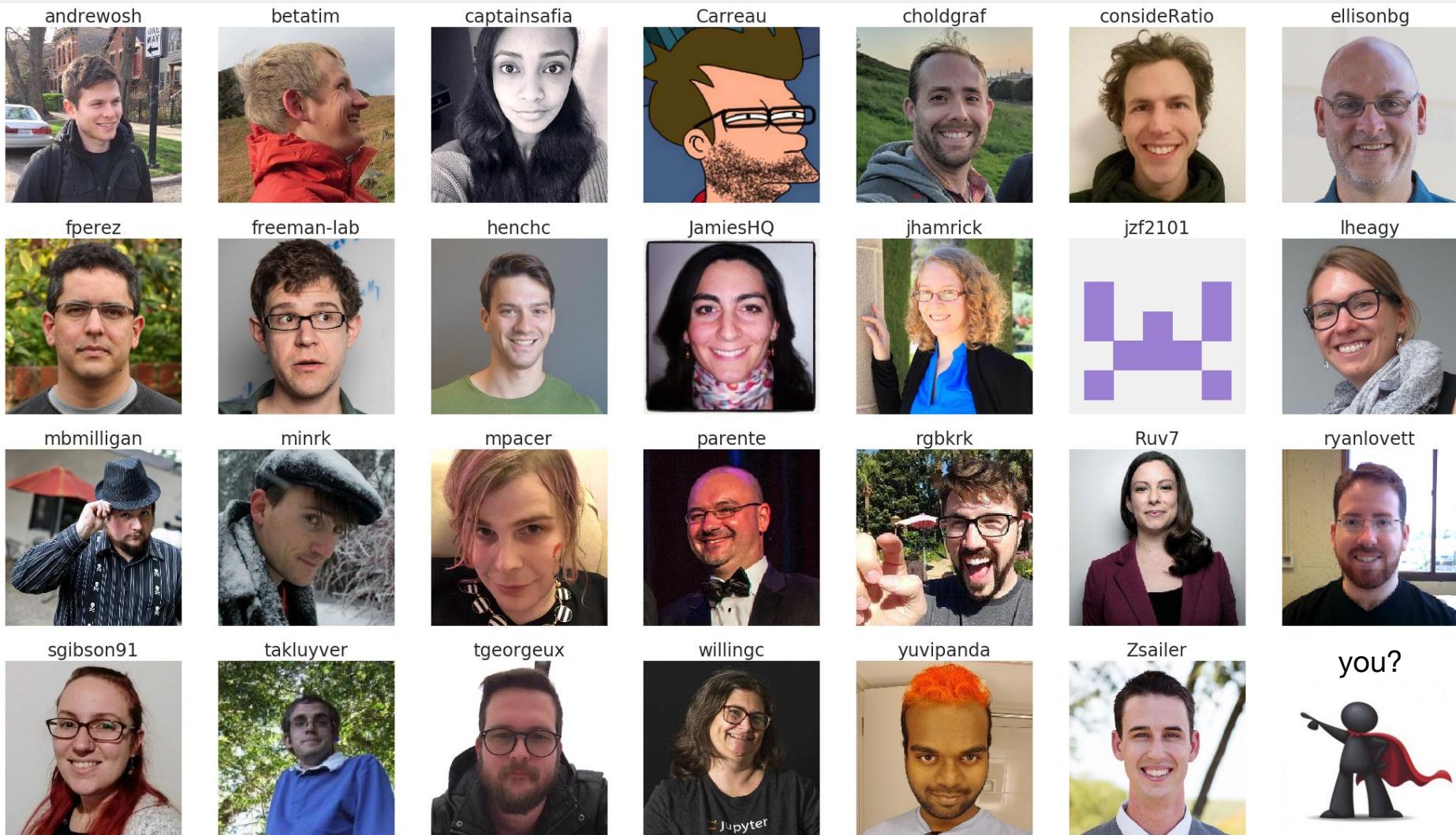




Sharing reproducible computations on Binder

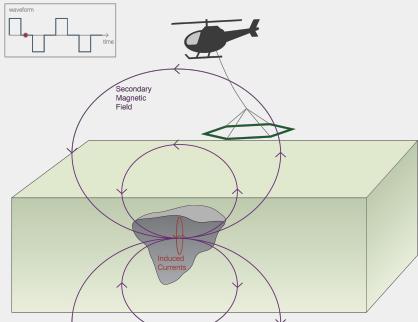
Lindsey Heagy, UC Berkeley & Project Jupyter





hello (a bit about me)

geophysical inversions



The logo of the University of British Columbia (UBC) is located in the bottom right corner. It consists of the letters "UBC" in a bold, serif font, with three wavy lines underneath representing water, and a stylized mountain peak at the bottom.

THE UNIVERSITY
OF BRITISH COLUMBIA

open-source software

jupyter 1_TTM_VericalConductor_2D_forward_uneven_charged

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

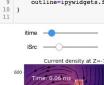
```
50
51
52     if showhi:
53         plt.show()
54
55     return [d for d in dati + dati]
```

View the current density through time

```
In [19]: 1 ipywidgets.interact(
2     plot_current,
3     plot_currents=ipywidgets.IntSlider(min=0, max=len(prk_timesteps), value=0),
4     idr=ipywidgets.IntSlider(min=0, max=len(prk_idrs), value=0),
5     k=ipywidgets.IntSlider(min=0, max=10, value=1),
6     showhi=ipywidgets.fixed(True),
7     showlo=ipywidgets.fixed(False),
8     outdir=ipywidgets.fixed(True)
9 )
10
11 )
```

Time 6
idr 5

Current density at $Z = 7.5m$



Time: 0.04 ms

x < 21m



z < 21m

Time: 0.04 ms

Current density (units: Am^-2)

```
from collections import Counter
@mpmathify('x+iy')
@precomp('StreamwiseSet at 0x1234567890abcdef')
@mpmathify('StreamwiseSet at 0x1234567890abcdef')
@mpmathify('StreamwiseSet at 0x1234567890abcdef')
@mpmathify('StreamwiseSet at 0x1234567890abcdef')

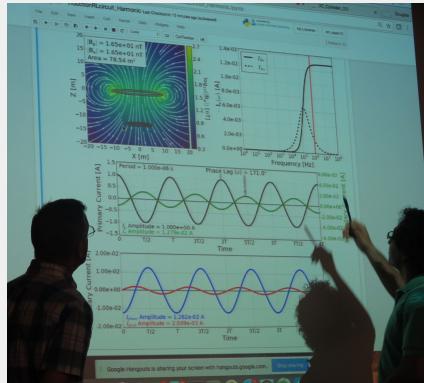
Out[19]: function  $\text{current\_plot}$ (currentsime, idrc, vlimNone, axNone, showhi, showlo, showhiTrue, outdir=True):
```

Out[20]: current_plot (currentsime, 10, None, None, showhi, True, showhiTrue, outdir=True)



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?)
^

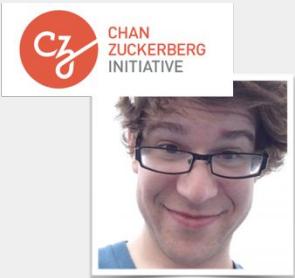


binder

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

mybinder.org

origins



github.com/freeman-lab

The screenshot shows the mybinder.org homepage. At the top is the "binder" logo, which consists of three overlapping circles in blue, orange, and grey. Below the logo is the word "binder" in a large, bold, dark grey sans-serif font. Underneath the title is a subtitle: "Turn a Git repo into a collection of interactive notebooks". A descriptive text follows: "Have a repository full of Jupyter notebooks? With Binder, open those notebooks in an executable environment, making your code immediately reproducible by anyone, anywhere." Below this is a form titled "Build and launch a repository". It contains fields for "GitHub repository name or URL" (with a dropdown menu for "GitHub"), "Git branch, tag, or commit", "Path to a notebook file (optional)" (with a "File" dropdown and a "launch" button), and "Copy the URL below and share your Binder with others:" (with a text input field and a clipboard icon). At the bottom is a section for "Copy the text below, then paste into your README to show a binder badge:" followed by a "launch binder" button.



+

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

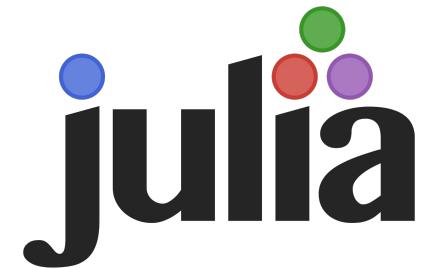
open science

(how?)

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

complete set of instructions

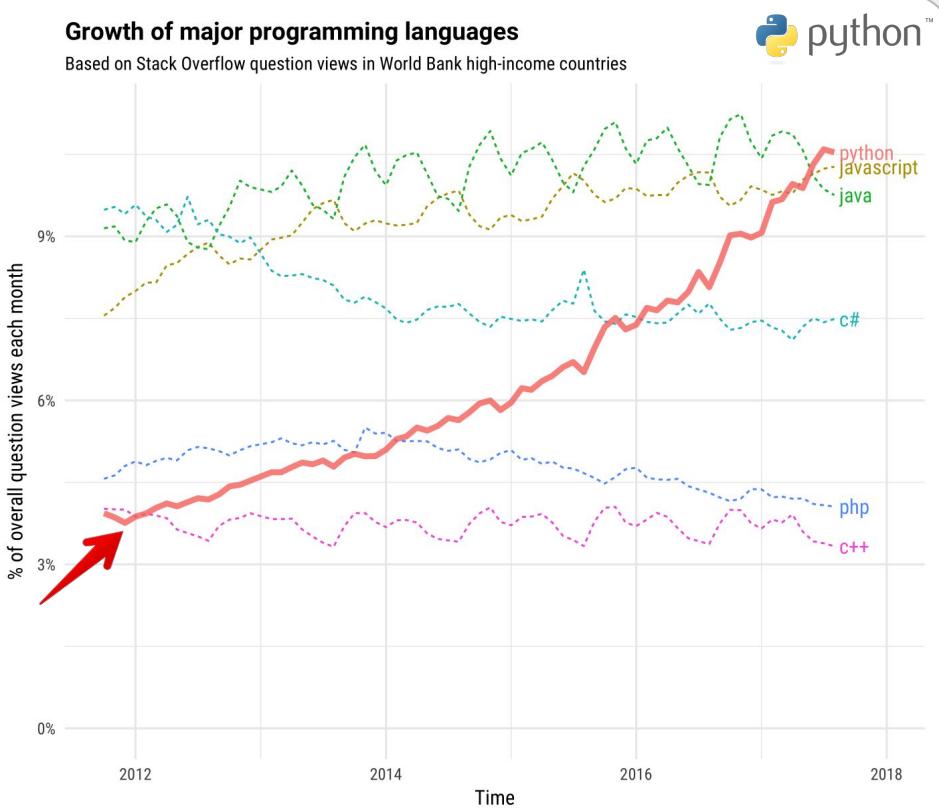
open-source languages are the raw material



complete set of instructions

Growth of major programming languages

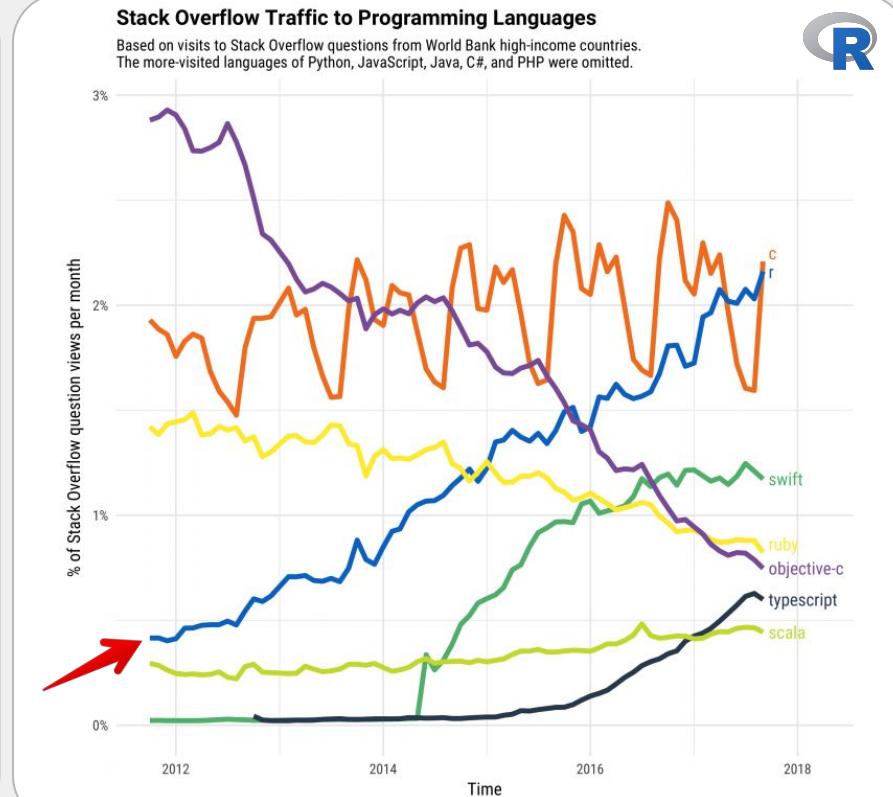
Based on Stack Overflow question views in World Bank high-income countries



<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.



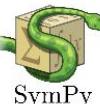
<https://stackoverflow.blog/2017/10/10/impressive-growth-r/>

complete set of instructions

mature ecosystems of tools



IPython



pandas



StatsModels
Statistics in Python



scikits-image
image processing in python



complete set of instructions

web-native interfaces for interacting with code



complete set of instructions

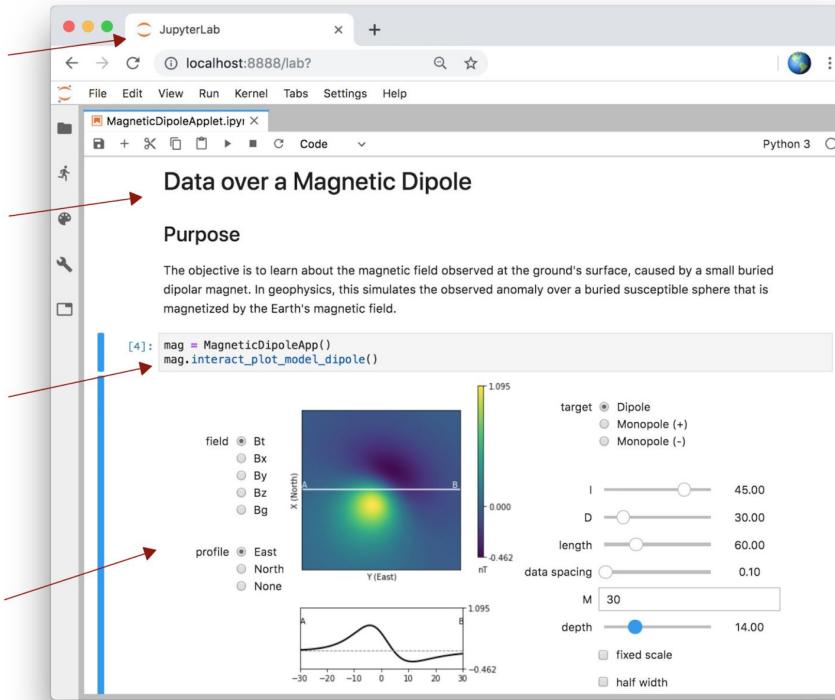
capture the steps

Runs in web browser

Text, comments, equations

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

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



complete set of instructions

capture the steps: what is a notebook?

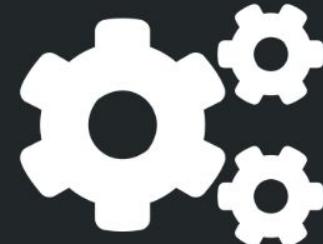
A document



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

The screenshot shows a GitHub repository page for 'binder-examples / r'. The repository title is 'Using R with Jupyter / RStudio on Binder'. It has 41 commits, 2 branches, 0 releases, and 5 contributors. The latest commit was made by betatim on April 15, 2018. A red arrow points to the commit log entry for 'index.ipynb'.

File	Description	Time Ago
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

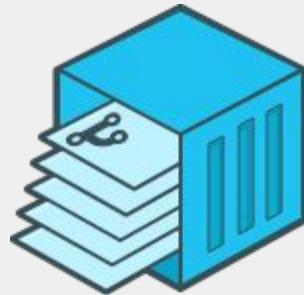
<https://github.com/binder-examples/r>

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

python™

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

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
```

Line 1, Column 1

Line 1, Column 1

Spaces: 2

R

```
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
```

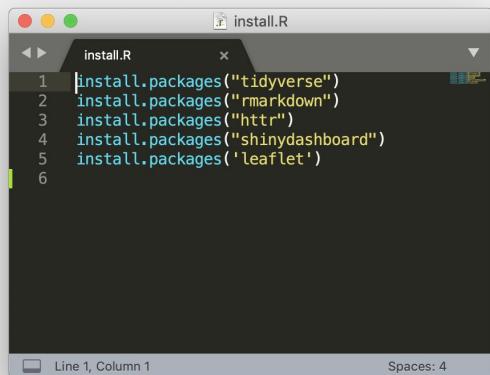
Line 1, Column 1

Line 1, Column 1

Spaces: 4

complete development environment

repo2docker

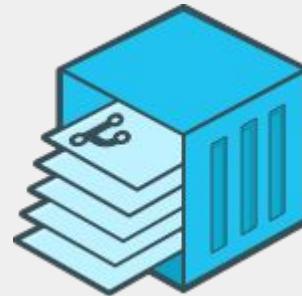


A screenshot of a terminal window titled "install.R". The window contains the following R code:

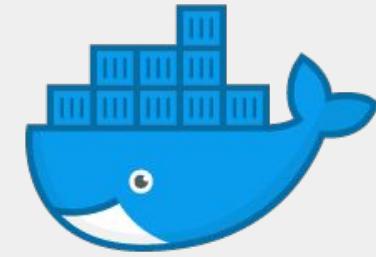
```
install.packages("tidyverse")
install.packages("rmarkdown")
install.packages("httr")
install.packages("shinydashboard")
install.packages('leaflet')
```

The terminal status bar at the bottom shows "Line 1, Column 1" and "Spaces: 4".

dependencies



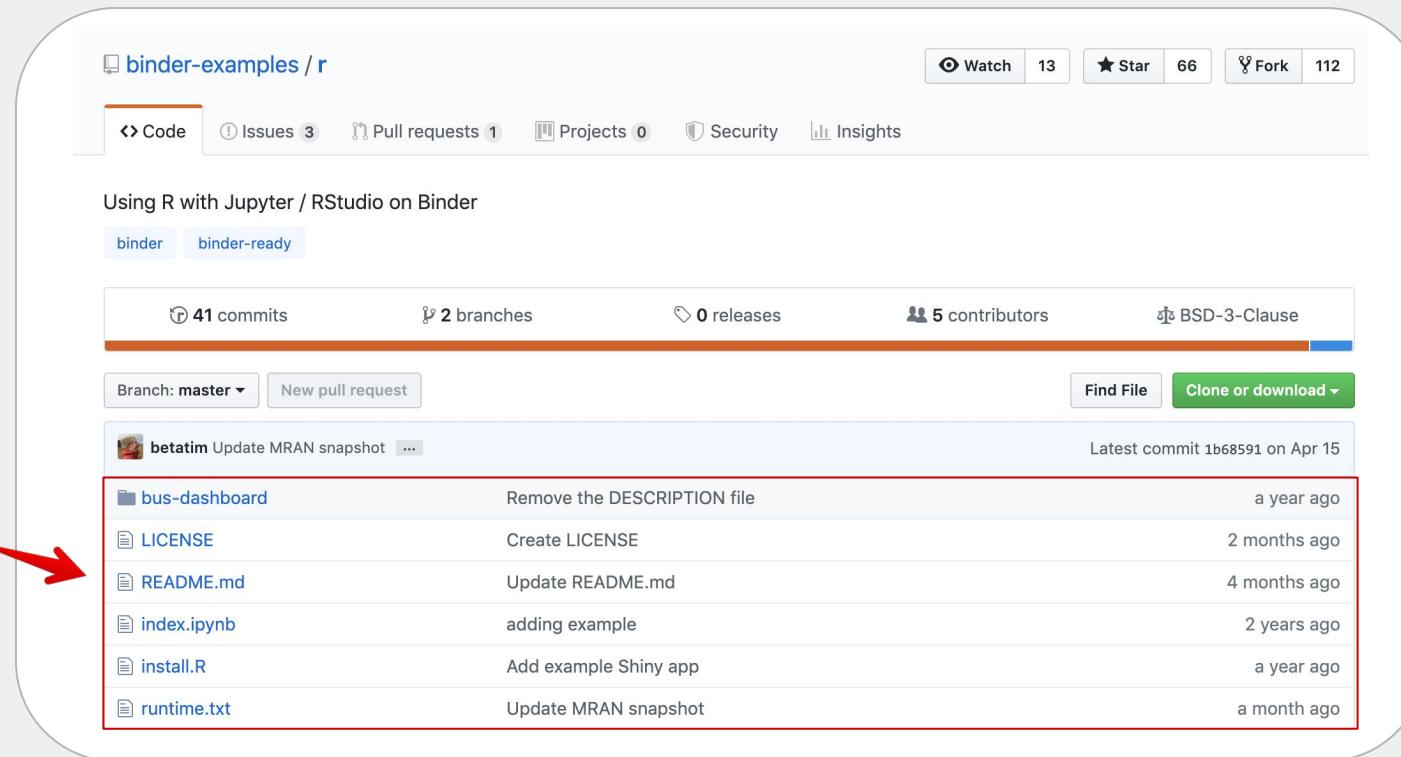
repo2docker



container file

complete development environment

define dependencies



binder-examples / r

Code Issues 3 Pull requests 1 Projects 0 Security Insights

Watch 13 Star 66 Fork 112

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

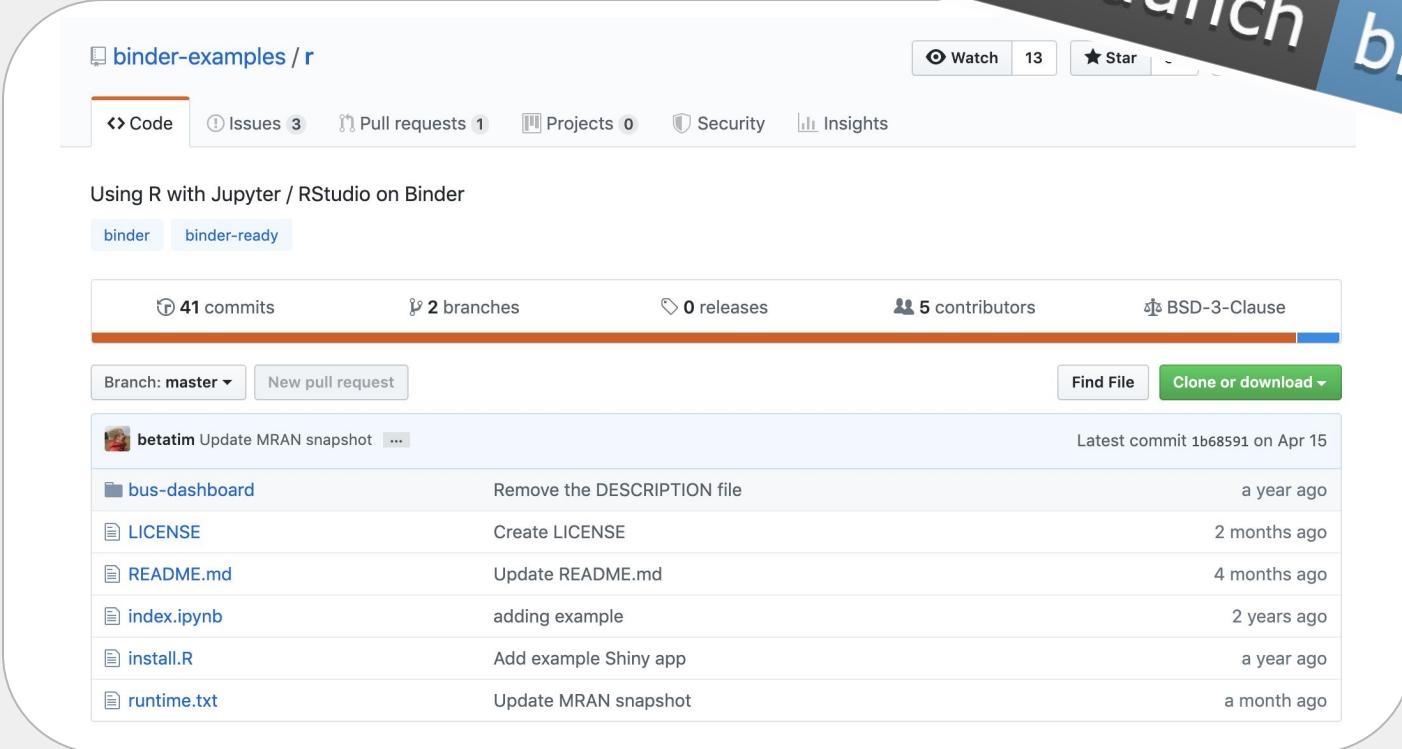
open science

(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



The screenshot shows a GitHub repository page for 'binder-examples / r'. At the top right, there is a large blue 'Launch binder' button with a 'binder' logo icon. The repository page includes standard GitHub navigation like 'Code', 'Issues 3', 'Pull requests 1', 'Projects 0', 'Security', and 'Insights'. Below the navigation, it says 'Using R with Jupyter / RStudio on Binder'. A 'binder' badge is present, along with a 'binder-ready' badge. Key statistics are displayed: 41 commits, 2 branches, 0 releases, 5 contributors, and BSD-3-Clause license. The commit history lists several recent changes, such as 'Update MRAN snapshot' by betatim, 'Remove the DESCRIPTION file' by bus-dashboard, and 'Create LICENSE' by betatim. The commits are timestamped from 'a year ago' to '2 years ago'.

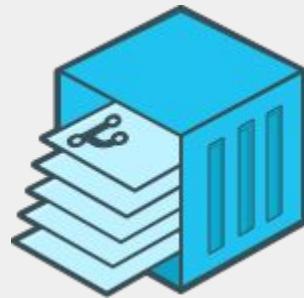
Commit	Description	Time Ago
bus-dashboard	Remove the DESCRIPTION file	a year ago
LICENSE	Create LICENSE	2 months ago
README.md	Update README.md	4 months ago
index.ipynb	adding example	2 years ago
install.R	Add example Shiny app	a year ago
runtime.txt	Update MRAN snapshot	a month ago

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:

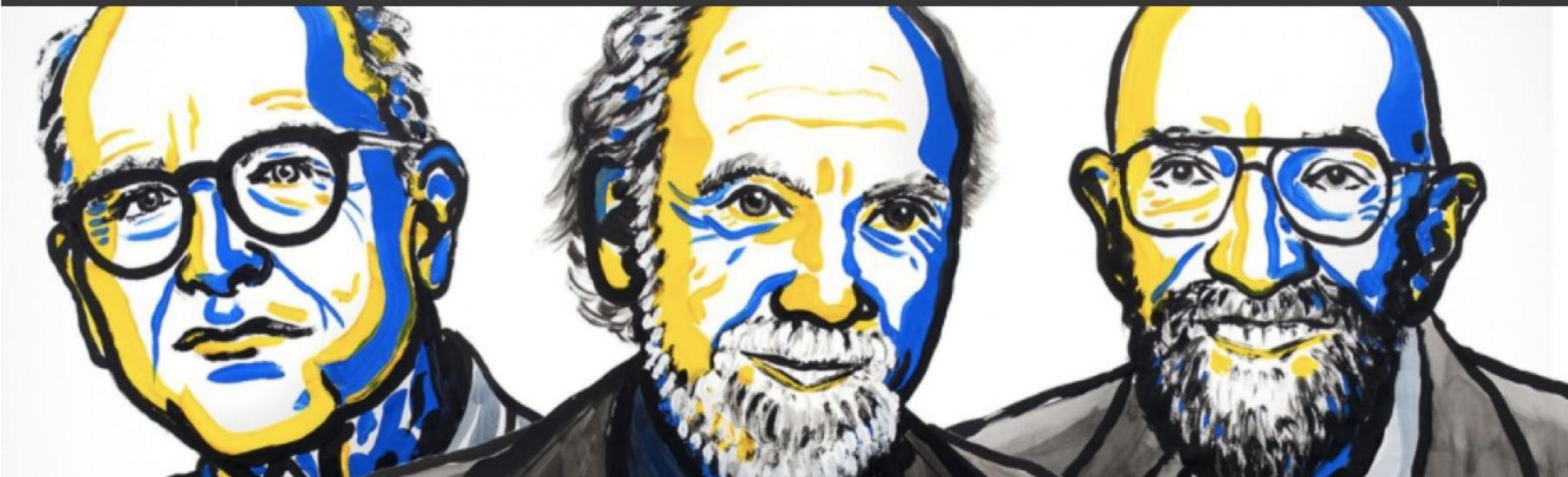
Fill in the fields to see a URL for sharing your Binder.



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

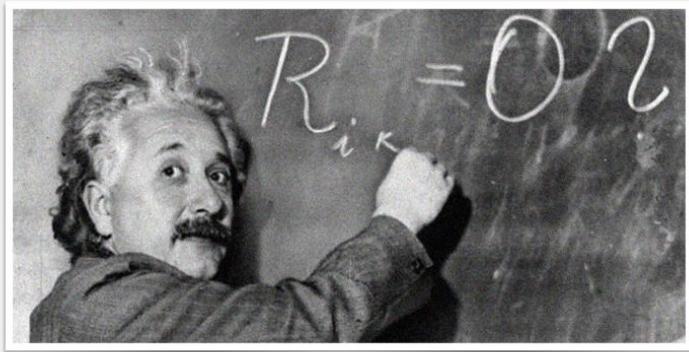
 launch binder



[START HERE](#)[NEWS](#)[EVENTS](#)[ACADEMICS](#)[RESEARCH & FACULTY](#)[ADMINISTRATION](#)

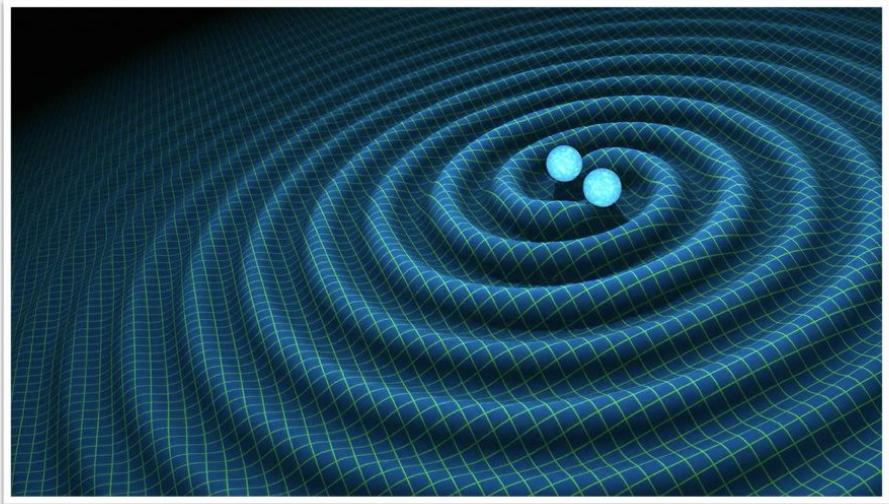
2017 NOBEL PRIZE IN PHYSICS AWARDED FOR DISCOVERY OF GRAVITATIONAL WAVES

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:

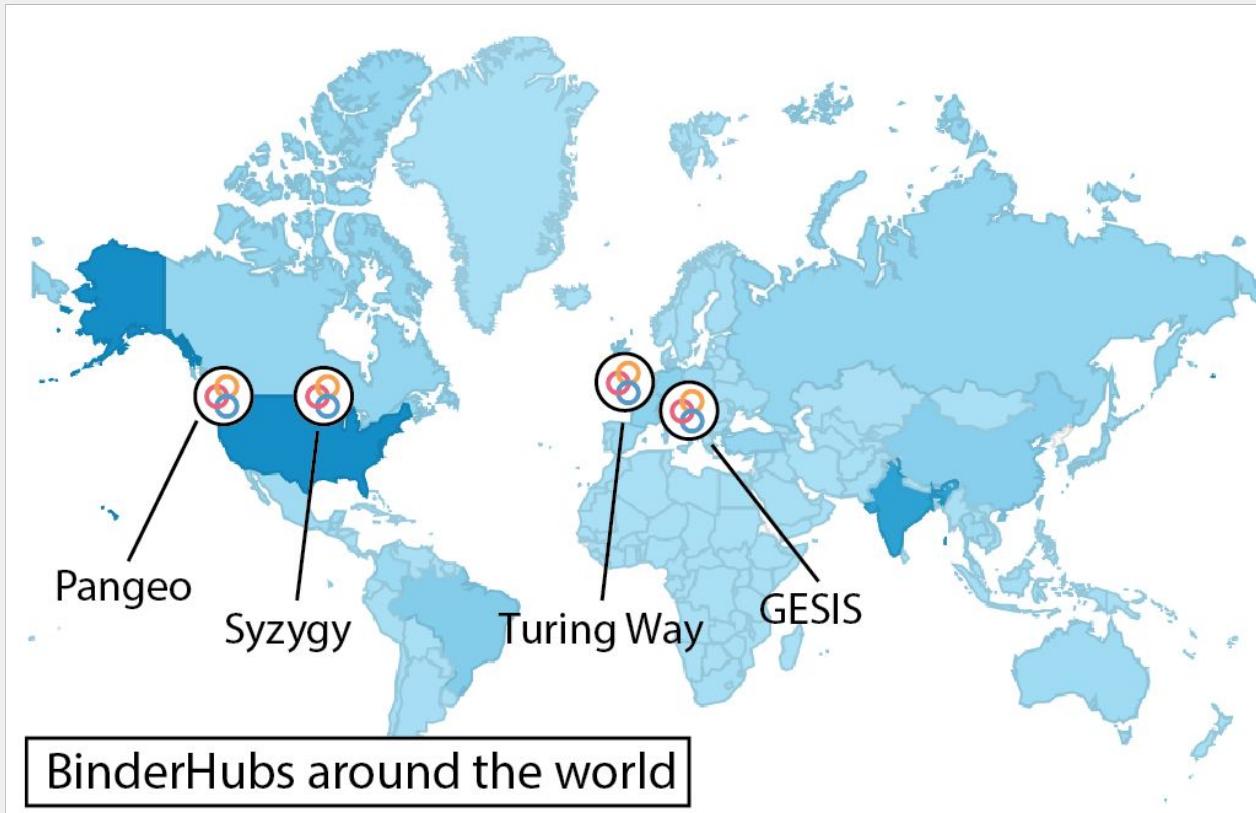
- ~ 1/1000 proton over 4 km.
- Sensitivity ~ $1e-21$
- Milky Way: $1e+21m$ across!

demo time



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

Future





thank you!

try out binder:

mybinder.org

connect with the community:

discourse.jupyter.org



ALFRED P. SLOAN
FOUNDATION

GORDON AND BETTY
MOORE
FOUNDATION