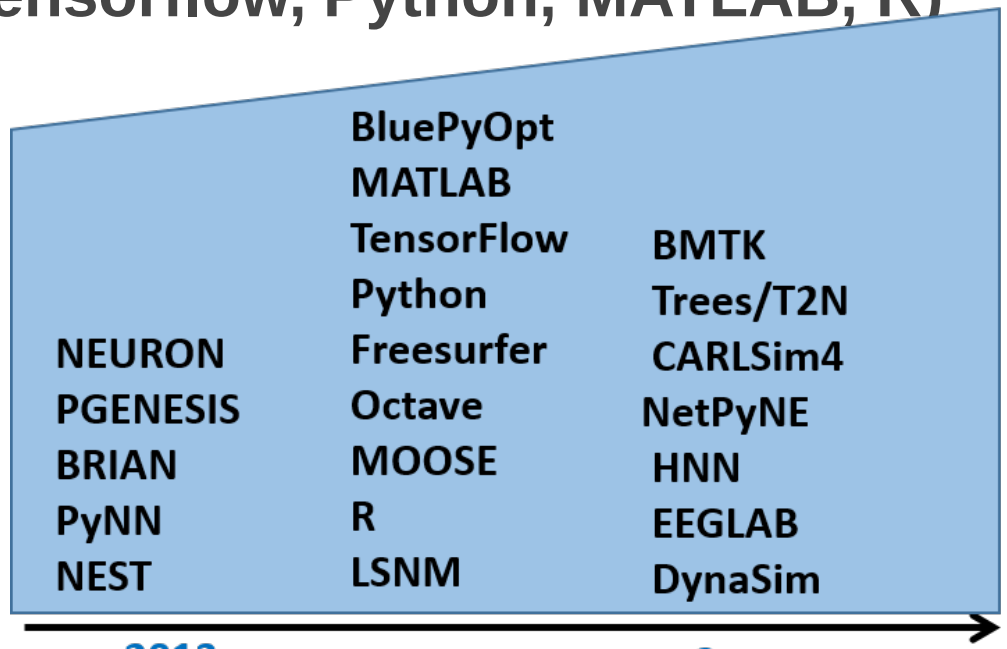


Figure 1. NSG Users Web frontend and REST Programmatic Access.

Tools

- Biological neuronal networks (NEURON, PGENESIS, MOOSE, NEST etc.)
- Experimental data analysis: EEG analysis (EEGLAB, HNN), fMRI (Freesurfer)
- Other analysis tools (Tensorflow, Python, MATLAB, R)
- A neuromorphic HW in the future
- Based on user request we continue to add new tools and pipelines



Current NSG tools, pipelines, software.

NSG XSEDE Allocations

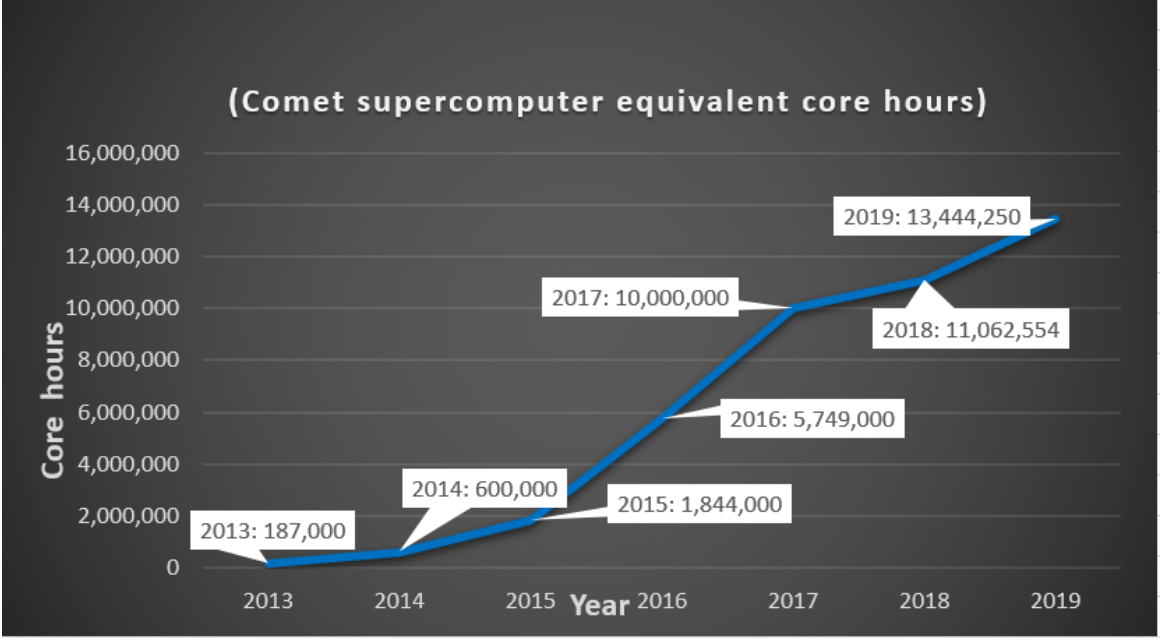


Figure 1. NSG total XSEDE computing allocation by years.

NSG # of Users

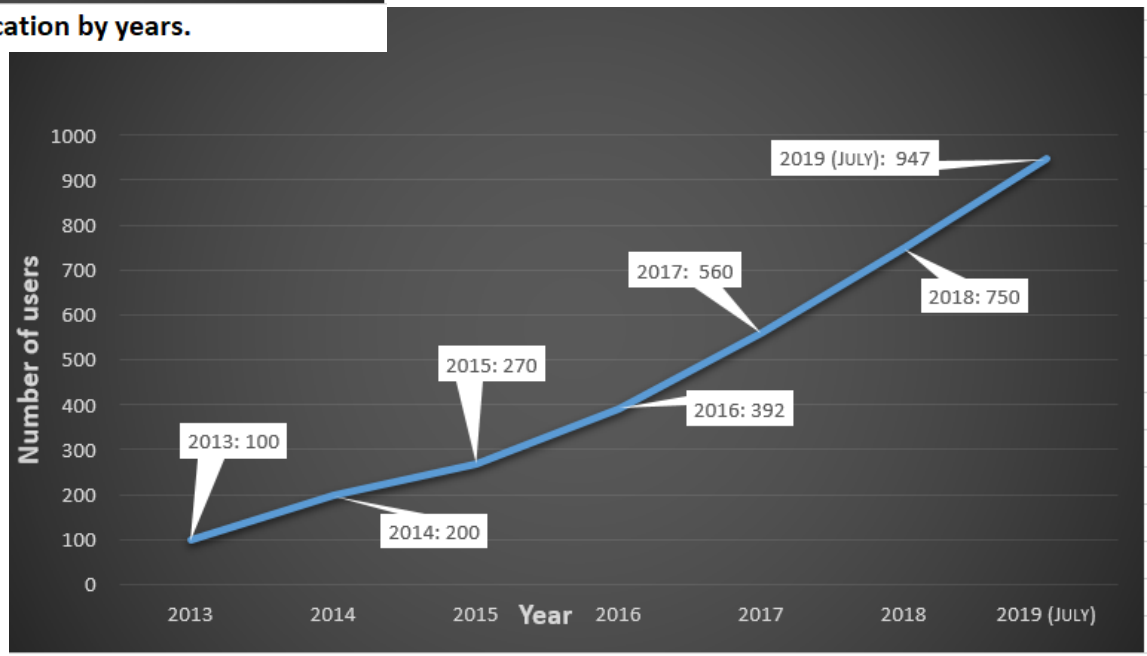


Figure 2. Growth in number of NSG users.

Evolving NSG

- **Tool Dissemination**
- **Education and Training**
  - NEURON Summer course
  - NIH funded Computational Neuroscience Training Course (U. Missouri)
  - SFN, CNS workshops
  - NSF funded Cyberinfrastructure Neuroscience training
- **Collaborative Environment**

Resources available for researchers and developers….

Building a gateway - Challenges

- **Building a gateway requires different types of expertise**
  - Software developer, Graphic designer, Security expert …..
  - But projects cannot always afford to hire these specialists
  - Short term hires are difficult as well

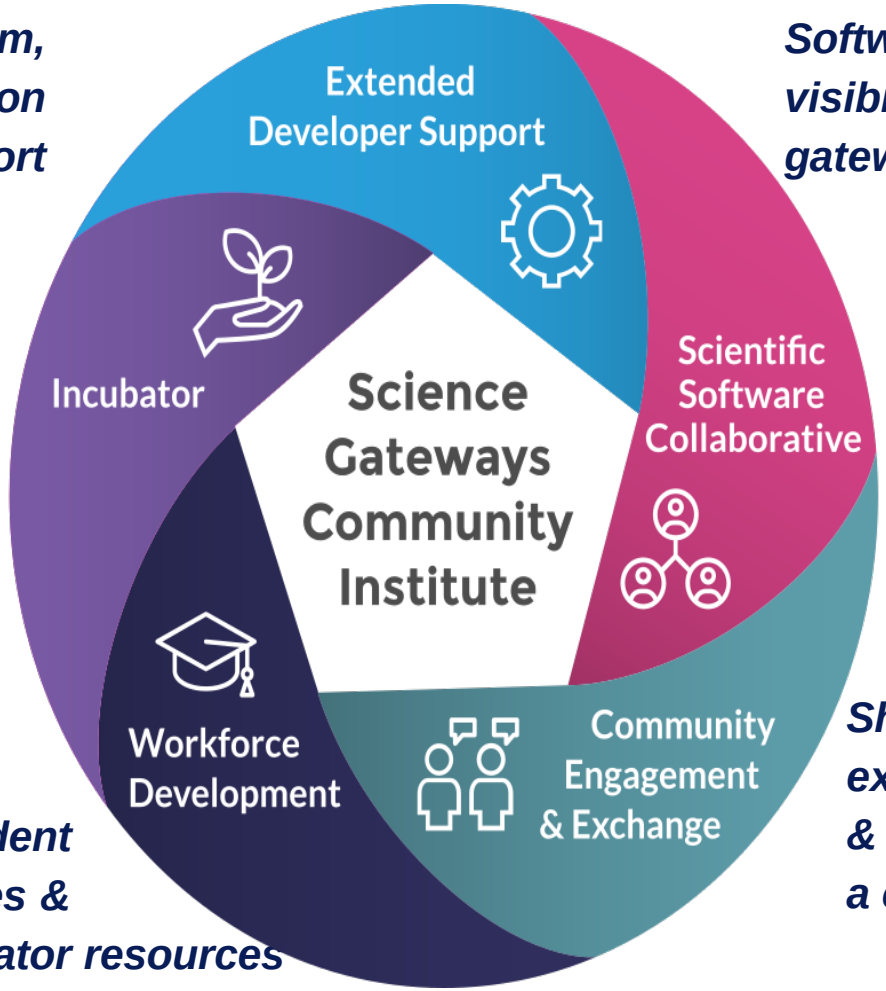
Science Gateways Community Institute

Designed to help researchers build gateways more effectively

Longer-term, hands-on support

Software & visibility for gateways

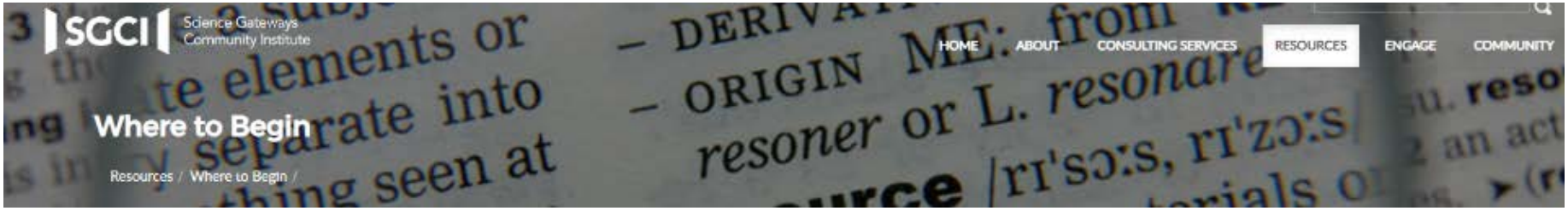
Diverse expertise on demand

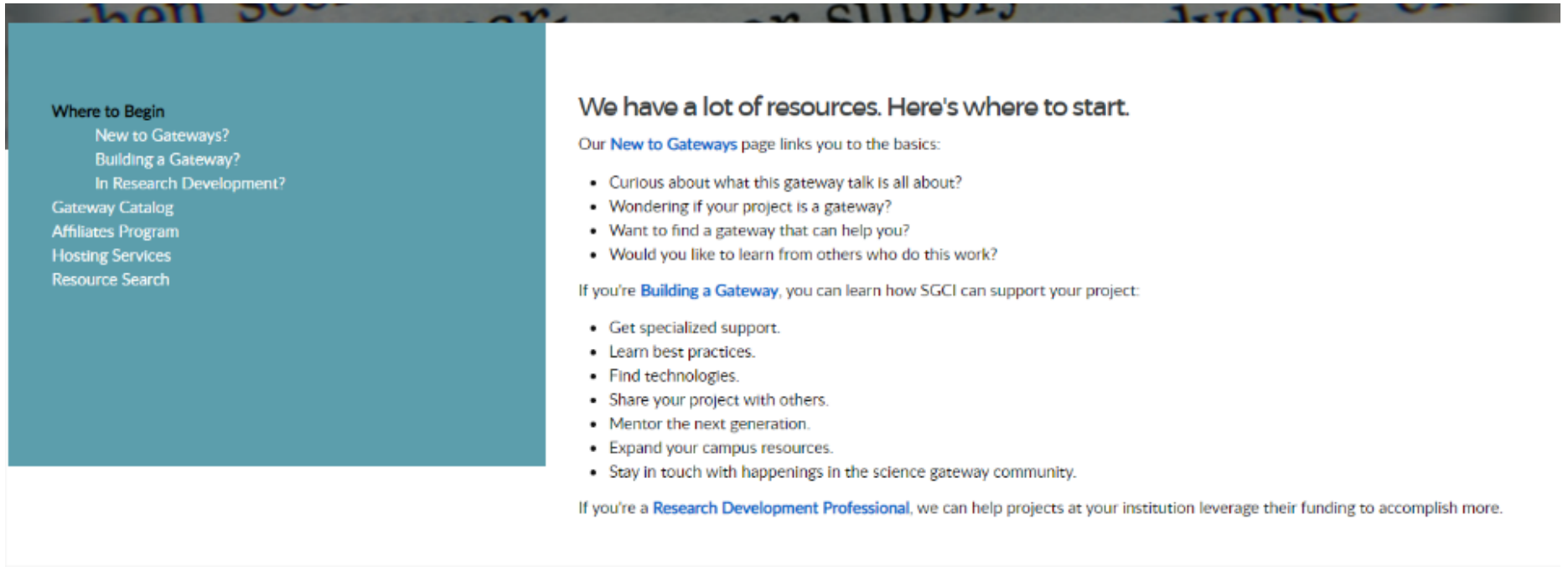


Services currently available at no cost to researchers.

Student opportunities & educator resources

Sharing experiences & knowledge as a community


SGCI
Science Gateways Community Institute
HOME
ABOUT
CONSULTING SERVICES
RESOURCES
ENGAGE
COMMUNITY
Where to Begin
Resources / Where to Begin /


Where to Begin
New to Gateways?
Building a Gateway?
In Research Development?
Gateway Catalog
Affiliates Program
Hosting Services
Resource Search
We have a lot of resources. Here's where to start.
Our New to Gateways page links you to the basics:
Curious about what this gateway talk is all about?
Wondering if your project is a gateway?
Want to find a gateway that can help you?
Would you like to learn from others who do this work?
If you're Building a Gateway, you can learn how SGCI can support your project:
Get specialized support.
Learn best practices.
Find technologies.
Share your project with others.
Mentor the next generation.
Expand your campus resources.
Stay in touch with happenings in the science gateway community.
If you're a Research Development Professional, we can help projects at your institution leverage their funding to accomplish more.

Thoughts and Tactics for Success - To Build a Science Gateway

1: identify a user population in need
2: commit to responding to users' needs
3: let user behavior/needs drive improvements
4: with limited resources, prioritization is key
5: stay in touch with your community
6: embrace customer service

If you are a researcher/end user

- **Science Gateway information at XSEDE, SGCI**
  - https://sciencegateways.org
  - http://www.xsede.org
- **Follow good practices and provide feedback**
- **Cite the gateway used in your research in your publication, annual reports**

Thanks…

email: majumdar@sdsc.edu