

Git

An Introduction to the Version Control System

Oliver Henrich

Email: oliver.henrich@strath.ac.uk

Strathclyde Research Software Engineers

**The University of Strathclyde:
the place of useful learning**

Table of Contents

- 1. Introduction to Git
 - 1.1 Automated Version Control
 - 1.2 Setting Up Git
 - 1.3 Creating A Git Repository
 - 1.4 Tracking Changes
 - 1.5 Exploring History
 - 1.6 Creating Remote Repositories on GitHub
 - 1.7 Collaborating On GitHub
 - 1.8 Conflicts

1. Introduction to Git

In this section

- ▶ Automated Version Control
- ▶ Setting Up Git
- ▶ Creating A Git Repository
- ▶ Tracking Changes
- ▶ Exploring History
- ▶ Creating Remote Repositories on GitHub
- ▶ Collaborating On GitHub
- ▶ Conflicts

1.1. Automated Version Control

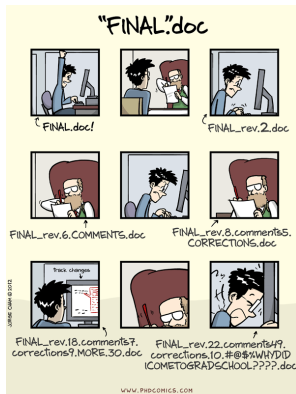
Why version control?

Problems

- ▶ Multiple, nearly identical versions of the same document
- ▶ Tracking changes is not an option for source code
- ▶ No protection against accidental deletion

Version control systems

- ▶ **start** with a base version of the document
- ▶ **record** changes you make each step of the way
- ▶ can **revert** to any previous versions if necessary
- ▶ **never lose** a previous state of your document
- ▶ allow many people to **work in parallel**



1.1. Automated Version Control

Sequential changes



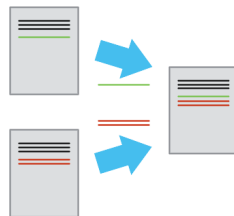
Start at the base document, apply each change, arrive at the more recent version

Diverging versions



Two users can make independent sets of changes on the same document

Merge versions



Incorporate two sets of changes into the same base document

1.2. Setting Up Git

Git comes in many different forms. We use **Git on the command line**.

- ▶ It's the only place you can run **all** Git commands.
- ▶ If you know the command line version, you can figure out how to run the GUI version.
- ▶ Everyone has the same command line tools.

On Linux

```
$ sudo dnf install git-all (Fedora)
```

```
$ sudo apt install git-all (Debian)
```

On Mac

```
$ brew install git (Homebrew)
```

```
$ port install git (Macports)
```

part of XCode IDE

On Windows

Git for Windows: <https://git-scm.com/download/win>

GitHub Desktop: <https://desktop.github.com>

Git Chocolatey: <https://chocolatey.org/packages/git>

1.2. Setting Up Git

When we use Git on a new computer for the first time, we need to configure a few things.

Here are a few examples of configurations we will set as we get started with Git:

- ▶ your name and email address
- ▶ and that we want to use these settings globally (i.e. for every project)

On a command line, Git commands are written as **git verb options**, where **verb** is what we actually want to do and **options** is additional optional information which may be needed for the verb.

So here is how Dracula sets up his new laptop:

```
$ git config --global user.name "Vlad Dracula"  
$ git config --global user.email  
    "vlad@tran.sylvan.ia"
```

1.2. Setting Up Git

Quite helpful and important are the commands to look up the manual

For a general overview of a range of Git commands

```
$ git --help
```

For a overview of specific Git command (here command = verb)

```
$ git verb -h
```

For an in-depth manual of specific Git command

```
$ git verb --help
```

For more information please check the **Git online documentation** at

<https://git-scm.com/docs>

1.3. Creating A Git Repository

Create a directory in your current working directory

```
$ mkdir git_exercise
```

Change to the new directory

```
$ cd git_exercise
```

Create the Git repository

```
$ git init
```

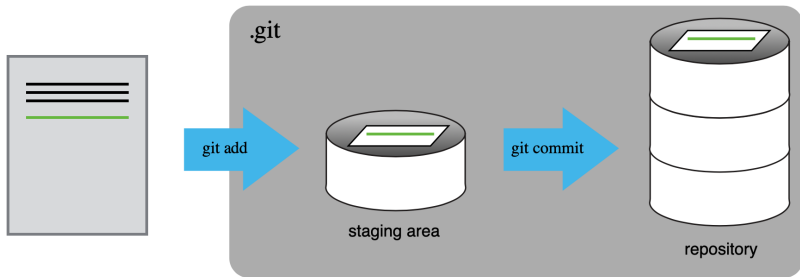
Note

An invisible file `.git` is created that stores all the history and dependencies of the repository.

1.4. Tracking Changes

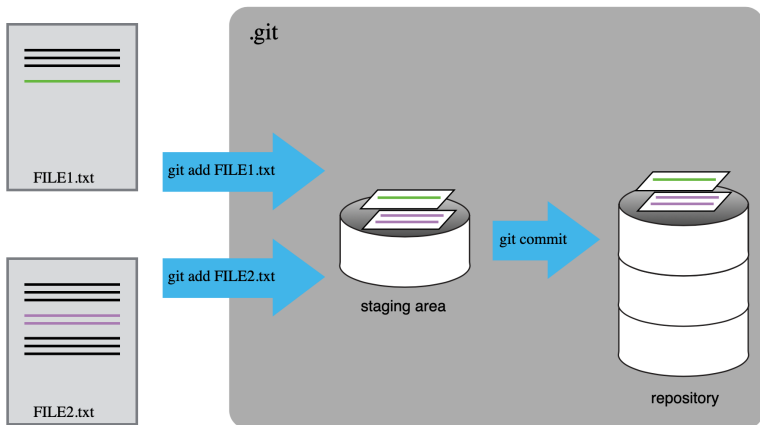
Think of Git as taking snapshots of your project in a two-step process:

- ▶ `git add` specifies **what** will go into a snapshot
- ▶ `git commit` **takes the actual snapshot** and records it permanently



1.4. Tracking Changes

It is of course possible to add multiple files (or changes thereof) before committing these, i.e. taking the snapshot.



1.4. Tracking Changes

Check the status

```
$ git status
```

Output if working directory / repository is brand new

```
On branch master
No commits yet
nothing to commit (create/copy files and use "git
    add" to track)
```

Now create a new file `hello_world.py` in your working directory that prints 'Hello World!'

1.4. Tracking Changes

When you check the status again, you will now see there is an untracked file.

```
$ git status
```

Output if working directory differs, but new file is not yet being tracked

```
On branch master
```

```
No commits yet
```

```
Untracked files:
```

```
(use "git add <file>..." to include in what will  
be committed)  
hello_world.py
```

```
nothing added to commit but untracked files present  
(use "git add" to track)
```

1.4. Tracking Changes

Add your new file and start the tracking.

```
$ git add hello_world.py
```

Check the status again

```
$ git status
```

Output if working directory differs and file is being tracked

```
On branch master
```

```
No commits yet
```

```
Changes to be committed:
```

```
(use "git rm --cached <file>..." to unstage)
```

```
new file: hello_world.py
```

1.4. Tracking Changes

Now commit the change and record it permanently.

```
$ git commit -m 'Added new file'
```

Output of `git commit`

```
[master (root-commit) b3557d4] Added new file
1 file changed, 8 insertions(+)
create mode 100644 hello_world.py
```

1.4. Tracking Changes

Check the status with `git status` and you get the following message:

```
On branch master
nothing to commit, working tree clean
```

You can check the log to see your commits. The most recent appears first.

```
$ git log
```

```
commit 249156049502e47d839735c34e31830885bc5092
(HEAD -> master)
Author: Oliver Henrich
<ohenrich@users.noreply.github.com>
Date: Wed Sep 2 16:56:07 2020 +0100
    Added new file
```


1.5. Exploring History

When working with repos, you often want to **review changes before committing them** or **revert to a previous version** of the file.

Add an additional line to the previous `hello_world.py` file. Check the **differences between your local and remote repository** with

```
$ git diff
```

Additional line `'print('Hello Scotland!')'` in local file (+)

```
diff --git a/hello_world.py b/hello_world.py
index 73fb7c3..e6f9107 100644
--- a/hello_world.py
+++ b/hello_world.py
@@ -1,2 @@
 print('Hello World!')
+print('Hello Scotland!')
```

1.5. Exploring History

Commit the change

```
$ git add hello_world.py  
$ git commit -m 'Added additional line'
```

and check the differences again.

```
$ git diff
```

There are no differences anymore as you committed your change.

Now check the status again.

```
$ git status
```

Output of `git status`

```
On branch master  
nothing to commit, working tree clean
```

1.5. Exploring History

Add another line, commit the change again and check your commit log.

Output of `git log`

```
commit 908944eb711c90f5bd46297639b34d8fc70993f0
(HEAD -> master)
Author: Oliver Henrich
<ohenrich@users.noreply.github.com>
Date: Wed Sep 2 16:59:00 2020 +0100
    Added another additional line

commit 28f46c36b5729ab26ca719cc1468b1a6e734d597
Author: Oliver Henrich
<ohenrich@users.noreply.github.com>
Date: Wed Sep 2 16:58:15 2020 +0100
    Added additional line

commit 249156049502e47d839735c34e31830885bc5092
Author: Oliver Henrich
<ohenrich@users.noreply.github.com>
Date: Wed Sep 2 16:56:07 2020 +0100
    Added new file
```

All commits have a unique ID, but Git knows a simple way to address them:

The **last commit** appears at the top and is marked with **HEAD**.

The **two previous commits** are not marked, but can be conveniently addressed with **HEAD~1** and **HEAD~2**.

1.5. Exploring History

If we want to see what the differences are between the current version (**HEAD**) and version two commits ago, we can issue for instance

```
$ git diff HEAD~2
```

Additional two lines marked as different in local file (+)

```
diff --git a/hello_world.py b/hello_world.py
index 73fb7c3..547a19b 100644
--- a/hello_world.py
+++ b/hello_world.py
@@ -1,3 @@
 print('Hello World!')
+print('Hello Scotland!')
+print('Hello Glasgow!')
```

1.5. Exploring History

Assume you want to obtain the previous version without the additional line.

First you need to **check the log for the ID of the previous commit**.

Output of `git log`

```
commit 28f46c36b5729ab26ca719cc1468b1a6e734d597
Author: Oliver Henrich
       <ohenrich@users.noreply.github.com>
Date: Wed Sep 2 16:58:15 2020 +0100
    Added additional line
```

The commit ID is the long bit starting 28f46c36b5. . .

It is **usually sufficient to specify only 7 digits**.

Use the `git checkout` command to retrieve a previous version.

```
$ git checkout 28f46c36b5 hello_world.py
```

Note: Do not forget the filename at the end as this will 'detach the HEAD'!

To retrieve the latest version again use

```
$ git checkout master hello_world.py
```

1.5. Exploring History

The previous command will not revert the commit (check e.g. `git log`).

To **revert an erroneous commit**, first look for its **ID** and use `git revert`.

Output of `git log`

```
commit 908944eb711c90f5bd46297639b34d8fc70993f0
(HEAD -> master)
Author: Oliver Henrich
<ohenrich@users.noreply.github.com>
Date: Wed Sep 2 16:59:00 2020 +0100
    Added another additional line
```

```
commit 28f46c36b5729ab26ca719cc1468b1a6e734d597
Author: Oliver Henrich
<ohenrich@users.noreply.github.com>
Date: Wed Sep 2 16:58:15 2020 +0100
    Added additional line
```

We want to revert the commit
starting 908944eb71...

We want the current version to
be this one.

```
$ git revert 908944eb71
```

This creates a new commit with the previous version of the file.

1.5. Exploring History

Revert "Added another additional line"

This reverts commit 908944eb711c90f5bd46297639b34d8fc70993f0.

```
# Please enter the commit message for your changes. Lines starting
# with '#' will be ignored, and an empty message aborts the commit.
#
# On branch master
# Changes to be committed:
#   modified:   hello_world.py
#
```

A dialogue window opens that asks you for a message providing a template.

```
commit 15f36c3bd31f594504756326df6b3baeb2d0982c (HEAD -> master)
Author: Oliver Henrich <ohenrich@users.noreply.github.com>
Date: Thu Sep 3 17:29:03 2020 +0100
    Revert "Added another additional line"
    This reverts commit 908944eb711c90f5bd46297639b34d8fc70993f0.
```

```
commit 908944eb711c90f5bd46297639b34d8fc70993f0
Author: Oliver Henrich <ohenrich@users.noreply.github.com>
Date: Wed Sep 2 16:59:00 2020 +0100
    Added another additional line
```

```
commit 28f46c36b5729ab26ca719cc1468b1a6e734d597
Author: Oliver Henrich <ohenrich@users.noreply.github.com>
Date: Wed Sep 2 16:58:15 2020 +0100
    Added additional line
```

```
commit 249156049502e47d839735c34e31830885bc5092
Author: Oliver Henrich <ohenrich@users.noreply.github.com>
Date: Wed Sep 2 16:56:07 2020 +0100
    Added new file
```

Your commit log has now an extra entry.

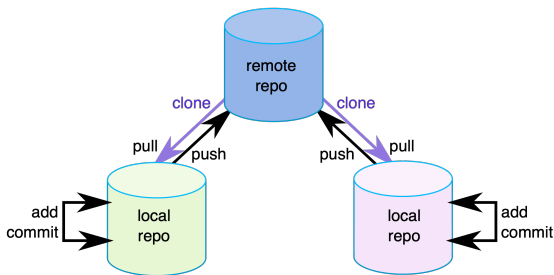
1.6. Creating Remote Repositories on GitHub

One of the main reasons for using repositories is also to **collaborate with other people** and **work on the same code**. This is done through a **remote repository**.

Pulling retrieves from the remote repo.

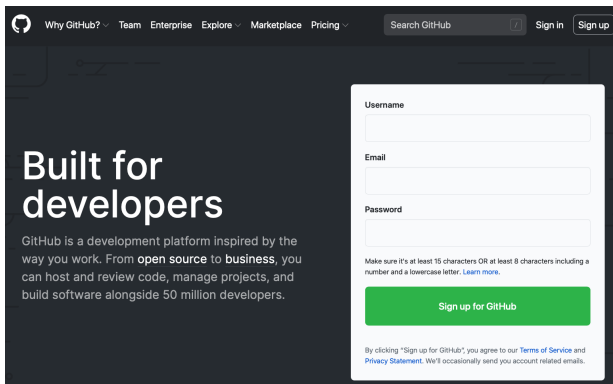
Pushing deposits into the remote repo.

Cloning checks out a **private copy** of the remote repo.



1.6. Creating Remote Repositories on GitHub

We will use **GitHub** available at <https://github.com>



The screenshot shows the GitHub homepage with a dark background. On the left, the text "Built for developers" is prominently displayed, followed by a description of GitHub as a development platform. On the right, there is a white sign-up form with fields for Username, Email, and Password. Below the Password field is a note about password requirements and a link to "Learn more". A large green button labeled "Sign up for GitHub" is at the bottom of the form. At the very bottom of the form, there is a small line of text about agreeing to the Terms of Service and Privacy Statement.

Why GitHub? Team Enterprise Explore Marketplace Pricing Search GitHub Sign in Sign up

Built for developers

GitHub is a development platform inspired by the way you work. From **open source** to **business**, you can host and review code, manage projects, and build software alongside 50 million developers.

Username

Email

Password

Make sure it's at least 15 characters OR at least 8 characters including a number and a lowercase letter. [Learn more.](#)

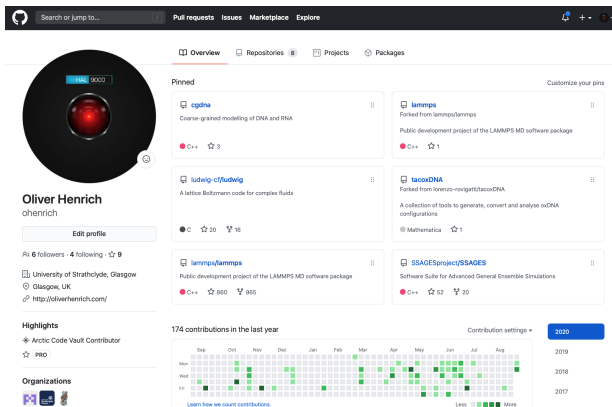
Sign up for GitHub

By clicking "Sign up for GitHub", you agree to our [Terms of Service](#) and [Privacy Statement](#). We'll occasionally send you account related emails.

Pick a username and sign up now with one of your email addresses.

1.6. Creating Remote Repositories on GitHub

Once you have an account, go to your profile page. This is in the drop-down menu under the avatar in the top right-hand corner.



The screenshot shows the GitHub profile page for Oliver Henrich. The profile includes a circular avatar with a red light effect, the name "Oliver Henrich", and the username "ohenrich". Below the name is an "Edit profile" button. The profile statistics show 6 followers, 4 following, and 9 repositories. The "University of Strathclyde, Glasgow" is listed as the organization, with the location "Glasgow, UK" and the website "http://oliverhenrich.com/".

The "Highlights" section lists "Arctic Code Vault Contributor" and "PJO". The "Organizations" section shows logos for "Arctic Code Vault" and "PJO".

The "Pinned" section displays four repositories:

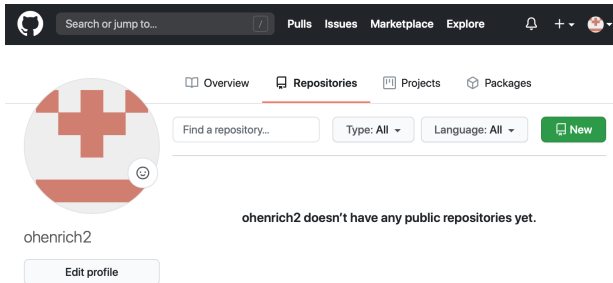
- cgdna**: Coarse-grained modelling of DNA and RNA. C++ 3 stars.
- ludwig-ct/ludwig**: A lattice Boltzmann code for complex fluids. C 20 stars, 16 forks.
- lammmps/lammmps**: Public development project of the LAMMPS MD software package. C++ 860 stars, 860 forks.
- SSAGESproject/SSAGES**: Software Suite for Advanced General Ensemble Simulations. C++ 52 stars, 20 forks.

The "174 contributions in the last year" section shows a calendar grid with green squares indicating contributions. The grid spans from September to August, with contributions visible in most months. The year 2020 is selected, and the years 2019, 2018, and 2017 are also visible. A link "Learn how we count contributions." is provided.

1.6. Creating Remote Repositories on GitHub

On GitHub, create your first repository.

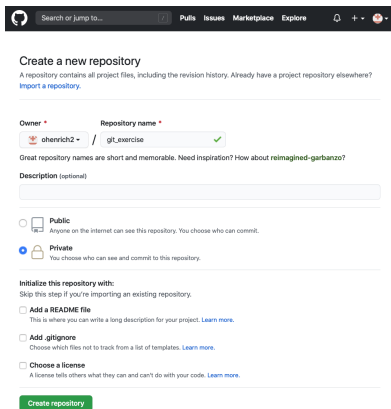
Go to **Repositories**



and click on the **New** button.

1.6. Creating Remote Repositories on GitHub



Then choose a name for your repo and create it.



Create a new repository

A repository contains all project files, including the revision history. Already have a project repository elsewhere? [Import a repository.](#)

Owner ^{*} Repository name ^{*}

 ohenrich2 ^{*} / git_exercise 

Great repository names are short and memorable. Need inspiration? How about [reimagined-garbanzo](#)?

Description (optional)

☐ Public
Anyone on the internet can see this repository. You choose who can commit.

☒ Private
You choose who can see and commit to this repository.

Initialize this repository with:
Skip this step if you're importing an existing repository.

☐ Add a README file
This is where you can write a long description for your project. [Learn more.](#)

☐ Add .gitignore
Choose which files not to track from a list of templates. [Learn more.](#)

☐ Choose a license
A license tells others what they can and can't do with your code. [Learn more.](#)

[Create repository](#)

There are a number of options:

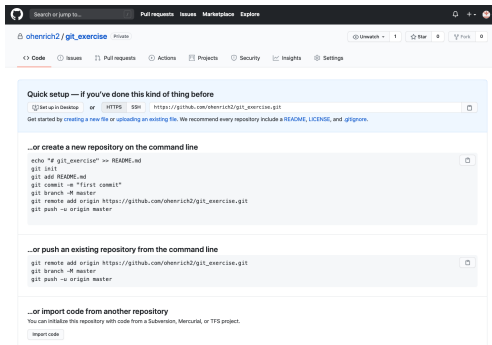
- ▶ **Import a repository:** This option is only available if you have another *remote* repo unlike the one you just created
- ▶ **Public / Private:** Choose private so only you can interact with it. You can give others access later on.
- ▶ **Initialize this repository with:** This adds certain files. Don't do this as we want to import your existing project.

Press the "Create repository" button.

1.6. Creating Remote Repositories on GitHub

The next thing you get is a useful overview of what you might want to do next. It contains already the correct code with your individual username, repo URL, etc.

We are going to do option 2:



The screenshot shows the GitHub repository page for 'ohenrich2/git_exercise'. The page is titled 'Quick setup — if you've done this kind of thing before'. It provides three options for setting up the repository:

- Setup on Desktop**: A dropdown menu showing the repository URL: `https://github.com/ohenrich2/git_exercise.git`.
- SSH**: A dropdown menu showing the repository URL: `https://github.com/ohenrich2/git_exercise.git`.
- HTTPS**: A dropdown menu showing the repository URL: `https://github.com/ohenrich2/git_exercise.git`.

Below these options, there are three sections for creating or pushing a repository from the command line:

- ...or create a new repository on the command line**: A code block showing the following commands:

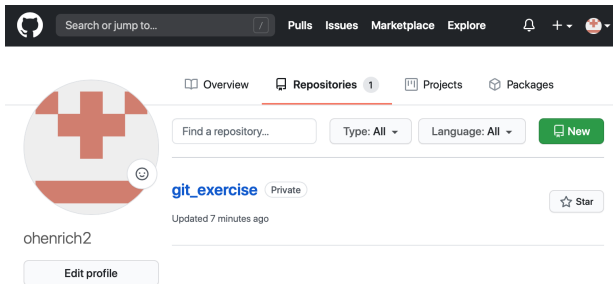
```
echo "# git_exercise" >> README.md
git init
git add README.md
git commit -m "first commit"
git branch -M master
git remote add origin https://github.com/ohenrich2/git_exercise.git
git push -u origin master
```
- ...or push an existing repository from the command line**: A code block showing the following commands:

```
git remote add origin https://github.com/ohenrich2/git_exercise.git
git branch -M master
git push -u origin master
```
- ...or import code from another repository**: A section with a link to 'Import code'.

**"push an existing repository
from the command line"**

1.6. Creating Remote Repositories on GitHub

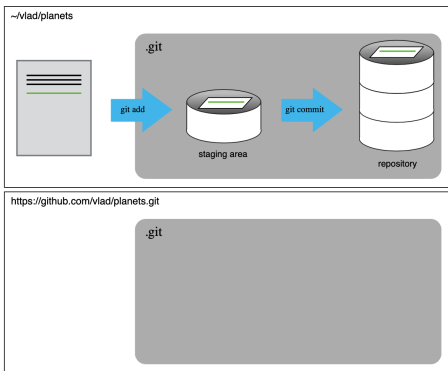
But let's first go back to your profile page and check your "Repositories".



When you browse the content of your repo you will see it is completely empty.

1.6. Creating Remote Repositories on GitHub

Your current situation looks like this



We will now **export your existing local** repo to the **newly created remote** repo on GitHub.

1.6. Creating Remote Repositories on GitHub

The full documentation is available on <https://docs.github.com/en>, under **GitHub.com** → **Importing your projects** → **Adding an existing project to GitHub using the command line**.

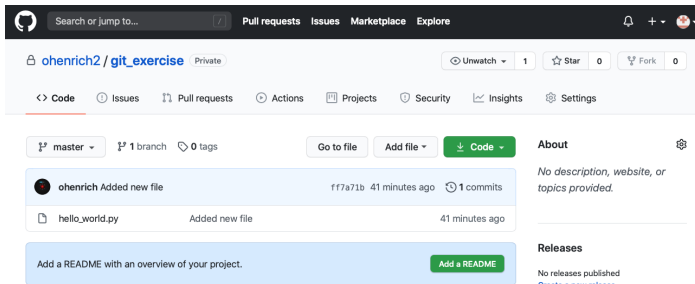
On GitHub, look up the URL of your remote repo.

On the command line change to the directory of your local repo and issue the following sequence of commands replacing username, etc accordingly.

```
$ cd git_exercise
$ git remote add origin
  https://github.com/username/repo_name.git
$ git branch -M master
$ git push -u origin master
```


1.6. Creating Remote Repositories on GitHub

On GitHub check what's in your remote repository.



The screenshot shows the GitHub interface for a repository named 'git_exercise' by user 'ohenrich2'. The repository is private. The main content area shows a commit by 'ohenrich' titled 'Added new file' with commit hash 'ff7a71b' and '1 commit' made '41 minutes ago'. Below this, a file named 'hello_world.py' is listed as 'Added new file' '41 minutes ago'. There are buttons for 'Go to file', 'Add file', and 'Code'. On the right, there are sections for 'About' (no description), 'Releases' (no releases published), 'Packages' (no packages published), and 'Languages' (Python 100.0%).

Hooray!

Here is your local code making a first appearance on GitHub!

1.7. Collaborating On GitHub

In [Section 2.6](#) we **exported** an existing *local repository* to your GitHub profile to **create a remote repository**.

The inverse process of **importing** an existing *remote repository* from GitHub is called **cloning**.

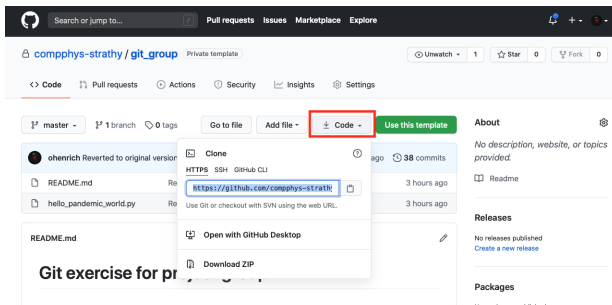
Cloning produces a **local copy of the remote repository** on your machine. It requires

- ▶ the **URL of the remote repository**
- ▶ the `git clone` command

You can modify the local copy and **push the changes to the remote repository** on GitHub to share them with your collaborators.

1.7. Collaborating On GitHub

First navigate to the repository that you want to clone and click on the "Code" button



Copy the URL, e.g. by pressing the button next to it.

1.7. Collaborating On GitHub

On your command line in your working directory issue the following command replacing **URL** with the actual URL:

```
$ git clone URL
```

Output of `git clone URL`

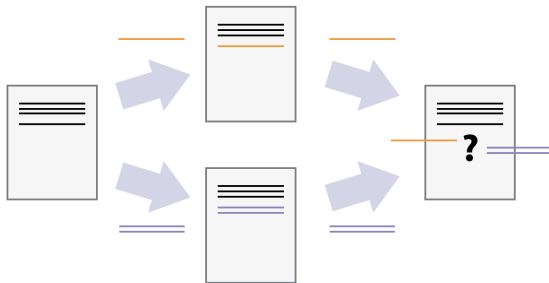
```
Cloning into 'git_group'...
remote: Enumerating objects: 109, done.
remote: Counting objects: 100% (109/109), done.
remote: Compressing objects: 100% (79/79), done.
remote: Total 109 (delta 33), reused 96 (delta 27), pack-reused 0
Receiving objects: 100% (109/109), 14.32 KiB | 4.77 MiB/s, done.
Resolving deltas: 100% (33/33), done.
```

You can clone the repository it into a different name than the default name (the name on GitHub), e.g. **blablabla** by adding this after the URL.

```
$ git clone URL blablabla
```

1.8. Conflicts

Conflicts emerge when **several** people work on the **same** code and **update** the **remote repo**.



The **orange line** and the **purple line** are approximately at the same position in the file.

1.8. Conflicts

Let's look at our Hello World! example.

Your collaborator has checked in the following version:

```
print('Hello World!')
print('Hello Scotland!')
print('Hello City of
      Glasgow!')
```

Your version differs in the last line:

```
print('Hello World!')
print('Hello Scotland!')
print('Hello Greater
      Glasgow!')
```

Output of `git push origin master`

```
To https://github.com/compphys-strathclyde/git_exercise.git
! [rejected] master -> master (fetch first)
error: failed to push some refs to 'https://github.com/compphys-strathclyde/git_exercise.git'
hint: Updates were rejected because the remote contains work that you do
hint: not have locally. This is usually caused by another repository pushing
hint: to the same ref. You may want to first integrate the remote changes
hint: (e.g., 'git pull ...') before pushing again.
hint: See the 'Note about fast-forwards' in 'git push --help' for details.
```

1.8. Conflicts

What we have to do is **pull the changes from the remote repo** on GitHub into your local repo.

Git **tries to merge them automatically** into your local copy, and **if successful this can be pushed** to the remote repo on GitHub.

```
$ git pull origin master
```

Output of `git pull origin master`

```
remote: Enumerating objects: 5, done.
remote: Counting objects: 100% (5/5), done.
remote: Compressing objects: 100% (2/2), done.
remote: Total 3 (delta 0), reused 0 (delta 0), pack-reused 0
Unpacking objects: 100% (3/3), 682 bytes | 682.00 KiB/s, done.
From https://github.com/compphys-strathclyde/git_exercise
 * branch      master    -> FETCH_HEAD
   dbd3016..5b9c121 master -> origin/master
Auto-merging hello_world.py
CONFLICT (content): Merge conflict in hello_world.py
Automatic merge failed; fix conflicts and then commit the result.
```

1.8. Conflicts

Check what Git has done to your local file.

```
print('Hello World!')
print('Hello Scotland!')
<<<<<<< HEAD
print('Hello Greater Glasgow!')
=====
print('Hello City of Glasgow!')
>>>>>>> 5b9c121bac
```

Our change is preceded by
<<<<<<< HEAD.

Git inserted **=====** as
separator between the
conflicting changes.

The end of the content
downloaded from GitHub is
marked with **>>>>>>>.**

We need to **remove** these markers, **reconcile** the changes and **check in a new version.**

1.8. Conflicts

```
print('Hello World!')  
print('Hello Scotland!')  
print('Hello Greater Glasgow!')  
print('Hello City of Glasgow!')
```

We remove the markers and keep both lines.

Lets' first check the status.

Output of `git status` after editing

```
git status  
On branch master  
Your branch and 'origin/master' have diverged,  
and have 2 and 1 different commits each, respectively.  
(use "git pull" to merge the remote branch into yours)  
  
You have unmerged paths.  
(fix conflicts and run "git commit")  
(use "git merge --abort" to abort the merge)  
  
Unmerged paths:  
(use "git add <file>..." to mark resolution)  
  both modified: hello_world.py
```

We are using the last option.

1.8. Conflicts

```
$ git add hello_world.py  
$ git commit -m 'Resolved conflict'
```

Output of `git commit`

```
[master c7c8fb8] Resolved conflict
```

```
$ git push origin master
```

Minimise the number of conflicts by using this workflow:

1. Update local `git pull origin master`
2. Make changes
3. Stage changes `git add your_edited_file.py`
4. Commit changes `git commit -m "Your commit message"`
5. Update remote `git push origin master`

1.8. Conflicts

Git is extremely powerful, but also very complex.

Excellent **Git tutorial materials** are available on the Software Carpentry website:

<http://swcarpentry.github.io/git-novice>



software carpentry

Teaching basic lab skills
for research computing