Intro to Unix (shell)

Dr. Kate Hertweck
University of Texas at Tyler
Department of Biology
Fall 2014

Twitter: @k8hert
Research website: https://sites.google.com/site/k8hertweck/

If you have a Windows laptop (and feel comfortable doing so), download and install Git SCM:
http://msysgit.github.io
(flash drives with program also available)

Goals and agenda

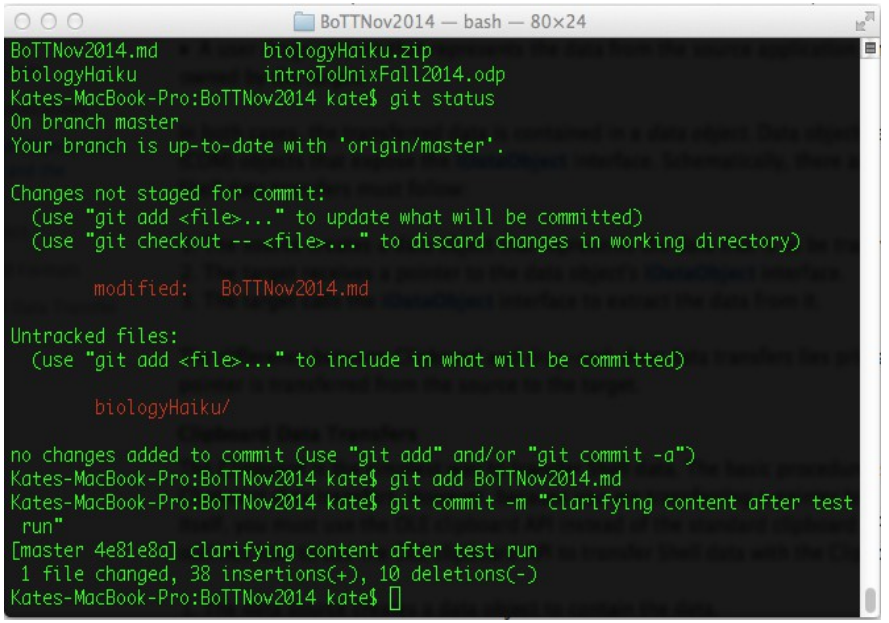
- Figure out why we're here
- Test some unix commands
  - File management
  - Finding things
- Be able to work and learn on your own later

What did I just install?

- Git (https://github.com)
  - Version control system
  - Code repository
- Bash (Unix shell)
  - Command processor
  - A different way to talk to your computer (type command, hit enter: command processes)
- On Mac computers: Go/Utilities/Terminal (only replaces bash shell)

Why did I just install it?

- Reproducibility: doing the same thing over and over (or remembering later what you did)
- Scalability: good for repeated tasks
- To get drunk on the power of command line interfaces

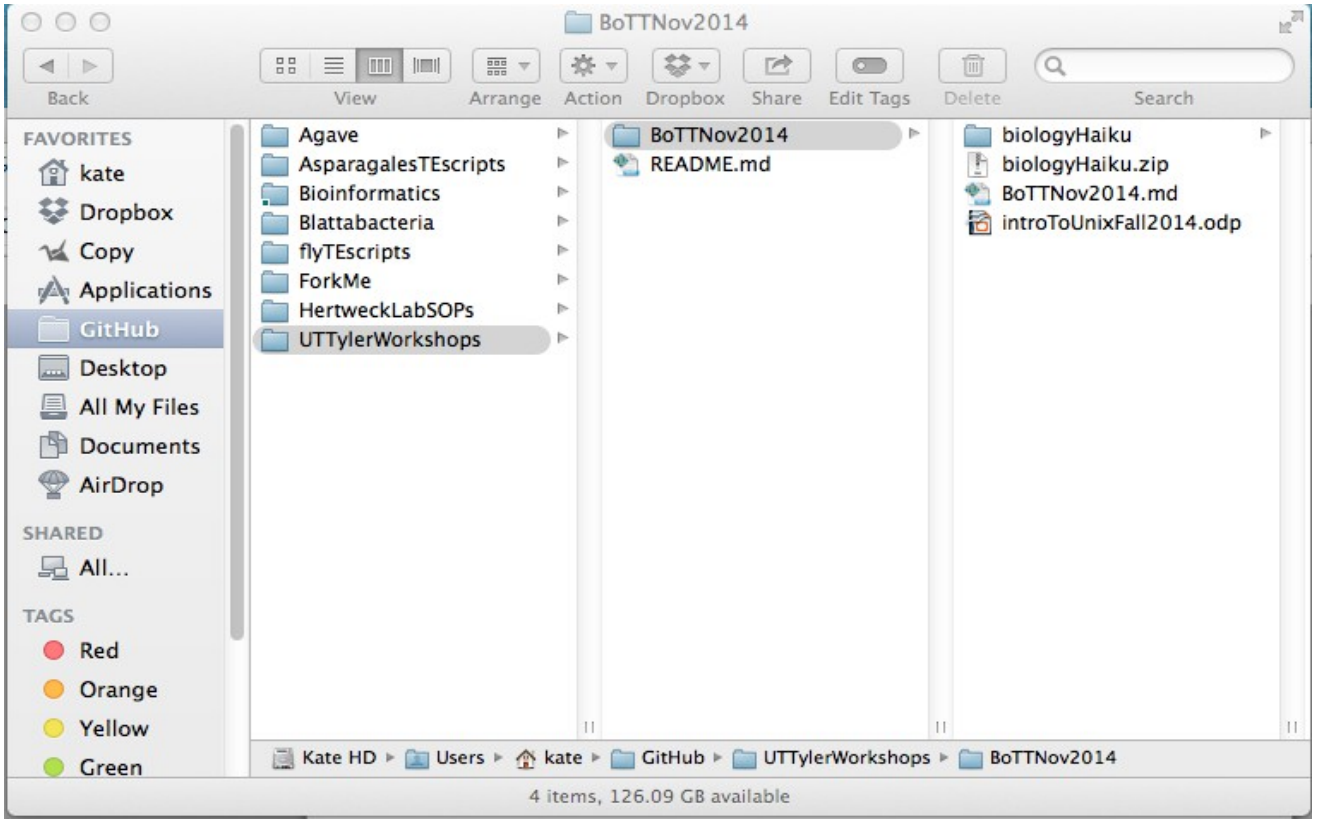


What can I really learn in an hour?

- Unix is accessible for entry-level learners
- Use shell scripting as backbone for pipelines and algorithms
- Skills are portable to other programming languages

Navigation and file management

- Where am I in the computer?
- How do I organize files?



Navigation and file management

- Paths
- Files, directories, filenames
- Commands
  - `ls` list files in a directory
  - `cd` change directory
  - `mkdir` make a new directory
  - `mv` move a file
  - `cp` copy a file
  - `rm` remove a file

A few shortcuts

- Tab completion: for filename entry
- Drag and drop files: enter path (Mac only?)
- Arrow up: recall old commands
- Control C: new line/stop command from running
- `--help`: get help (Mac: also `man command`)

Download this small “data” file:
http://bit.ly/1tm4k7G; click on “View Raw”
(Also available in Kate's GitHub folder)

Finding things

- `find` to locate files and directories
- `grep` to search text within files
  - http://regexpal.com

Other important commands to explore

- `sort` to sort lines in a file
- `uniq` to remove duplicates
- `wc` to count words, lines, characters
- `cut` to remove portions of a line
- `cat` to concatenate files
- `head, tail` print first or last lines in a file

Software

- Text editor: Notepad++ http://notepad-plus-plus.org
  - Write and save scripts independently of shell
- Shell: Cygwin https://www.cygwin.com
  - A larger, more recently updated set of unix tools
- GitHub student developer pack: https://education.github.com/pack
  - Free access to other tools

Resources

- Software Carpentry lessons:
  - Introducing the Shell
  - Files and Directories
  - Finding Things
- Unix cheat sheet
- Stuck and don't know what to do?
  - Google
  - Stack Overflow

Want to learn more?

- Spring 2015: BIOL 4306 and 4106, Bioinformatics lecture + lab
- Undergrad minor in Bioinformatics and Genomics
- Fall 2015: Bioinformatics for Research/Adv Bioinformatics (graduate level)

- Slides and outline available: https://github.com/k8hertweck/UTTylerWorkshops/

- Other topics?
  - `git` for version control and collaboration
  - `awk`
  - Pipes, filters, redirects
  - Loops
  - Shell scripts