

Software For Data Science

SETH RUSSELL – SETH.RUSSELL@CUANSCHUTZ.EDU

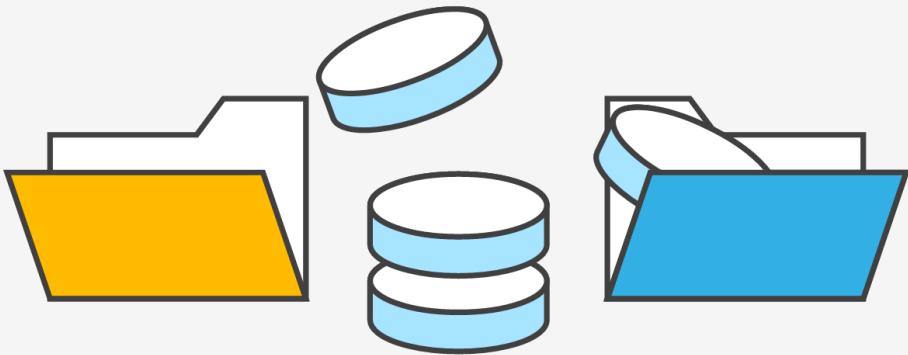
ACCORDS DATA SCIENCE PROGRAM, D2V ANALYTICS CORE

[HTTPS://GITHUB.COM/MAGIC-LANTERN/BIOS6642](https://github.com/magic-lantern/bios6642)

Version Control

“Did you think the lion was sleeping
because he didn't roar?”

– Friedrich Schiller, Die Verschwörung des Fiesco zu Genua



Why Git for your organization

Git for developers / Git for marketing / Git for product management / Git for designers / Git for customer support / Git for human resources / Git for anyone managing a budget

Why Git Rules the world (of software)

Common Terms

- ▶ git checkout
- ▶ git status
- ▶ git add
- ▶ git commit
- ▶ git push
- ▶ git pull
- ▶ git branch
- ▶ git log

Why Git Rules the world

Save Time

Git is lightning fast. And although we're talking about only a few seconds per command, it quickly adds up in your work day. Use your time for something more useful than waiting for your version control system to get back to you.

Work Offline

What if you want to work while you're on the move? With a centralized VCS like Subversion or CVS, you're stranded if you're not connected to the central repository. With Git, almost everything is possible simply on your local machine: make a commit, browse your project's complete history, merge or create branches... Git let's you decide where and when you want to work.

<https://www.git-tower.com/learn/git/ebook/en/command-line/appendix/why-git>

Why Git Rules the world

Undo Mistakes

People make mistakes. A good thing about Git is that there's a little "undo" command for almost every situation. Correct your last commit because you forgot to include that small change. Revert a whole commit because that feature isn't necessary, anymore. And when the going gets tough you can even restore disappeared commits with the Reflog - because, behind the scenes, Git rarely really deletes something. This is peace of mind.

Don't Worry

Git gives you the confidence that you can't screw things up - and this is a great feeling. In Git, every clone of a project that one of your teammates might have on his local computer is a fully usable backup. Additionally, almost every action in Git only adds data (deleting is very rare). That means that losing data or breaking a repository beyond repair is really hard to do.

<https://www.git-tower.com/learn/git/ebook/en/command-line/appendix/why-git>

Why Git Rules the world

Don't Mix Things Up

Separation of concerns is paramount to keeping track of things. While you're working on feature A, nothing (and no-one) else should be affected by your unfinished code. What if it turns out the feature isn't necessary anymore? Or if, after 10 commits, you notice that you took a completely wrong approach? Branching is the answer for these problems. And while other version control systems also know branches, Git is the first one to make it work as it should: fast & easy.

<https://www.git-tower.com/learn/git/ebook/en/command-line/appendix/why-git>



Welcome to GitHub Desktop

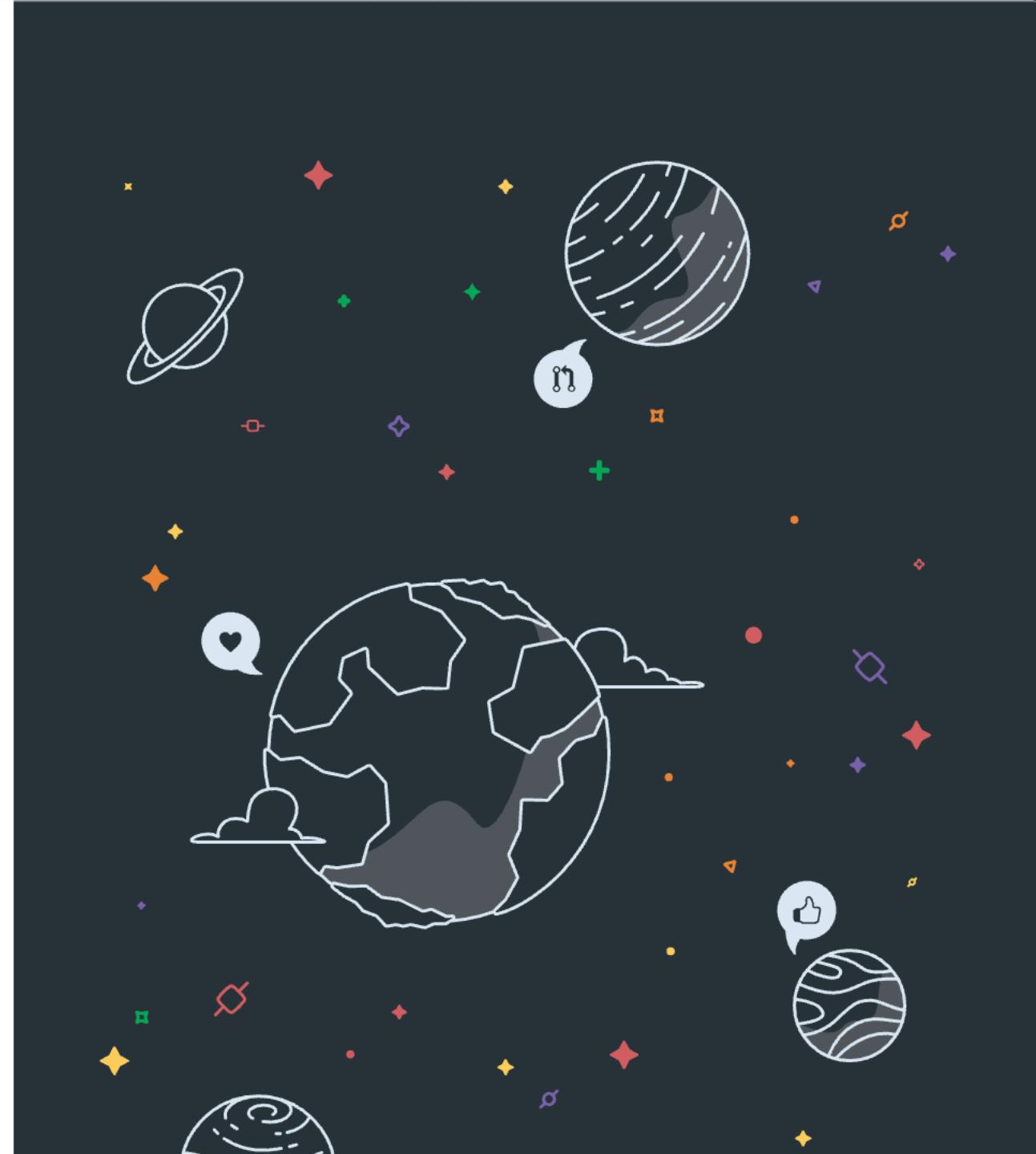
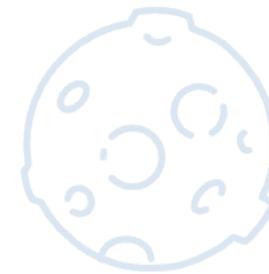
GitHub Desktop is a seamless way to contribute to projects on GitHub and GitHub Enterprise Server. Sign in below to get started with your existing projects.

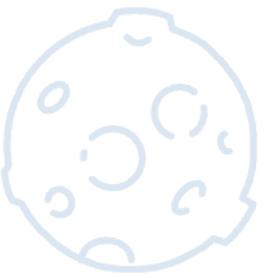
New to GitHub? [Create your free account.](#)

[Sign in to GitHub.com](#)

[Sign in to GitHub Enterprise Server](#)

[Skip this step](#)

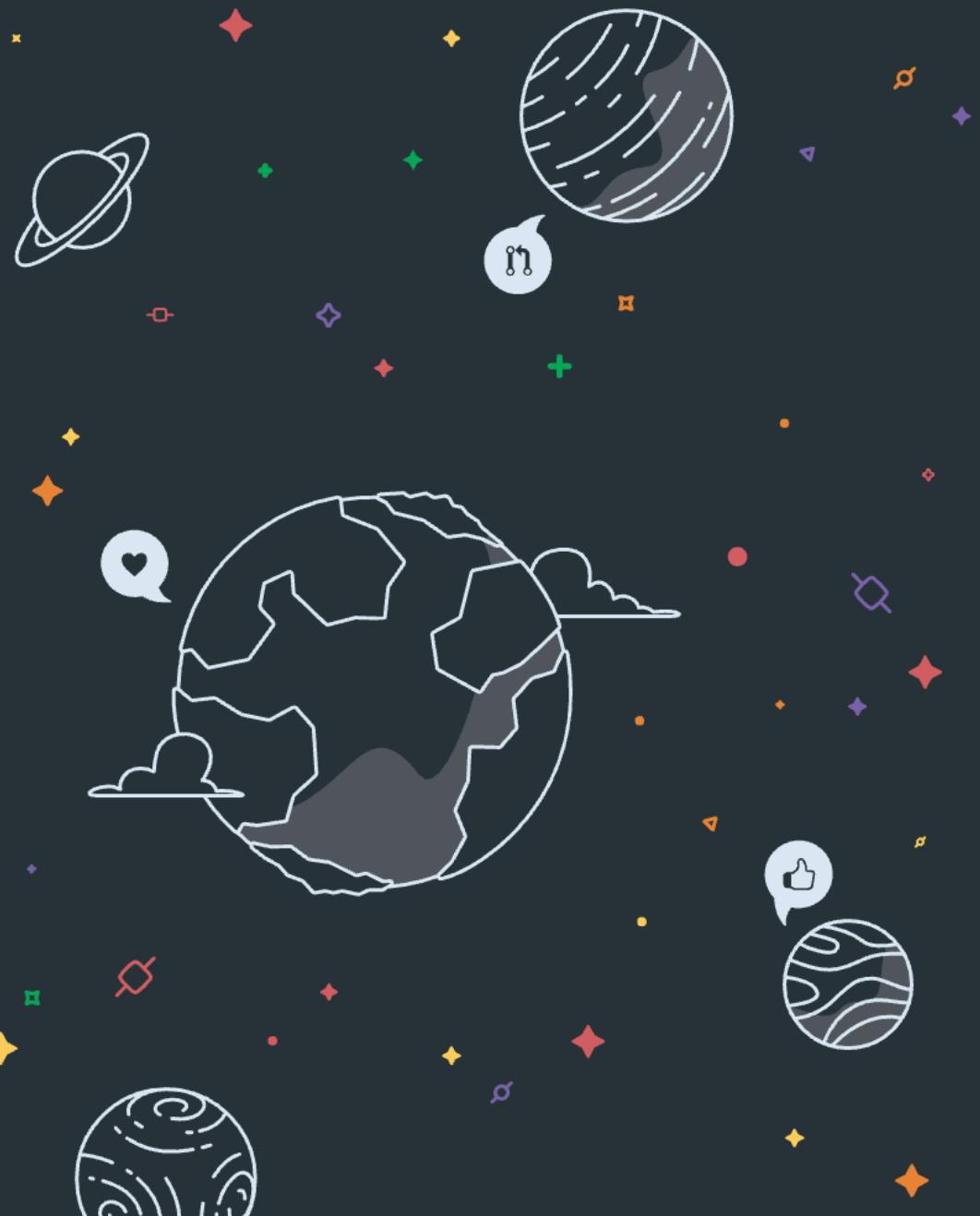




Sign in to GitHub.com

Username or email address

Password

[Forgot password?](#)[Sign in using your browser ↗](#)

Your Identity

The first thing you should do when you install Git is to set your user name and email address. This is important because every Git commit uses this information, and it's immutably baked into the commits you start creating:

```
$ git config --global user.name "John Doe"  
$ git config --global user.email johndoe@example.com
```

Configure Git

This is used to identify the commits you create. Anyone will be able to see this information if you publish commits.

Name

Seth Russell

Email

magic-lantern@users.noreply.github.com

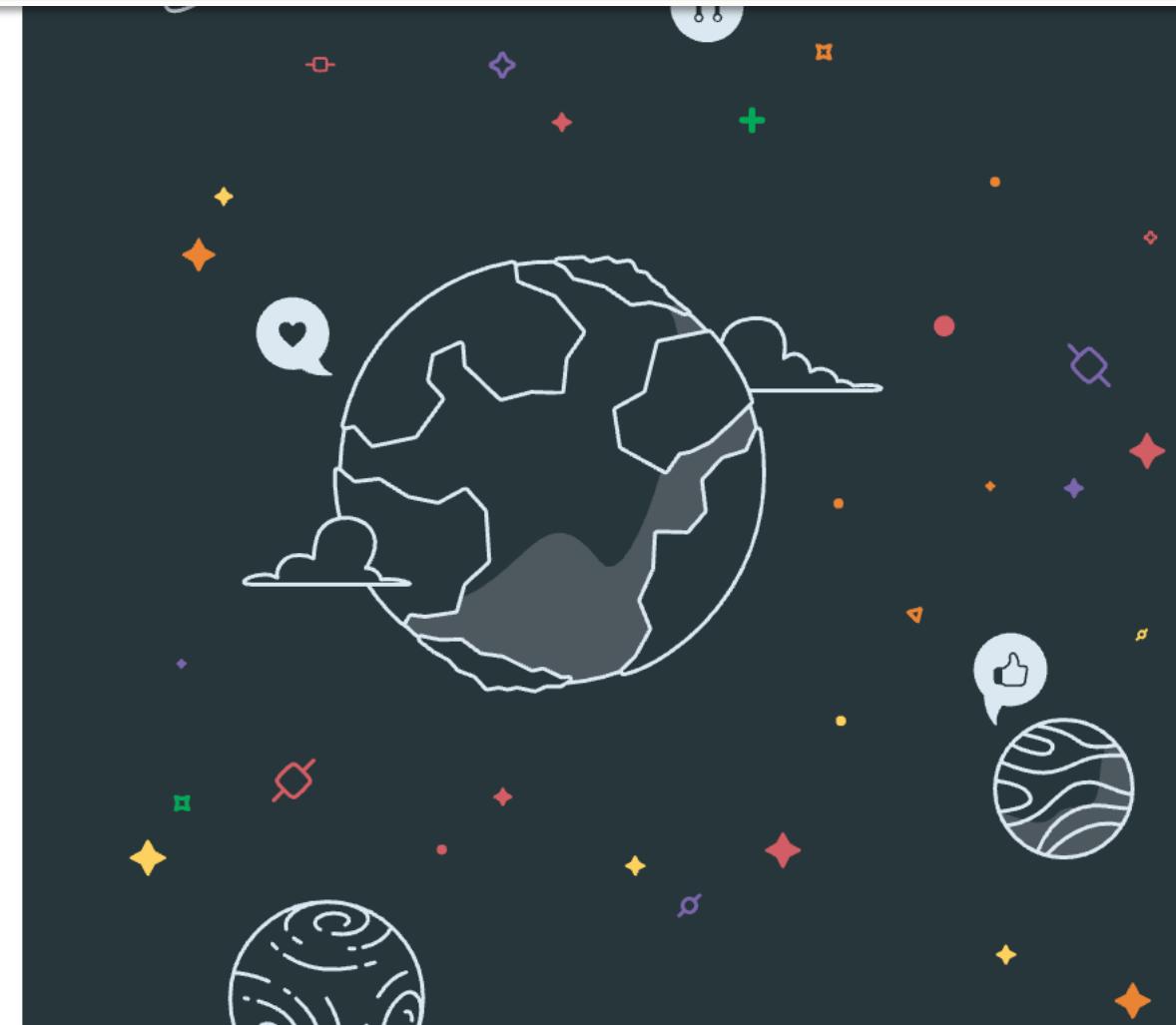
Continue

Cancel

Example commit

Fix all the things

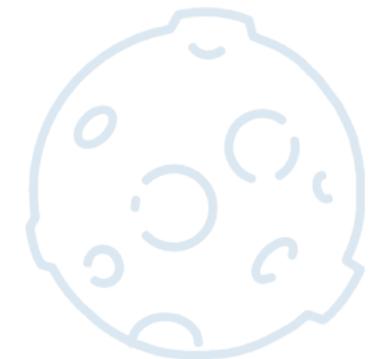
 Seth Russell committed 30 minutes ago





Let's get started!

Add a repository to GitHub Desktop to start collaborating



Create a Tutorial Repository...

Clone a Repository from the Internet...

Create a New Repository on your Hard Drive...

Add an Existing Repository from your Hard Drive...

Filter your repositories

Your Repositories

- balinterdi/rarwe-code
- callahantiff/Abra-Collaboratory
- magic-lantern/abc-test
- magic-lantern/Abra-Collaboratory
- magic-lantern/altair_exploration
- magic-lantern/AntimicrobialCDS
- magic-lantern/becausalfrence
- magic-lantern/bios6645
- magic-lantern/biostatistics-seminar-2017
- magic-lantern/bleeding-edge



Let's get started!

Add a repository to GitHub Desktop to

 Create a Tutorial Repository...

 Clone a Repository from the Internet

 Create a New Repository on your computer

 Add an Existing Repository from your computer

 **ProTip!** You can drag & drop an archive here to add it to Desktop

Create a New Repository ×

Name

Description

Local Path Choose...

Initialize this repository with a README

Git Ignore Python

License Apache License 2.0

Cancel Create Repository



Current Repository
testing

Current Branch
master

Publish repository
Publish this repository to GitHub

Changes

History



0 changed files

No local changes

There are no uncommitted changes in this repository. Here are some friendly suggestions for what to do next.



Publish your repository to GitHub

This repository is currently only available on your local machine. By publishing it on GitHub you can share it, and collaborate with others.

[Publish repository](#)

Always available in the toolbar for local repositories or

Open the repository in your external editor

Select your editor in [Preferences](#)

[Open in Visual Studio Code](#)

Repository menu or

Summary (required)

Description

View the files of your repository in Finder

Repository menu or

[Show in Finder](#)

Changes 1

History



runme.py

 1 changed file runme.py

		@@ -1 +1,2 @@
	1	-print('hello world') ↵
	1	+print('hello world')
	2	+print('new line') ↵



Update runme.py

Description

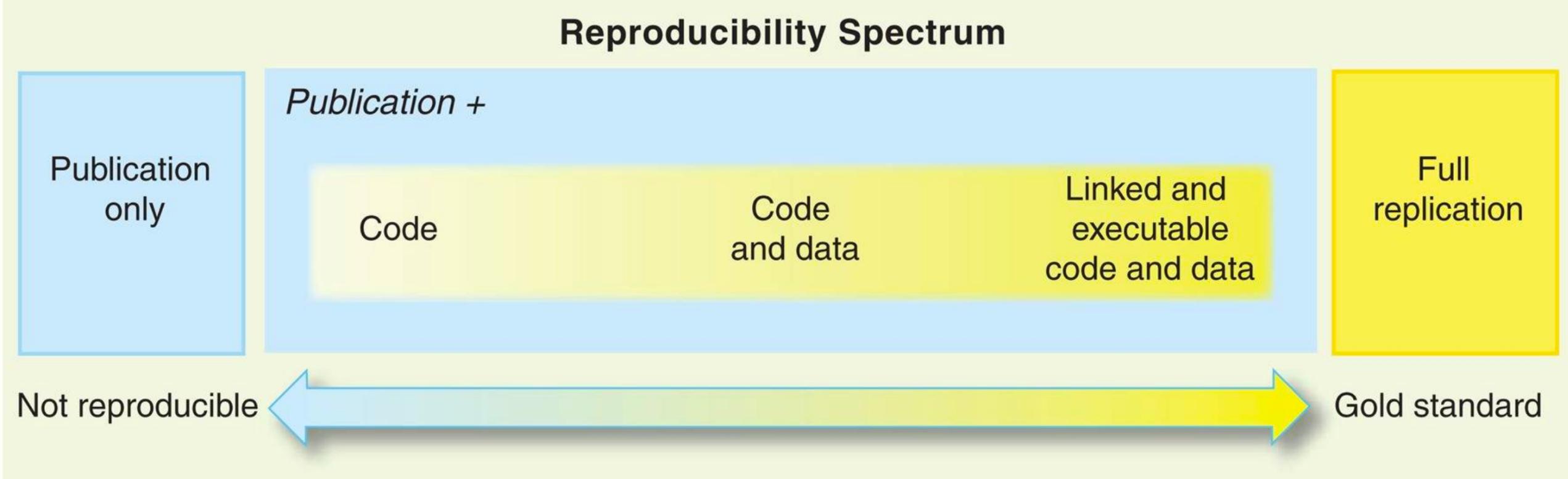
Commit changes

Version Control

What deep wounds ever closed
without a scar?

Lord Byron

Git and Reproducible Science



Peng RD. Reproducible research in computational science. *Science*. 2011 Dec 2;334(6060):1226-7. doi: 10.1126/science.1213847. PMID: 22144613; PMCID: PMC3383002.