Version Control in Science, Why and How?

. . . could the benefits outweigh the costs?

Adrian Radillo[1], David White[2], Joris Vincent[3]

[1] *Postdoc in Gold Lab*
[2] *Ph.D. candidate in Burge Lab*
[3] *Postdoc in Brainard Lab*

May 8, 2019

Version Control Software

git = Free Software that allows ...

Version Control Software

git = Free Software that allows ...

- tracking changes in a folder and its subfolders

Version Control Software

git = Free Software that allows ...

- tracking changes in a folder and its subfolders
- easy handling and sharing of code and its history

Version Control Software

git = Free Software that allows ...

- tracking changes in a folder and its subfolders
- easy handling and sharing of code and its history

GitHub + GitLab + BitBucket = cloud-based platforms

- hosting of git projects (private or public)

Version Control Software

git = Free Software that allows ...

- tracking changes in a folder and its subfolders
- easy handling and sharing of code and its history

GitHub + GitLab + BitBucket = cloud-based platforms

- hosting of git projects (private or public)
- project management tools geared towards collaboration

How many times did the following happen in your lab?

How many times did the following happen in your lab?

- Different versions of the same code “sit” on the same computer, or in distinct computers.

How many times did the following happen in your lab?

- Different versions of the same code “sit” on the same computer, or in distinct computers.
- Two people using the same code on the same computer “step on each other’s toes”.

How many times did the following happen in your lab?

- Different versions of the same code “sit” on the same computer, or in distinct computers.
- Two people using the same code on the same computer “step on each other’s toes”.
- It takes weeks for someone to reproduce a figure that they produced in the first place.

How many times did the following happen in your lab?

- One has many iterations of a figure in a folder, and it is unclear what code produced what figure.

How many times did the following happen in your lab?

- One has many iterations of a figure in a folder, and it is unclear what code produced what figure.
- Code isn't reused after a project ends.

Coding has become a part of the scientific methodology.

But can't we get away without it?

But can't we get away without it?

- Yes for tiny projects

But can't we get away without it?

- Yes for tiny projects
- But it remains a great tool to collaborate...

But can't we get away without it?

- Yes for tiny projects
- But it remains a great tool to collaborate...
- and to use other people's code!

Git also helps publishing

- Both *Nature* and *Science* strongly encourage the use of GitHub.

Git also helps publishing

- Both *Nature* and *Science* strongly encourage the use of GitHub.
- *Nature* expects researchers to maintain code used in publications.

Git also helps publishing

- Both *Nature* and *Science* strongly encourage the use of GitHub.
- *Nature* expects researchers to maintain code used in publications.

Incorporating git in your workflow will make this a breeze!

So, what's the real cost???

So, what's the real cost???

- Every coder needs to learn the basics of git.
  (4 hours with a teacher)

So, what's the real cost???

- Every coder needs to learn the basics of git.
  (4 hours with a teacher)
- The whole lab needs to adopt a new coding culture
  (months)

So, what's the real cost???

- Every coder needs to learn the basics of git.
  (4 hours with a teacher)
- The whole lab needs to adopt a new coding culture
  (months)
- Everyone spends a little extra time on their code.
  15% at the beginning,

So, what's the real cost???

- Every coder needs to learn the basics of git.
  (4 hours with a teacher)
- The whole lab needs to adopt a new coding culture
  (months)
- Everyone spends a little extra time on their code.
  15% at the beginning, then 5%,

So, what's the real cost???

- Every coder needs to learn the basics of git.
  (4 hours with a teacher)
- The whole lab needs to adopt a new coding culture
  (months)
- Everyone spends a little extra time on their code.
  15% at the beginning, then 5%, then -10%!

Where to start

Where to start

- Seek advice!

Where to start

- Seek advice!
- Provide a 4-hour tutorial to any newcomer.

Where to start

- Seek advice!
- Provide a 4-hour tutorial to any newcomer.
- Establish lab guidelines and resources.

Where to start

- Seek advice!
- Provide a 4-hour tutorial to any newcomer.
- Establish lab guidelines and resources.
- Designate a code administrator for each project.

Where to start

- Seek advice!
- Provide a 4-hour tutorial to any newcomer.
- Establish lab guidelines and resources.
- Designate a code administrator for each project.
- Take advantage of a hardware renewal to implement the new strategy.

What we will offer you today

Beginner Tutorial Get git, GitHub and GitKraken up and running. Learn and practice basic workflows.

What we will offer you today

Beginner Tutorial Get git, GitHub and GitKraken up and running. Learn and practice basic workflows.

Intermediate Tutorial For those who already use git, we'll dive into more technical details that crop up often.

What we will offer you today

Beginner Tutorial Get git, GitHub and GitKraken up and running. Learn and practice basic workflows.

Intermediate Tutorial For those who already use git, we'll dive into more technical details that crop up often.

Demos & Discussion We will demonstrate daily use cases of git. We can also discuss open issues.

Your questions!