Leiden University
Medical Center

Code and data management with Git

Introduction to Version Control

Course Outline

- Introduction
- Git basics
- Practical 1
- The git commit graph
- Practical 2
- Lunch break
- Colaboration: remote repositories
- Practical 3

`https://git.lumc.nl/courses/gitcourse`

Introduction Outline

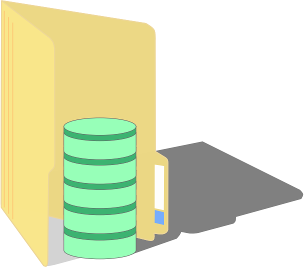
- What is version control?
  - Definition
  - General features
  - Step by step example
  - Distributed repositories
  - Why to use/not use version control?
- Git
- Web-based remotes
  - GitLab vs GitHub

What is version control?

Definition

The **management of changes** *to documents, computer programs, large web sites, and other collections of information.*
— Wikipedia

A system that **records changes** *to a file or set of files over time so that you can recall specific versions later.*
— https://git-scm.com/



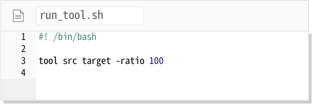
What is version control?

General features

- Keep track of your files in an orderly manner.
  - Hide old versions.
  - Easily access old versions.
  - Check the differences between current and older versions.
- Enables collaboration.
  - Record who made changes and when.

What is version control?

Step by step

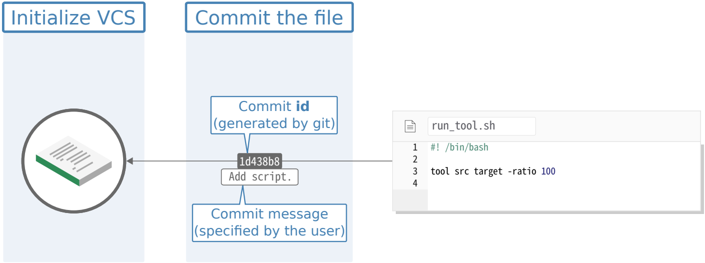


run_tool.sh

```
#! /bin/bash

tool src target -ratio 100

```

What is version control?

Step by step



Initialize VCS

Commit the file

Commit **id**
(generated by git)

1d438b8

Add script.

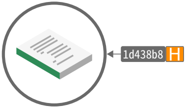
Commit message
(specified by the user)

run_tool.sh

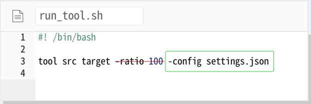
```
#! /bin/bash

tool src target -ratio 100

```

What is version control?

Step by step

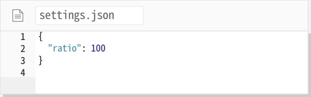


Edit the existing file



run_tool.sh

```
#! /bin/bash

tool src target ~~-ratio 100~~ -config settings.json

```



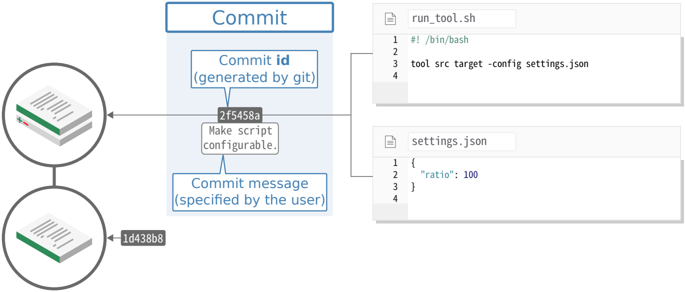
settings.json

```
{
  "ratio": 100
}

```

Create new file

What is version control?

Step by step



run_tool.sh

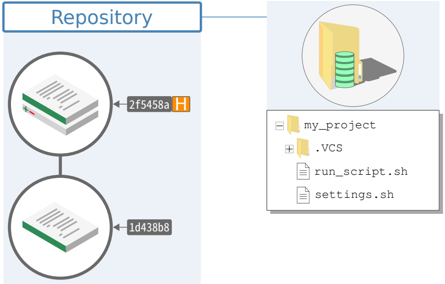
```
#! /bin/bash

tool src target -config settings.json

```

settings.json

```
{
  "ratio": 100
}

```

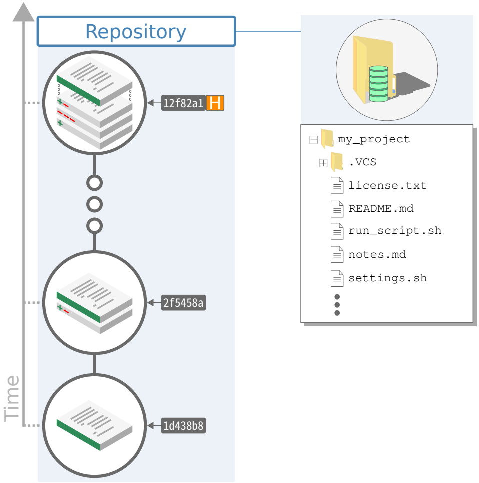
What is version control?

Step by step



What is version control?

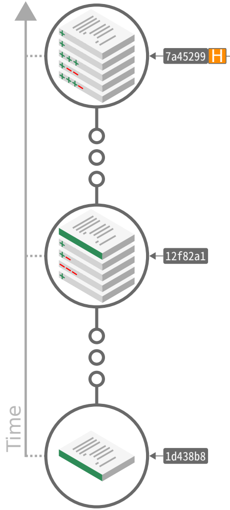
Step by step

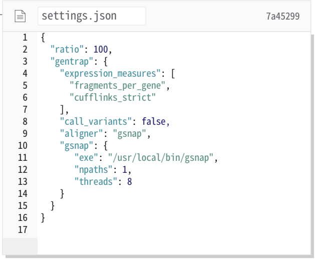


my_project
.VCS
license.txt
README.md
run_script.sh
notes.md
settings.sh

What is version control?

Step by step

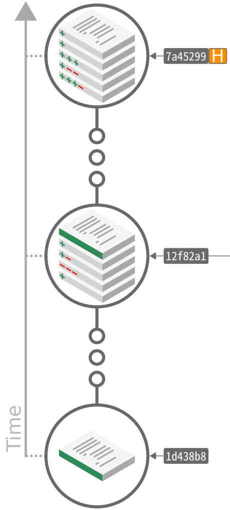


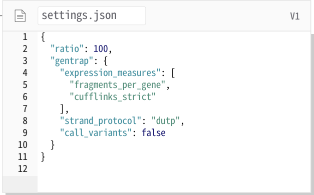


```
{
  "ratio": 100,
  "gentrap": {
    "expression_measures": [
      "fragments_per_gene",
      "cufflinks_strict"
    ],
    "call_variants": false,
    "aligner": "gsnap",
    "gsnap": {
       "exe": "/usr/local/bin/gsnap",
       "npaths": 1,
       "threads": 8
    }
  }
}
```

What is version control?

Step by step

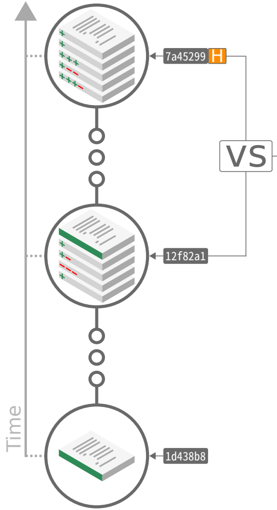


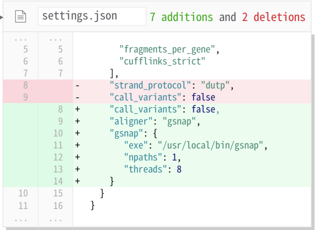


```
{
  "ratio": 100,
  "gentrap": {
    "expression_measures": [
      "fragments_per_gene",
      "cufflinks_strict"
    ],
    "strand_protocol": "dutp",
    "call_variants": false
  }
}
```

What is version control?

Step by step

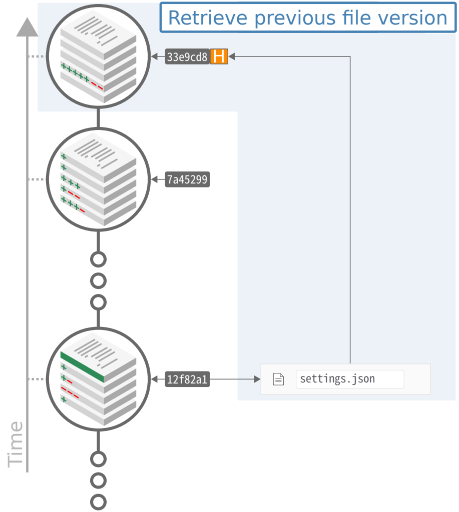




```
 ...   ...
  5     5           "fragments_per_gene",
  6     6           "cufflinks_strict"
  7     7         ],
  8          -    "strand_protocol": "dutp",
  9          -    "call_variants": false
        8    +    "call_variants": false,
        9    +    "aligner": "gsnap",
       10    +    "gsnap": {
       11    +        "exe": "/usr/local/bin/gsnap",
       12    +        "npaths": 1,
       13    +        "threads": 8
       14    +    }
 10    15       }
 11    16     }
 ...   ...
```

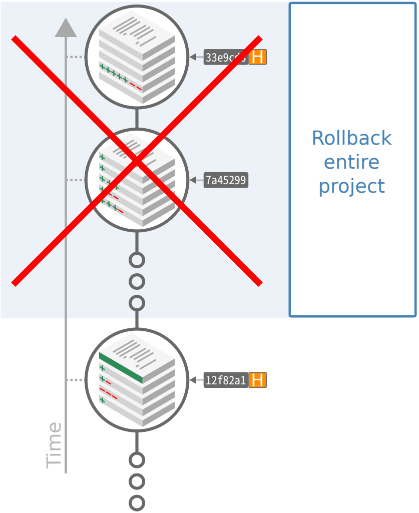
What is version control?

Step by step



What is version control?

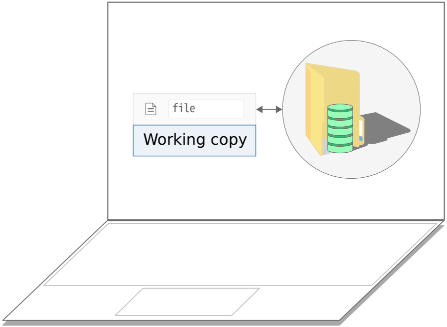
Step by step



Rollback entire project

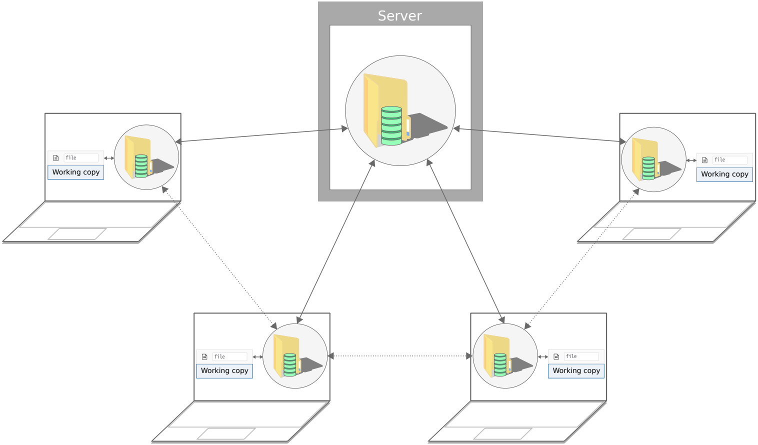
What is version control?

Local repository



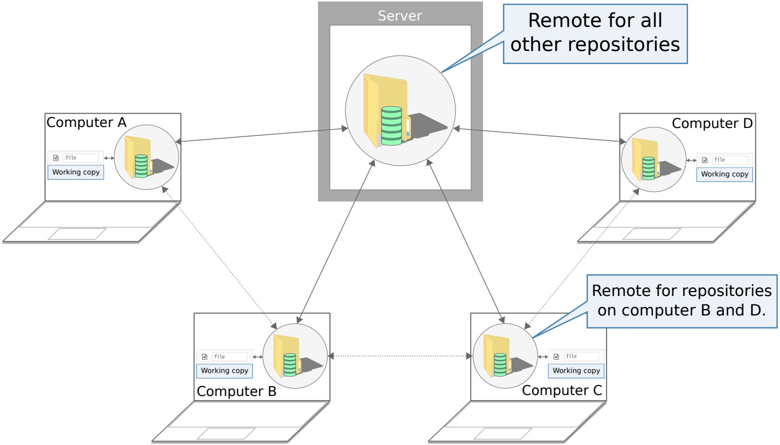
What is version control?

Distributed repositories



What is version control?

Distributed repositories



Remote for all other repositories

Remote for repositories on computer B and D.

What is version control?

Why should I use it?

For a single user:

- Revert files to a previous state.
- Revert the entire project back to a previous state.
- Review changes made over time.
- Backup.

What is version control?

Why should I use it?

For a single user:

- Revert files to a previous state.
- Revert the entire project back to a previous state.
- Review changes made over time.
- Backup.

For multiple users:

- A reliable way to share files between people/computers.
- Allow multiple people working on the same project at the same time.
- Conflict resolution.
- See who made which changes at which time.

What is version control?

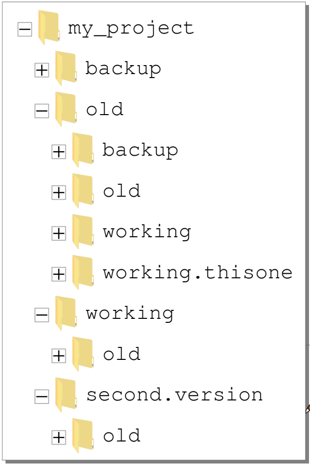
Why should I not use it?

- I have my own system.

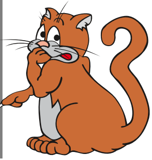
What is version control?

Why should I not use it?

- I have my own system.



```
⊟ my_project
  ⊞ backup
  ⊟ old
    ⊞ backup
    ⊞ old
    ⊞ working
    ⊞ working.thisone
  ⊟ working
    ⊞ old
  ⊟ second.version
    ⊞ old
```

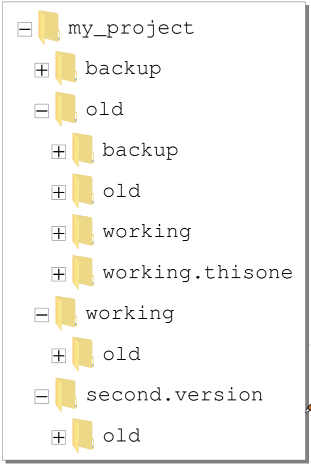


What is version control?

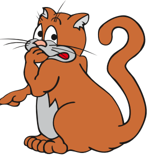
Why should I not use it?

A list of common excuses:

- I have my own system.
- It is too much work.
- I am the only one working on this project.
- This code will not be used by anyone else.
- The bugs can be tracked forever.
- ...



```
⊟ my_project
  ⊞ backup
  ⊟ old
    ⊞ backup
    ⊞ old
    ⊞ working
    ⊞ working.thisone
  ⊟ working
    ⊞ old
  ⊟ second.version
    ⊞ old
```



What is version control?

Why should I not use it?

A list of common excuses:

- I have my own system.
- It is too much work.
- I am the only one working on this project.
- This code will not be used by anyone else.
- The bugs can be tracked forever.
- . . .

Eventually leading to:

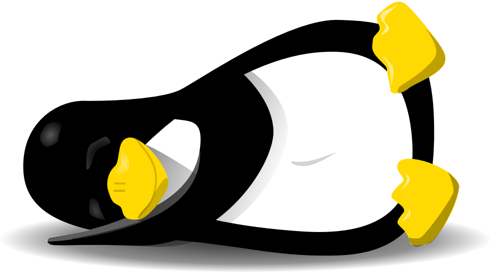
- I'm too busy rewriting the code I accidentally deleted.

Git

The name

Git (n): A person who is deemed to be despicable or contemptible.
— WordNet.

I'm an egotistical bastard, and I name all my projects after myself. First "Linux", now "git". — Linus Torvalds.



Git

History

Designed to replace the commercial package *BitKeeper*.

- Speed.
- Simple design.
- Strong support for non-linear development
  - Thousands of parallel branches.
- Fully distributed.
- Able to handle large projects like the Linux kernel efficiently.

Installation

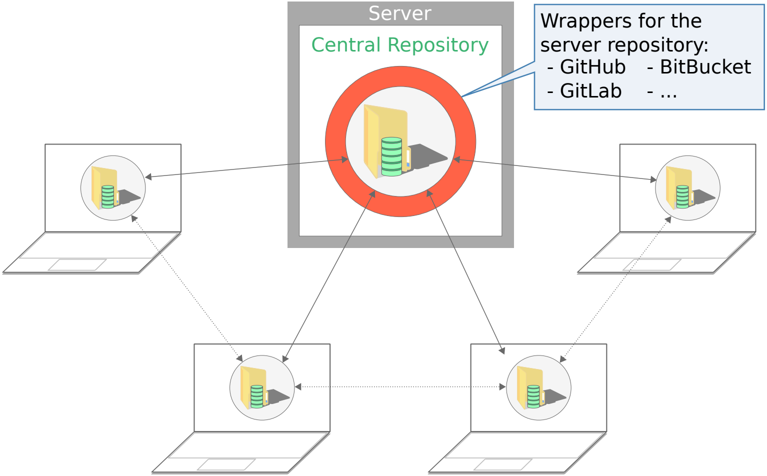
- Linux
  - `sudo apt-get install git` (Ubuntu)
  - `sudo yum install git` (Fedora)
- Windows
  - git-scm.com/download/win
- Mac
  - git-scm.com/mac

Once installed you can check Git version:

- `git --version`

Remotes

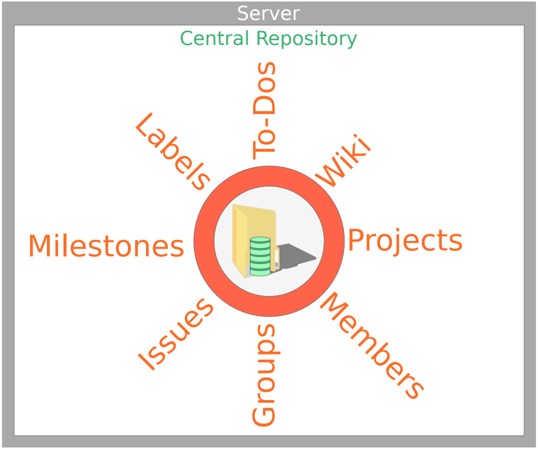
Web-based



Wrappers for the server repository:
- GitHub - BitBucket
- GitLab - ...

Remotes

Web-based



Remotes

Web-based

GitHub:

- Free for public repositories. Paid for private ones.
- Hosting provided for free.

GitLab:

- Self hosting - own installation.
- No hosting provided.
- Basic features for free. Paid advanced ones.

Acknowledgements

Martijn Vermaat
Wibowo Arindrarto
Szymon Kiełbasa
Jeroen Laros

https://git.lumc.nl/humgen/gitcourse