

CONTRIBUTING TO COMMUNITY OPEN SOURCE PROJECTS

Simon J. L. Billinge ^{1,2}

¹Columbia University

²Brookhaven National Laboratory

August 19, 2018

A BIT ABOUT MYSELF

- Professor at Columbia, joint appt at Brookhaven National Lab

A BIT ABOUT MYSELF

- Professor at Columbia, joint appt at Brookhaven National Lab
- Materials scientist working on the Nanostructure inverse problem

A BIT ABOUT MYSELF

- Professor at Columbia, joint appt at Brookhaven National Lab
- Materials scientist working on the Nanostructure inverse problem
- User x-ray, neutron and electron scattering methods

A BIT ABOUT MYSELF

- Professor at Columbia, joint appt at Brookhaven National Lab
- Materials scientist working on the Nanostructure inverse problem
- User x-ray, neutron and electron scattering methods
- algorithm and software development are central to our work

A BIT ABOUT MYSELF

- Professor at Columbia, joint appt at Brookhaven National Lab
- Materials scientist working on the Nanostructure inverse problem
- User x-ray, neutron and electron scattering methods
- algorithm and software development are central to our work
- Main software projects:

A BIT ABOUT MYSELF

- Professor at Columbia, joint appt at Brookhaven National Lab
- Materials scientist working on the Nanostructure inverse problem
- User x-ray, neutron and electron scattering methods
- algorithm and software development are central to our work
- Main software projects:
 - ① DIFFPY - a pluggable framework for building refinement programs

A BIT ABOUT MYSELF

- Professor at Columbia, joint appt at Brookhaven National Lab
- Materials scientist working on the Nanostructure inverse problem
- User x-ray, neutron and electron scattering methods
- algorithm and software development are central to our work
- Main software projects:
 - ① DIFFPY - a pluggable framework for building refinement programs
 - ② XPDACQ/XPDAN - framework software for analyzing data on the fly

A BIT ABOUT MYSELF

- Professor at Columbia, joint appt at Brookhaven National Lab
- Materials scientist working on the Nanostructure inverse problem
- User x-ray, neutron and electron scattering methods
- algorithm and software development are central to our work
- Main software projects:
 - ① DIFFPY - a pluggable framework for building refinement programs
 - ② XPDACQ/XPDAN - framework software for analyzing data on the fly
 - ③ REGOLITH - software for organizing your group and your academic career

BAD OLD DAYS

Bad old days

BAD OLD DAYS

Bad old days

- 1 Student finds a bug or has a feature request

BAD OLD DAYS

Bad old days

- 1 Student finds a bug or has a feature request
- 2 Student emails developer

BAD OLD DAYS

Bad old days

- ① Student finds a bug or has a feature request
- ② Student emails developer
- ③ When he/she has time, they will get to it

BAD OLD DAYS

Bad old days

- ① Student finds a bug or has a feature request
- ② Student emails developer
- ③ When he/she has time, they will get to it
- ④ Email back a fixed version

BRAVE NEW WORLD

Brave new world

BRAVE NEW WORLD

Brave new world

- Teams of programmers all over the world working on the code

BRAVE NEW WORLD

Brave new world

- Teams of programmers all over the world working on the code
- How is that possible without disaster?

SUMMARY OF TALK

- Our life, working in teams of student/post-doc programmers, has been quite simply revolutionized by

SUMMARY OF TALK

- Our life, working in teams of student/post-doc programmers, has been quite simply revolutionized by
 - ① GIT
 - ② GITHUB

SUMMARY OF TALK

- Our life, working in teams of student/post-doc programmers, has been quite simply revolutionized by
 - ① GIT
 - ② GITHUB
- Learn these workflows.....you won't regret it

SUMMARY OF TALK

- Our life, working in teams of student/post-doc programmers, has been quite simply revolutionized by
 - ① GIT
 - ② GITHUB
- Learn these workflows.....you won't regret it
- If you can, set up your own projects this way.....you won't regret it

REQUIREMENTS

What has to happen to community supported code development to work?:

Scalable code quality practices

REQUIREMENTS

What has to happen to community supported code development to work?:

Scalable code quality practices

- code readability/uniformity (pep8 for Python)

REQUIREMENTS

What has to happen to community supported code development to work?:

Scalable code quality practices

- code readability/uniformity (pep8 for Python)
- developer docs (docstrings for Python)

REQUIREMENTS

What has to happen to community supported code development to work?:

Scalable code quality practices

- code readability/uniformity (pep8 for Python)
- developer docs (docstrings for Python)
- unit tests

REQUIREMENTS

What has to happen to community supported code development to work?:

Scalable code quality practices

- code readability/uniformity (pep8 for Python)
- developer docs (docstrings for Python)
- unit tests
- continuous integration (CI)

REQUIREMENTS

What has to happen to community supported code development to work?:

Scalable code quality practices

- code readability/uniformity (pep8 for Python)
- developer docs (docstrings for Python)
- unit tests
- continuous integration (CI)
- well defined release workflow

Let's take a poll

NUMPY

The screenshot shows the GitHub repository for Numpy. The repository is named 'numpy / numpy' and has 416 watches, 7,966 stars, and 2,945 forks. It has 18,513 commits, 17 branches, 141 releases, and 678 contributors. The main repository URL is <https://www.numpy.org/>.

The commit history table shows the following entries:

Commit	Description	Time Ago
mattip and charis	BUG: check return value of _buffer_format_string (#11754)	Latest commit 3026928 14 hours ago
cirdec	TST: Add Python 3.7 to CI testing (#11598)	26 days ago
github	Make fixes requested by @rgommers	19 days ago
benchmarks	MAINT: Produce a more readable repr of argument packs in benchmark	2 months ago
branding/icons	add .gitattributes and fix line endings	8 years ago
doc	Merge pull request #11528 from eric-wieser/polynomial-ipython-latex	5 days ago
numpy	BUG: check return value of _buffer_format_string (#11754)	14 hours ago
tools	TST: add C code line coverage to codecov report	10 days ago
appveyor.yml	TST: Add Python 3.7 to CI testing (#11598)	26 days ago
codecov.yml	TST: inspire setup.py files for codecov reports	11 days ago

Numpy!

NUMPY



Want to submit an issue to numpy/numpy?

[Dismiss](#)If you have a bug or an idea, read the [contributing guidelines](#) before opening an issue.

Filters ▾

Labels

Milestones

New issue

① 1,695 Open ✓ 4,721 Closed

Author ▾

Labels ▾

Projects ▾

Milestones ▾

Assignee ▾

Sort ▾

① parsing final comma in dtype string shorthand

#11770 opened 2 hours ago by mamikonyan

① a['f0'].dtype != a.dtype['f0']

#11769 opened 2 hours ago by mamikonyan

💬 1

① dot between the same two vectors should use nrm2

#11768 opened 2 hours ago by jakirkham

💬 3

① Exporting mtrand.RandomState to public Cython API 23 - Wish List

#11764 opened 13 hours ago by jakirkham

① Weird behavior with FFT on array views

#11762 opened 15 hours ago by david-hoffman

💬 4

① garbled formulae in NumPy documentation

#11753 opened a day ago by kfahnke

💬 5

① BUG: repr of structured array containing generic datetime crashes 00 - Bug

#11752 opened a day ago by eric-wieser

💬 1

① np.testing.assert_almost_equal Issue in Python 3.5 for numpy==1.15.0 06 - Regression

💬 17

NUMPY



Dismiss

If you have a bug or an idea, read the [contributing guidelines](#) before opening an issue.

Filters ▾

is:issue is:open

Labels

Milestones

New issue

① 1.695 Open ✓ 4.721 Closed

Aug 07 ▾

Labels ▼

Products ▾

Milestones ▾

Assignee ▾

Sort ▼

- ❗ parsing final comma in dtype string shorthand

#11770 opened 2 hours ago by mamikonyan

❗ `a['f0'].dtype != a.dtype['f0']`

#11769 opened 2 hours ago by mamikonryan

1

- ! dot between the same two vectors should use `nrm2`

#11768 opened 2 hours ago by jakirkham

3

Exporting mtrand.RandomState to public Cython API 23 - Wish List

#11764 opened 13 hours ago by jakirkham

⚠ Weir behavior with FFT on array views

#11762 opened 15 hours ago by david-hoffman

4

🚫 garbled formulae in NumPy documentation

#11753 opened a day ago by kfjahnke

5

① BUG: repr of structured array containing generic datetime crashes 00 - Bug

00 - Bug

#11752 opened a day ago by eric-wieser

1

np.testing.assert_almost_equal Issue in Python 3.5 for numpy=1.15.0 06 - Regression

06 - Regression

17

NUMPY



Want to submit an issue to numpy/numpy?

[Dismiss](#)

If you have a bug or an idea, read the [contributing guidelines](#) before opening an issue.

Filters ▾

🔍 is:issue is:open

Labels

Milestones

New issue

① 1,695 Open ✓ 4,721 Closed

Author ▾

Labels ▾

Projects ▾

Milestones ▾

Assignee ▾

Sort ▾

① parsing final comma in dtype string s

#11770 opened 2 hours ago by mamikonyan

① a['f0'].dtype != a.dtype['f0']

#11769 opened 2 hours ago by mamikonyan

① dot between the same two vectors sh

#11768 opened 2 hours ago by jakirkham

① Exporting mtrand.RandomState to p

#11764 opened 13 hours ago by jakirkham

① Weird behavior with FFT on array vie

#11762 opened 15 hours ago by david-hoffman

① garbled formulae in NumPy documen

#11753 opened a day ago by kfahnke

① BUG: repr of structured array contain

#11752 opened a day ago by eric-wieser

① np.testing.assert_almost_equal Issue

Filter by label

Filter labels

component: other

component: swig

component: tools

difficulty: Easy

difficulty: Expert

difficulty: Intermediate

Embedded

Patch

priority: high

priority: highest

1

3

4

5

1

17

NUMPY



Want to submit an issue to numpy/numpy?

[Dismiss](#)

If you have a bug or an idea, read the [contributing guidelines](#) before opening an issue.

Filters ▾

🔍 is:issue is:open

Labels

Milestones

New issue

1,695 Open ✓ 4,721 Closed

Author ▾

Labels ▾

Projects ▾

Milestones ▾

Assignee ▾

Sort ▾

🔍 parsing final comma in dtype string s

#11770 opened 2 hours ago by mamikonyan

🔍 a['f0'].dtype != a.dtype['f0']

#11769 opened 2 hours ago by mamikonyan

🔍 dot between the same two vectors s

#11768 opened 2 hours ago by jakirkham

🔍 Exporting mtrand.RandomState to p

#11764 opened 13 hours ago by jakirkham

🔍 Weird behavior with FFT on array v

#11762 opened 15 hours ago by david-hoffma

🔍 garbled formulae in NumPy docum

#11753 opened a day ago by kfjahnke

🔍 BUG: repr of structured array contain

#11752 opened a day ago by eric-wieser

🔍 np.testing.assert_almost_equal Issue

Filter by label

Filter labels

component: Other

component: swig

component: tools

difficulty: Easy

difficulty: Expert

difficulty: Intermediate

Embedded

Batch

priority: high

priority: highest

1

3

4

5

1

17

Issue

MATPLOTLIB

① 38 Open ✓ 27 Closed	Author ▾	Labels ▾	Projects ▾	Milestones ▾	Assignee ▾	Sort ▾
① DOC: Add more explanatory comments to gallery examples		Documentation	good first issue			5
#11654 opened on Jul 13 by story645						
① ticklabels font using latex		Documentation	good first issue	text/usetex		10
#11586 opened on Jul 6 by photoniker	👤 v3.1					
① Matplotlib.pyplot.quiver does not accept X, Y, U, V and C as keyword arguments		API consistency				2
good first issue						
#11526 opened on Jun 28 by niwatori1217	👤 v3.1					
① Examples not shown in API docs for many methods.		Documentation	good first issue			23
#10974 opened on Apr 6 by jklymak	👤 v3.1					
① Add documentation for set_rlim		Documentation	good first issue	polar		3
#10704 opened on Mar 7 by dstansby	👤 v3.1					
① Make documentation builds (more) reproducible		Difficulty: Medium	Documentation	good first issue		2
#10702 opened on Mar 6 by tacaswell	👤 v2.2-doc					
① Most warnings calls do not set the stacklevel		good first issue				6
#10643 opened on Mar 1 by anntzer	👤 v3.1					
① Add a __repr__ for Figure without the object address ?		Wishlist	good first issue	third-party integration		15
#9372 opened on Oct 11, 2017 by Carreau	👤 needs sorting					
① matplotlibrc parameter for legend title font size		Wishlist	good first issue	rcparams		4
#9201 opened on Sep 18, 2017 by andrewfowlie	👤 needs sorting					
① docstring of parameter Fs of matplotlib.pyplot.specgram() inconsistent with default value		good first issue				10
#9100 opened on Aug 26, 2017 by andrewh7	👤 needs sorting					



LET'S JUMP IN

- ① OK, So let's help out. Everyone can do a little bit.

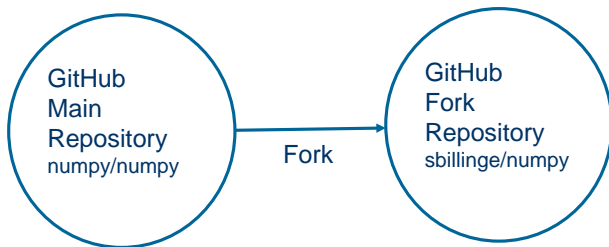
LET'S JUMP IN

- ① OK, So let's help out. Everyone can do a little bit.
- ② How do I get involved?

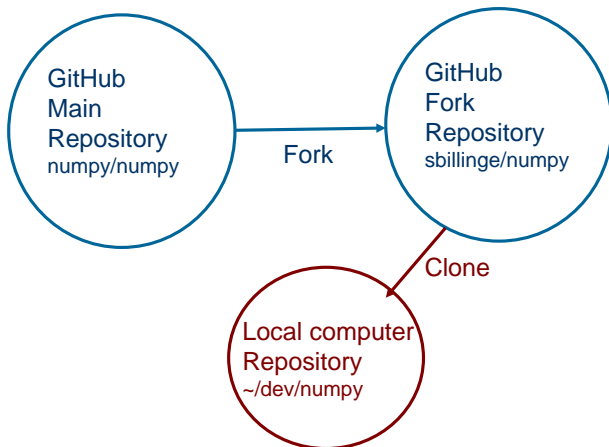
NUMPY GITHUB WORKFLOW



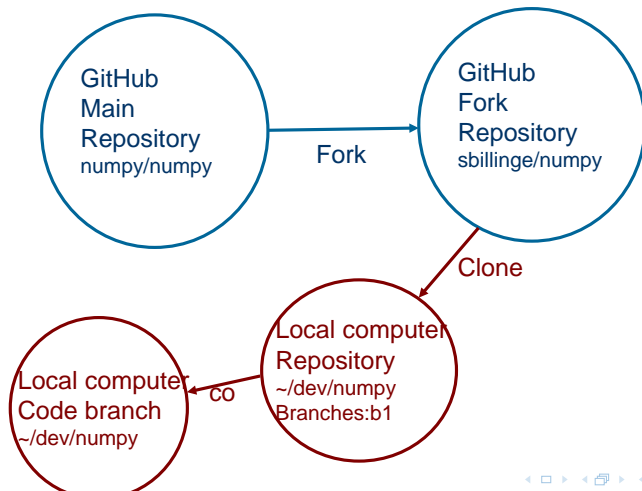
NUMPY GITHUB WORKFLOW



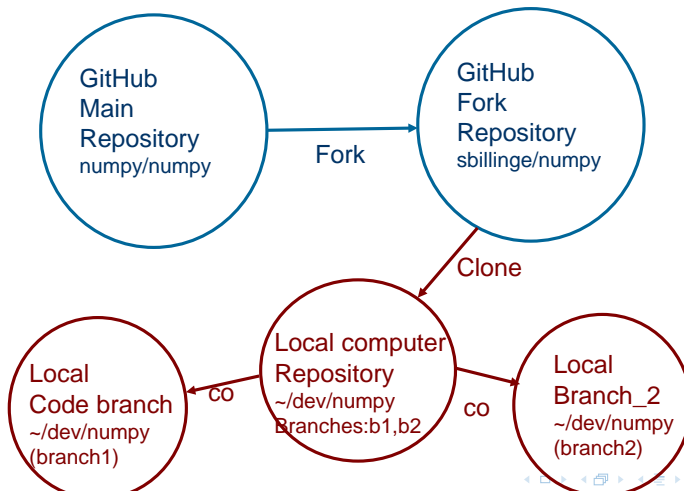
NUMPY GITHUB WORKFLOW



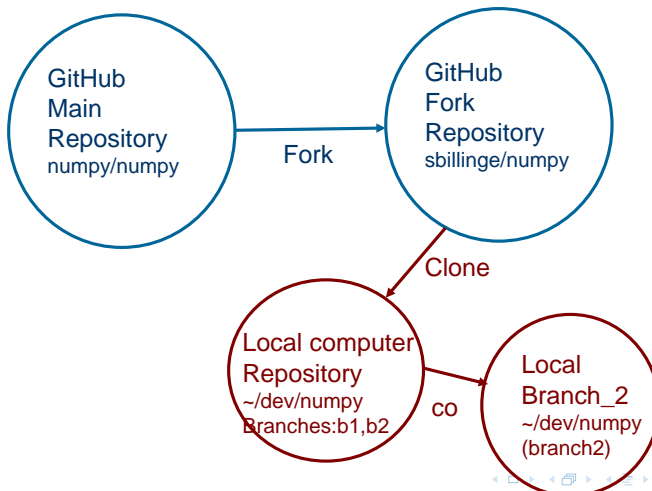
NUMPY GITHUB WORKFLOW



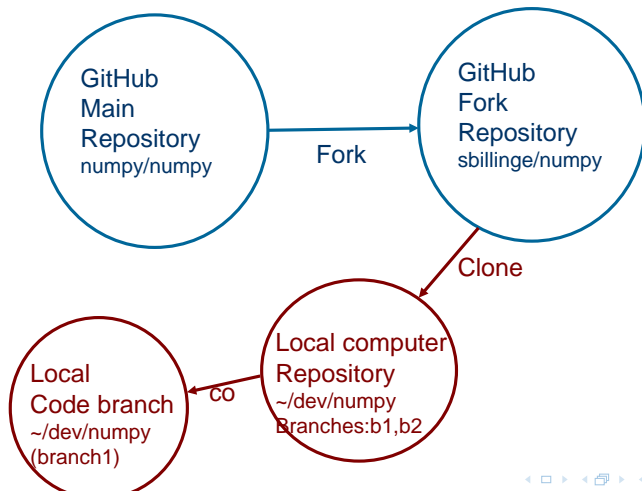
NUMPY GITHUB WORKFLOW



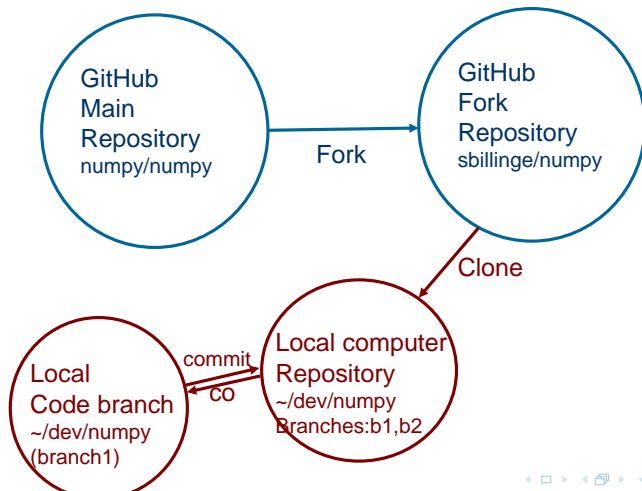
NUMPY GITHUB WORKFLOW



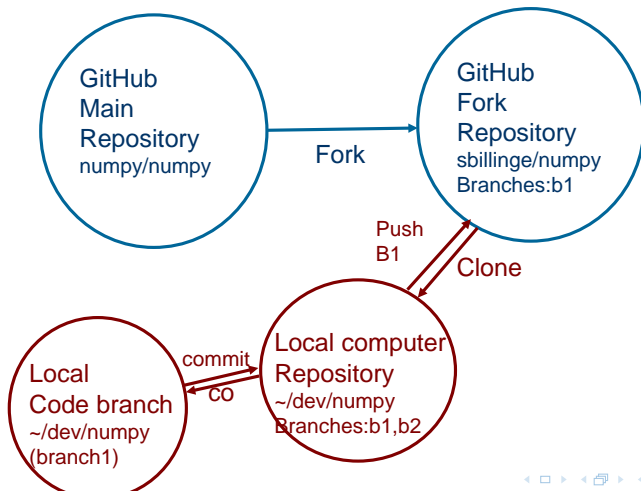
NUMPY GITHUB WORKFLOW



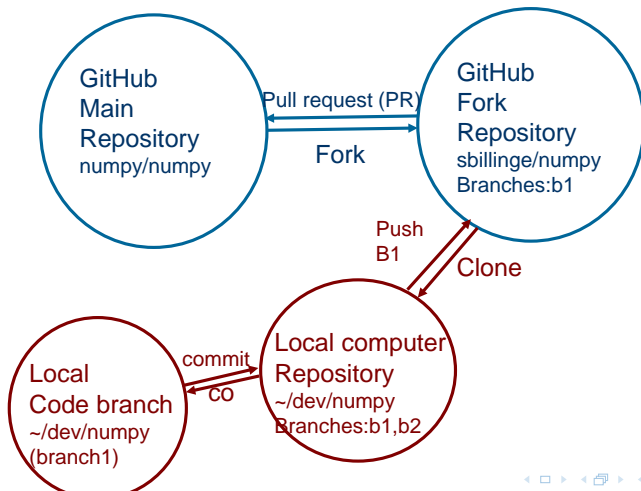
NUMPY GITHUB WORKFLOW



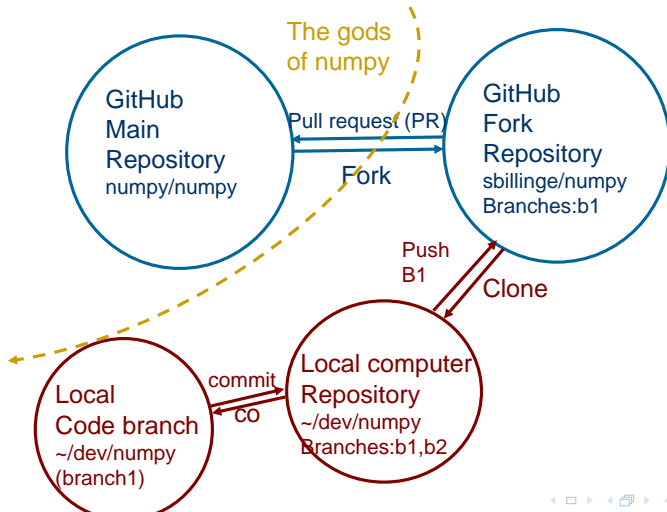
NUMPY GITHUB WORKFLOW



NUMPY GITHUB WORKFLOW



NUMPY GITHUB WORKFLOW



NUMPY GITHUB WORKFLOW

