2017-02-06
Rick O. Gilmore
2017-01-29

Today's topics

- The whats, whys and hows of version control

What is version control

- "management of changes to documents, computer programs, large web sites, and other collections of information." https://en.wikipedia.org/wiki/Version_control
- Track changes an example
- Wikipedia another

Why do version control

Why do version control

- Revert to previous version
- Keep record of who made what changes when
- Transparency

How to do version control

- Edit, Save, Attach
- Consistent file-naming (and incrementing) conventions
    - gilmore-etal-nature-2017-01-28-1319.docx
    - gilmore-etal-nature-2017-01-28-1400.docx
    - ... etc.

Pros and cons

- Pros
    - No special software needed
    - Works with every file type
    - Good for any file type
- Cons
    - Not everyone uses the same file naming conventions
    - Dealing with conflicts
    - File management

How to do version control

- Google Docs
  - Gilmore, Diaz, Wyble, & Yarkoni (in press). Progress Toward Openness, Transparency, and Reproducibility in Cognitive Neuroscience. *Annals of the New York Academy of Sciences.* https://docs.google.com/document/d/1BOL_AdvMceJluyPsL9aop1iBqxOla15fXcstPFHe9e0/edit?usp=sharing

Pros and cons

- Pros
  - Works with docs and sheets
  - Unlimited, automatic track changes
  - Stored in cloud
  - No special naming conventions
- Cons
  - NYAS Editor wanted Word doc with track changes!
  - Not all types of files
  - Not everyone likes/knows how to use Google docs

How to do version control

- Box.com
  - Enterprise version stores 100 versions

Your browser does not support the video tag.

Pros and cons

- Pros
  - Keep the same file name
  - Let Box do automatic versioning
  - Shared file space
  - Any kind of file Box supports
- Cons
  - 100 versions
  - Penn State enterprise license

How to do version control

- Microsoft OneDrive
- Pros
  - Multiple platforms (like Box)
  - Multiple file types
  - Hallquist lab uses for non-code
- Cons
  - ?

How to do version control

- Open Science Framework
  - Integrates with Box, Dropbox, GitHub
- Pros
  - Free, open source, devoted to open science
  - Many different file types
- Cons
  - Version control depends on storage sources
  - Wyble lab reports some frustrations

How to do version control

- Git
  - free, open source version control system
  - created by Linus Torvalds, creator of the Linux operating system, to manage that project's software development
- Web-based git repositories
  - GitHub
  - GitLab
  - Bitbucket

Cons

- Designed by and for software developers
- Text files
- Wonky
- Longevity of web-based repos?
- User have to specifically push (upload) and pull (download) files

Pros

- Great for R code, data
- Great for Jupyter notebooks (later in course)
- GitHub exploits the power of the web
- Gilmore, Hallquist, Wyble use

Distributed version control

Distributed version control

- Everyone works on their own local copies
- Pull before they start work
- Save and push when they're finished