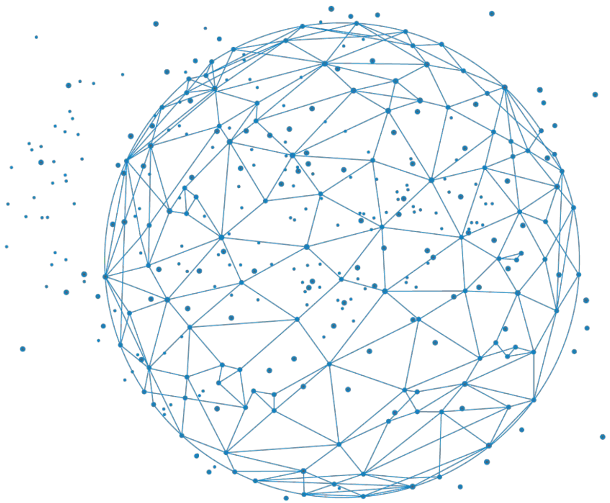
2025 June 5th



Introduction to Git

ROUSSEAU Baptiste
SEILER Julien

IFB INSTITUT FRANÇAIS DE BIOINFORMATIQUE

elixir FRANCE

cnrs INRAE Inserm cea

Plan

- What is Git?
- Key concepts
- How to use the Git command?
- The Burger project
  - Create a repository
  - Add files
  - Commit changes
  - Modify files
  - Check status
  - Revert changes
  - Use branches

The **French Institute of Bioinformatics**, or **IFB**, is the national bioinformatics infrastructure that provides support, deploys services, organizes training, and carries out innovative developments for the life sciences communities.

The IFB offers **a national-scale computing infrastructure** in addition to the services provided regionally, known as the Core Cluster.

The Core Cluster is a computing infrastructure co-managed by system administrators and bioinformaticians from 5 IFB platforms.


4640 cores (9280 threads)


62TB of RAM

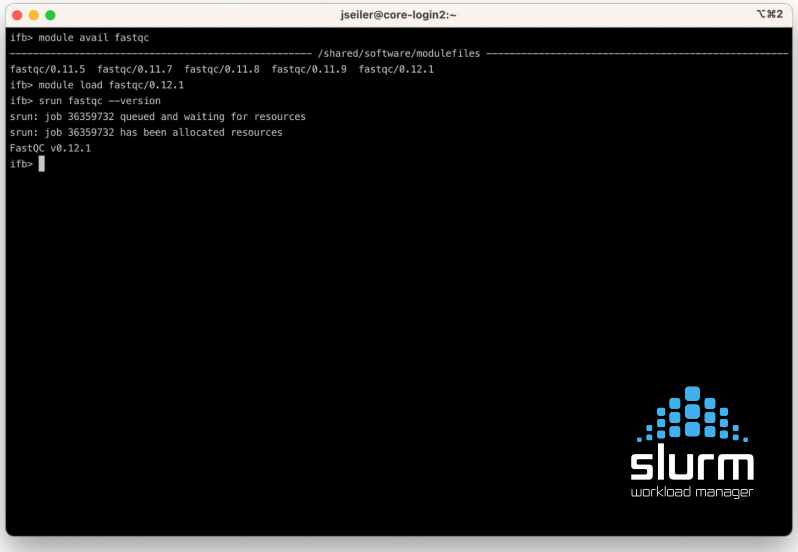

42 GPUs


2PB of scratch storage

2 PB of project storage (with backup)


more than 600 tools

The Core Cluster offers **three modes** of access to its resources.

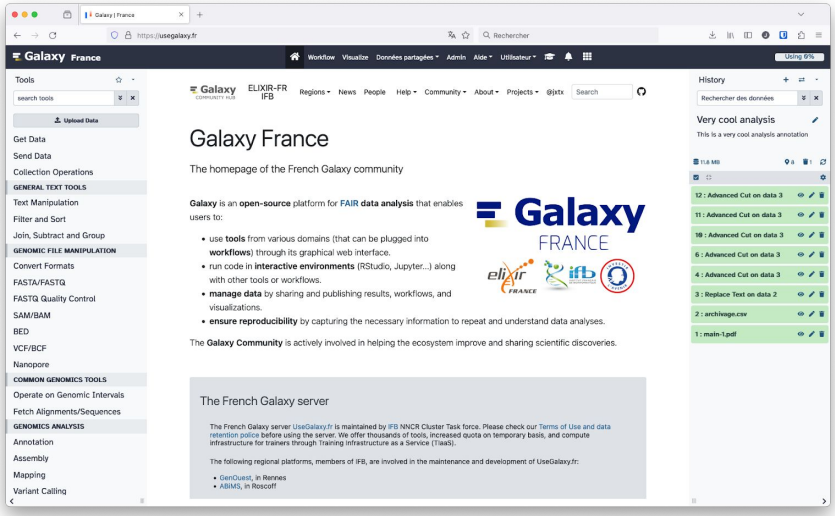


SSH

core.cluster.france-bioinformatique.fr

A connection node for launching SLURM tasks

The Core Cluster offers **three modes** of access to its resources.



Galaxy

https://usegalaxy.fr

French instance of Galaxy

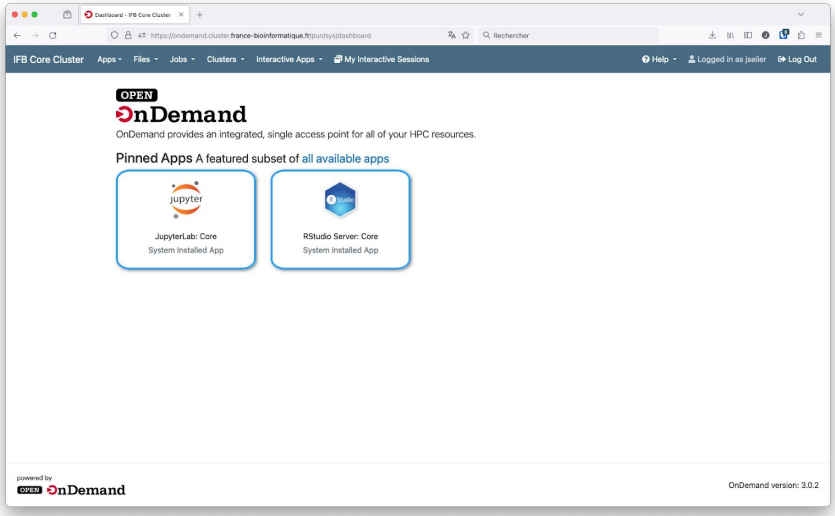
3,174 tools available

6 thematic subdomains:

- Workflow4Metabolomics
- ProteoRE
- Covid19
- Metabarcoding
- Met4J
- MNHN

TiaaS : Training infrastructure as a Service

The Core Cluster offers **three modes** of access to its resources.



Open OnDemand

https://ondemand.cluster.france-bioinformatique.fr/

Web portal for accessing computing resources

Interactive applications:

- JupyterLab
- RStudio
- XFCE Desktop
- And more to come

File manager

SLURM task launching and monitoring

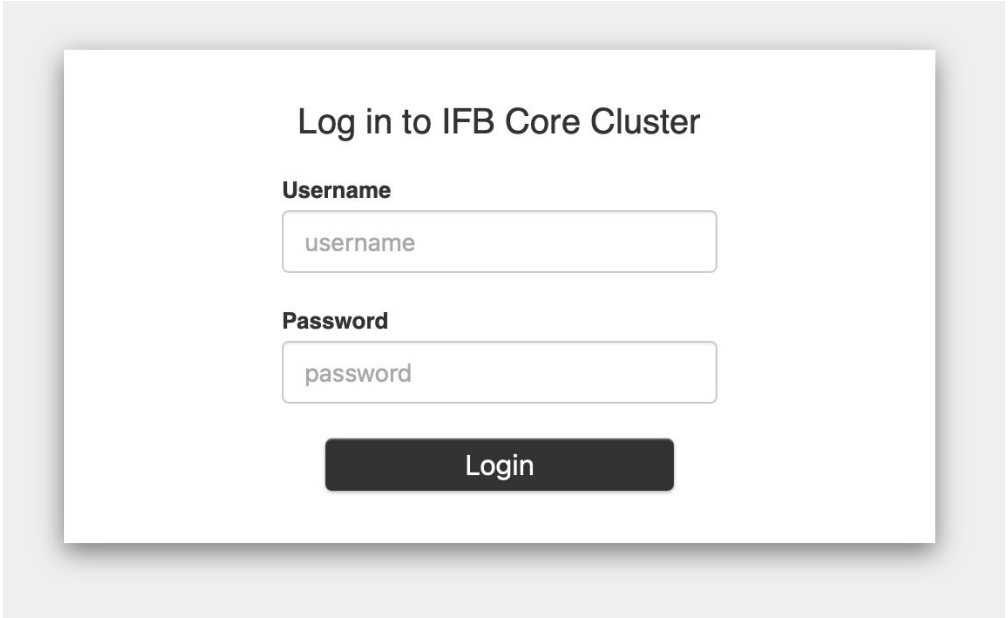
This is **an interactive training**

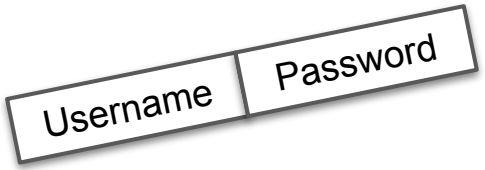
- You will learn to use **`git`** by working with it
- You only need a web browser and an Internet connection to follow this training.
- Each time you see this icon, you will have something to do

https://ondemand.cluster.france-bioinformatique.fr/

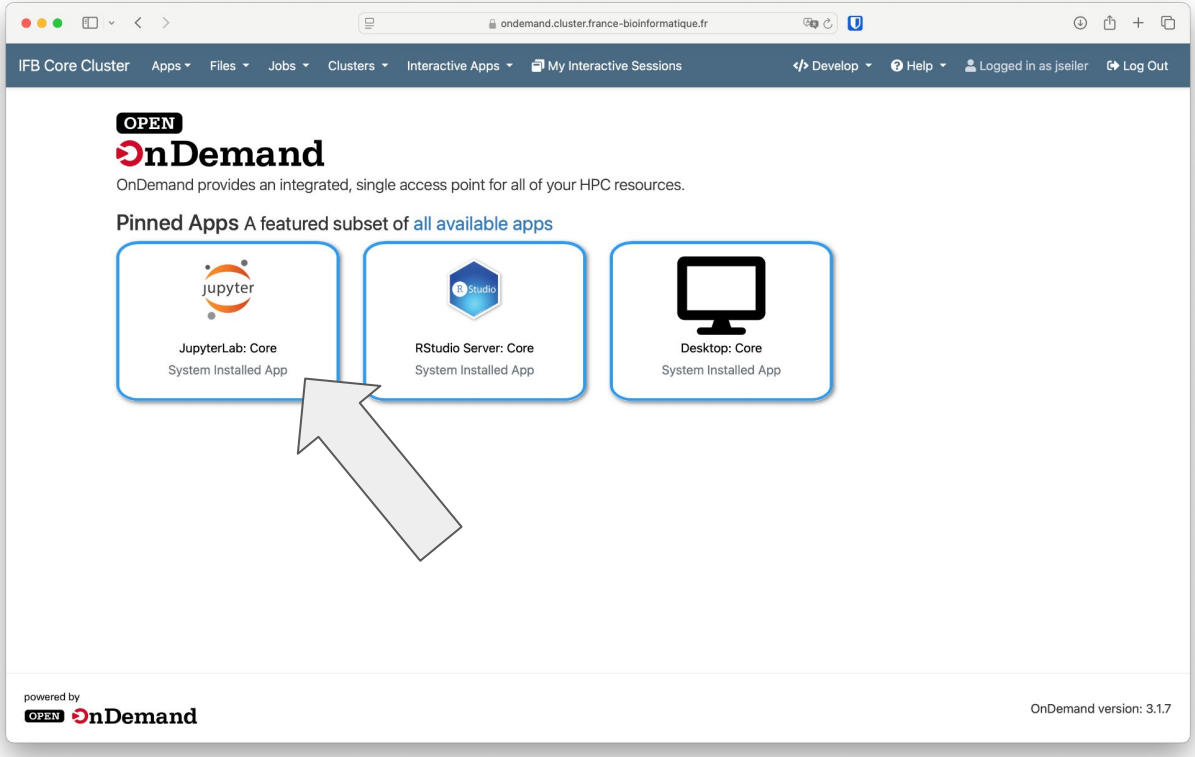
CC BY NC SA

This is **an interactive training**

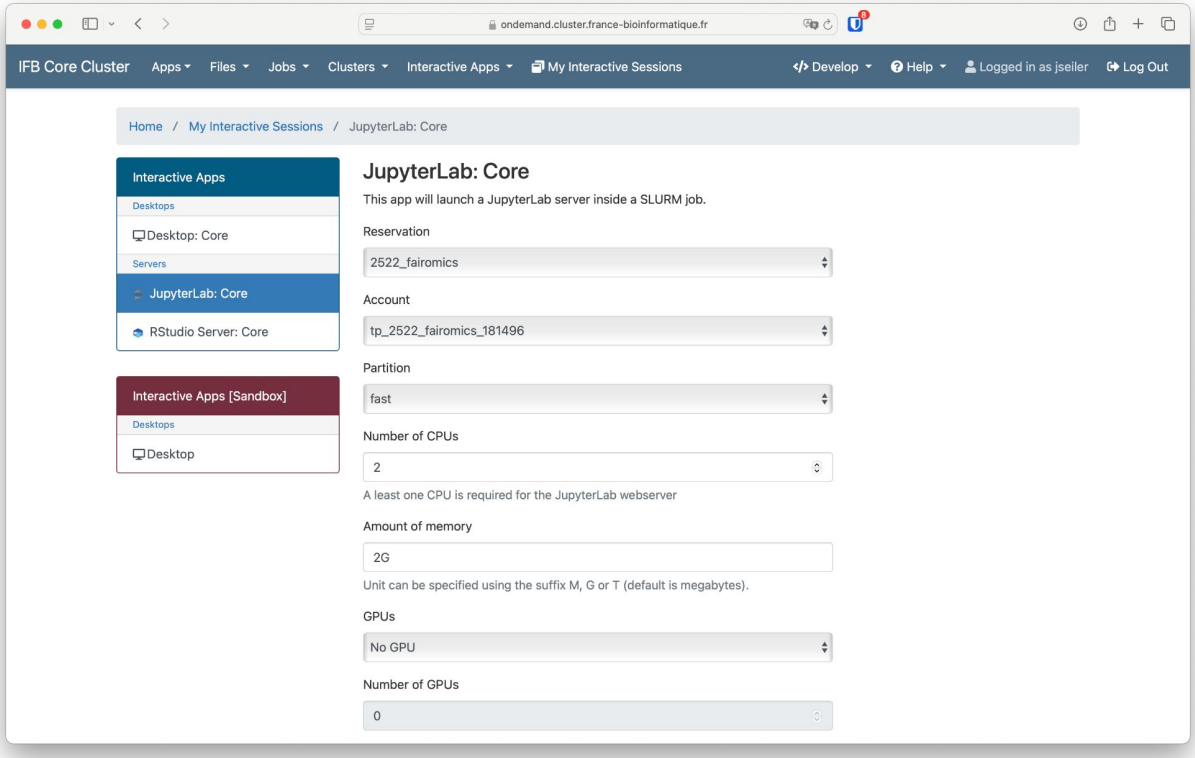




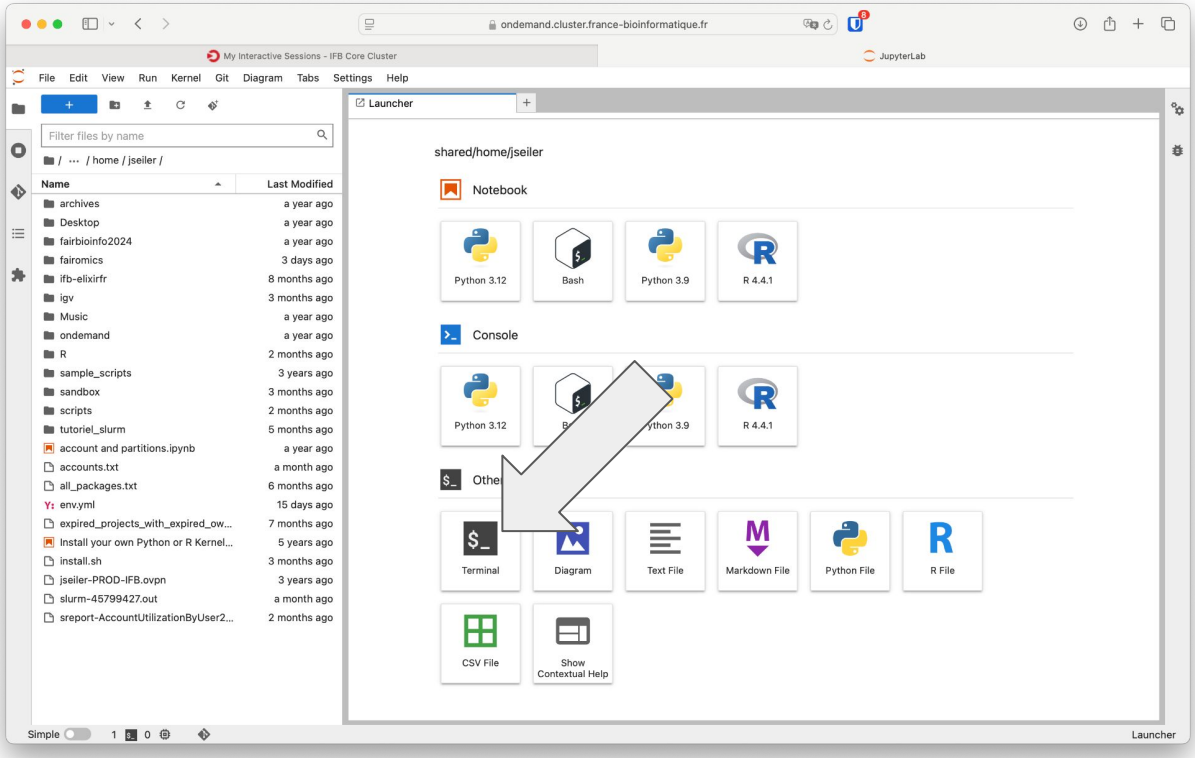
This is **an interactive training**



This is **an interactive training**



This is **an interactive training**



Git is

a command line tool	*command line (Git)* *developed in C, Bash and Perl* *Open Source (GNU GPL 2)*
a version control system	- track changes to a file/folder or a set of files/folder - navigation in the history of modifications - sharing of changes
decentralized	no need for a server
and distributed	multi-user

CC BY NC SA

Where does Git come from?

Git was invented in 2005 by Linus Torvalds.

Git is the successor of many similar tools like cvs or subversion.

git means “unpleasant person”
(Linus like to name his projects after himself…)

A few examples

- Follow the steps of modification of a program
- Test a complex change and be able to go back easily
- Working with others on a project
- Invite collaborators on a project
- Contribute to OpenSource projects

CC BY NC SA

How to use Git?

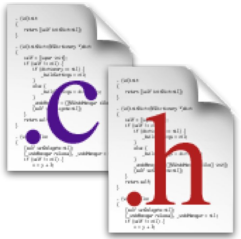
- Git is **a command line tool**
- the command is **git**
- you can use it from shell terminal

How does Git work?

For each project you track with Git, Git maintains a repository at the root of the project in a **`.git`** folder

The **`git`** command let you interact with this repository.

Working directory



Your working folder contains the files and folders that make up your project.

Git can modify these files to update them or present them to you at different versions of the project through its index.

Git repository

.git

Your Git repository contains the entire history of your project. All file versions, all modifications, etc.

This is the **`.git`** folder at the root of your working folder.

Okay… but **what is an index?**

A Git repository will allow you to track the **history of changes** in your project.

Each change is first recorded in **an index** (or indexed) to form **a collection of changes**.

This collection of changes is then validated (or committed) in your repository

Each commit is a new version or revision in your project history.

CC BY NC SA

To use the latest version of Git, load the command using **`module`**

```
$ module load git
$ git --version
git version 2.40.1
```


CC BY NC SA

In your console, type **`git`**

```
$ git
usage: git [--version] [--help] [-C <path>] [-c <name>=<value>]
           [--exec-path[=<path>]] [--html-path] [--man-path] [--info-path]
           [-p | --paginate | -P | --no-pager] [--no-replace-objects] [--bare]
           [--git-dir=<path>] [--work-tree=<path>] [--namespace=<name>]
           <command> [<args>]

These are common Git commands used in various situations:

start a working area (see also: git help tutorial)
   clone             Clone a repository into a new directory
   init              Create an empty Git repository or reinitialize an existing one



work on the current change (see also: git help everyday)
   add               Add file contents to the index
   mv                Move or rename a file, a directory, or a symlink
   restore           Restore working tree files
   rm                Remove files from the working tree and from the index
   sparse-checkout   Initialize and modify the sparse-checkout

[…]

'git help -a' and 'git help -g' list available subcommands and some
concept guides. See 'git help <command>' or 'git help <concept>'
to read about a specific subcommand or concept.
See 'git help git' for an overview of the system.
```


CC BY NC SA

The **`git`** program allows you to run commands to manage your Git repository

```
$ git <command> <arguments>
```


CC BY NC SA

Defining your identity

Your identity will be associated with the changes you make in your repositories

It is defined in the file `~/.gitconfig` or `%USERPROFILE%\.gitconfig`

```
$ git config --global user.name "Your name"
$ git config --global user.email your@email
$ git config --global init.defaultBranch main
```


CC BY NC SA

To illustrate this training, we will work on a burger recipes project.

Create a **`burgers`** folder in your home folder

```
$ mkdir burgers
```

Go to your burgers folder and run the **`git init`** comand

```
$ cd burgers
$ git init
Initialized empty Git repository in /Users/seilerj/burgers/.git
$ ls -a
.       ..      .git
```


CC BY NC SA

`git init <working dir>`

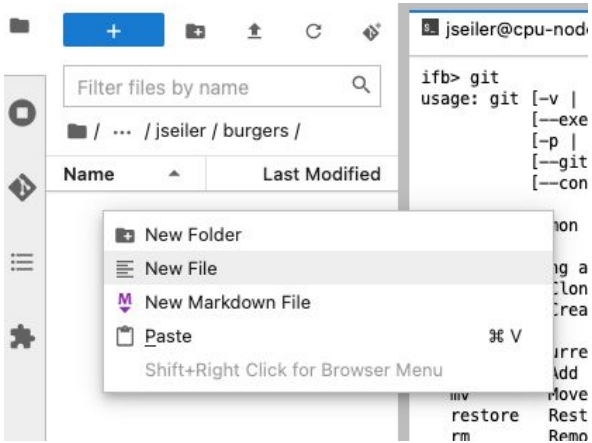
creates a Git repository

If the specified working dir does not exist, it will be created.

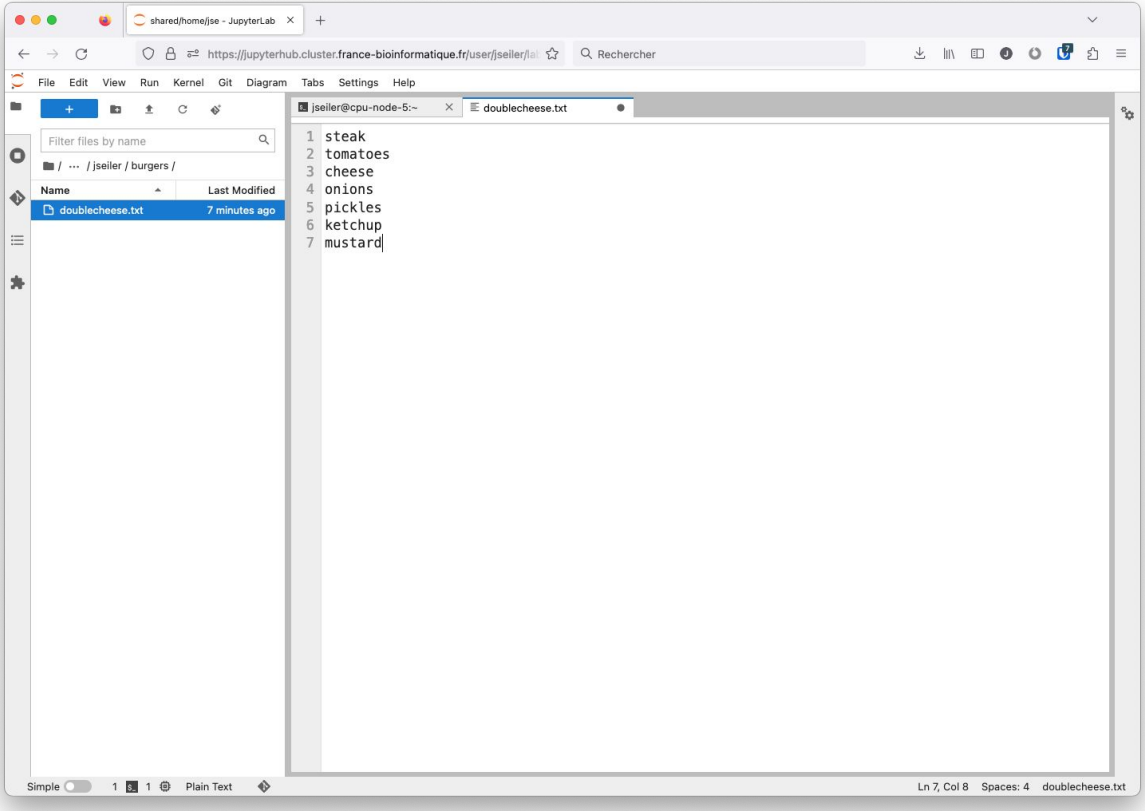
Without parameters, the command creates a Git repository for the **current folder**.

Let’s cook a burger

Create the file **`doublecheese.txt`** in the burgers folder and write down the list of ingredients to make a double cheese.



Add a file to your repository



Index a file in your Git repository

Index the doublecheese.txt file in your repo with the **`git add doublecheese.txt`** command

```
$ git add doublecheese.txt
```

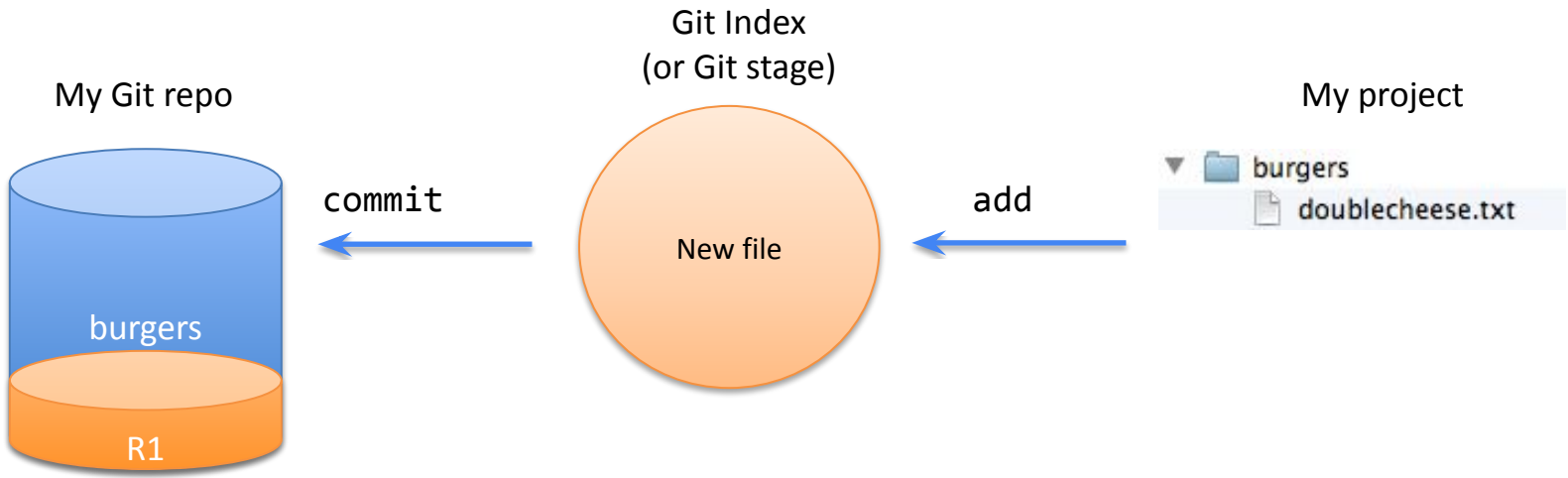
Save changes

Validate this modification in order to save it in your repository with the **`git commit`** command

```
$ git commit -m "Birth of the double cheese"
[main (root-commit) bb0188d] Birth of the double cheese
 1 file changed, 7 insertions(+)
 create mode 100644 doublecheese.txt
```


CC BY NC SA

Add a file to your repository

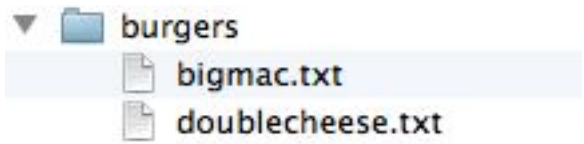


Modify a file

There is no tomatoes in the double cheese!

Correct the file `doublecheese.txt`

Add the file bigmac.txt to your burgers project



steak
salad
tomatoes
onions
pickle
ketchup
mustard

Where are we now?

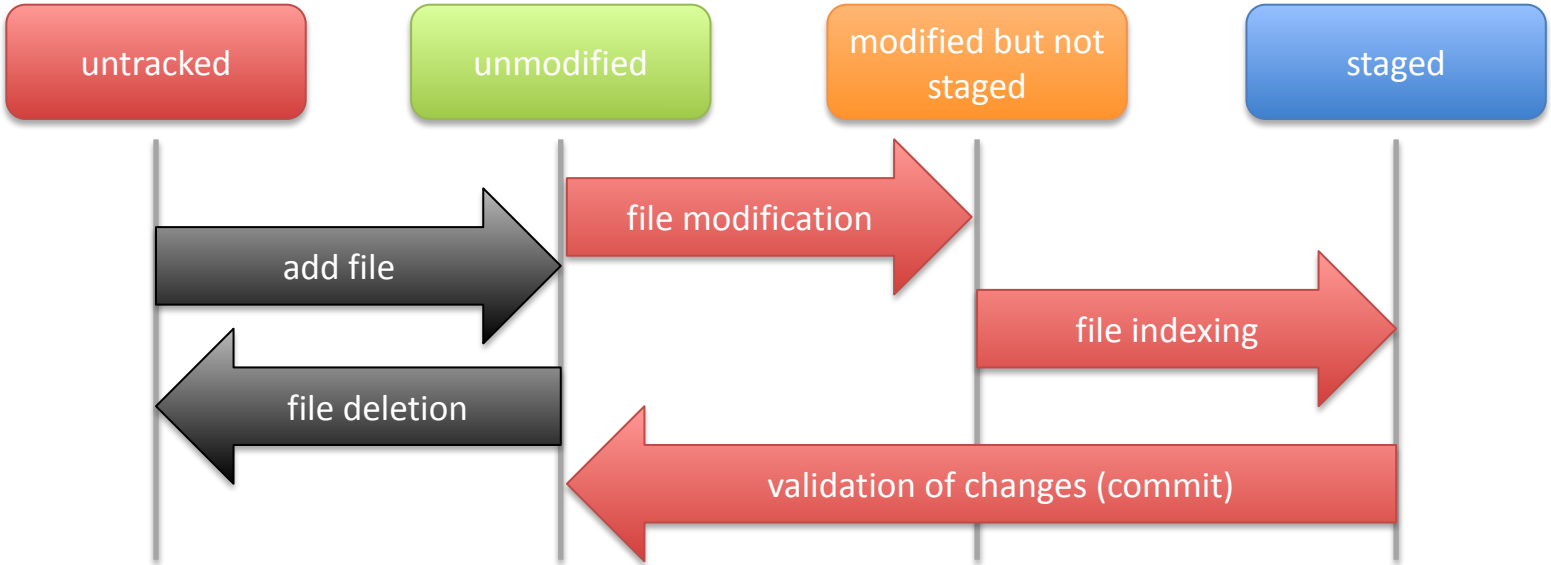
Check what has changed in your project with the `git status` command

```
$ git status
On branch main
Changes not staged for commit:
  (use "git add <file>..." to update what will be committed)
  (use "git restore <file>..." to discard changes in working directory)
        modified:   doublecheese.txt

Untracked files:
  (use "git add <file>..." to include in what will be committed)
        .ipynb_checkpoints/
        bigmac.txt

no changes added to commit (use "git add" and/or "git commit -a")
```

Files **status** in a Git repository



Index the change on the doublecheese.txt file with the `git add` command

```
$ git add doublecheese.txt
```


CC BY NC SA

Check what has changed in your project with the `git status` command

```
$ git status
On branch main
Changes to be committed:
  (use "git restore --staged <file>..." to unstage)
   modified:   doublecheese.txt

Untracked files:
  (use "git add <file>..." to include in what will be committed)
   .ipynb_checkpoints/
        bigmac.txt
```


CC BY NC SA

Add a new file to the repo

Add the new file bigmac.txt with the **`git add`** command

```
$ git add bigmac.txt
```

Check what has changed in your project with the **`git status`** command

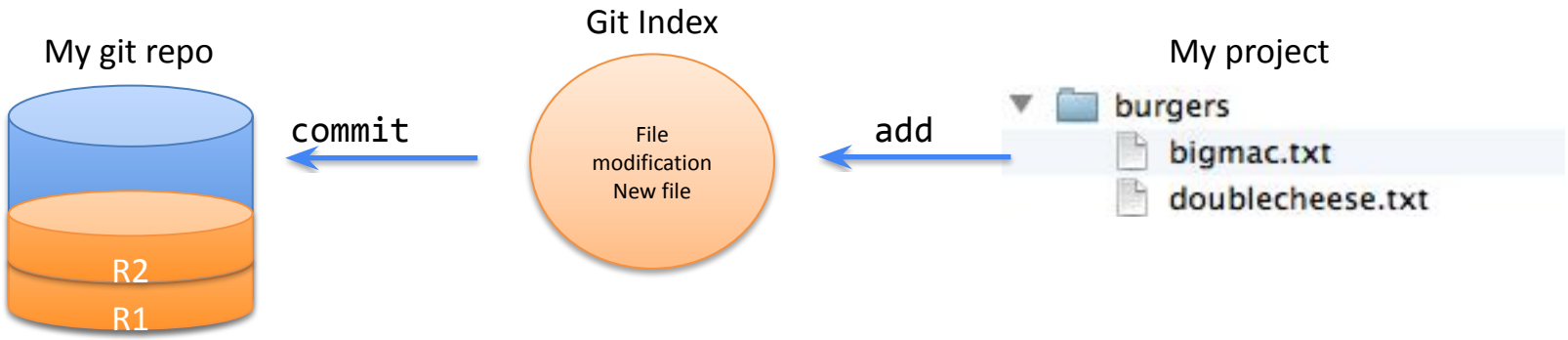
```
$ git status
On branch main
Changes to be committed:
  (use "git restore --staged <file>..." to unstage)
        new file:   bigmac.txt
        modified:   doublecheese.txt

Untracked files:
  (use "git add <file>..." to include in what will be committed)
        .ipynb_checkpoints/
```

Run the `git commit` command

```
$ git commit -m "Add the big mac and correct the double cheese"
[main 299a6b2] Add the big mac and correct the double cheese
 2 files changed, 7 insertions(+), 1 deletion(-)
 create mode 100644 bigmac.txt
```

A new revision is here



Edit the `bigmac.txt` file to add a bad ingredient.

Retrieve **the previous version of a file**

Restore the last valid version of the file with the `git checkout` command

```
$ git checkout bigmac.txt
```

Did you say revision number?

Each commit to your repository creates a new revision or version of your project.

Each revision is actually **a set of changes**.

Use the `git log` command to view the revisions of your local repository

```
$ git log
commit 299a6b210eed54e9f4c164b85ecbcb9ed899e6eb (HEAD -> main)
Author: Julien SEILER <seilerj@igbmc.fr>
Date:   Tue May 16 12:14:35 2023 +0200

    Add the big mac and correct the double cheese

commit bb0188df5bf0c3cb3a152e52e22df1249d52e2be
Author: Julien SEILER <seilerj@igbmc.fr>
Date:   Tue May 16 11:49:37 2023 +0200

    Birth of the double cheese
```

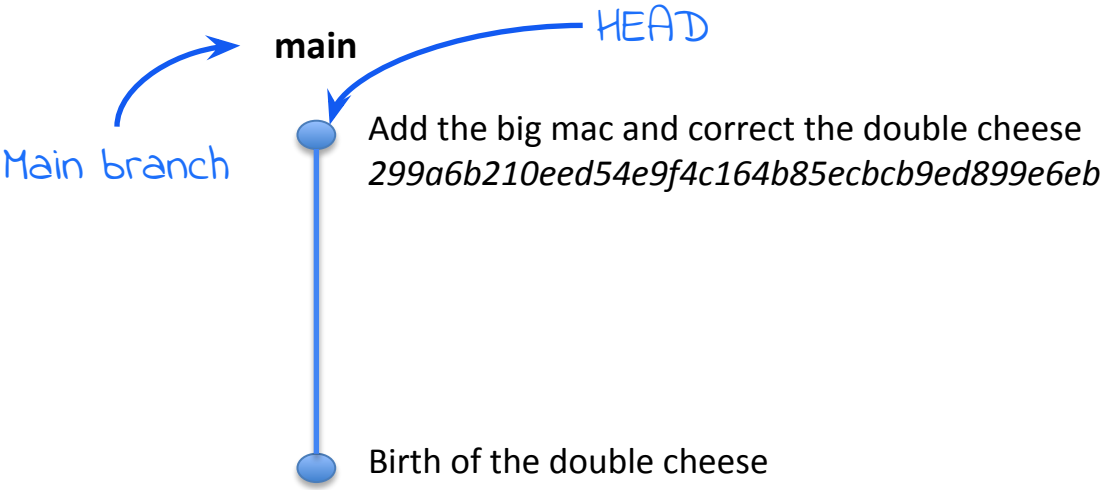

CC BY NC SA

Revert to the initial version of your project (just the doublecheese.txt) using the `git checkout <id rev>` command

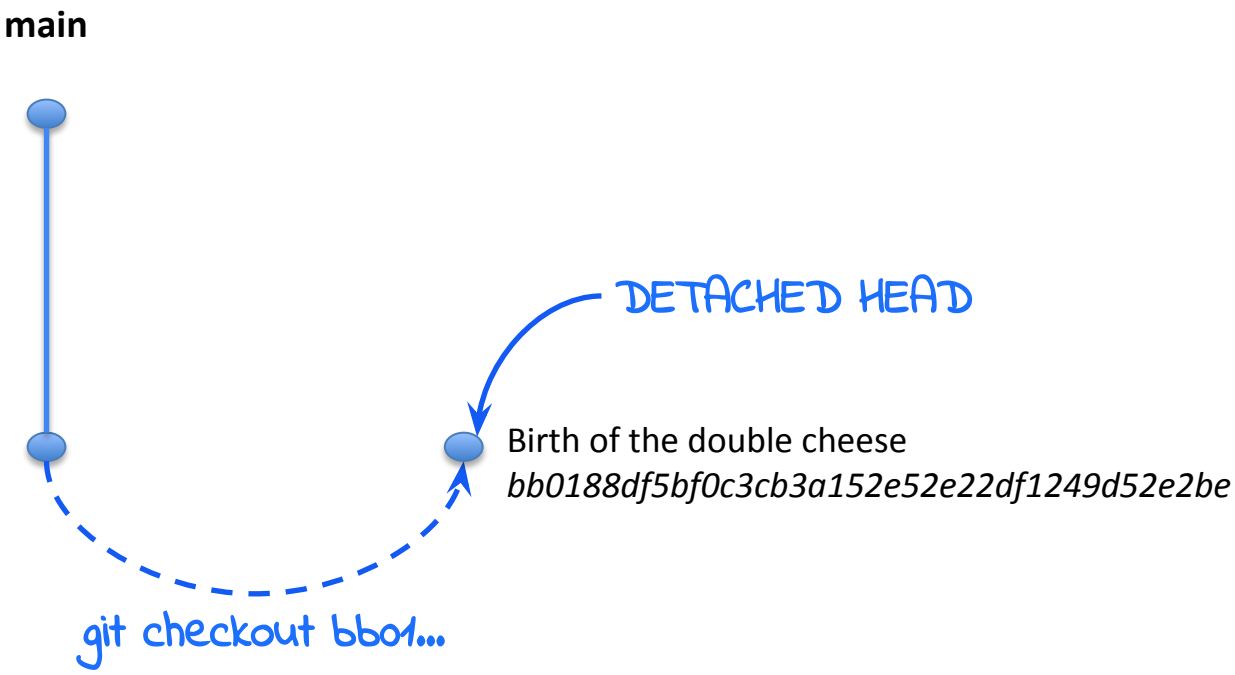
```
$ git checkout bb0188df5bf0c3cb3a152e52e22df1249d52e2be
Note: switching to 'bb0188df5bf0c3cb3a152e52e22df1249d52e2be'.

You are in 'detached HEAD' state. You can look around, make experimental
changes and commit them, and you can discard any commits you make in
this
[…]
$ ls
doublecheese.txt
```

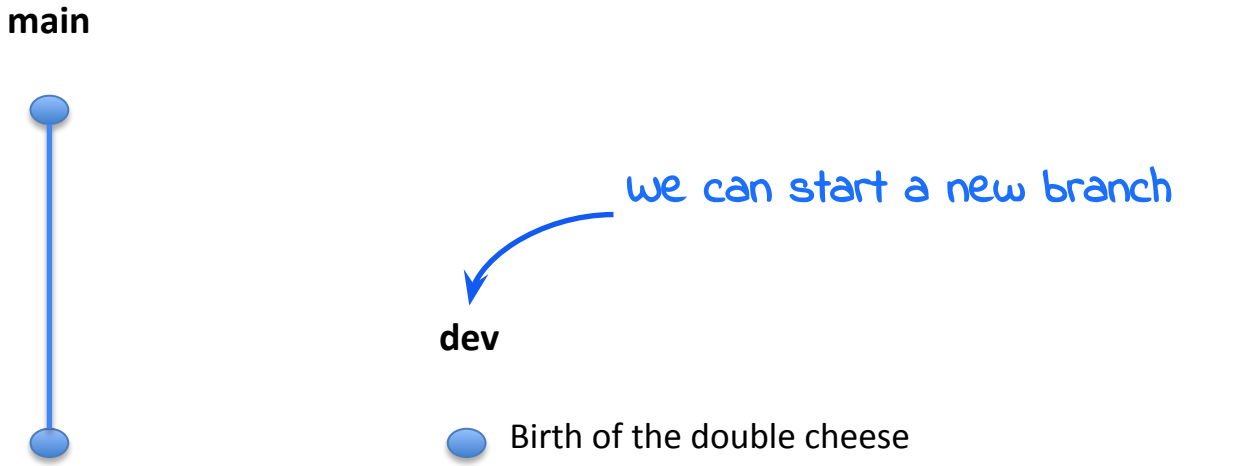
Introduction to the concept of **branch**

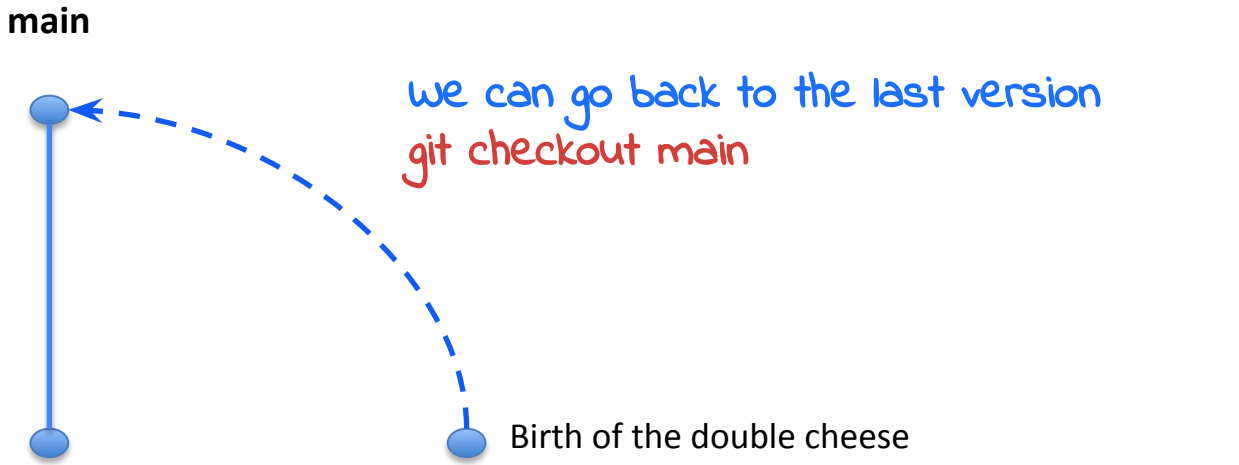


Introduction to the concept of **branch**



Introduction to the concept of **branch**



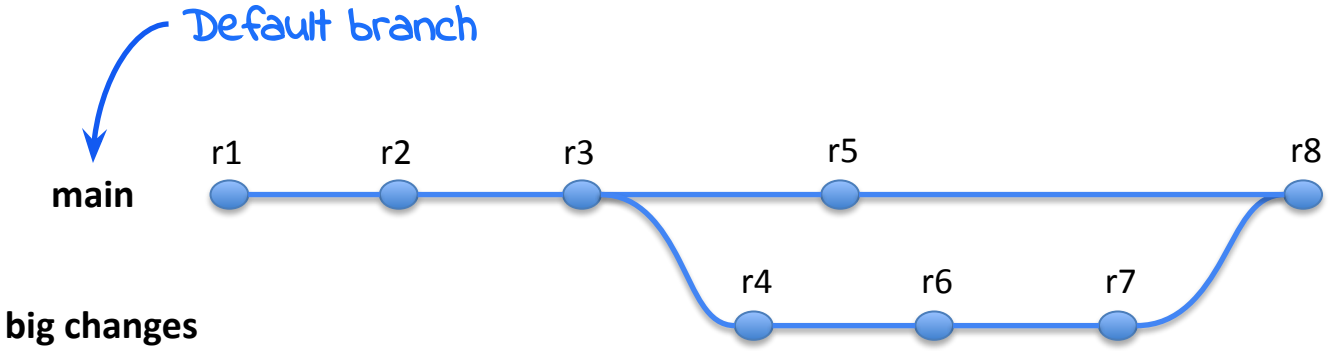

main
we can go back to the last version
git checkout main
Birth of the double cheese
bb0188df5bf0c3cb3a152e52e22df1249d52e2be

Use the command `git checkout main`

```
$ git checkout main
Previous HEAD position was bb0188d Birth of the double cheese
Switched to branch 'main'
$ ls
bigmac.txt   doublecheese.txt
```

Going further with the **branches**

Going further with the **branches**



Our fast-food restaurant wants to go **ORGANIC**.
We need to change all our recipes!!!

That's a « big change »...

Create a new `organic` branch on your repository

```
$ git branch organic
```

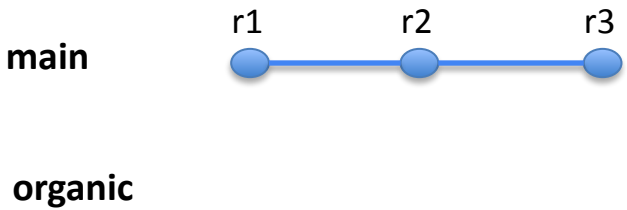
Consult the branches available on your repository

```
$ git branch
* main
  organic
```

Use the **`git checkout`** command to switch branches.

```
$ git checkout organic
Switched to branch 'organic'
```

Going further with the **branches**

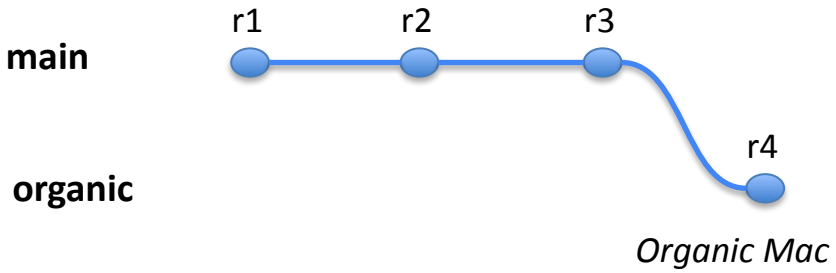


Modify the Big Mac recipe and commit the change

organic steak
organic salad
organic tomatoes
organic onions
organic pickles
organic ketchup
organic mustard

```
$ git commit -a -m "organic mac"
[organic a7a6f6e] organic mac
 1 file changed, 7 insertions(+), 7 deletions(-)
```

Going further with the **branches**

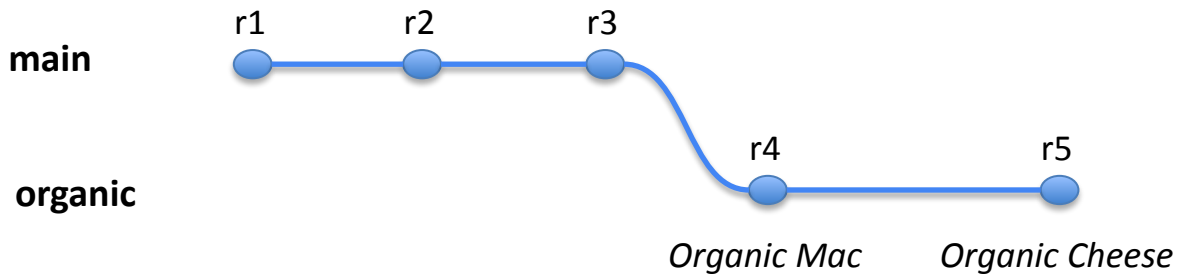


Modify the Double Cheese recipe and commit the change

organic steak
organic cheese
organic onions
organic pickles
organic ketchup
organic mustard

```
$ git commit –a –m  "organic cheese"
[organic 1ed7510] organic cheese
 1 file changed, 6 insertions(+), 6 deletions(-)
```

Going further with the **branches**



Let's go back to our **main branch**

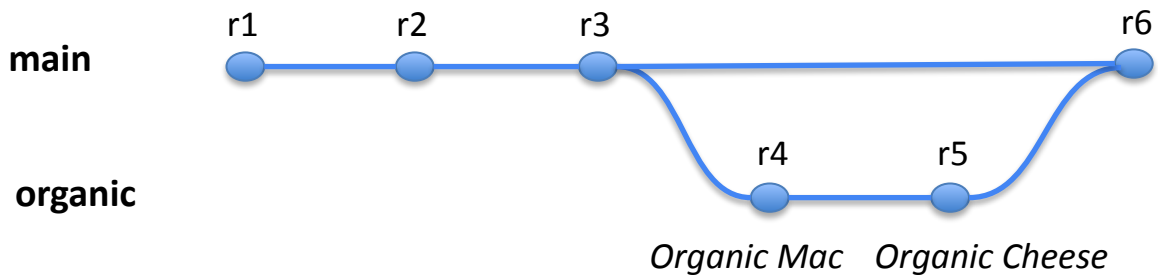
Return to the main branch with the `git checkout` command

```
$ git checkout main
Switched to branch 'main'
```

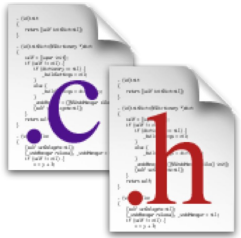
We can retrieve the changes saved on the organic branch with the `git merge` command

```
$ git merge organic
Updating 299a6b2..1ed7510
Fast-forward
 bigmac.txt       | 14 +++++++-------
 doublecheese.txt | 12 ++++++------
 2 files changed, 13 insertions(+), 13 deletions(-)
```

Going further with the **branches**



Working directory



checkout
merge

commit

git repository

.git

CC BY NC SA

The Git commands to manage your local repository

git command	Description
`init`	Creating a repository for a project/folder
`add`	Indexing of a modification or addition of a file or folder
`rm`	Deleting a file or folder
`mv`	Moving a file or folder
`status`	Visualization of the repository status
`diff`	Viewing changes between two revisions or between a revision and the current version
`checkout`	Retrieving a file from the repository
`log`	Consultation of the list of revisions (commits) registered on the repository

Exercise

1. Delete the bigmac and save the change in your repository
2. Add a burger and save the change to your repository
3. Restore the bigmac to your working folder

Exercise

1. Delete the bigmac and save the change in your repository

```
$ git rm bigmac.txt
$ git commit –m “bye bye big mac“
[main 916c075] bye bye big mac
 1 file changed, 7 deletions(-)
 delete mode 100644 bigmac.txt
```

2. Add a burger and save the change to your repository

```
$ git add newburger.txt
$ git commit -m "add a new burger"
[main 26f9033] add a new burger
 1 file changed, 1 insertion(+)
 create mode 100644 newburger.txt “
```

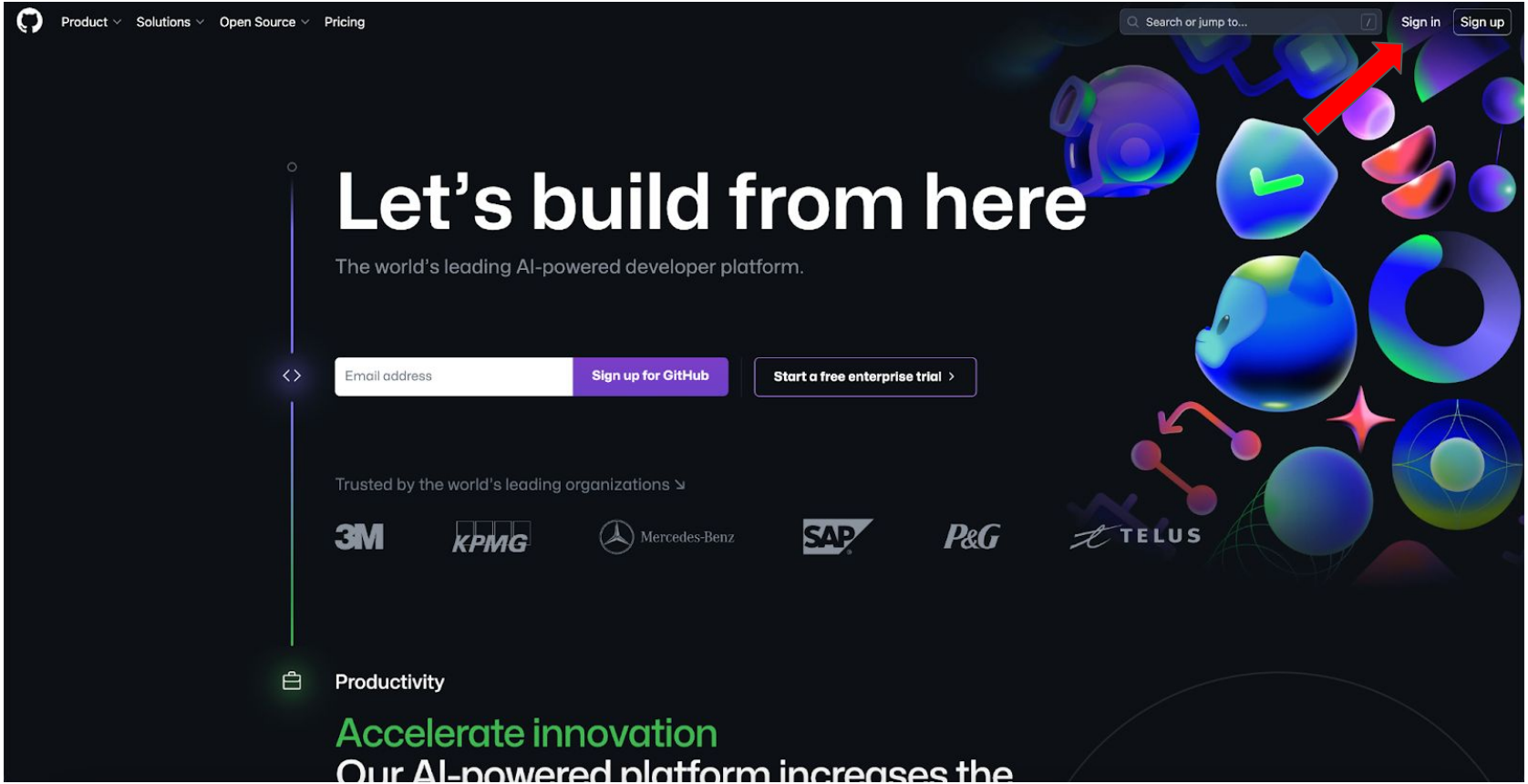
3. Restore the bigmac to your working folder

```
$ git checkout <rev> bigmac.txt
Updated 1 path from 69a5a31
$ git commit -a -m "return of the big mac"
[main 9a4a1c6] return of the big mac
 1 file changed, 7 insertions(+)
 create mode 100644 bigmac.txt
```

<rev> is the last revision at which the bigmac.txt was present

To share our work and be able to collaborate with other people, we need to use a “code forge”

Connect to Github



Connect to Github

Sign in to GitHub

Username or email address

Password Forgot password?

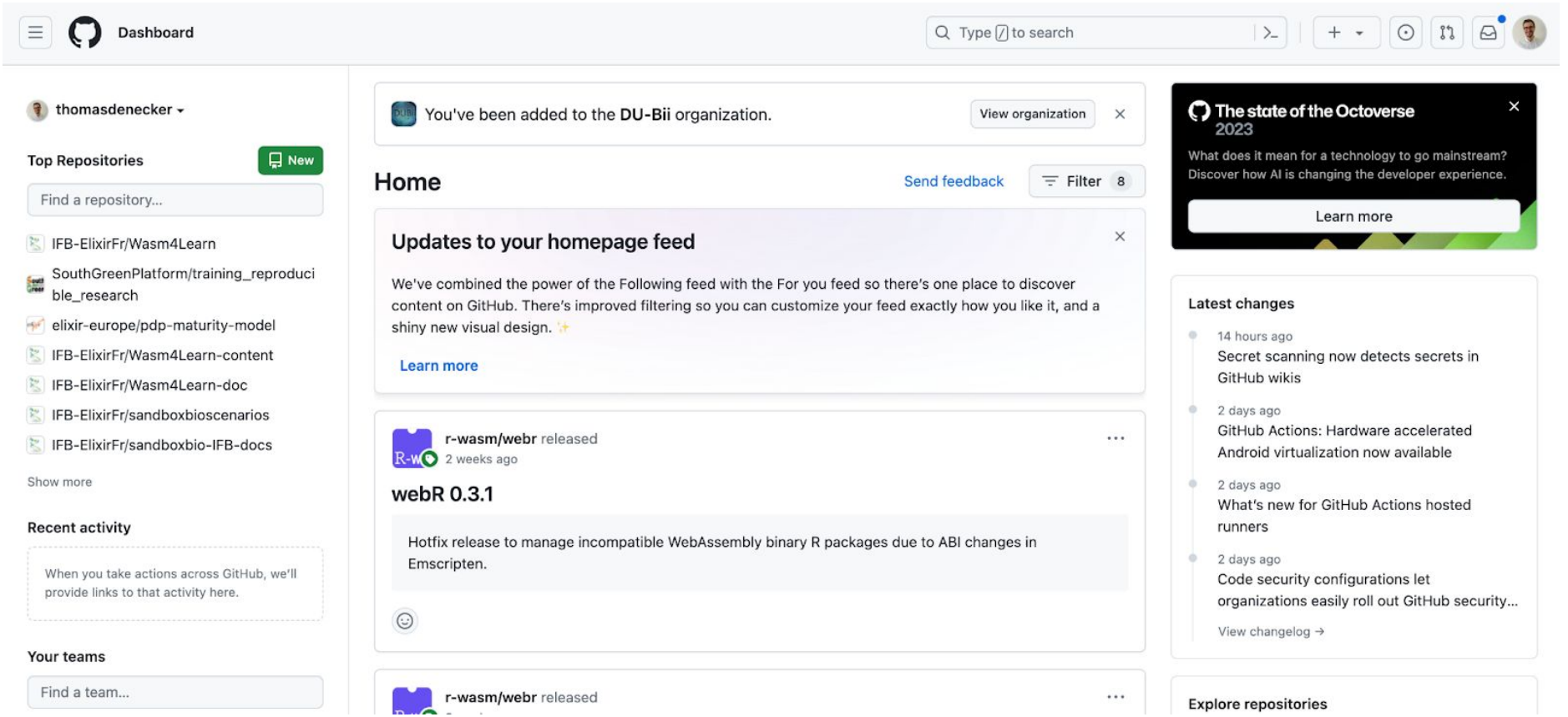
Sign in

Sign in with a passkey

New to GitHub? Create an account

Terms Privacy Docs Contact GitHub Support Manage cookies Do not share my personal information

Your Dashboard



Generating a new SSH key

In a terminal in Jupyter

```
$ ssh-keygen -t ed25519 -C "your_email@example.com"
$ Generating public/private ALGORITHM key pair.
$ Enter a file in which to save the key
(/home/YOU/.ssh/id_ALGORITHM):[Press enter]
$ Enter passphrase (empty for no passphrase): [Type a passphrase]
$ Enter same passphrase again: [Type passphrase again]
```

https://docs.github.com/fr/authentication/connecting-to-github-with-ssh/generating-a-new-ssh-key-and-adding-it-to-the-ssh-agent

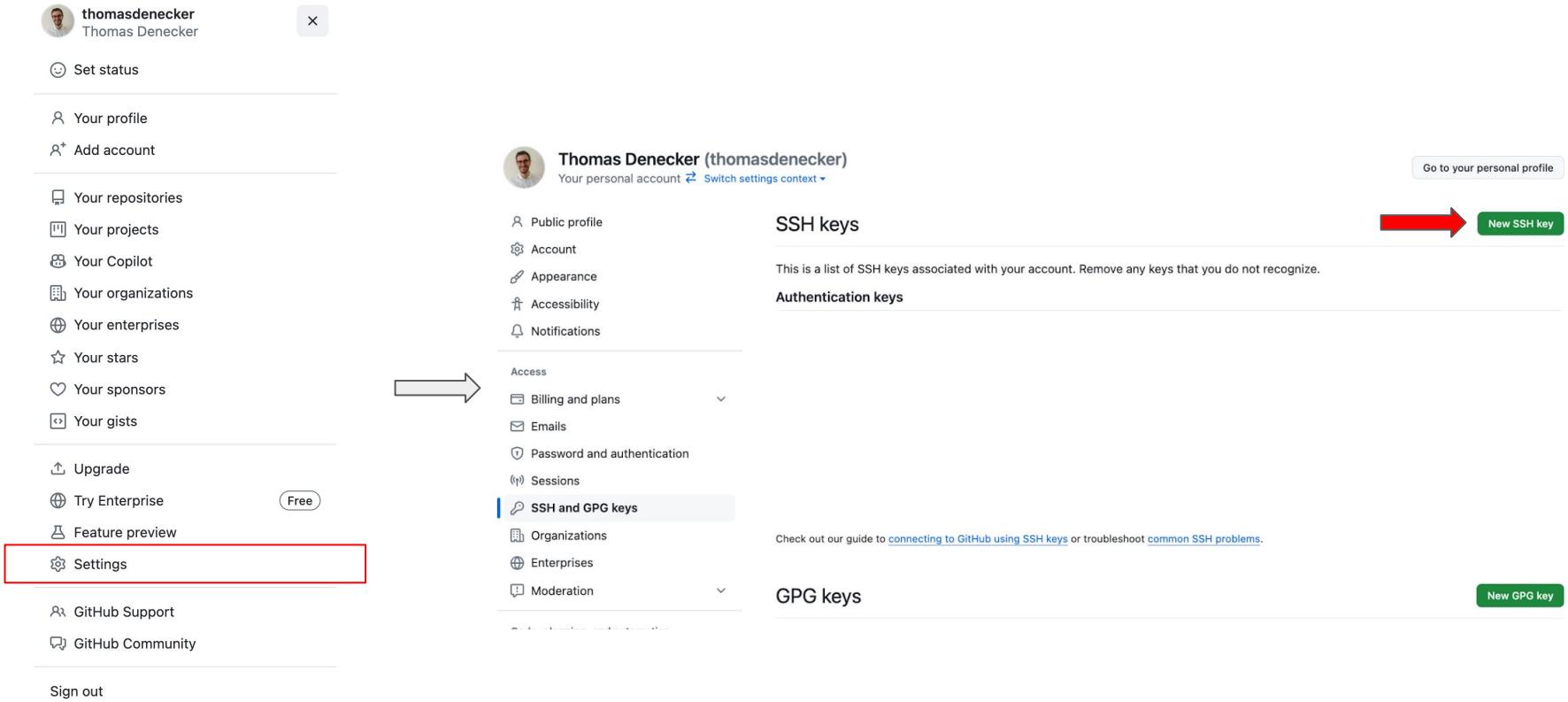
In a terminal in Jupyter

```
$ ls -al ~/.ssh
total 28
drwx------  2 tdenecker tdenecker  4096 Apr  4 10:40 .
drwx------ 30 tdenecker tdenecker 12288 Apr  4 10:37 ..
-rw-------  1 tdenecker tdenecker   419 Apr  4 10:40 id_ed25519
-rw-r--r--  1 tdenecker tdenecker   106 Apr  4 10:40 id_ed25519.pub
-rw-r--r--  1 tdenecker tdenecker  1063 Jul 13  2023 known_hosts

$ cat ~/.ssh/id_ed25519.pub
```

Copy the output

Add your SSH key



Add your SSH key

Thomas Denecker (thomasdenecker)
Your personal account ⇄ Switch settings context ▾

Go to your personal profile

- Public profile
- Account
- Appearance
- Accessibility
- Notifications

Access

- Billing and plans
- Emails
- Password and authentication
- Sessions
- **SSH and GPG keys**
- Organizations
- Enterprises
- Moderation

Code, planning, and automation

- Repositories

Add new SSH Key

Title

Key type

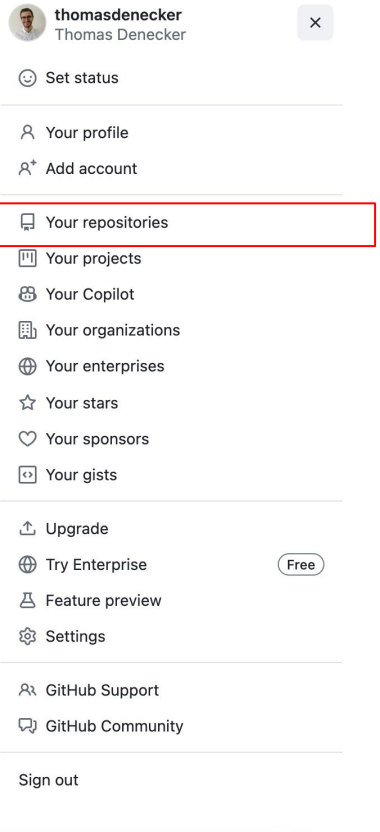
Authentication Key

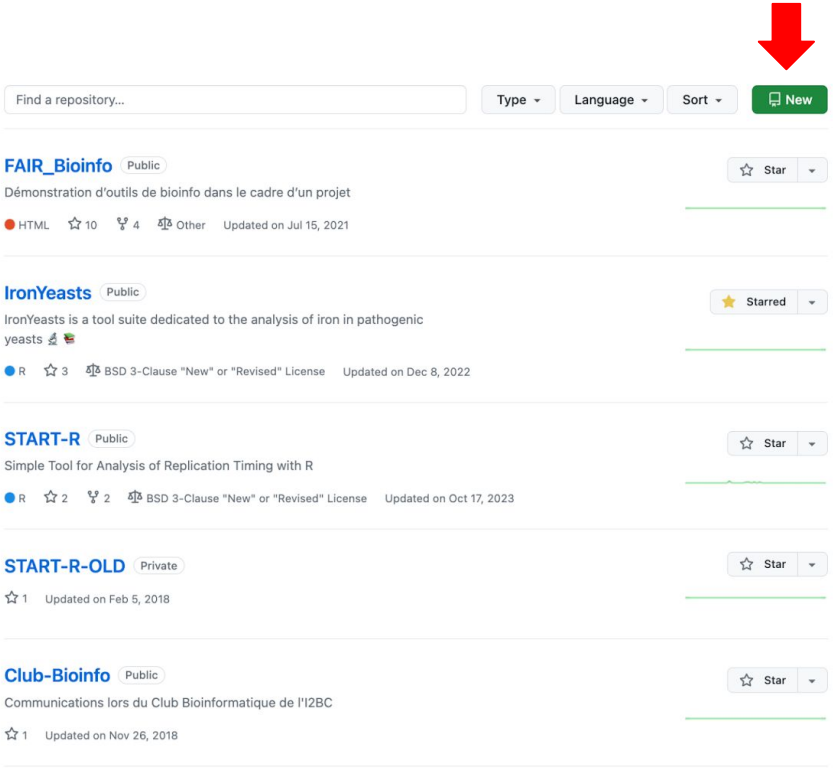
Key

Begins with 'ssh-rsa', 'ecdsa-sha2-nistp256', 'ecdsa-sha2-nistp384', 'ecdsa-sha2-nistp521', 'ssh-ed25519', 'sk-ecdsa-sha2-nistp256@openssh.com', or 'sk-ssh-ed25519@openssh.com'

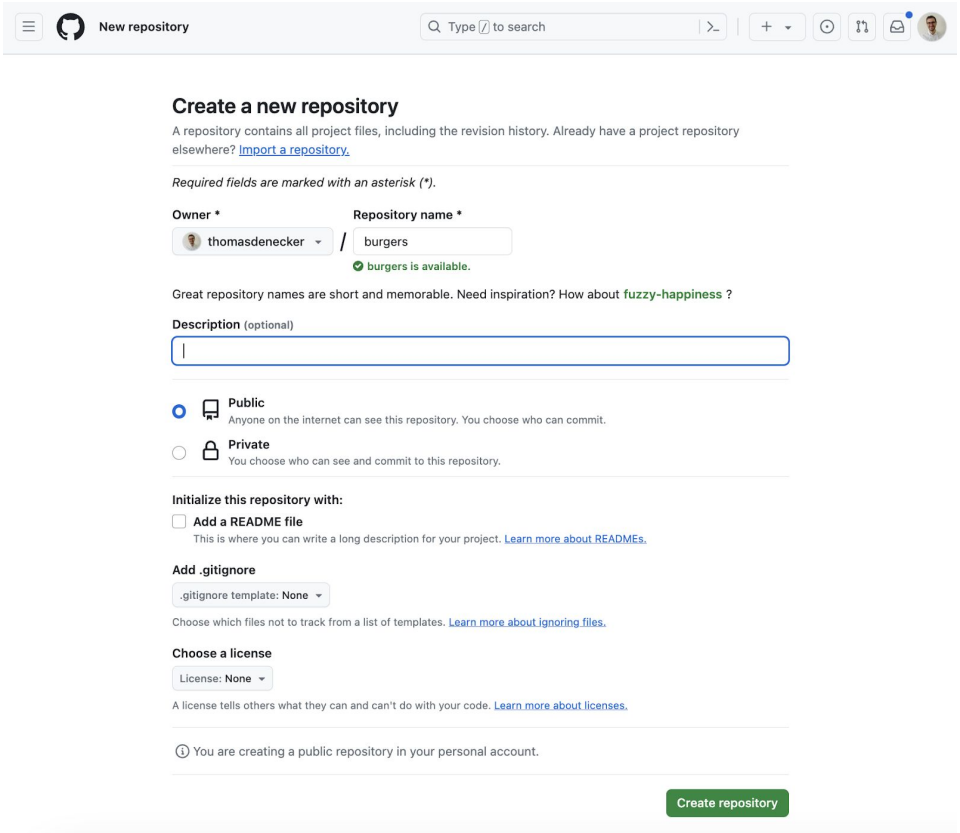
Add SSH key

Create a new repository





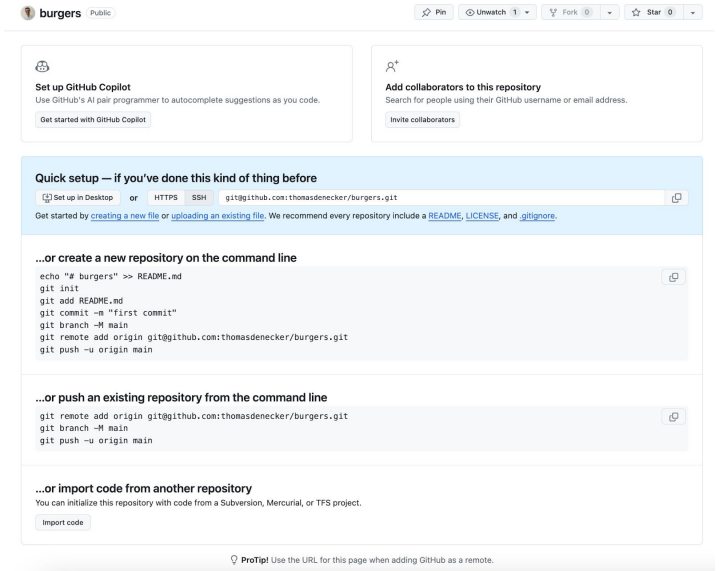
Create a new repository



`git remote`

Push an existing repository from the command line

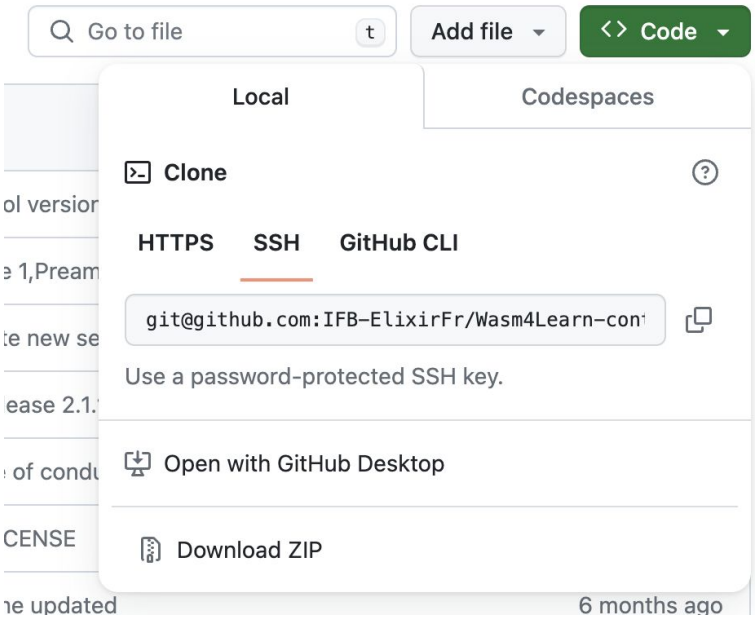
```
$ git remote add origin git@github.com:<repository_URL>/burgers.git
$ git push -u origin main
```



`git clone`

Clone a remote repository to your local machine.

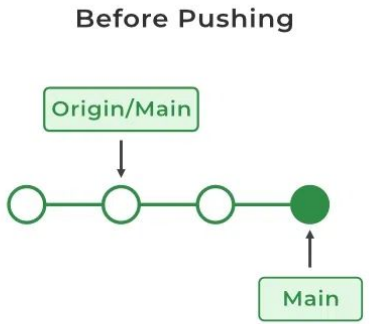
```
git clone <repository_URL>
```

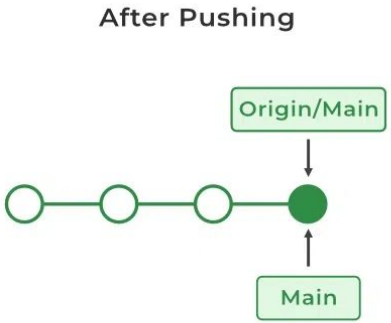


`git push`

Send local commits to the remote repository.

```
git push <remote_name> <branch_name>
```

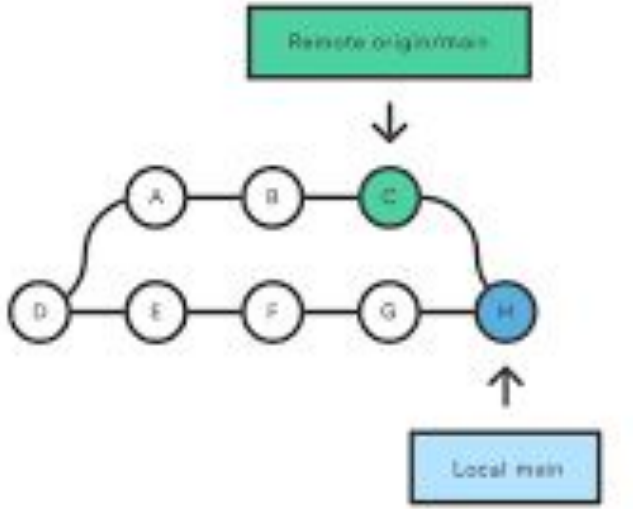




`git push`

Fetch the latest changes from the remote repository and merge them into your local branch.

```
git pull <remote_name> <branch_name>
```



`git remote`

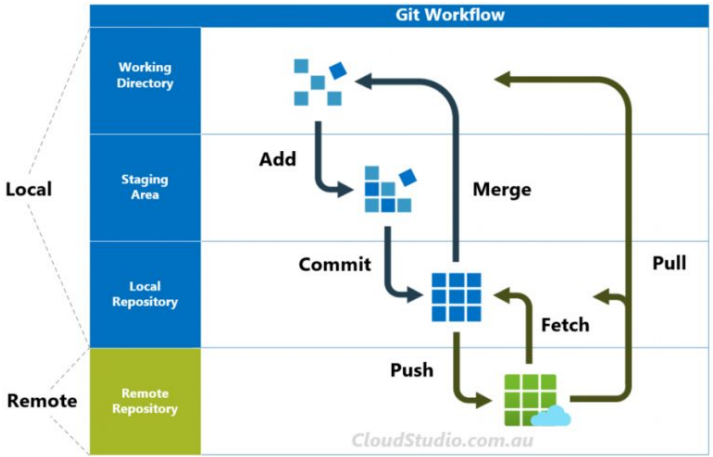
Show or add remote repositories.

```
git remote -v
git remote add <remote_name> <remote_URL>
```

`git fetch`

Fetch branches and commits from the remote repository, but do not merge the changes into your local branch.

```
git fetch <remote_name>
```



Thank you for your attention !

IFB
INSTITUT FRANÇAIS DE BIOINFORMATIQUE

elixir FRANCE

cnrs INRAE Inserm cea