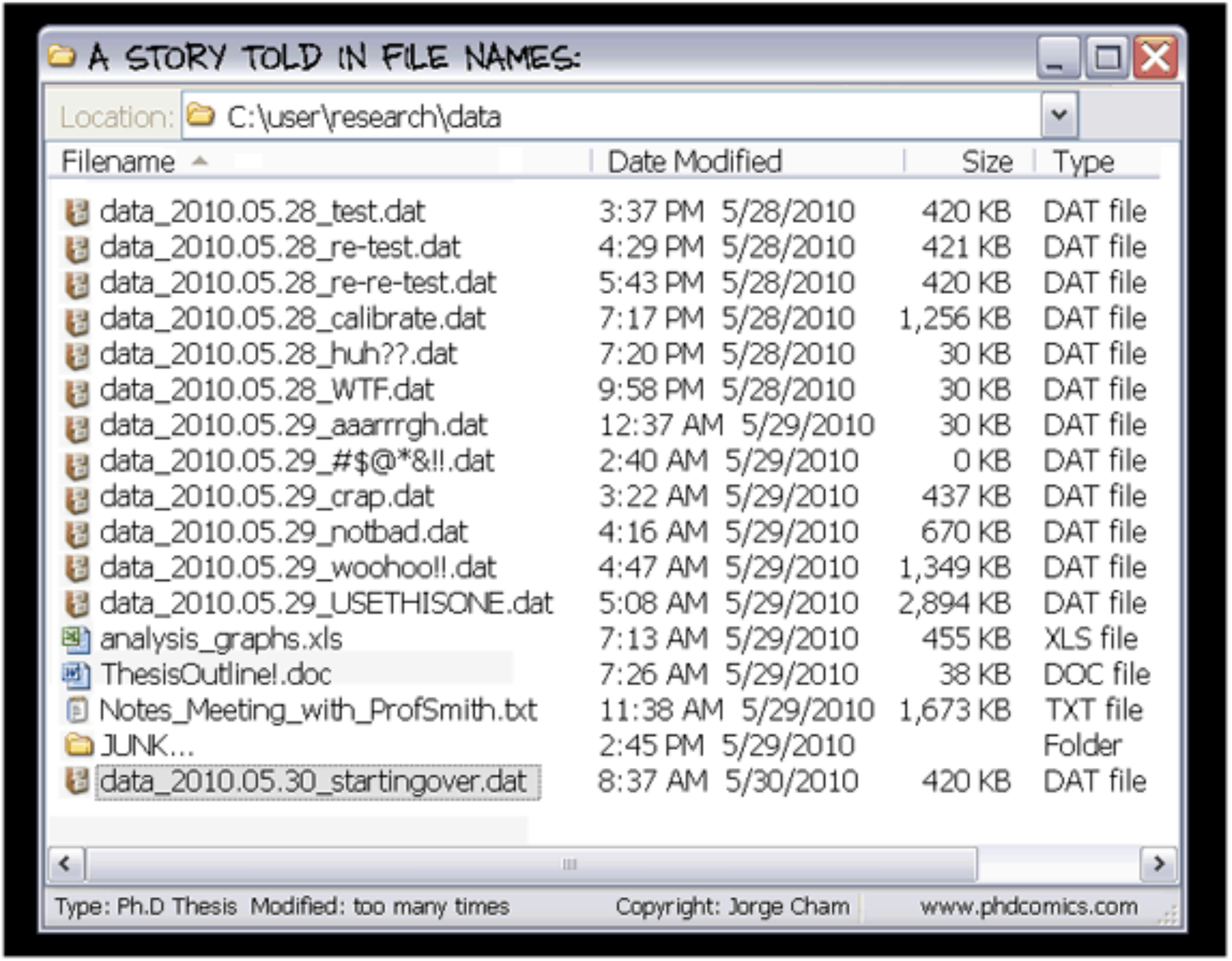
Intro to Version Control

NR995 Module 9
2017 Fall

Version Control

- History of your files & changes
- Why use version control?
- What's your system?
  - Sync to dropbox or google drive
  - Periodically copy folder to external drive
  - Rename revisions

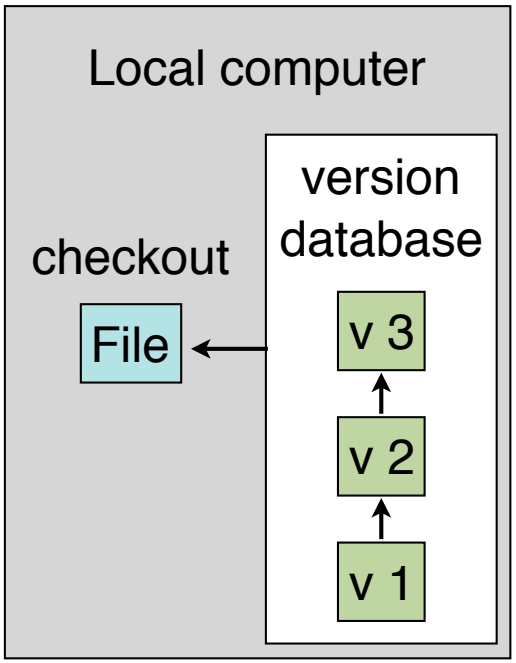
"FINAL".doc


A STORY TOLD IN FILE NAMES:
Location: C:\user\research\data
Filename
Date Modified
Size
Type
data_2010.05.28_test.dat 3:37 PM 5/28/2010 420 KB DAT file
data_2010.05.28_re-test.dat 4:29 PM 5/28/2010 421 KB DAT file
data_2010.05.28_re-re-test.dat 5:43 PM 5/28/2010 420 KB DAT file
data_2010.05.28_calibrate.dat 7:17 PM 5/28/2010 1,256 KB DAT file
data_2010.05.28_huh??.dat 7:20 PM 5/28/2010 30 KB DAT file
data_2010.05.28_WTF.dat 9:58 PM 5/28/2010 30 KB DAT file
data_2010.05.29_aaarrrgh.dat 12:37 AM 5/29/2010 30 KB DAT file
data_2010.05.29_#$@*&!!.dat 2:40 AM 5/29/2010 0 KB DAT file
data_2010.05.29_crap.dat 3:22 AM 5/29/2010 437 KB DAT file
data_2010.05.29_notbad.dat 4:16 AM 5/29/2010 670 KB DAT file
data_2010.05.29_woohoo!!.dat 4:47 AM 5/29/2010 1,349 KB DAT file
data_2010.05.29_USETHISONE.dat 5:08 AM 5/29/2010 2,894 KB DAT file
analysis_graphs.xls 7:13 AM 5/29/2010 455 KB XLS file
ThesisOutline!.doc 7:26 AM 5/29/2010 38 KB DOC file
Notes_Meeting_with_ProfSmith.txt 11:38 AM 5/29/2010 1,673 KB TXT file
JUNK... 2:45 PM 5/29/2010 Folder
data_2010.05.30_startingover.dat 8:37 AM 5/30/2010 420 KB DAT file
Type: Ph.D Thesis Modified: too many times
Copyright: Jorge Cham
www.phdcomics.com

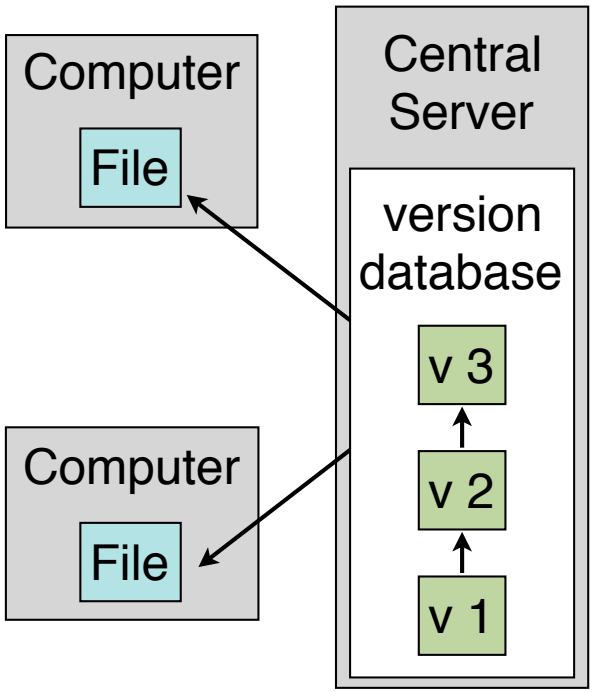
Version Control

- Version control systems record changes to a file over time
- Long history with programmers

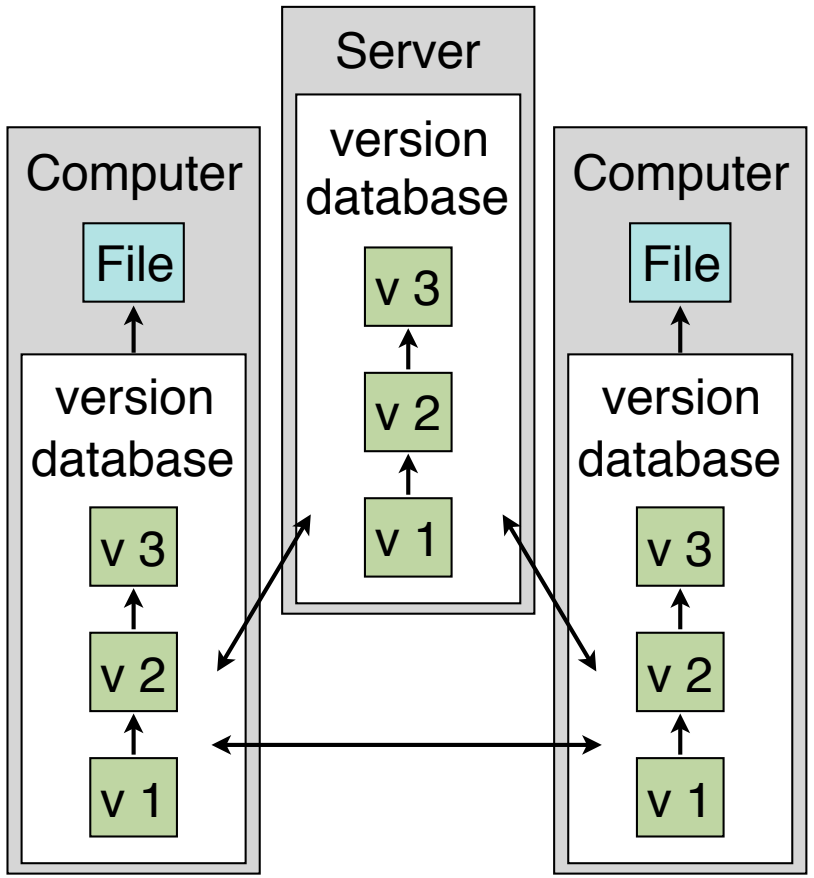
Local VCS



Centralized VCS

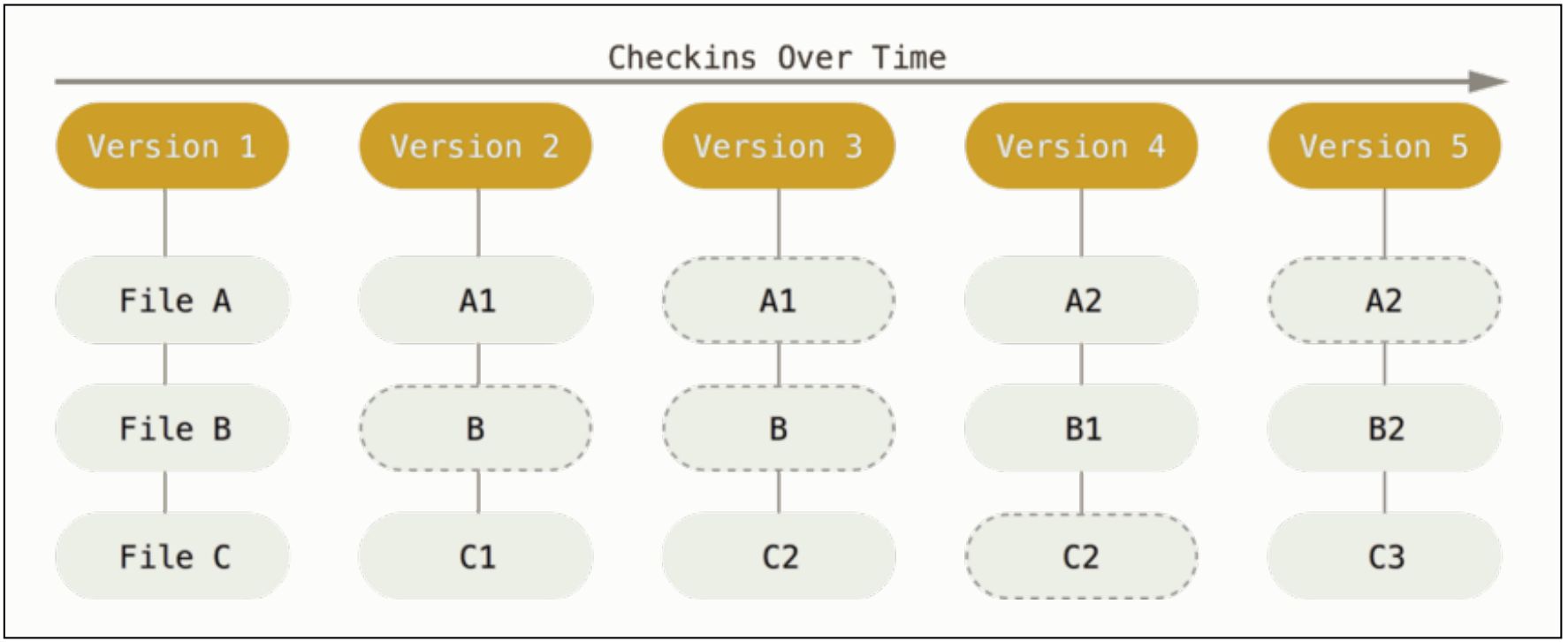


Distributed VCS



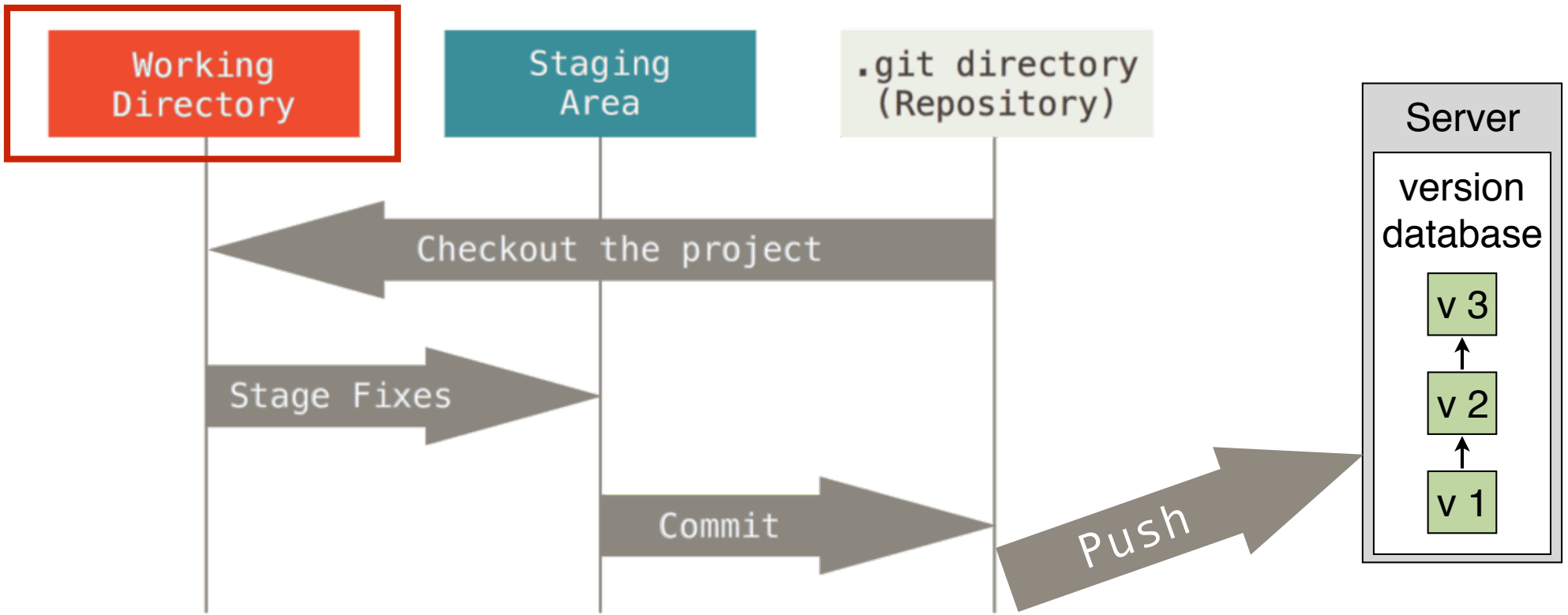
Git (git-scm.com)

- Distributed VCS
- Full version database on each computer
- Stores snapshots of all tracked files in directory
- Store locally, push to remote server (e.g., GitHub)



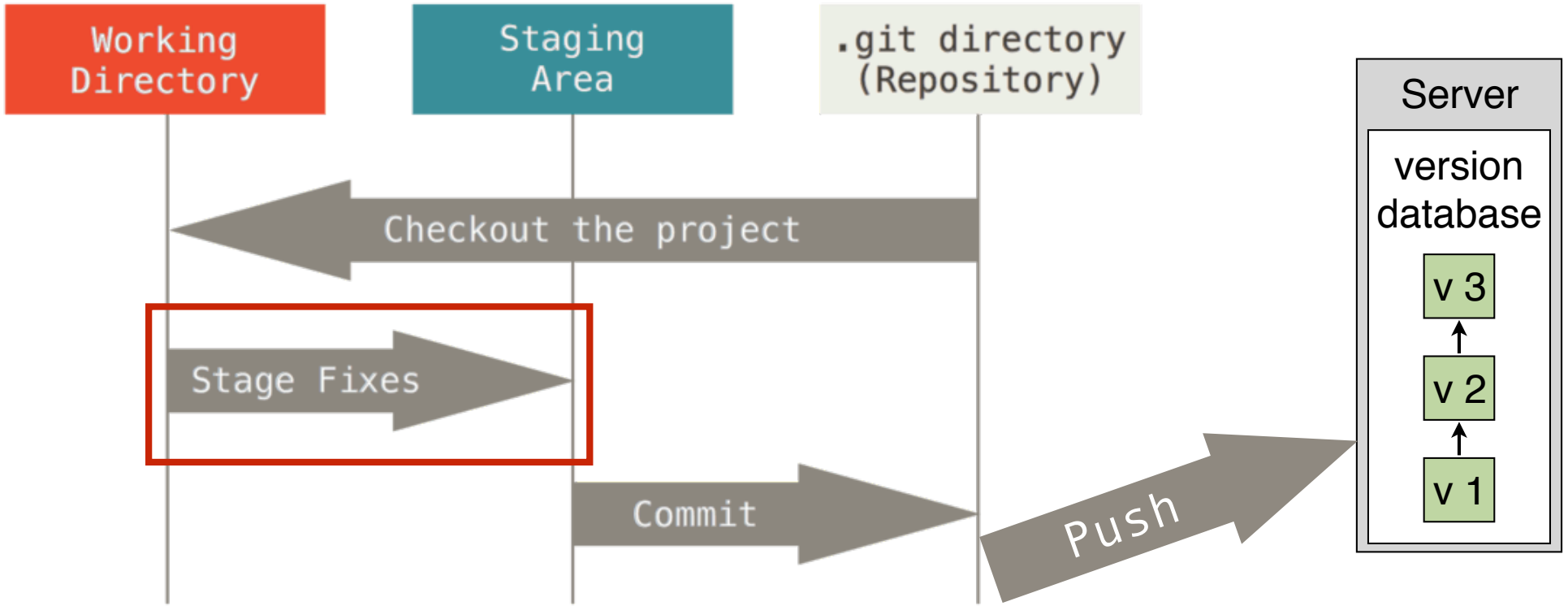
Git workflow

1. Work on a file, save changes



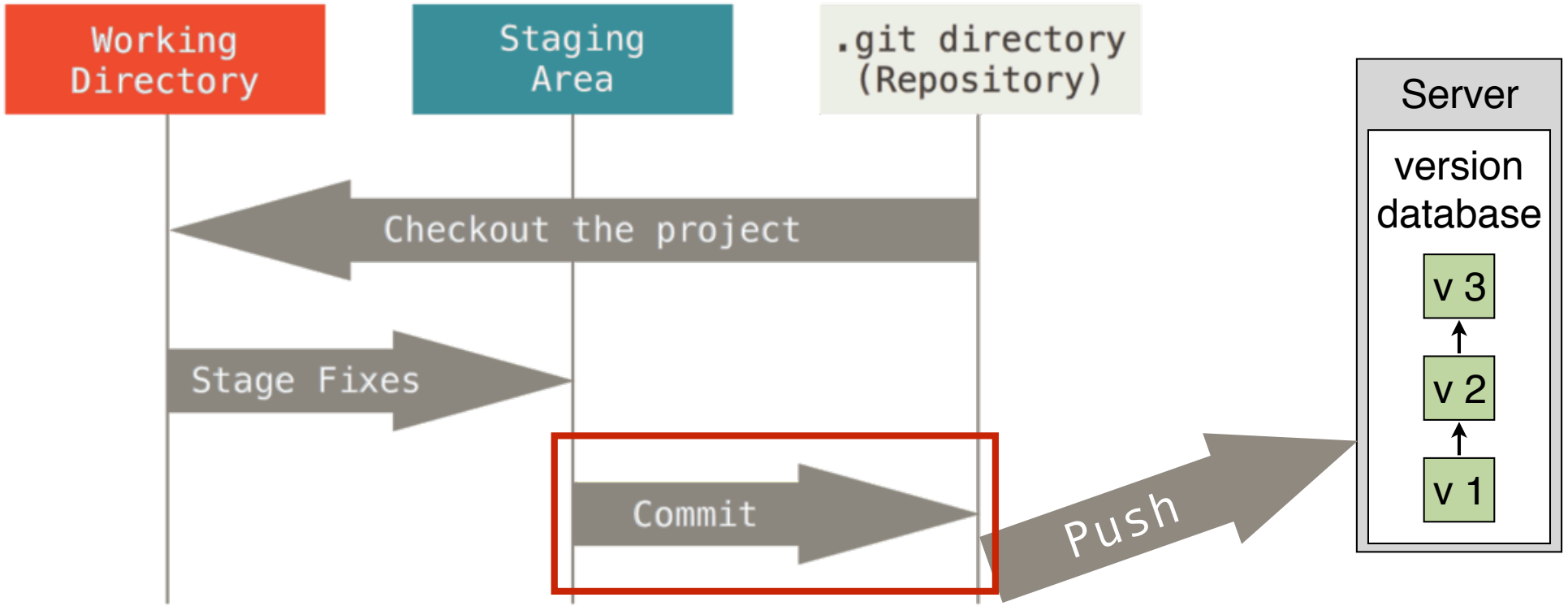
Git workflow

1. Work on a file, save changes
2. Stage changes



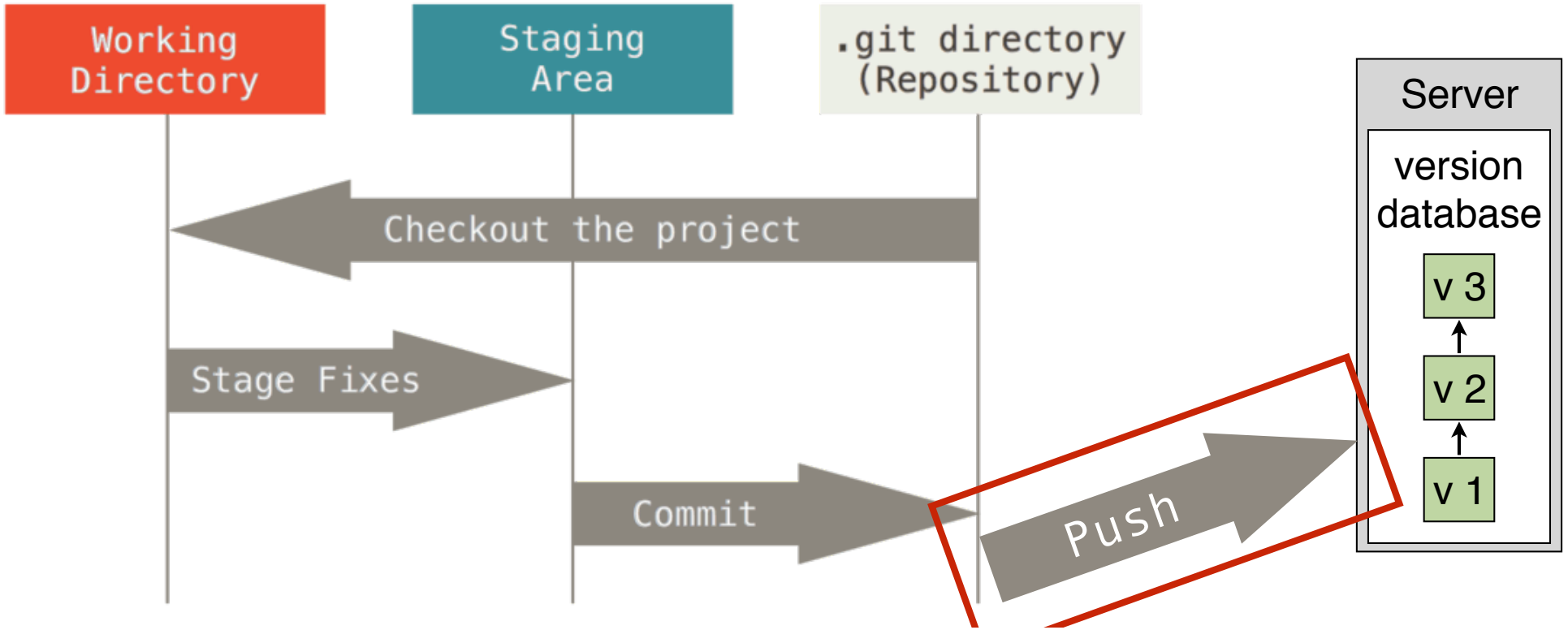
Git workflow

1. Work on a file, save changes
2. Stage changes
3. Commit staged changes



Git workflow

1. Work on a file, save changes
2. Stage changes
3. Commit staged changes
4. Push commit to remote repository



Benefits of Git

- Record of all file changes
  - Actual files, time stamps, commit messages
- Collaboration
  - Merge work, parallel branches, user history
- Integration with RStudio
- Reproducibility & transparency
- Organization