EPSC Abstracts
Vol. 12, EPSC2018-1058-2, 2018
European Planetary Science Congress 2018
© Author(s) 2018

The Use of Git in Planetary Science Research

Alessandro Frigeri
Istituto di Astrofisica e Planetologia Spaziali, INAF, Rome, Italy (Alessandro.Frigeri@iaps.inaf.it, Phone:+39-06-4993-4227);

Abstract

The accessiblity to planetary data is increasing with the development of new technologies and strategies for data distribution across the network.

This forces the remote sensing data analyst and planetary science researcher to update its own digital working environment towards flexible solutions.

The management of changes to documents, computer programs, and other collections of information is critical to keep track of developments and refinement within a research project.

A family of softwares generically called Version Control Systems are focused on supporting this task. Initially common among the software developers, in the last decade these system are becoming popular across the scientific community.

Among the different Version Control System, Git is gaining popularity as it adapts particularly well to a wide of different use cases from software developments, numerical experiment setup and scientific production [1].

The flexibility of Git comes at the price of a quite steep learning curve due to the absence of a standard git workflow, which has to be decided on the base of the specific project's needs. Here we explore some use cases which can turn useful in the field of Planetary Science.

1 Introduction

Git was created in 2005 by Linus Torvalds, the main developer of the Linux kernel. As the main pupil of its creator, Git is distributed as free open source software [2].

Being a Version Control System, Git shares the basic concepts of this software family: it manages *changes*, which are called *revisions* to any kind of digital information, where the *changes* are associated to the *time* and a *person*. Simple version control systems ranges from file naming convention to the Dropbox or Google Drive services.

Git is part of a group of more advanced softwares, where the process of tracking changes allows to precisely overview which changes have been applied, when, where and by whom. A Git tracked project can be run off line, for example for a own project, or on line where collaborator can contribute asynchronously.

2 The Git workflow

There is no standardized process on how to interact with Git, and this is probably the main reason this system is extremely flexible. Anyway, a workflow necessary. It is commonly created and shared with the working team in case of collaborative projects, or simply self-decided for own projects.

A common workflow, which is similar to older concurrent versioning systems, is the *Centralized Workflow*, where all the updates are committed to a central codebase.

Another workflow is the *Feature Branch Workflow* where different features are developed individually in different branches, leaving the master branch untouched. This is extremely useful for continuous integration, as the master branch contains a codebase which is never broken.

3 Git in planetary science

GitHub is the online platform which allows to create Git repositories online. Several institutions involved in planetary research have chosen GitHub as versioning platform for some projects. ESA and NASA have their own github accounts and numerous projects are developed on this platform. The OpenPlanetary initiative[3] is using GitHub for the document tracking and collaborative developments.

The main field of application of GitHub Among the current GitHub repositories devoted to planetary science are software and document development and numerical experiments setup.

4. Conclusions

Git represents an efficient way to track versions of a wide range of digital information. The absence of a standardized approach means that a workflow must be developed form time to time in function of the specific idea or project we want to keep under version control. This hides the most appreciated characteristic of Git: the extreme flexibility of the system which can be used both for small own, offline projects, or for very large collaborative projects where online platform like GitHub enable different subjects to asynchronously work at the same project.

Building experience and use cases on the use of Git will facilitate collaboration among researchers which is important in science in general and critical in planetary science.

References

[1] Perkel, J.: Democratic Databases: Science on GitHub. Nature News 538 (7623): 127. 2016

[2] Stallman, R.: The GNU Manifesto. In Computers, Ethics, & Society, Oxford University Press, Inc. 1990

[3] Manaud, N. et al: OpenPlanetary: An Open Science Community and Framework for Planetary Scientists and Developers, this meeting, EPSC2018-89.