

# OneSciencePlace: Sustainability Measures for a Science Gateway Framework

Sandra Gesing, David R. Benham and the HUBzero Team  
sandra.gesing@nd.edu

2019 Collegeville Workshop on Sustainable Scientific Software  
July 24, 2019

# Technology-Enhanced Research

- 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

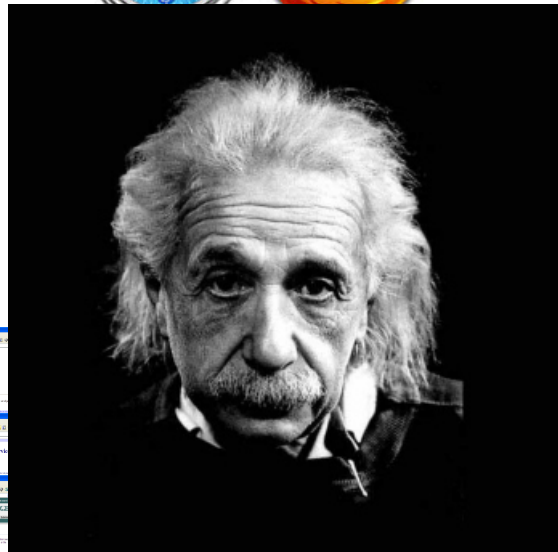
*We need end-to-end solutions that provide **broad access to advanced resources** and allow **all** to tackle today's challenging science questions*

**➔ Science Gateways**

Data and compute-intensive problems

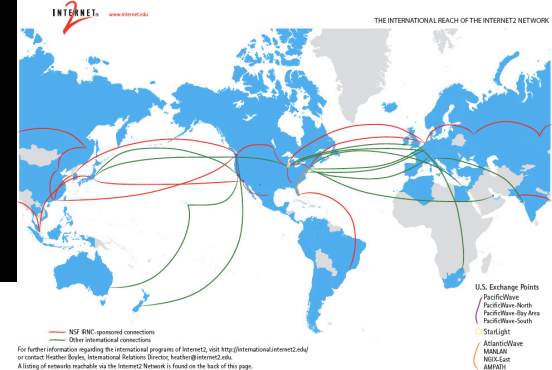


Web-based agile frameworks



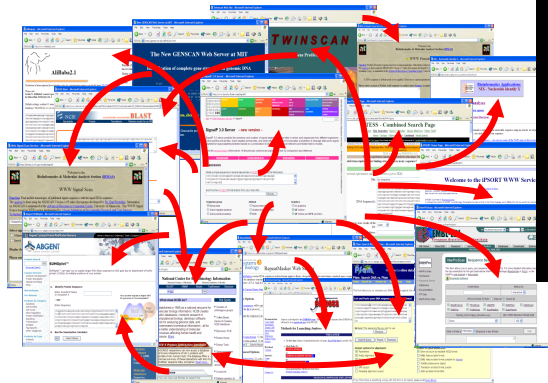
Users generally not IT specialists

Distributed data and computing infrastructures



High-speed networks

Tools and workflow engines



Data and compute-intensive problems

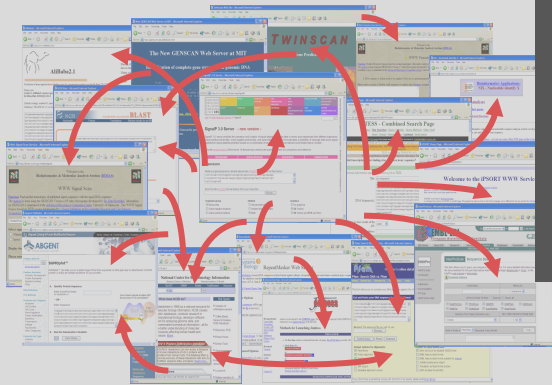
Web-based agile frameworks



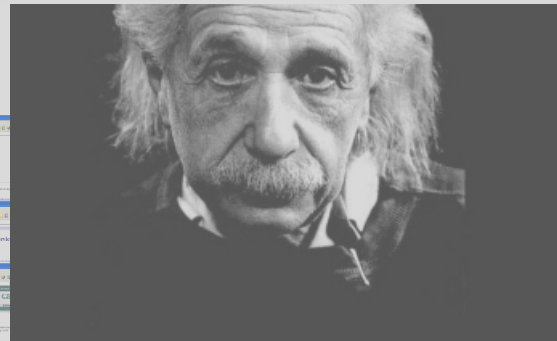
Distributed data and computing infrastructures



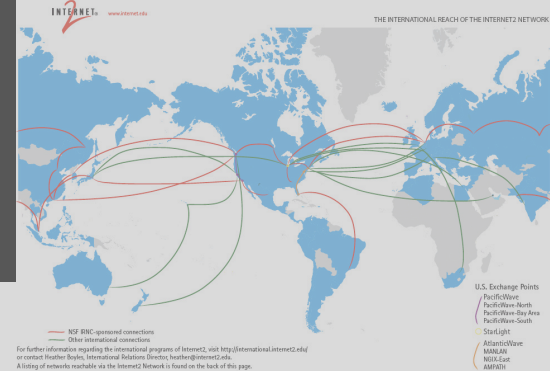
**Need for science gateways!**



Tools and workflow engines



Users generally not IT specialists

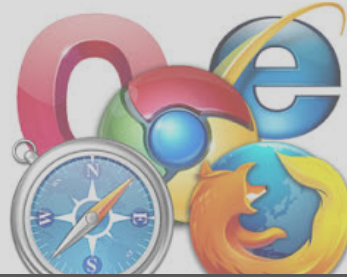


High-speed networks

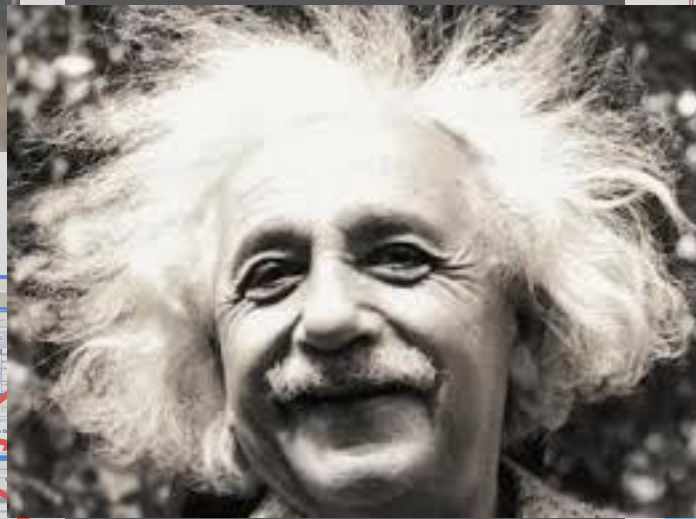


Data and compute-intensive problems

Web-based agile frameworks

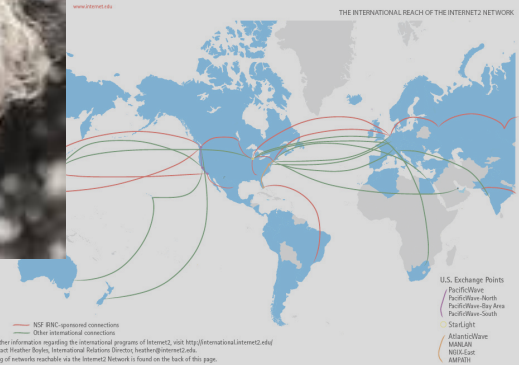


Distributed data and computing infrastructures

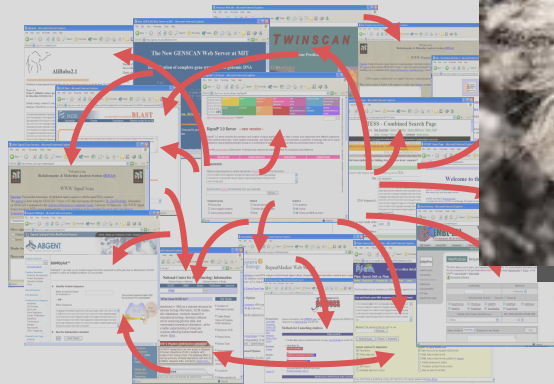


Users generally not IT specialists

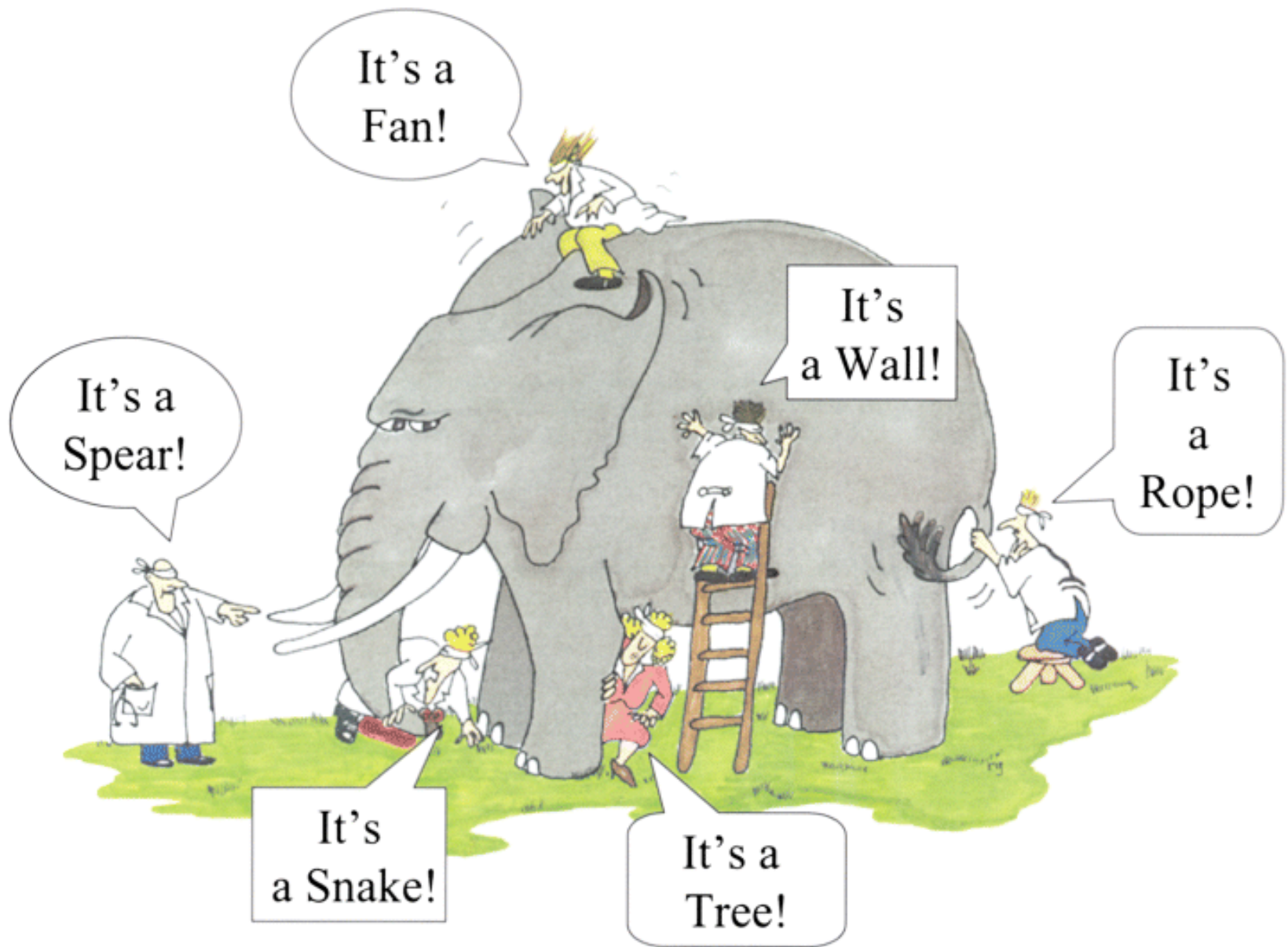
High-speed networks

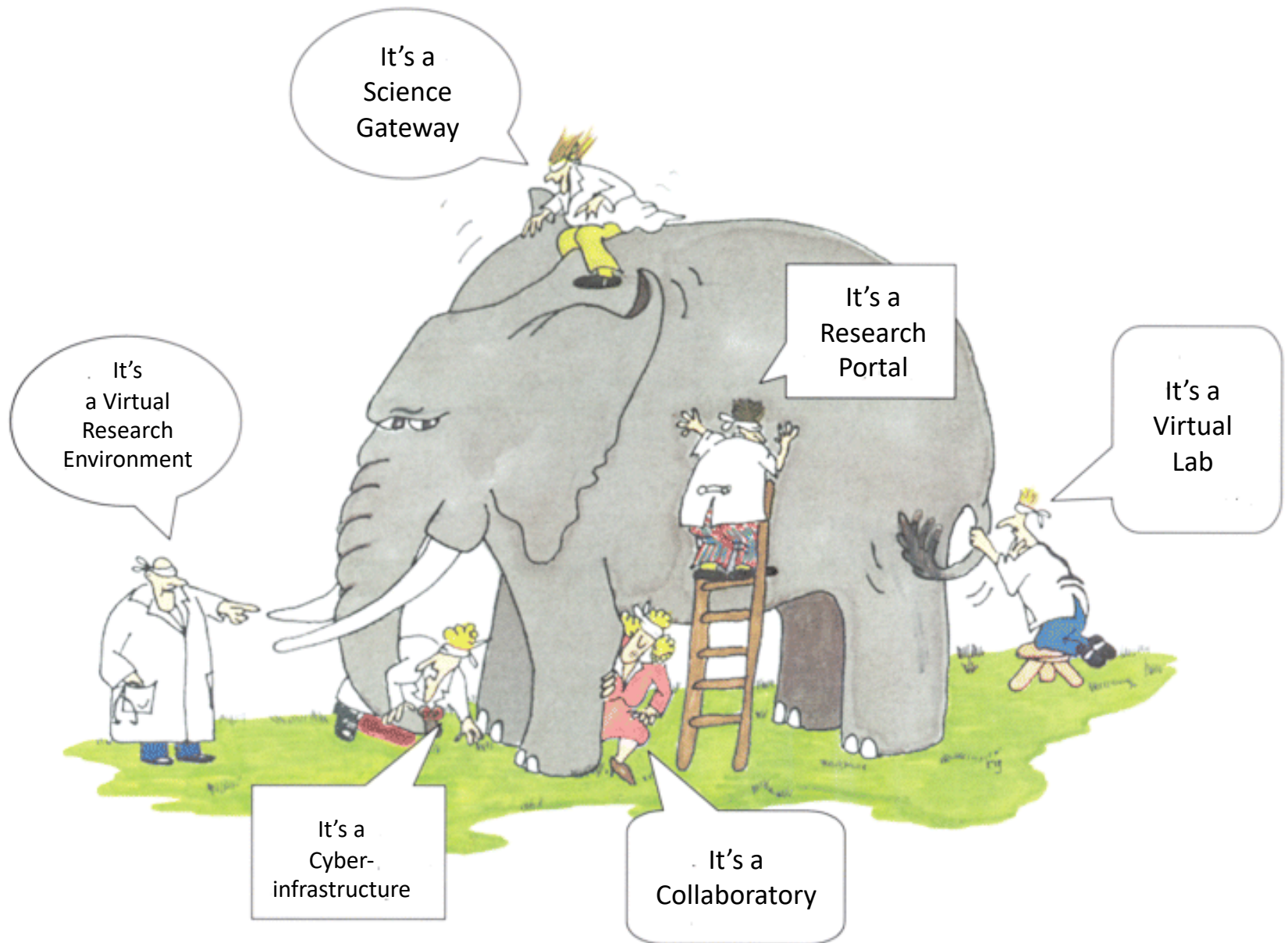


For further information regarding the international programs of Internet2, visit <http://international.internet2.edu/> or contact Heather Boyles, International Relations Director, [hboyles@internet2.edu](mailto:hboyles@internet2.edu). A listing of networks reachable via the Internet2 Network is found on the back of this page.

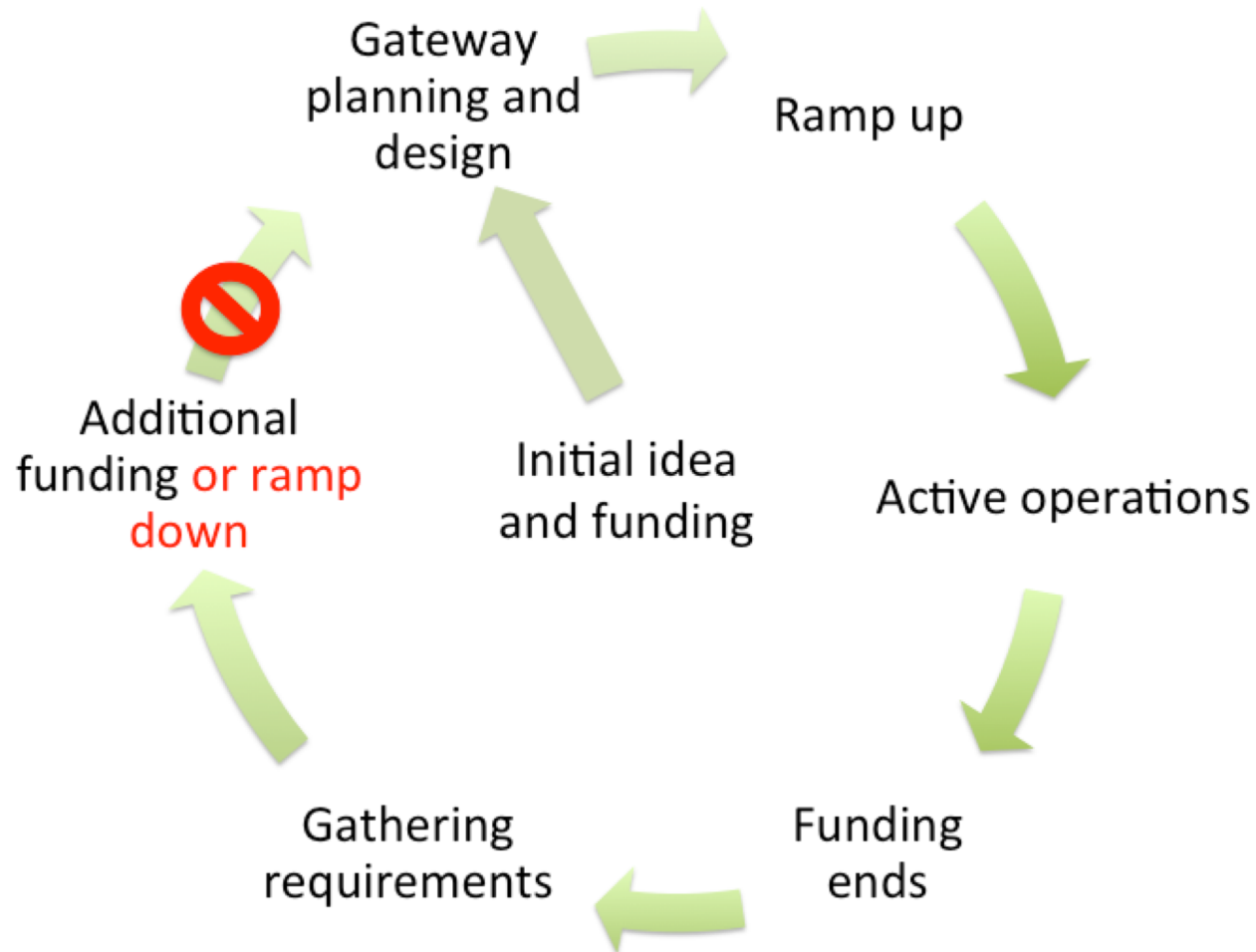


Tools and workflow engines





# Lifecycle of a Science Gateway



# HUBzero History: Operating Model Path

- 1996: PUNCH (precursor to nanoHUB)
- 2002: NSF funding for nanoHUB
- 2007: Spin-out of HUBzero from nanoHUB
- 2015: Diversification, self-funded sustainability
- Composition:
  - Hubs hosted by Purdue (for a variety of institutions)
  - Foundation members running their own hubs
  - Open source hubs



# HUBzero History: Key Team Enablers

- 1996: Vision (Mark Lundstrom)
- 2002: Vertical Expansion (Gerhard Klimeck)
- 2007: Horizontal Expansion (Michael McLennan)
- 2015: Business Transition (Michael Zentner)
- Composition:
  - Hubs hosted by Purdue (for a variety of institutions)
  - Foundation members running their own hubs
  - Open source hubs

# HUBzero Origins

NANO *is* **HUGE**

LARGEST NANOTECHNOLOGY ONLINE RESOURCE

500+

simulation tools

1.4M

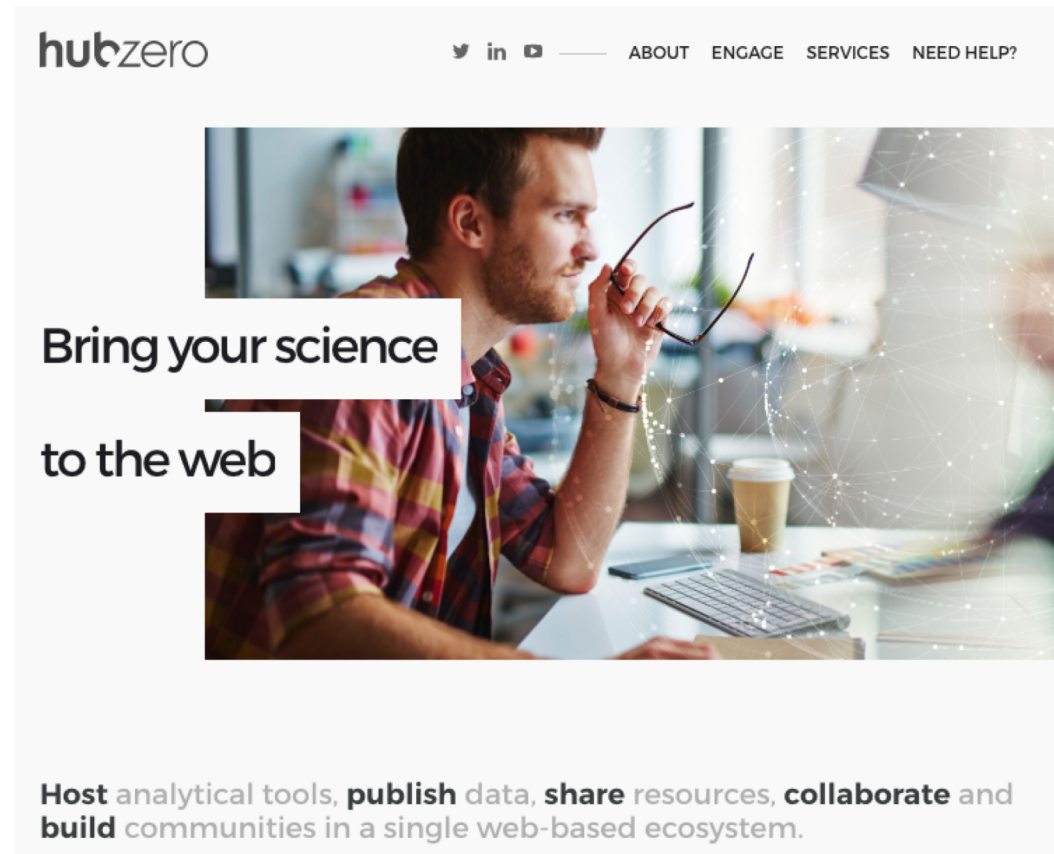
users

5500+

resources

# HUBzero Today

- Open source platform
- Foundation
- Development
- Hosting services
- 60+ Hubs
- 1.5 million visitors worldwide



# HUBzero

- HUBzero® (<https://hubzero.org>) is an open source software platform for building powerful websites that host analytical tools, publish data, share resources, collaborate and build communities in a single web-based ecosystem.

- A hub is a place for you to:

Build communities

Use & Publish tools

Jupyter Notebooks

Linux Apps

Windows Apps

RStudio

Web Apps

Publish Research



# The Platform

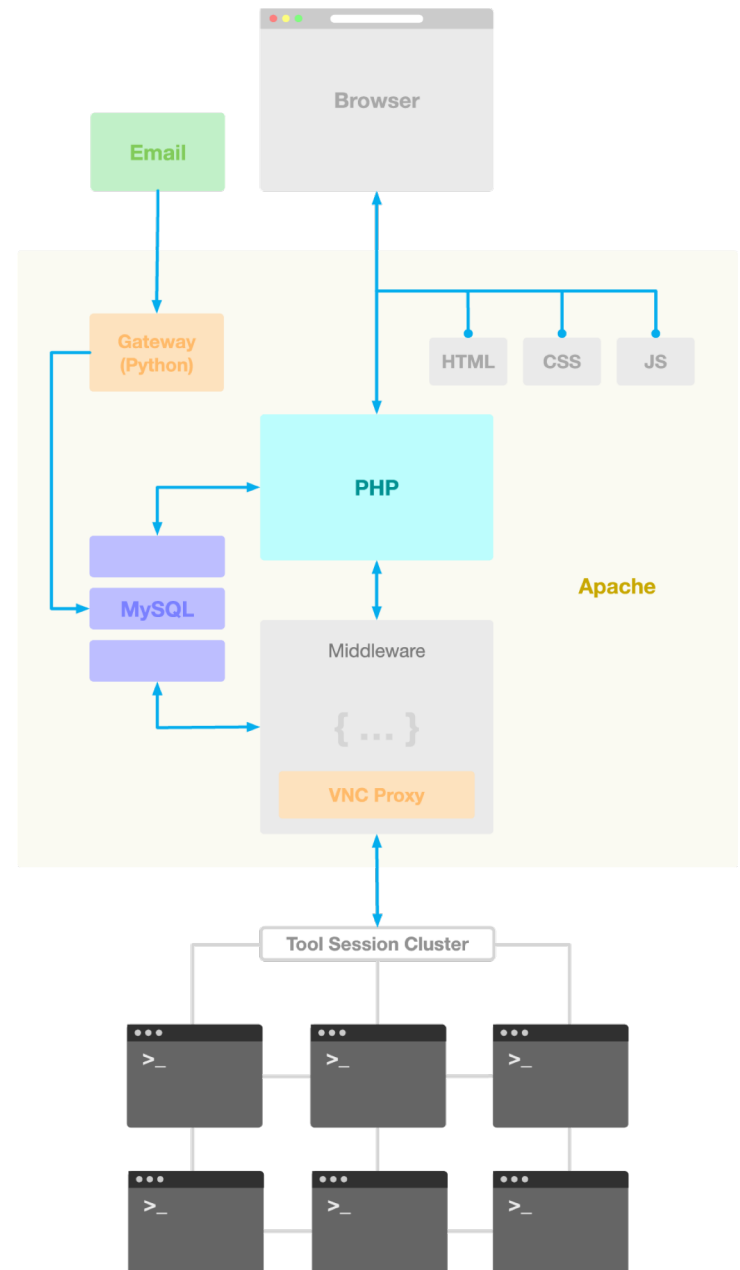
- LAMP stack

PHP & MySQL: Content Management System

Python: Middleware Library & Email Gateway

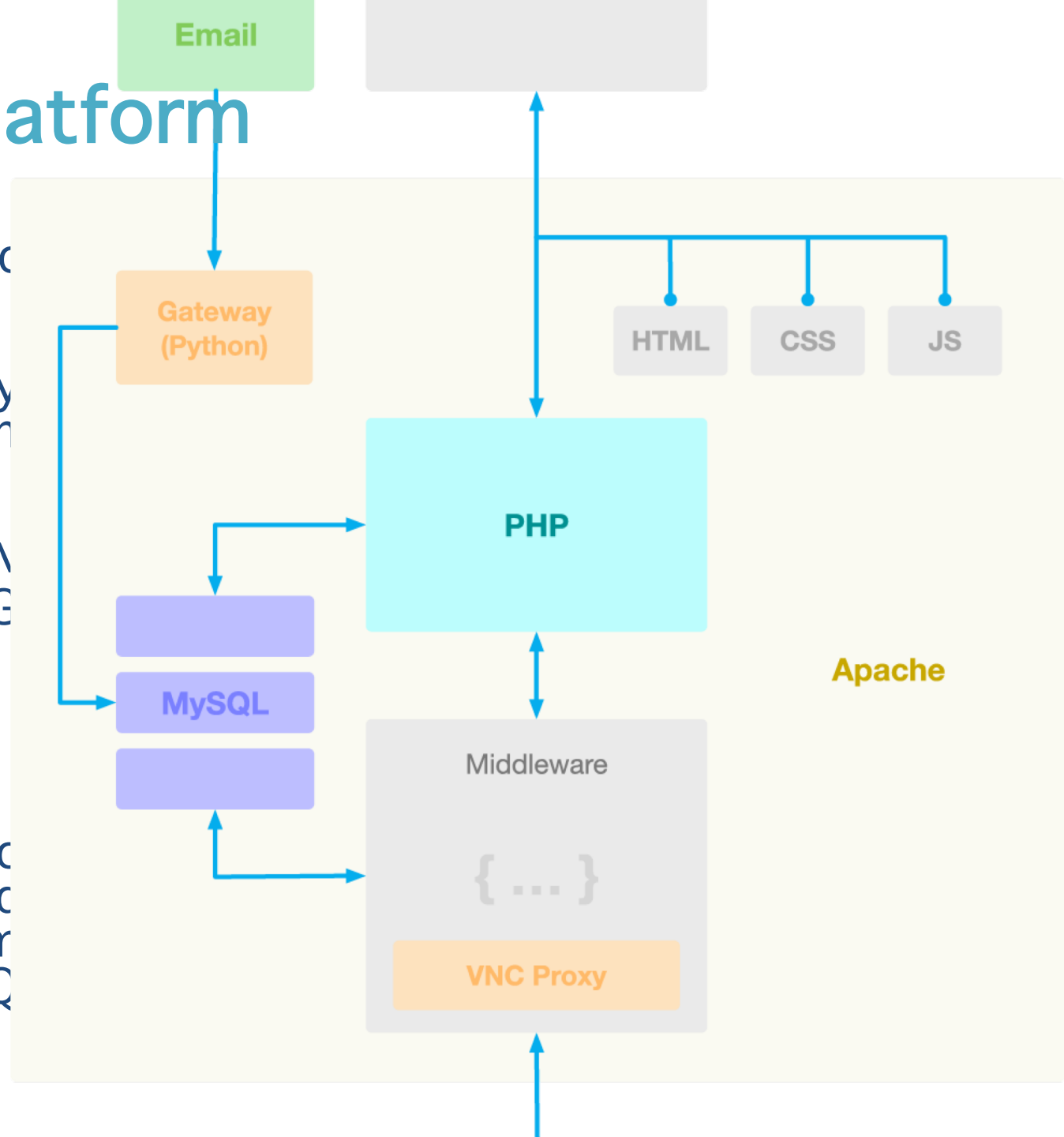
OpenVZ: Tool containers

Other technologies employed: Docker, Node.js, Web Components, RabbitMQ, Redis





Other technologies employed  
Web Components  
RabbitMQ



# Collaborations & Integrations

InCommon®

 THOMSON REUTERS

scistarter

 slack

 Dropbox

facebook

 GitLab

twitter 

 GitHub

Google+

PRESSFORWARD 

ORCID

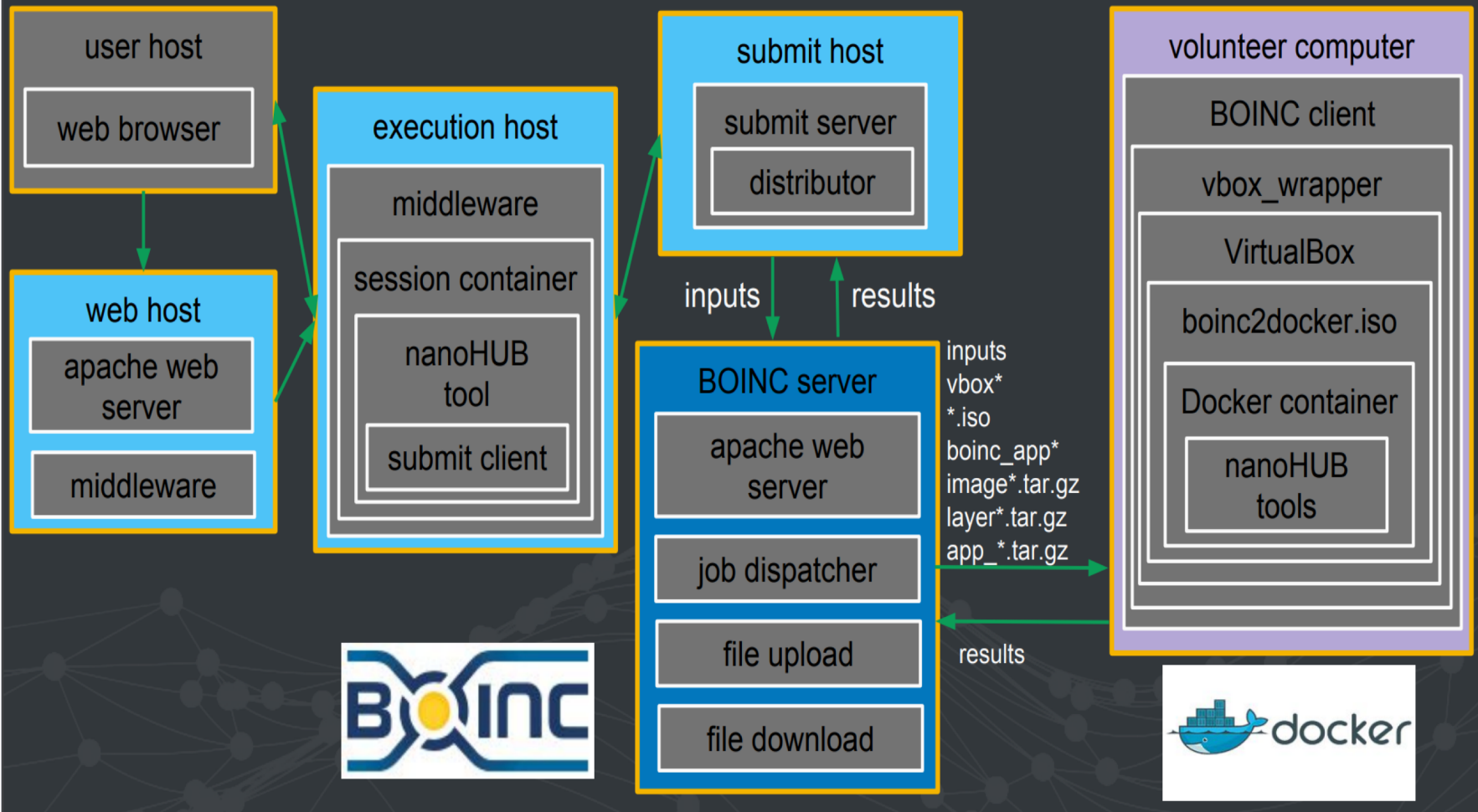
iRODS

Linked  in

Solr 

 globus

# ARCHITECTURE - nanoHUB/BOINC



# HUBzero Gateways

## REMEDI

<https://catalyzecare.org>

An online platform dedicated to the improvement of healthcare services by applying engineering principles in all aspects of service delivery.

## PURR

<https://purrr.purdue.edu>

A research collaboration and data management solution for Purdue University researchers and their collaborators.

## NCIPhub

<https://nciphub.org>

Community research and collaboration in cancer research and informatics for sharing resources, hosting online communities, and collaboration.

## DiaGrid

<https://diagrid.org>

A scientific computing gateway for instant access to research tools and a distributed network of HPC, HTC and data resources.

## nanoHUB

<https://nanohub.org>

The largest online nanotechnology facility, hosting online simulation tools and resources for use in education and research.

## CDMhub

<https://cdmhub.org>

Increase the number and use of simulation tools through education and evaluation of existing and emerging simulation tools.

# OneSciencePlace – Aiming at Sustainability for Science Gateways

- Long-term community-hosting concept
- Flexible financial models
- Resource usage tracking
- Cross-site embeddable tools
- Scalability
- Interoperability



# Thanks!

## Questions?

sandra.gesing@nd.edu

hubzero

<https://hubzero.org/>