

Git

Stefan Pinkernell

27.11.2019

Overview

- ◇ Why version control?
- ◇ The Git ecosystem
- ◇ Git concepts
- ◇ Git 101
- ◇ Using Git for Jupyter notebooks

- ◇ Demo

Why version control?

- ◇ Log your work
- ◇ Switch back to old versions
- ◇ Manage parallel versions (development branches)
- ◇ Share your code on different computers
- ◇ Share your code with other people
- ◇ Reduce chaos

The Git ecosystem

GUI

- git-cola
- git-gui
- GitAhead
- gitk
- Guitar
- Magit
- QGit
- ...



git

IDE Support

Eclipse, PyCharm, Spyder,...



GitLab

Atlassian



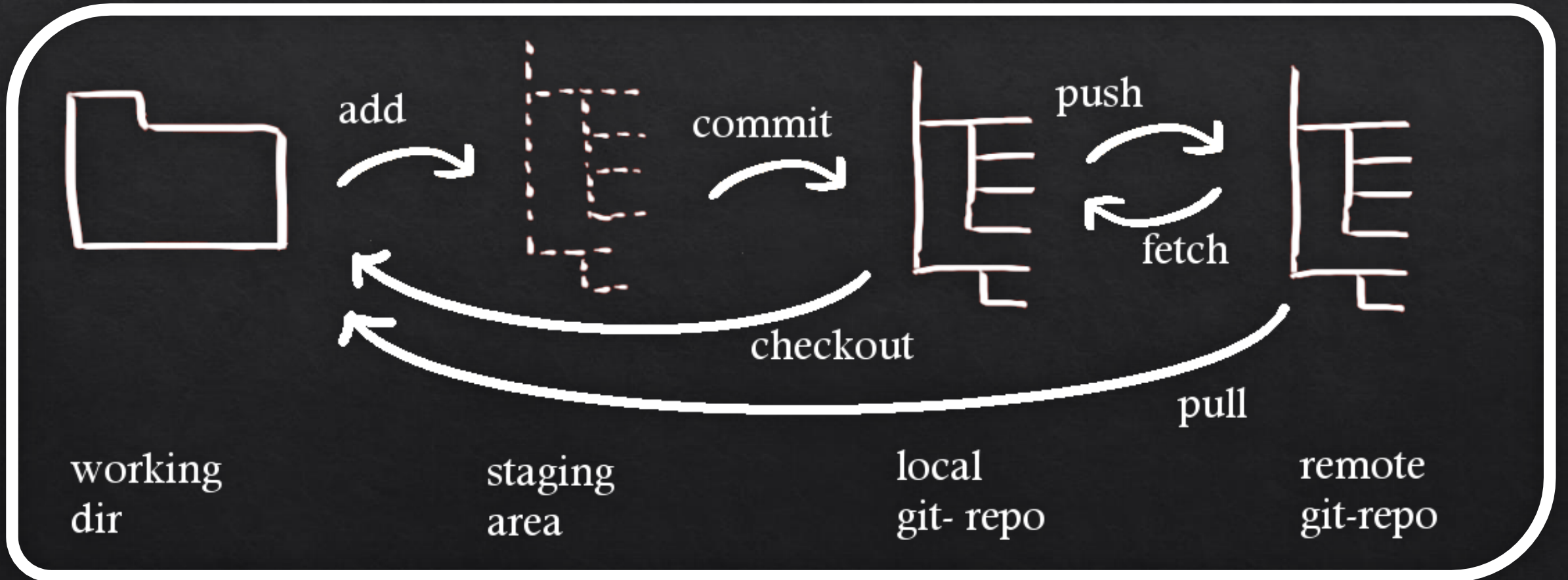
Gitea

...and many more!

Git concepts

- ◇ Distributed system
 - ◇ Every copy has the complete version history
 - ◇ Cryptographic authentication of the history
- ◇ Non-linear development
 - ◇ Branching system
 - ◇ Merging system
- ◇ Use of well established protocols: HTTP, FTP, SSH

Git - Workflow



Git – Installation & basic setup

- ◇ Git is available for Mac OS X, Windows, and Linux/Unix:

→ <https://git-scm.com>

- ◇ Git is preinstalled on AWI linux servers
- ◇ Git is used by a command line interface

- ◇ Basic setup

```
git config --global user.email stefan.pinkernell@awi.de
```

```
git config --global user.name „Stefan Pinkernell“
```

Git – Init (create a new repository)

Navigate to your project folder and create a local, empty git repository.

This will add a .git folder.

```
cd ~/myGitProject
```

```
git init
```


Git – Clone (existing repository)

Copy a complete repository from a server to the target system.

```
git clone https://github.com/o2a-data/o2a-data-pangaea.git
```



o2a-data / o2a-data-pangaea

Watch 0

Star 0

Fork 1

Code

Issues 0

Pull requests 0

Projects 0

Security

Insights

Join GitHub today

Dismiss

GitHub is home to over 40 million developers working together to host and review code, manage projects, and build software together.

Sign up

Provides examples on how to download, perform basic analysis and visualize datasets and collections archived in PANGAEA.de using Jupyter Notebook <https://www.awi.de/data-flow>

2 commits

1 branch

0 packages

0 releases

1 contributor

Apache-2.0

Branch: master

New pull request

Find file

Clone or download

StefanPinkernell added content

data

added content

.gitignore

added content

LICENSE

Initial commit

README.md

added content

Clone with HTTPS

Use Git or checkout with SVN using the web URL.

https://github.com/o2a-data/o2a-data-p

Open in Desktop

Download ZIP

Git – Add & commit

- ◇ Suggest changes (=adding to staging area)

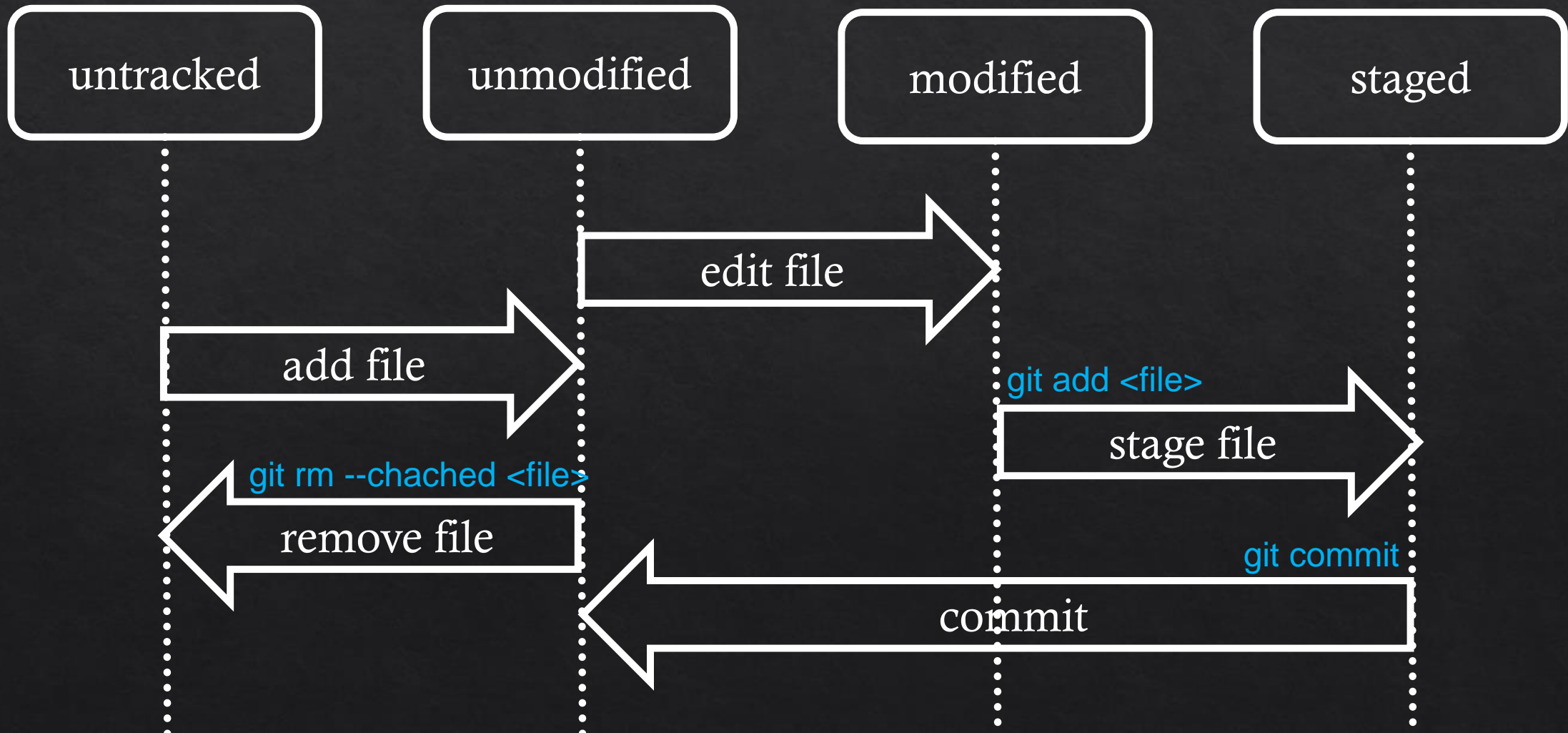
```
git add file1.py file2.py
```

```
git add .
```

- ◇ Confirm changes (=adding to local repository)

```
git commit -m „my commit message“
```

File status



Git – Publish the repository

- ◆ Publish local branch (master) to remote repository (origin)

```
git push origin master
```

- ◆ Connect a remote repository

```
git remote add <repoName> <repoServer>
```

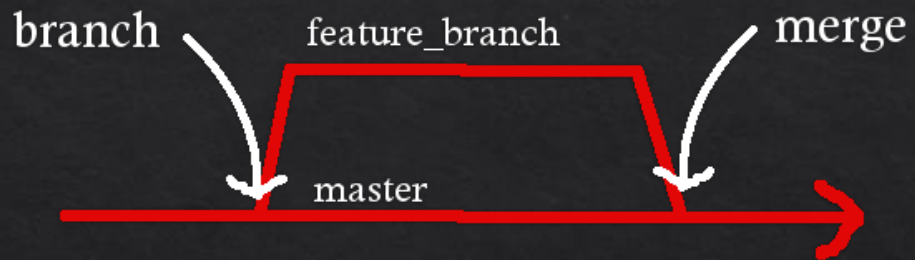
- ◆ Show info about repositories

```
git remote -v
```

```
origin https://github.com/o2a-data/o2a-data-pangaea.git (fetch)
```

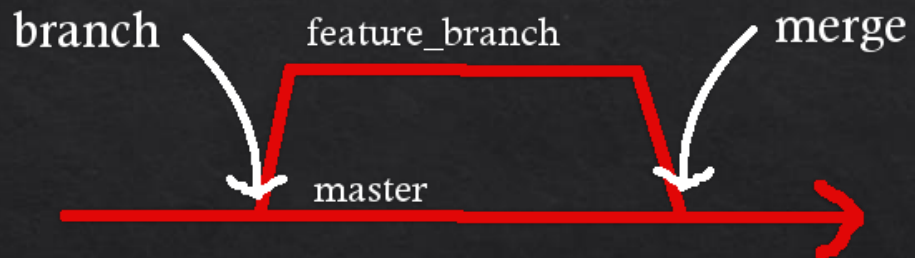
```
origin https://github.com/o2a-data/o2a-data-pangaea.git (push)
```

Git - Branches



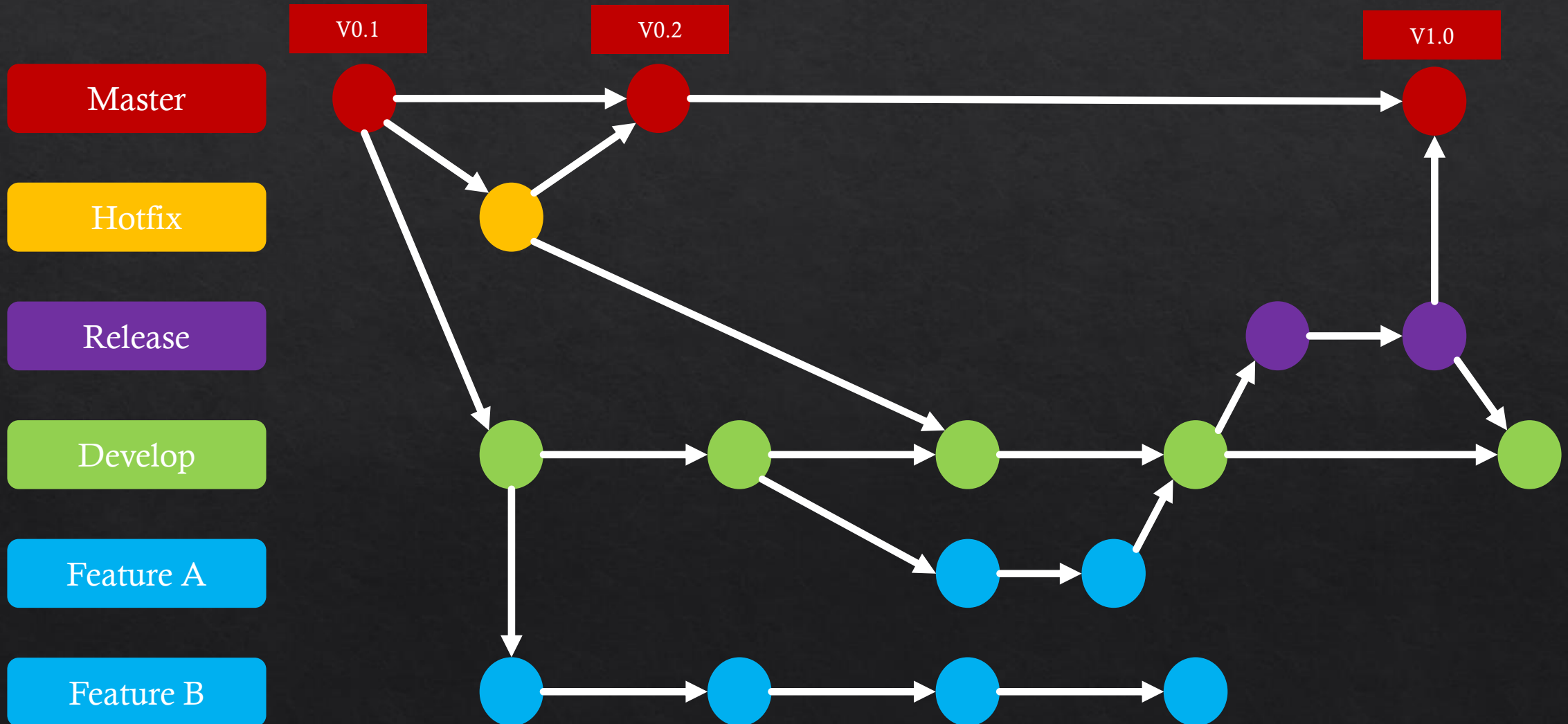
- ◇ Create feature_branch and change to it
`git checkout -b feature_branch`
- ◇ Switch between branches
`git checkout master`
`git checkout feature_brach`
- ◇ Delete the feature branch
`git branch -d feature_branch`
- ◇ Push feature_branch to remote repository
`git push origin feature_branch`

Git – Update & merge



- ◇ Switch to master branch
`git checkout master`
- ◇ Update the local repository
`git pull`
- ◇ Merge feature_branch into master
`git merge feature_branch`
- ◇ Resolve conflicts
`git add conflict_file`
`git commit -m „merged feature_branch“`

Gitflow



Git – Undoing things

- ◇ Replace file in working directory by version of the local repository

`git checkout -- filename`

- ◇ Remove all local changes and get status of remote repository

`git fetch origin`

`git reset --hard origin/master`

- ◇ ...

Infos about the git repository

git status

On branch master

Your branch is up to date with 'origin/master'.

nothing to commit, working tree clean

git help

git help branch

git help ...

git log

commit 9c6874ae8041ab54eaed6518097a0e8a364af976

Author: Giorgio Busatto <giorgio.busatto@awi.de>

Date: Thu Oct 24 15:22:43 2019 +0200

Added first preprocessing step for deep learning.

commit 3f8a4d423e06d35b2ec5f5f617e8b726ae68ee55

Author: Giorgio Busatto <giorgio.busatto@awi.de>

Date: Thu Oct 24 15:14:19 2019 +0200

Corrected image size in notebook. Converted to script.

commit 31e03469aadfd0e6f0c7365050bb66af5ff58228

Author: Stefan Pinkernell <stefan.pinkernell@awi.de>

Date: Wed Oct 23 14:05:41 2019 +0200

added classification example and data for whale acoustics

commit bd5f9ff28b398ad42d9e849589d2c4daa4395c48

Merge: 0492ed4 2dfd313

Author: Stefan Pinkernell <stefan.pinkernell@awi.de>

Git & Jupyter notebooks

- ◇ Jupyter Notebooks: document format based on Json files
- ◇ Changes in a document might result in massive changes in the Json file
- ◇ Git stores changes since last commit → can be a lot of overhead

Consider to reset the notebook before a commit

(Kernel → Restart and Clear Output)