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

SGCI

Science Gateways: Addressing Data Management Challenges

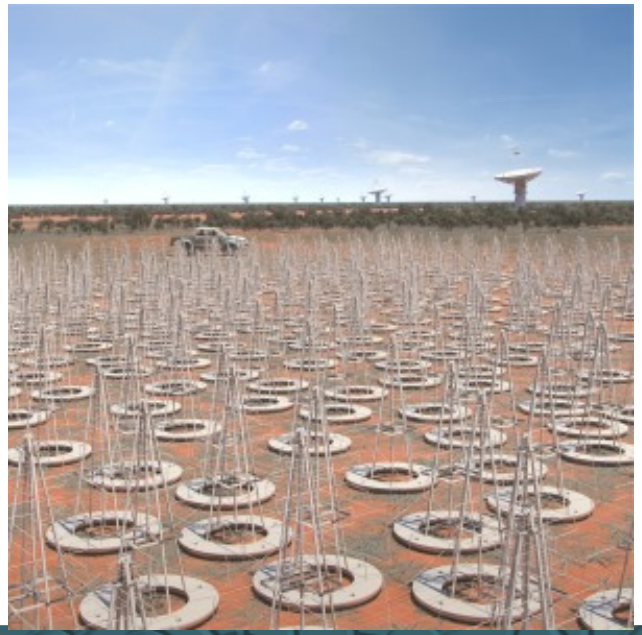
Sandra Gesing
Center for Research Computing, University of Notre Dame
sandra.gesing@nd.edu

7th National Data Service Consortium Workshop, Chicago
13 April 2017

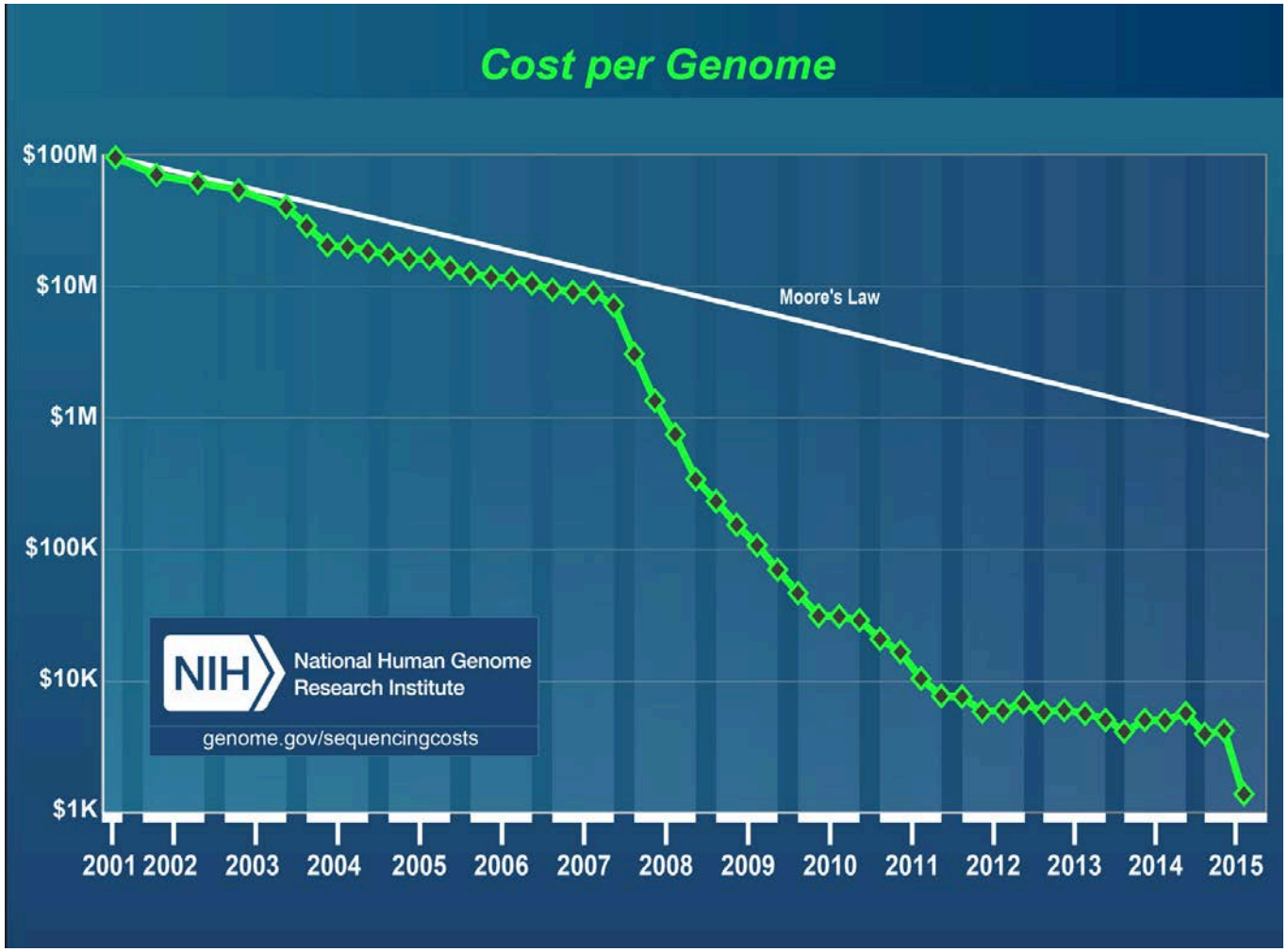
Award Number
ACI-1547611

Data

- Explosion in the quantity, variety and complexity of data
- Questions can be answered impossible to even ask about 10 years ago
- Costs far reduced (e.g., Human Genome project, 15 years, ~ $2 billion; today ~3 days, $1000)



Data



http://www.genome.gov/images/content/cost_per_genome_oct2015.jpg

Challenges

1. Meaningful data aggregation and analysis
2. Real-time analytics
3. Privacy and security demands
4. Lack of usability of solutions
5. Missing integration of data sources and instruments
6. Complicated US and European privacy laws on health data
7. Diversity of stakeholders

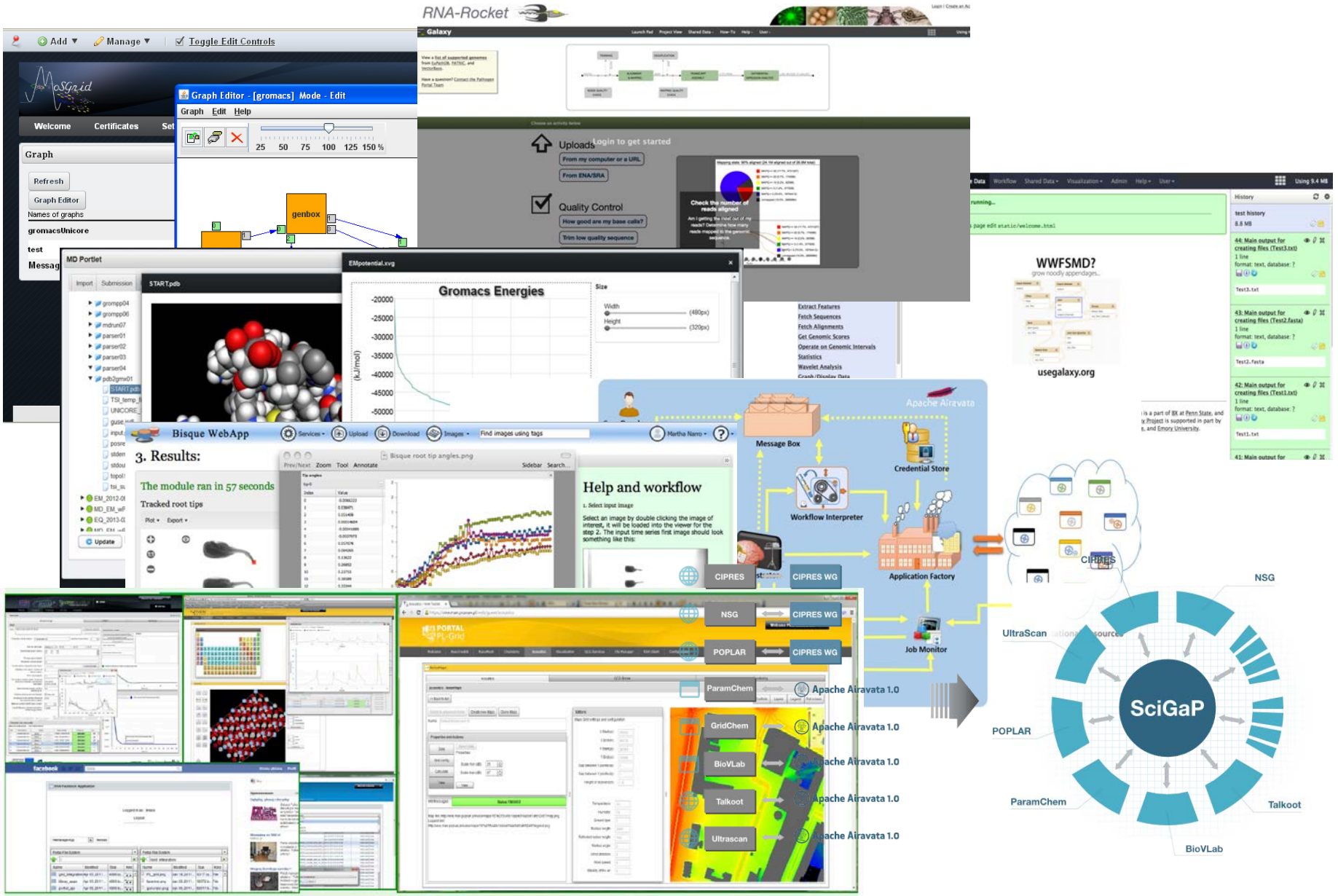
➔ ***Technical solutions can***

- address the first five challenges
- assist with data and measures for easing policies on health data
- support diverse user roles via easy-to-use end-to-end solutions

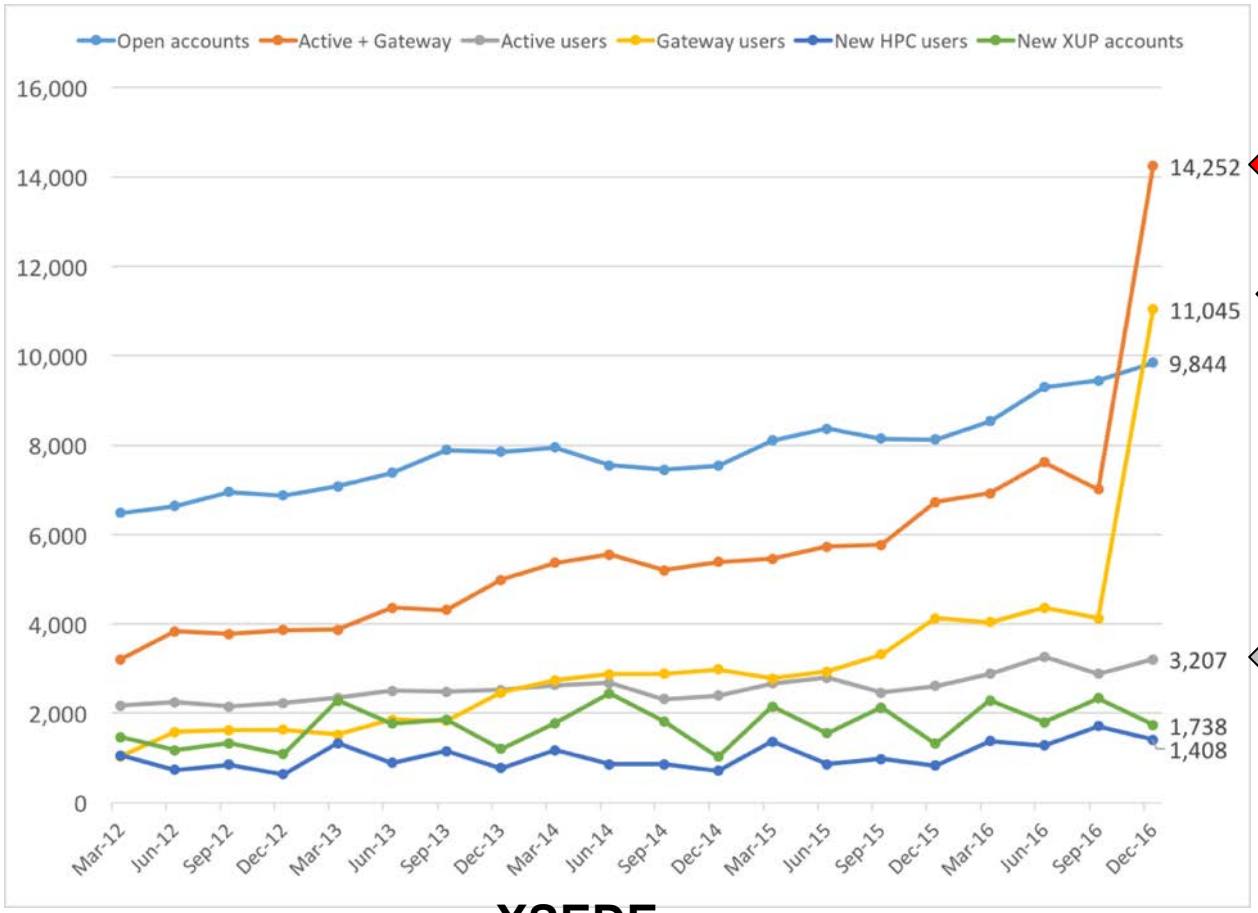
Science Gateways

- Increased complexity of
  - today's research questions
  - hardware and software
  - skills required
- Greater need for openness and reproducibility
  - Science increasingly driving policy questions
- Opportunity to integrate research with teaching
  - Better workforce preparation

We need end-to-end solutions that provide ***broad access*** *to advanced resources and allow* ***all*** *to tackle today's challenging science questions.*


RNA-Rocket
Galaxy
Add
Manage
Toggle Edit Controls
Welcome
Certificates
Graph
Refresh
Graph Editor
Names of graphs
gromacsUnicore
test
Graph Editor - [gromacs] Mode - Edit
Graph Edit Help
genbox
Uploads
Login to get started
From my computer or a URL
From ENA/SRA
Quality Control
How good are my base calls?
Trim low quality sequence
MD Portlet
Import
Submission
START.pdb
EMpotential.xvg
Gromacs Energies
WWFSMD?
usegalaxy.org
Bisque WebApp
3. Results:
The module ran in 57 seconds
Tracked root tips
Help and workflow
Apache Airavata
Message Box
Credential Store
Workflow Interpreter
Application Factory
Job Monitor
CIPRES
CIPRES WG
NSG
POPLAR
ParamChem
GridChem
BioVLab
Talkoot
Ultrascan
Apache Airavata 1.0
SciGaP
UltraScan
POPLAR
ParamChem
Talkoot
BioVLab

Gateway users are 77% of active XSEDE users in Q4 2016



XSEDE users

This is largely due to the CIPRES and I-TASSER gateways, but others are gaining

Science Gateways Survey 2014

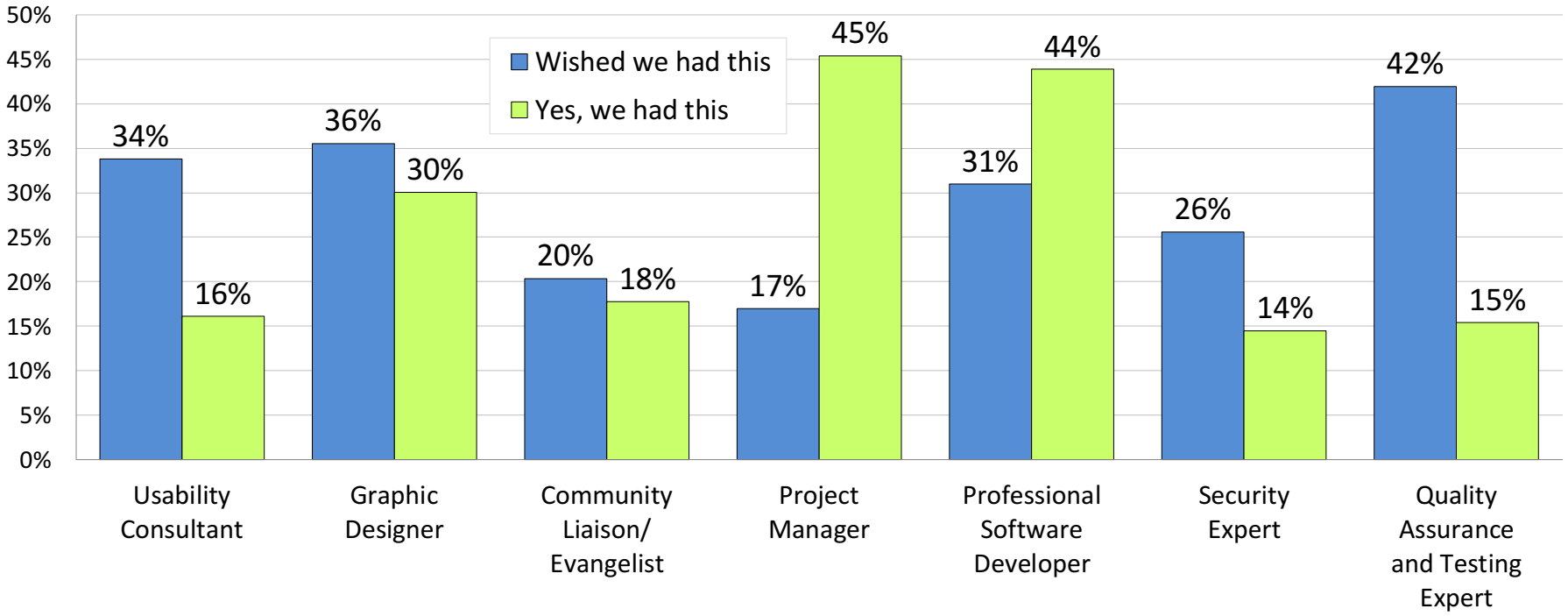
- sent out to 29,000 persons
- 4,957 responses from across domains
- 52% from life, physical or mathematical sciences
- 32% from computer and information sciences or engineering
- 45% develop data collections
- 44% develop data analysis tools

What services would be helpful?

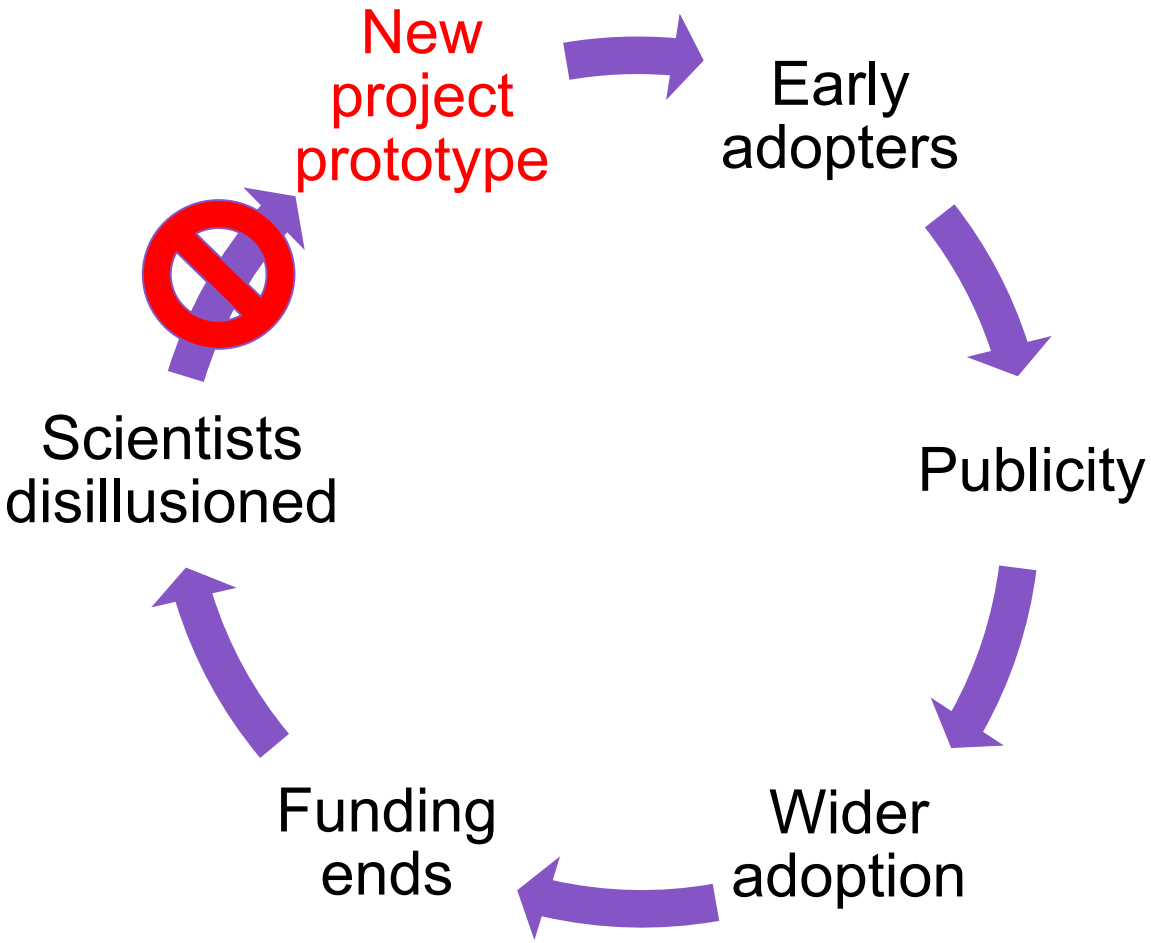
Proposed Service	% Interest
Evaluation, impact analysis, website analytics	72%
Adapting technologies	67%
Web/visual/graphic design	67%
Choosing technologies	66%
Usability Services	66%
Visualization	65%
Developing open-source software	64%
Support for education	64%
Community engagement mechanisms	62%
Keeping your project running	62%
Legal perspectives	61%
Managing data	60%
Computational resources	59%
Mobile technology	59%
Database structure, optimization, and query expertise	59%
Data mining and analysis	58%
Cybersecurity consultation	57%
Website construction	57%
Software engineering process consultation	53%
Source code review and/or audit	51%
High-bandwidth networks	45%
Scientific instruments or data streams	44%
Management aspects of a project	38%

Science Gateways Survey 2014

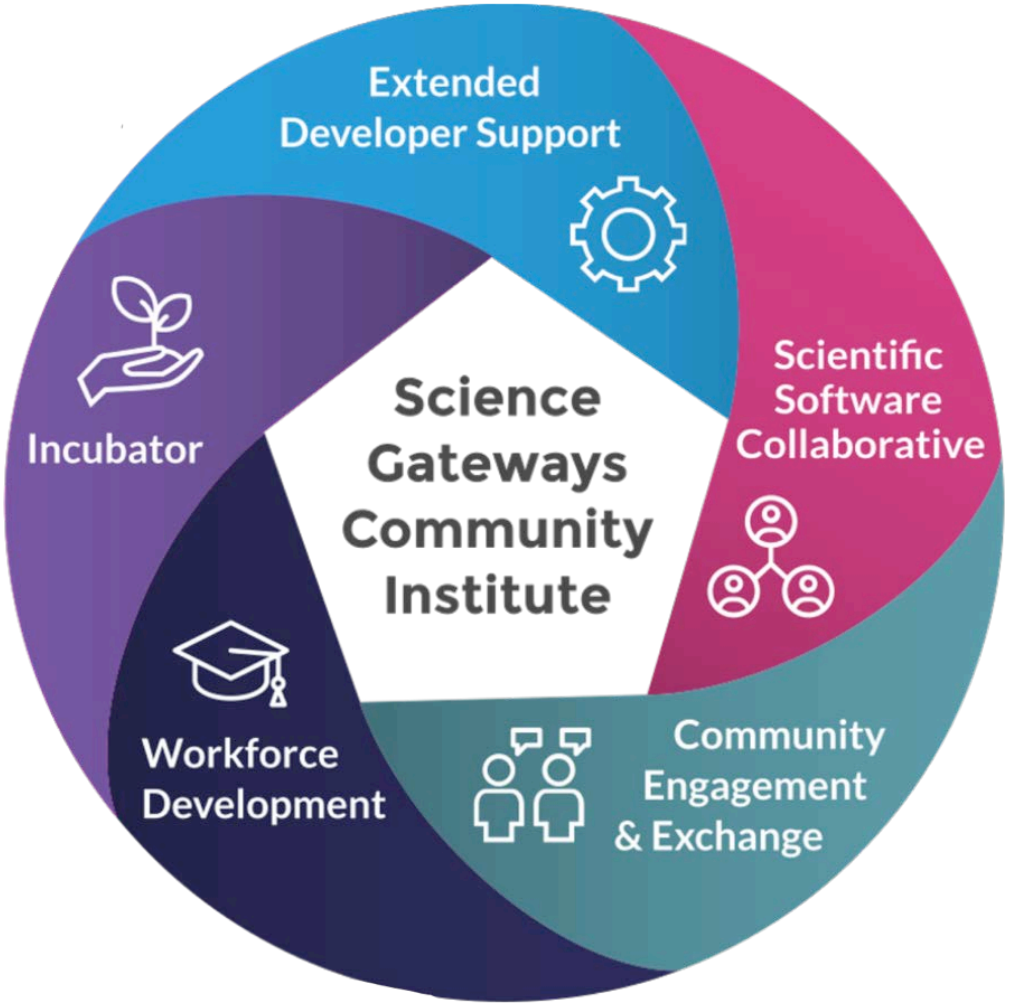
Well-designed gateways require a variety of expertise



Typical Lifecycle of a Science Gateway

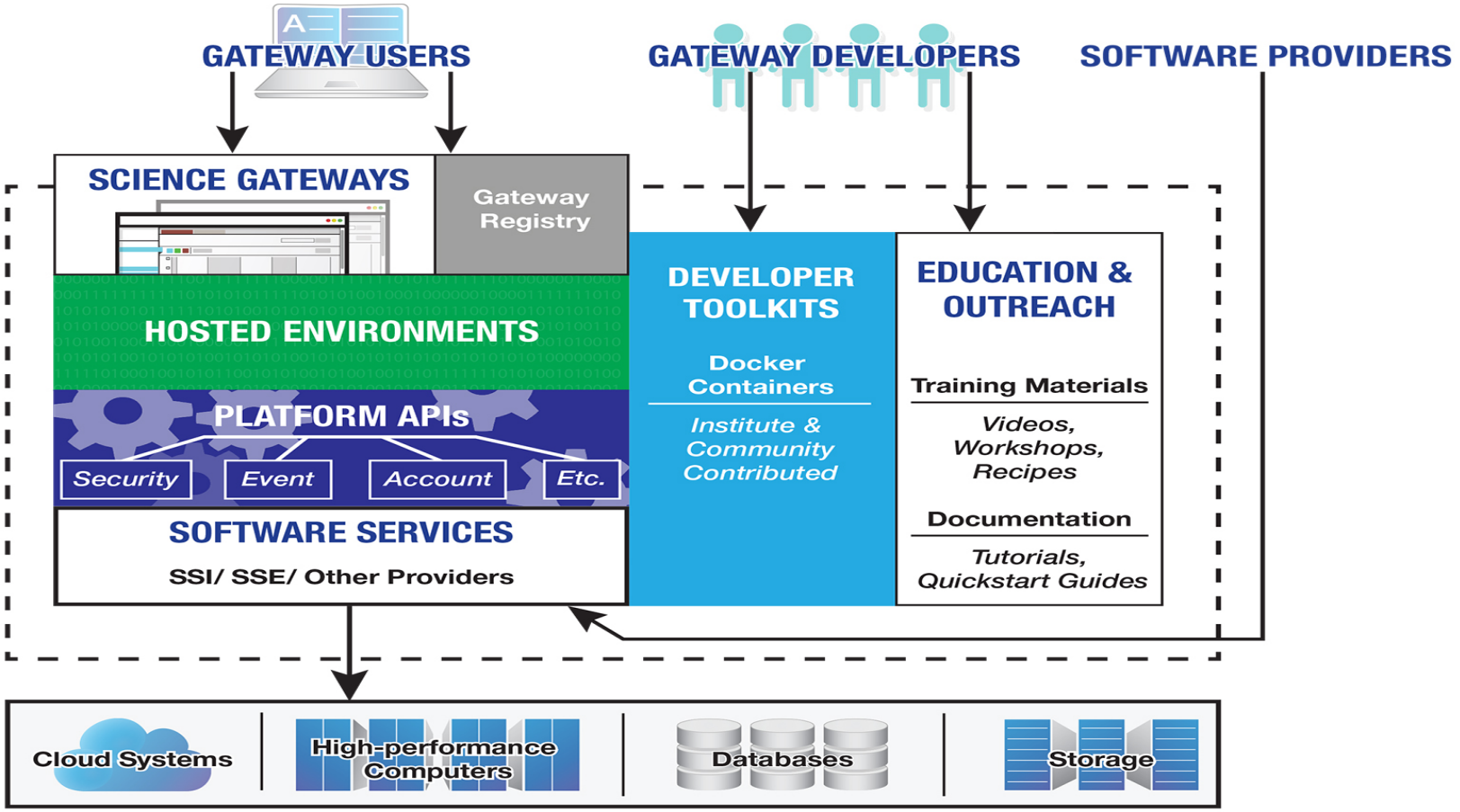


Science Gateways Community Institute



- Diverse expertise on demand
- Longer term support engagements
- Software and visibility for gateways
- Information exchange in a community environment
- Student opportunities and more stable career paths

Scientific Software Collaborative

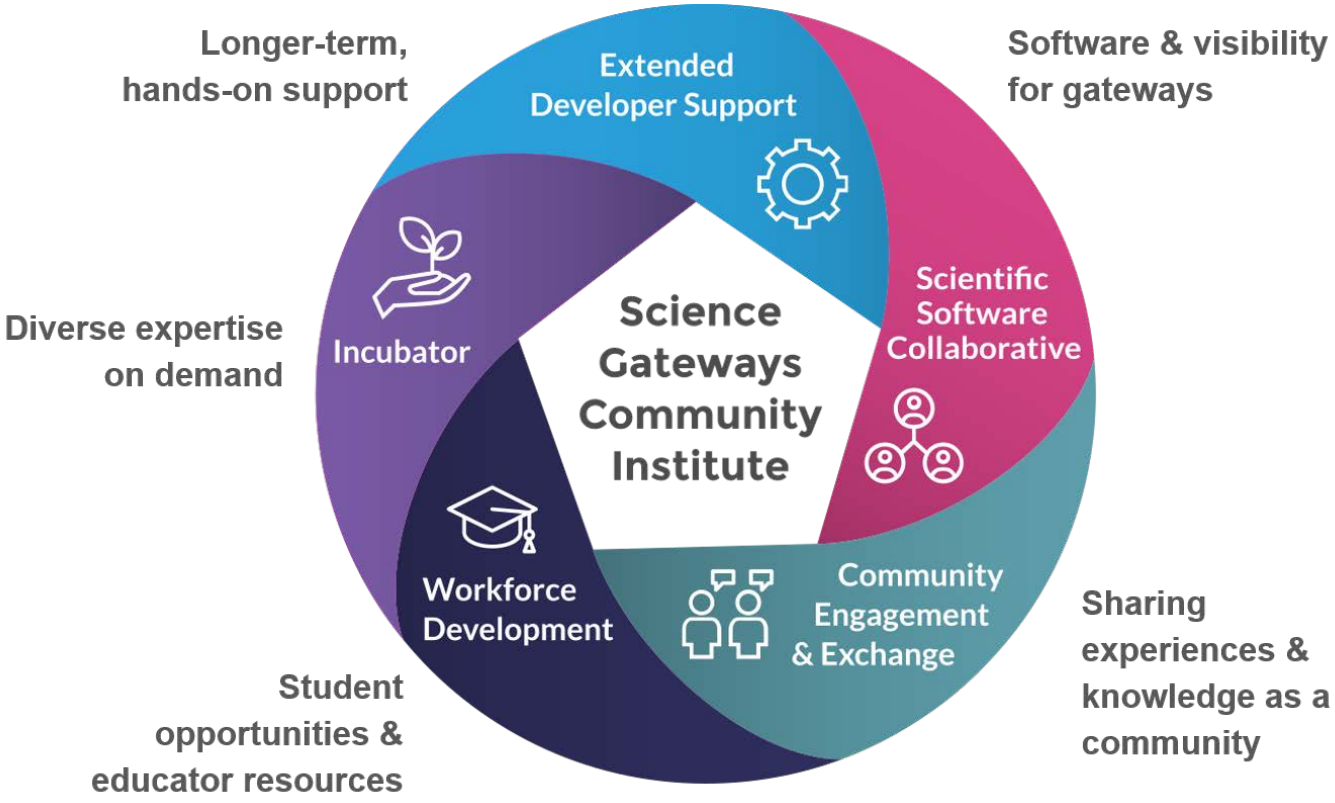


High-level design of the scientific software collaborative

Opportunities

- Get advice (sustainability, usability, cybersecurity, other) on your existing gateway
- Work with SGCI to build a gateway for you
- Find a working gateway or gateway development software in our catalog
- Learn how to set up a gateway group on your campus
- Keep up to date on gateway developments
  - webinar series, gateways in the news, google scholar feed, case studies, annual conference, blog posts
- Enroll a student in an internship program to learn gateway development
- Partner with SGCI on your own projects, technologies and events

Thanks for your attention!



help@sciencegateways.org

http://sciencegateways.org/