lecture02

February 12, 2023

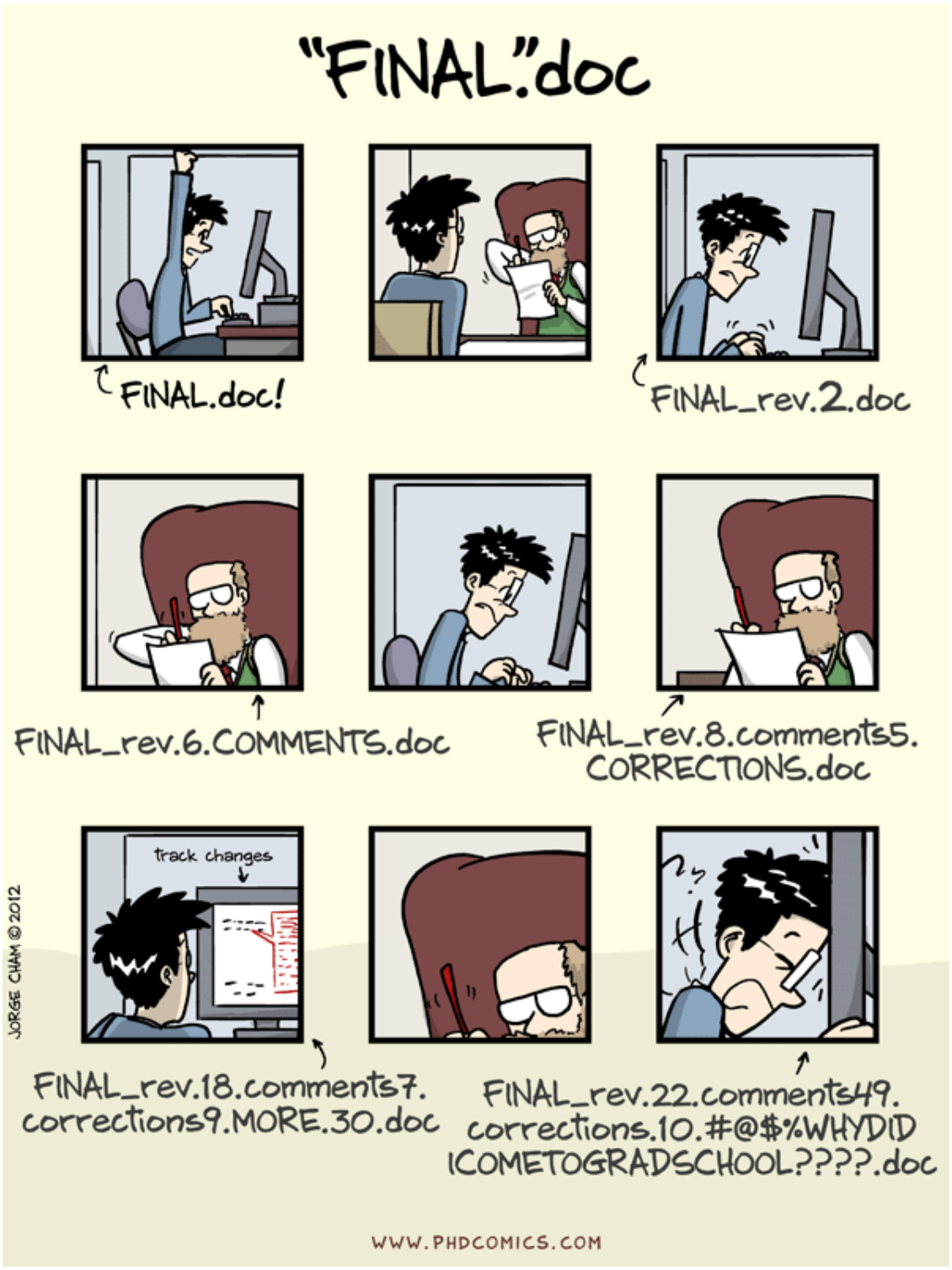
1 Gentle Introduction to GitHub

Ahmed Moustafa

2 Quick Review

https://www.menti.com/alnpj76xuxwu

https://www.mentimeter.com/app/presentation/alx5mivoqmqxen4s2jh158wwjovgfhrh


"FINAL".doc
FINAL.doc!
FINAL_rev.2.doc
FINAL_rev.6.COMMENTS.doc
FINAL_rev.8.comments5.
CORRECTIONS.doc
track changes
FINAL_rev.18.comments7.
corrections9.MORE.30.doc
FINAL_rev.22.comments49.
corrections.10.#@$%WHYDID
ICOMETOGRADSCHOOL????.doc
JORGE CHAM © 2012
WWW.PHDCOMICS.COM

3 Before & After

```
<IPython.core.display.Image object>
```

```
<IPython.core.display.Image object>
```

Source: Towards Data Science: Getting Started with Git and GitHub

4 What is Version Control?

Version control is a system that records changes to a file or set of files over time so that you can recall specific versions later

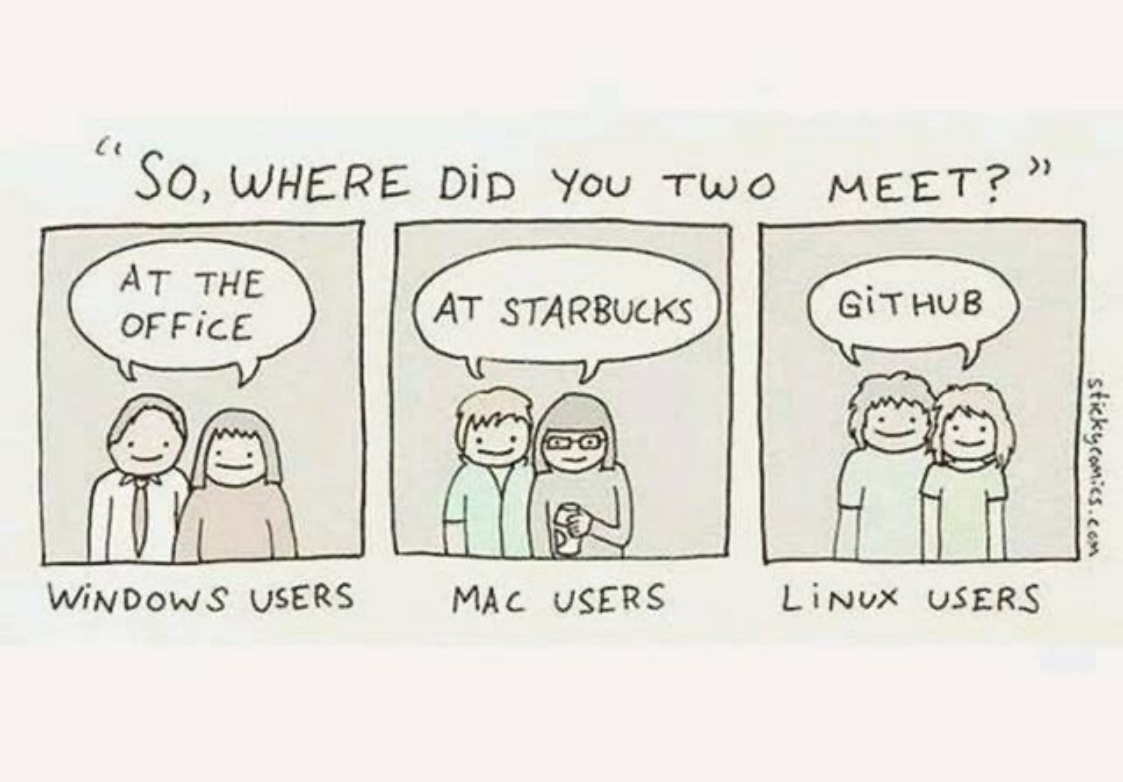
- **Tracks changes** made to files over time
- Enables **collaboration** and sharing of files
- Maintains a **history** of changes made to files
- Facilitates **revert** to a previous version of a file if necessary

5 Why is Version Control Important in Data Science?

- Manages the history and evolution of data science projects
- Facilitates collaboration and sharing among data scientists
- Helps maintain and organize different versions of code and data
- Enables tracking and **reproducibility** of data science projects

6 Git & GitHub

- **Git** is a **distributed** version control system that is used to manage and track changes made to files.
- **GitHub** is a web-based platform that provides hosting for Git repositories and facilitates collaboration among data scientists.



7 How to Obtain Git?

https://git-scm.com/

8 How Does Git Work?

<IPython.core.display.Image object>

Source: NeSI's Git: Reference Sheet

9 How Does Git Work? *local*

Command	Description
`clone`	Copies a remote repository into the current directory
`init`	Creates a new empty repo in the current directory
`add`	Adds files to the staging area
`status`	Lists changes in the working directory, and staged files
`commit`	Records everything in the staging area to the repository
`reset`	Removes all files from staging area (opposite of `add`)

10 How Does Git Work? *remote*

Command	Description
`fetch`	Gets status of `origin`

Command	Description
`pull`	Incorporates changes from `origin` into local repo
`push`	Incorporates changes from local repo into `origin`

11 Contributing to a Repository

- **Fork** a Repository: To contribute to a repository, you must first fork it to create a copy of the codebase on your own GitHub account
- Make Changes: Make the desired changes to the code and commit the changes to your local repository
- **Push** Changes: Push the changes to your forked repository on GitHub.
- Create a **Pull Request** (**PR**): Create a pull request to request the changes be merged into the original repository.

12 Contributing to a Repository

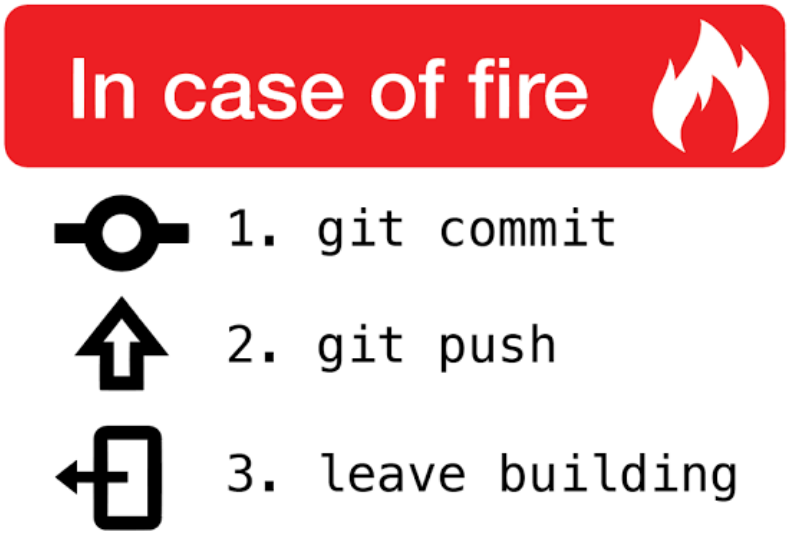
```
<IPython.core.display.Image object>
```

Source: edav.info: Chapter 6 GitHub/git Resources

13 GitHub Example

- Go to repository https://github.com/ahmedmoustafa/hello-world
- Fork the repo under your account
- Edit notebook `hello-world.ipynb`, using for example, Colab or Codespace
- Add a cell, write a `print` statement to print your full name
- Commit and push your changes
- Submit a pull request (PR)

14 GitHub Exercise

- Go to repository https://github.com/ahmedmoustafa/bug-in-the-code
- Fork the repo under your account
- Edit notebook `bug-in-the-code.ipynb`, using for example, Colab or Codespace
- There is a syntax error.
- Fix it.
- The **first** to submit a pull request with the **fix** will have have a **bonus**

15 So, as a general coding rule of thumb