Git and GitHub Research Collaboration: Day 2

Richard W. Evans [2]

[2]University of Chicago, Open Source Economics Laboratory, M.A. Program in Computational Social Science, and Open Research Group, Inc.

September 21, 2019
New York University and Schmidt Futures

Git review

Updating master

- `git fetch upstream`
- `git merge upstream/master`
- `git push origin master`

Adding changes

- `git add [filename]` or `git add -A`
- `git commit -m "message"`
- `git push origin [BranchName]`

What is git merge?

git merge

`git merge` combines file changes from another branch into the current branch

$$\text{git merge } [\text{remote}]/[\text{branch}]$$

or

$$\text{git merge } [\text{branch}]$$

merge conflict

`git merge` conflicts arise when new commits are in both branches that have differences

- I like Atlassian git merge conflict tutorial
- https://www.atlassian.com/git/tutorials/using-branches/merge-conflicts

Merge conflict

- Things you can do to resolve merge conflict
    - Complete burndown and reclone
    - use `git diff` to predict
    - `git reset -hard HEAD`
    - Resolve merge conflict

git diff

From https://stackoverflow.com/questions/8382019/how-do-i-git-diff-on-a-certain-directory

▲

43

▼

If you're comparing different branches, you need to use `--` to separate a git revision from a filesystem path. For example, with two local branches, `master` and `bryan-working` :

```
$ git diff master -- AFolderOfCode/ bryan-working -- AFolderOfCode/
```

Or from a local branch to a remote:

```
$ git diff master -- AFolderOfCode/ origin/master -- AFolderOfCode/
```

Resolve merge conflict

- Create a problem
  - Make new branch off master called newbranch
  - Make and commit update to master branch
  - Make different update and commit to newbranch
  - Change directory to newbranch
  - git merge master will create merge conflict
- Resolve problem with terminal text editor
  - vim [conflicted_file]
  - git add [conflicted_file]
  - git commit -m "merge message"

Data best practices

Stanford Libraries site

https://library.stanford.edu/research/data-management-services/data-best-practices

- File naming
- File formats
- Data versioning
- Creating metadata
- Helpful tools
- Working with sensitive data

AEA data policy revisited

AEA Data and Code availability policy

Python data differencing

- Python data and file differencing tools: difflib and filecmp
  (https://docs.python.org/2/library/difflib.html)
- Makes use of the cryptographic hash function
  (https://en.wikipedia.org/wiki/Cryptographic_hash_function)

Project 2: GE model

- Project 2: Overlapping Generations Model
- Multi-part project
- Lots of GitHub conversation