

OG-ZAF: An open-source collaborative policy modeling environment

Jason DeBacker¹ **Richard W. Evans**²

¹University of South Carolina, Department of Economics

²Abundance Institute, Open Research Group, Inc.

August 2, 2024

United Nations, Development Policy Seminar

Two famous Git and GitHub quotes

Two famous Git and GitHub quotes

The ability to effectively and efficiently scale up collaboration is the biggest driver of productivity growth in the world.

Richard W. Evans, August 2, 2024, Cape Town, South Africa

Two famous Git and GitHub quotes

The ability to effectively and efficiently scale up collaboration is the biggest driver of productivity growth in the world.

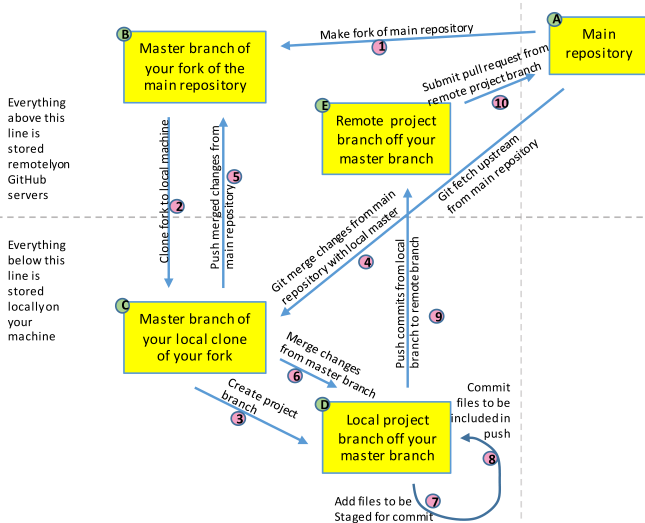
Richard W. Evans, August 2, 2024, Cape Town, South Africa

Two warnings that a seasoned Git and GitHub user should always give a new entrant to this type of version control and code collaboration are the following.

- *The learning curve is steep.*
- *The workflow is initially not intuitive.*

Richard W. Evans, Git-Tutorial, UN-OG-Training material, page 1.

Git and GitHub workflow



What are Git and GitHub?

Definition (**Git**)

Git is an open source distributed version control system (DVCS) software that resides on your local computer and tracks changes and the history of changes to all the files in a directory or repository. See the Git website <https://git-scm.com/> and the [Git Wikipedia entry](#) for more information.

What are Git and GitHub?

Definition (**GitHub**)

GitHub or **GitHub.com** is a cloud source code management service platform designed to enable scalable, efficient, and secure version controlled collaboration by linking local Git version controlled software development by users. GitHub's main business footprint is hosting a collection of millions of version controlled code repositories. In addition to being a platform for distributed version control system (DVCS), GitHub's primary features include code review, project management, continuous integration unit testing, GitHub actions, and associated web page (GitHub pages) and documentation hosting and deployment.

What are Git and GitHub?

Definition (**repository** or **repo**)

A **repository** or **repo** is a directory containing files that are tracked by a version control system. A local repository resides on a local machine. A remote repository resides in the cloud.

Wide usage

“[Octoverse 2023: The state of open source software](#)”. In 2023:

- “93% of developers use [GitHub] to build and deploy software everywhere”
- 420 million total projects
- 284 million public repositories
- 65 thousand public generative AI projects
- 4.5 billion total contributions to all projects on GitHub
- 98 million new projects started on GitHub

Wide usage

“The best indication of Git’s market dominance is a survey of developers by Stack Overflow. This found that 88.4% of 74,298 respondents in 2018 used Git (up from 69.3% in 2015). The nearest competitors were Subversion, with 16.6% penetration (down from 36.9%); Team Foundation Version Control, with 11.3% (down from 12.2%); and Mercurial, with 3.7% (down from 7.9%). In fact, so dominant has Git become that the data scientists at Stack Overflow didn’t bother to ask the question in their 2019 survey.”

Andy Favel, *[“The history of Git,”](#)* Feb. 2020

Git and GitHub brought us Linux

“The development of Git began on 3 April 2005. Torvalds announced the project on 6 April and became self-hosting the next day. The first merge of multiple branches took place on 18 April. Torvalds achieved his performance goals; on 29 April, the nascent Git was benchmarked recording patches to the Linux kernel tree at the rate of 6.7 patches per second. On 16 June, Git managed the kernel 2.6.12 release.

Wikipedia, “[Git](#),” Aug. 2024

Git and GitHub brought us Linux

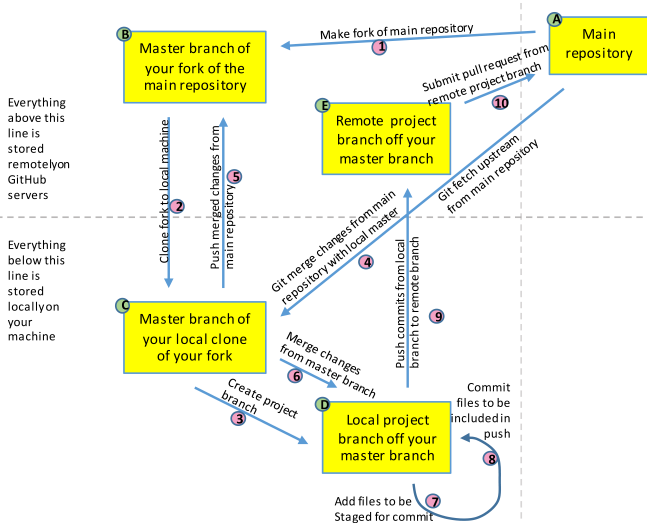
Show Linux contributors page: <https://github.com/torvalds/linux>

Alternatives to Git and GitHub

Alternatives to GitHub include GitLab and Bitbucket. Other alternatives are documented in this [March 2024 post](#) by Software Testing Help.

But GitHub has the largest user base and largest number of repositories.

Git and GitHub workflow



Git vs. Google Docs and Dropbox

- Google Docs and Dropbox can allow changes with no outside confirmation
- Google Docs and Dropbox comments and suggesting mode can pile up in a difficult way
- Google Docs history is difficult to order and search
- Dropbox can create version conflicts
- Cannot set hierarchical permissions easily

Git workflow example in the terminal

Show what workflow looks like in the terminal.

GitHub Issues

GitHub issues are a great place to make lists of things to update and fix and a great forum for discussing problems.

GitHub Issue AllStars

- OG-ZAF [Issue #18](#), “Compute lifetime earnings profiles”
- OG-USA [Issue #20](#), “Values of dataframe in psid_data_setup.py”
- OG-Core [Issue #828](#), “Generalize spline tax functions to 2D”
- OG-Core, [Issue #574](#), “New monotone tax function estimation and Euler equation solution”
- OG-Core, [Issue #234](#), “Estimate tax functions from microdata that are monotonic in labor and capital income”

Pull requests (PRs)

Pull requests are the submissions of code changes that await approval, testing, and merging.

Components of a PR

- Description of what the PR does
- Show exactly what changed
- Continuous integration (CI) automatic testing
- PR thread
 - OG-ZAF, [PR #42](#), “UN data portal ssl fix”
 - OG-ZAF, [PR #31](#), “Calibrate IO matrix”
 - OG-ZAF, [PR #28](#) “Update get_e_orig() with an adjustment by age calibration for ZAF”