



Sharing reproducible computations on Binder

Lindsey Heagy, UC Berkeley & Project Jupyter

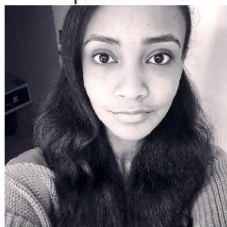
andrewsh



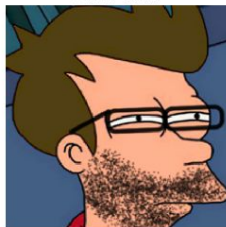
betatim



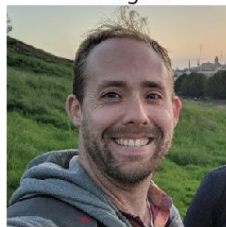
captainsafia



Carreau



choldgraf



consideratio



ellisonbg



fperez



freeman-lab



henhc



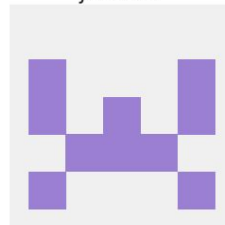
JamiesHQ



jhamrick



jzf2101



lheagy



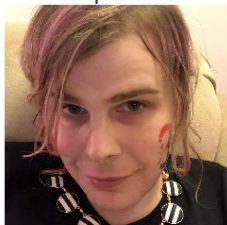
mbmilligan



minrk



mpacer



parente



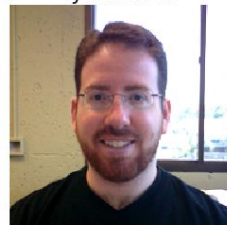
rgbkrk



Ruv7



ryanlovett



sgibson91



takluyver



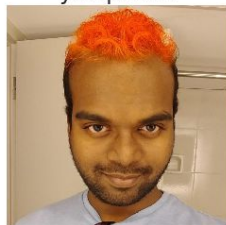
tgeorgeux



willingc



yuvipanda



Zsailer

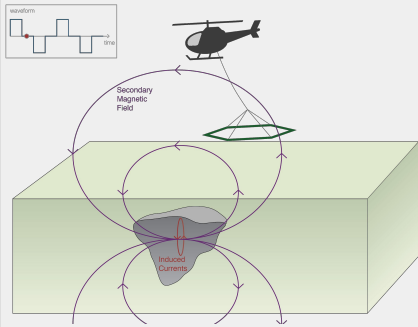


you?



👋 hello (a bit about me)

geophysical
inversions



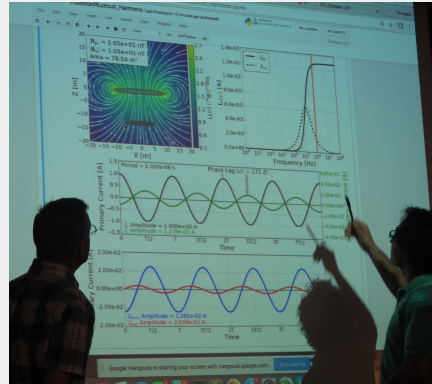
THE UNIVERSITY
OF BRITISH COLUMBIA

open-source
software



simpeg

open research &
education



GeoSci
.xyz

Jupyter, geoscience
+ data science



Berkeley
UNIVERSITY OF CALIFORNIA



a *community* of people and an *ecosystem* of open tools and standards for interactive computing



the science is the code

An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.

-- Buckheit and Donoho
WaveLab and Reproducible Research, 1995

the science is the code

An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the **complete software development environment** and the **complete set of instructions** which generated the figures.

-- Buckheit and Donoho
WaveLab and Reproducible Research, 1995

the science is the code

An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the **complete software development environment** and the **complete set of instructions** which generated the figures.

-- Buckheit and Donoho
WaveLab and Reproducible Research, 1995

^
(and a place to run the code?)



live, computational environments, running on the
cloud, built from your research repositories

mybinder.org



Turn a Git repo into a collection of interactive notebooks

Have a repository full of Jupyter notebooks? With Binder, open those notebooks in an executable environment, making your code immediately reproducible by anyone, anywhere.

Build and launch a repository

GitHub repository name or URL

GitHub ▾

Git branch, tag, or commit


Path to a notebook file (optional)

File ▾

launch

Copy the URL below and share your Binder with others:



Copy the text below, then paste into your README to show a binder badge:  launch binder

origins



github.com/freeman-lab



+



GitHub

+

explicit dependencies

+



docker



kubernetes



- creates reproducible containers from repositories ([repo2docker](#))
- generates user sessions that serve these containers ([JupyterHub](#))
- provides an interface to create, share, and use these sessions ([BinderHub](#))
- demonstrates the above as a free public service/tech demo ([mybinder.org](#))

open science

(why?)

- in order to collaborate
- to build on the work of others
- for others to build upon your work
- to make revisions to your paper when you hear back from reviewers in 8 months

open science

(how?)

- complete set of instructions
- complete development environment
- a place to run the code

open science

(how?)

- complete set of instructions
- complete development environment
- a place to run the code

complete set of instructions

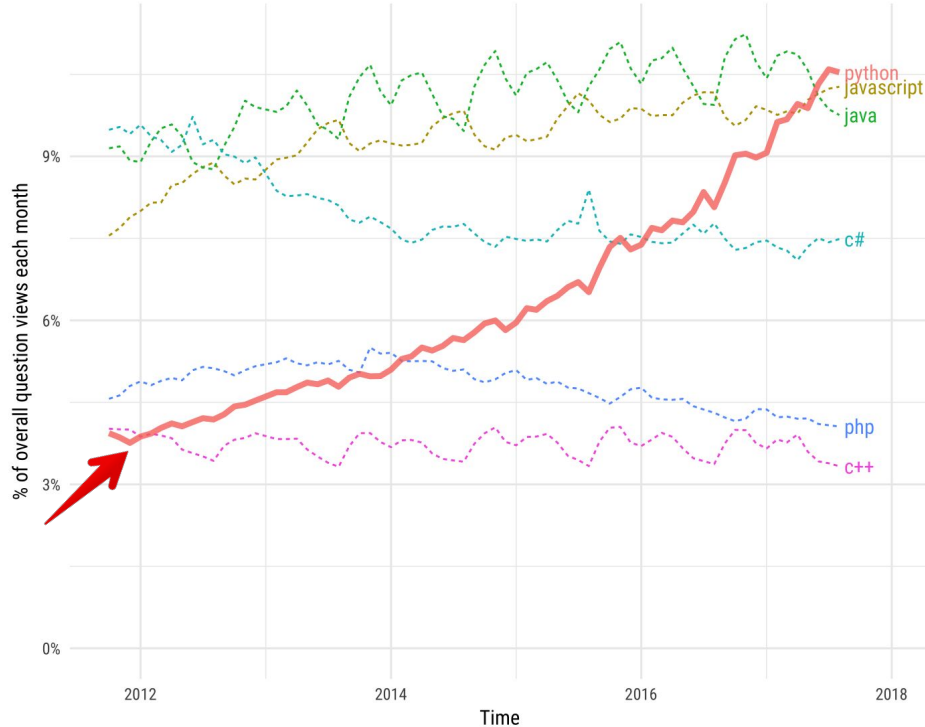
open-source languages are the raw material



complete set of instructions

Growth of major programming languages

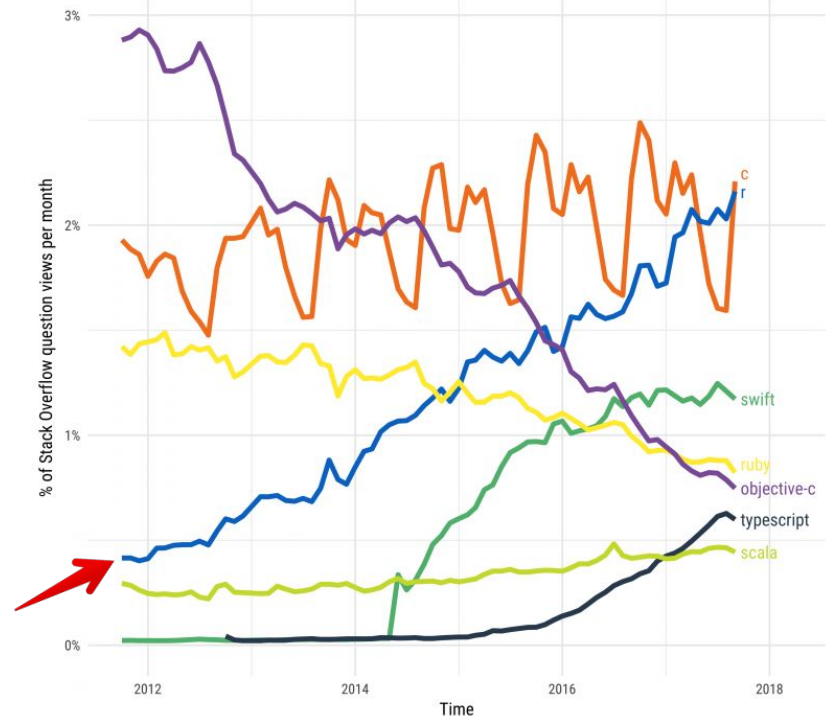
Based on Stack Overflow question views in World Bank high-income countries



<https://stackoverflow.blog/2017/09/06/incredible-growth-python/>

Stack Overflow Traffic to Programming Languages

Based on visits to Stack Overflow questions from World Bank high-income countries. The more-visited languages of Python, JavaScript, Java, C#, and PHP were omitted.



<https://stackoverflow.blog/2017/10/10/impressive-growth-r/>

complete set of instructions

mature ecosystems of tools



IPython



complete set of instructions

web-native interfaces for interacting with code



complete set of instructions

capture the steps



Runs in web browser

Text, comments, equations

Lines of code
(Python, Julia, R, Matlab, ...)

Output: values, images, plots,
tables, interactive widgets

The screenshot shows a JupyterLab interface in a web browser. The browser address bar shows 'localhost:8888/lab?'. The notebook title is 'MagneticDipoleApplet.ipynb'. The notebook content includes a title 'Data over a Magnetic Dipole', a 'Purpose' section with text about geophysics, a code cell with Python code, and a plot area with a heatmap and a line graph. The plot area includes interactive widgets for field selection, profile selection, and various sliders for parameters like current (I), depth (D), length, data spacing, and depth (depth).

complete set of instructions

capture the steps: what is a notebook?

A document



An interface



An environment



complete set of instructions

maintenance and sharing



- version control
- issue tracking
- licensing
- integrations with
 - testing services
 - documentation hosting
 - ...

complete set of instructions

maintenance and sharing

binder-examples / r

Watch 13 Star 66 Fork 112

Code Issues 3 Pull requests 1 Projects 0 Security Insights

Using R with Jupyter / RStudio on Binder

binder binder-ready

41 commits 2 branches 0 releases 5 contributors BSD-3-Clause

Branch: master New pull request Find File Clone or download

betatim Update MRAN snapshot Latest commit 1b68591 on Apr 15

bus-dashboard	Remove the DESCRIPTION file	a year ago
LICENSE	Create LICENSE	2 months ago
README.md	Update README.md	4 months ago
index.ipynb	adding example	2 years ago
install.R	Add example Shiny app	a year ago
runtime.txt	Update MRAN snapshot	a month ago

open science

(how?)

- complete set of instructions
- complete development environment
- a place to run the code

repo2docker



repo2docker deterministically build a docker image
from a repository with documented dependencies

complete development environment

define dependencies following community standards of practice



```
requirements.txt
1 numpy==1.16.*
2 matplotlib==3.*
3 seaborn==0.8.1
4 pandas
5

environment.yml
1 name: example-environment
2 channels:
3   - conda-forge
4 dependencies:
5   - numpy
6   - psutil
7   - toolz
8   - matplotlib
9   - dill
10  - pandas
11  - partd
12  - bokeh
13  - dask
14
```

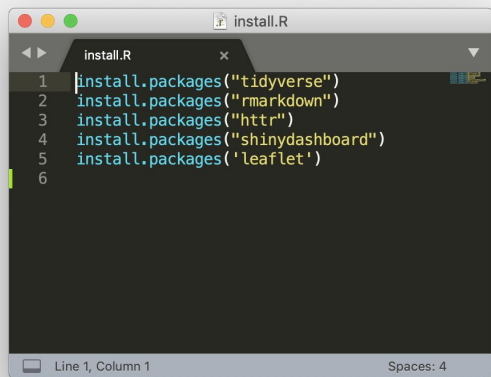


```
runtime.txt
1 r-2019-04-10

install.R
1 install.packages("tidyverse")
2 install.packages("rmarkdown")
3 install.packages("httr")
4 install.packages("shinydashboard")
5 install.packages('leaflet')
6
```


complete development environment

repo2docker

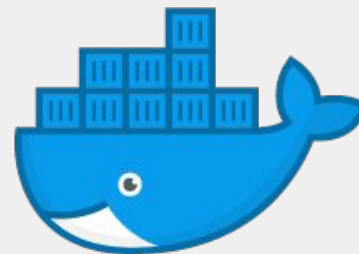


```
install.R
1 install.packages("tidyverse")
2 install.packages("rmarkdown")
3 install.packages("httr")
4 install.packages("shinydashboard")
5 install.packages("leaflet")
6
```

dependencies



repo2docker



container file

complete development environment

define dependencies

binder-examples / r

Watch 13 Star 66 Fork 112

Code Issues 3 Pull requests 1 Projects 0 Security Insights

Using R with Jupyter / RStudio on Binder

binder binder-ready

41 commits 2 branches 0 releases 5 contributors BSD-3-Clause

Branch: master New pull request Find File Clone or download

betatim Update MRAN snapshot Latest commit 1b68591 on Apr 15

bus-dashboard	Remove the DESCRIPTION file	a year ago
LICENSE	Create LICENSE	2 months ago
README.md	Update README.md	4 months ago
index.ipynb	adding example	2 years ago
install.R	Add example Shiny app	a year ago
runtime.txt	Update MRAN snapshot	a month ago

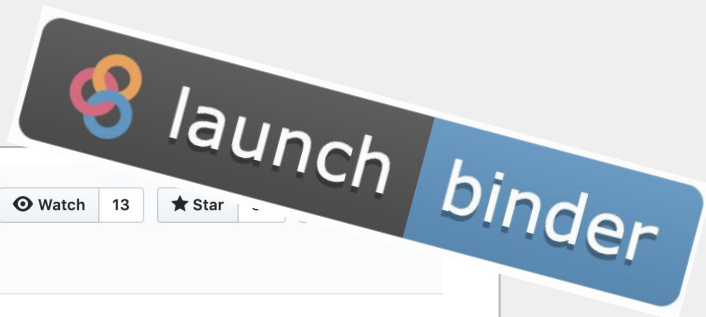
open science

(how?)

- complete set of instructions
- complete development environment
- a place to run the code

a place to run the code

binder: computational environments on the cloud



[binder-examples / r](#) Watch 13 Star

[Code](#) [Issues 3](#) [Pull requests 1](#) [Projects 0](#) [Security](#) [Insights](#)

Using R with Jupyter / RStudio on Binder

[binder](#) [binder-ready](#)

[41 commits](#) [2 branches](#) [0 releases](#) [5 contributors](#) [BSD-3-Clause](#)

Branch: [master](#) [New pull request](#) [Find File](#) [Clone or download](#)

betatim Update MRAN snapshot ... Latest commit 1b68591 on Apr 15

bus-dashboard	Remove the DESCRIPTION file	a year ago
LICENSE	Create LICENSE	2 months ago
README.md	Update README.md	4 months ago
index.ipynb	adding example	2 years ago
install.R	Add example Shiny app	a year ago
runtime.txt	Update MRAN snapshot	a month ago

a place to run the code

binder!



binder



repo2docker



JupyterHub
(next talk!)



Turn a Git repo into a collection of interactive notebooks

Have a repository full of Jupyter notebooks? With Binder, open those notebooks in an executable environment, making your code immediately reproducible by anyone, anywhere.

Build and launch a repository

GitHub repository name or URL

GitHub ▾

Git branch, tag, or commit

Path to a notebook file (optional)

File ▾

launch

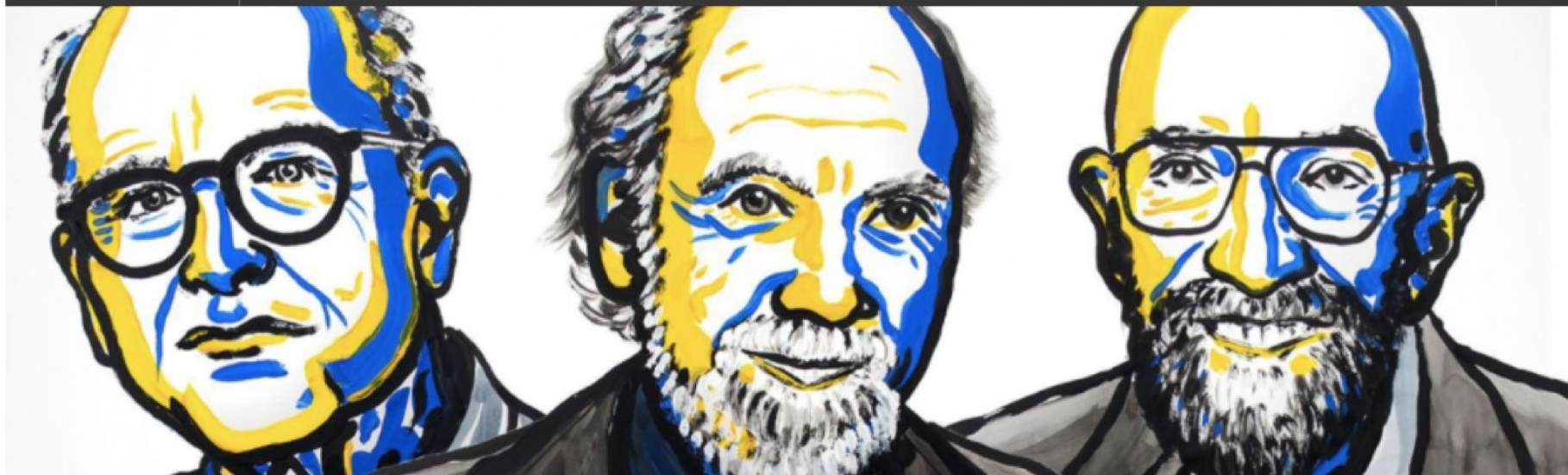
Copy the URL below and share your Binder with others:

Fill in the fields to see a URL for sharing your Binder.



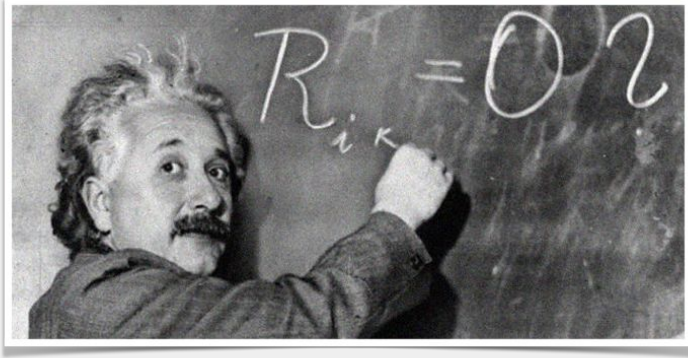
Copy the text below, then paste into your README to show a binder badge:  [launch binder](#)





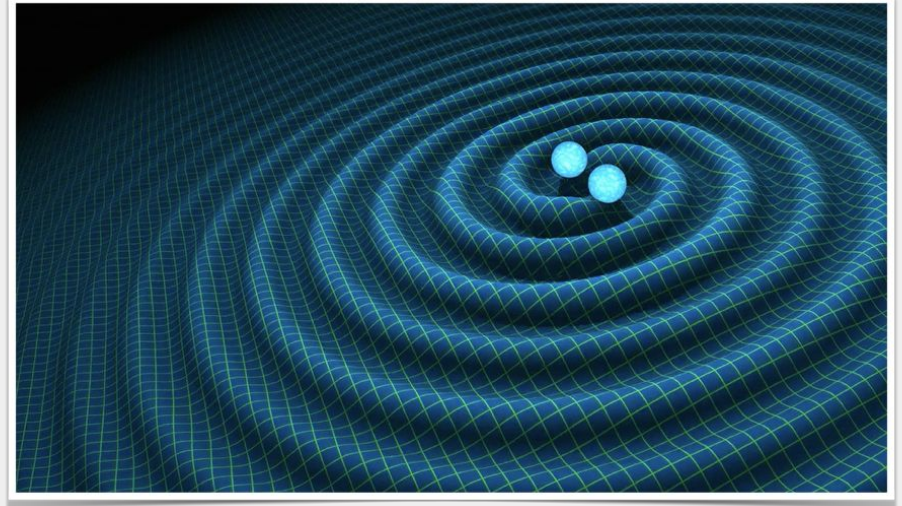
2017 NOBEL PRIZE IN PHYSICS AWARDED FOR DISCOVERY OF GRAVITATIONAL WAVES

as predicted by Einstein



$$R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}$$

Einstein's Field Equations of General Relativity
Annalen der Physik, 1916



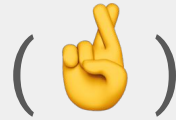
LIGO collaboration discovery: Sept 14, 2015



Detection problem:

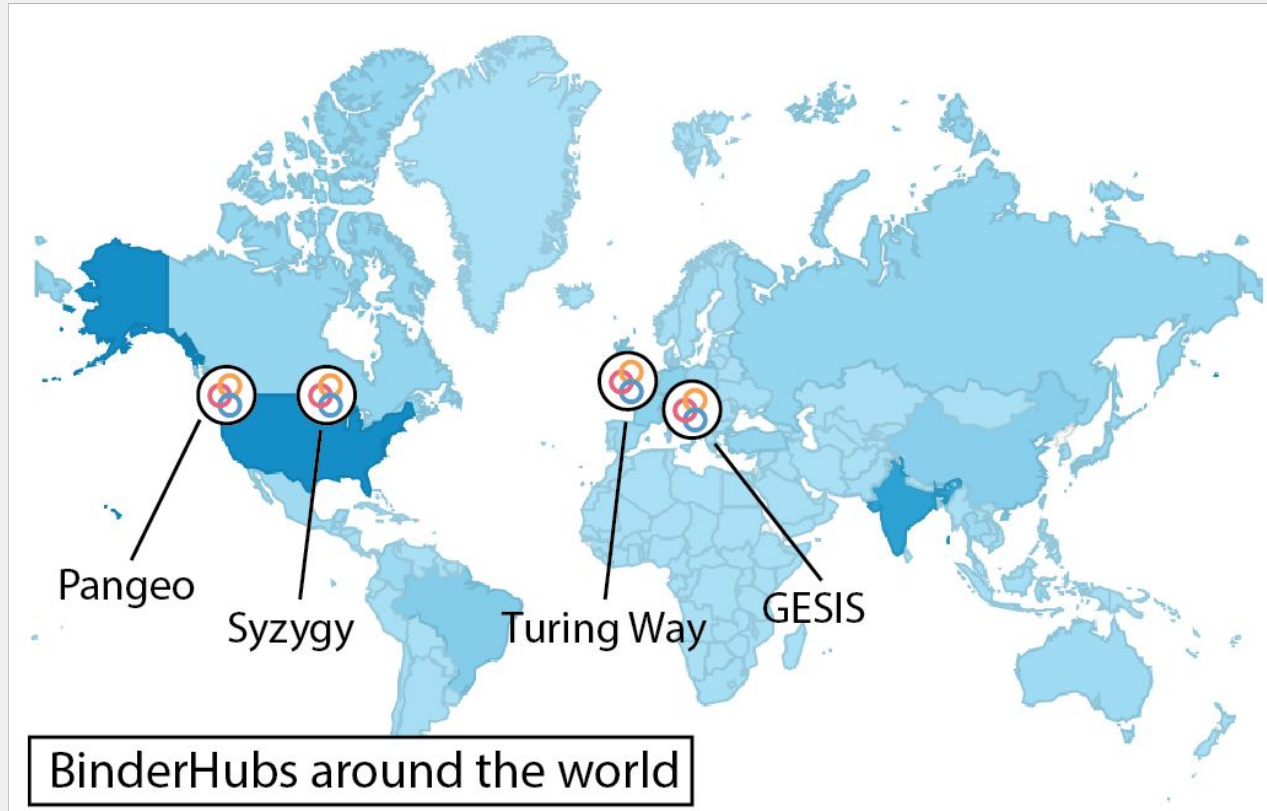
- $\sim 1/1000$ proton over 4 km.
- Sensitivity $\sim 1e-21$
- Milky Way: $1e+21$ m across!

demo time



<https://www.gw-openscience.org/tutorials/>
<https://github.com/binder-examples/r>

Future



thank you!



try out binder:

mybinder.org

connect with the community:

discourse.jupyter.org

THE LEONA M. AND HARRY B.
HELMSLEY
CHARITABLE TRUST



ALFRED P. SLOAN
FOUNDATION

GORDON AND BETTY
MOORE
FOUNDATION