

FAIR_bioinfo : Open Science and FAIR principles in a bioinformatics project

How to make a bioinformatics project more reproducible

C. Hernandez¹ T. Denecker² J. Seiler² G. Le Corguillé²
C. Toffano-Nioche¹

¹Institute for Integrative Biology of the Cell (I2BC)
UMR 9198, Université Paris-Sud, CNRS, CEA
91190 - Gif-sur-Yvette, France

²IFB Core Cluster taskforce

June 2021

General information

Practical information:

- Dates: June 28th - 30th
- Location: Institut des Systèmes Complexes, 113 rue Nationale, 75013-Paris
- Courses: 9:00 to 17:30
- Meal: 12:30-14:00
- Pauses: 10:30-11:00 + 15:30-16:00
- 2 days of courses + 1 day of course building

Round table:

- Teachers
- Learners

Ressources:



- GitLab
- L^AT_EX

Training schedule

Day 1:

- Introduction to reproducibility
- History management (3 Practical Sessions, git, GitHub)
- Control your development environment (1 PS, CONDA)
- Encapsulation (2 PS, docker)

Day 2:

- Workflow (2 PS, SNAKEMAKE)
- Traceability with notebooks (2 PS, jupyter, zenodo)
- IFB resources (2 PS, slurm, singularity)
- Sharing and disseminating (GitHub, zenodo)
- Conclusion

Day 3:

- Empowerment and improvement of resources

Table of contents

- 1 Introduction to reproducibility
- 2 History management
- 3 Control your development environment
- 4 Workflow
- 5 Tracability with Notebook
- 6 IFB resources
- 7 Sharing and dissemination
 - Showcase your work
 - Release
 - Obtain a DOI
- 8 Conclusion
- 9 3rd Day

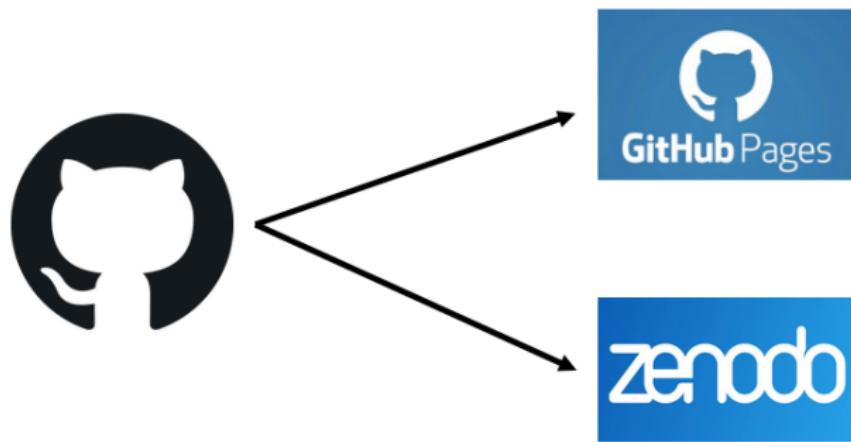


Sharing and disseminating your project

Sharing and disseminating

Goals of this session:

- Showcase your work
- Add a licence
- Create a release
- Obtain a DOI for the project



Showcase your work

Showcase your work



Showcase your work



Why ?

- Your project is simpler to share and find

Advantages

- Free hosting of static websites
- Able to convert Markdown into a website

Documentation : <https://pages.github.com/>

Showcase your work

In practice

From the main page of your repository, go to :

- "Settings" tab
- → "Options" (left hand side menu)
- → navigate to the "GitHub Pages" paragraph.

GitHub Pages

[GitHub Pages](#) is designed to host your personal, organization, or project pages from a GitHub repository.

Source

GitHub Pages is currently disabled. Select a source below to enable GitHub Pages for this repository. [Learn more](#).

[None ▾](#)

Theme Chooser

Select a theme to publish your site with a Jekyll theme using the `master` branch. [Learn more](#).

[Choose a theme](#)

Showcase your work

In practice

From the main page of your repository, go to "Settings" → "Options" → "GitHub Pages".

① Choose the source

GitHub Pages

GitHub Pages is designed to host your personal, organization, or project pages from a GitHub repository.

Source

GitHub Pages is currently disabled. Select a source below to enable GitHub Pages for this repository. [Learn more.](#)

None ▾

Save

Select branch

|| theme using the main branch. [Learn more.](#)

Select branch

main

✓ None

Danger Zone

Showcase your work

In practice

From the main page of your repository, go to "Settings" → "Options" → "GitHub Pages".

① Choose the source

GitHub Pages

GitHub Pages is designed to host your personal, organization, or project pages from a GitHub repository.

The screenshot shows the GitHub Pages settings interface. In the 'Source' section, there is a dropdown menu labeled 'Branch: main' and a 'Save' button. Below this, there is a 'Theme Chooser' section with a 'Choose a theme' button. A modal window titled 'Select folder' is open, showing two options: '/root' (which is checked) and '/docs'. The background of the page shows a message: 'GitHub Pages is currently disabled. Select a source below to enable GitHub Pages for this repository.' with a 'Learn more.' link.

Showcase your work

In practice

From the main page of your repository, go to "Settings" → "Options" → "GitHub Pages".

① Choose the source

GitHub Pages

GitHub Pages is designed to host your personal, organization, or project pages from a GitHub repository.

Source

GitHub Pages is currently disabled. Select a source below to enable GitHub Pages for this repository. [Learn more.](#)

 Branch: main ▾

 / (root) ▾

Save

Theme Chooser

Select a theme to publish your site with a Jekyll theme using the [main branch](#). [Learn more.](#)

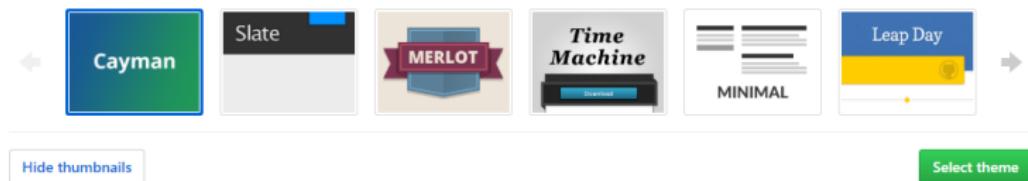
Choose a theme

Showcase your work

In practice

From the main page of your repository, go to "Settings" → "Options" → "GitHub Pages".

- ① Choose the source
- ② Choose the theme



Showcase your work

Convert Markdown into HTML !

```
# Welcome !

**Bienvenue à FAIR_bioinfo**

Vous trouverez ici des communications réalisées lors des sessions FAIR_bioinfo. Tous le contenu présenté existe déjà en anglais sur internet. Nous proposons donc des ressources pour les francophones.

*You will find here some communications made during the I2BC Bioinformatics Club. Communications will be mainly in French. All the content presented also exists in English on the Internet. Therefore, we propose here resources for Francophones.

**Informations pratiques**
- Quand ? : le dernier vendredi après midi
- Durée ? : 1h30 (questions incluses)
- Lieu ? : Salle de conférence A.Kalogeropoulos

**Objectifs**
L'objectif est de proposer et d'utiliser un panel d'outils permettant la réalisation d'un projet complet de bio-info en partant de rien et aboutissant à la création d'un conteneur (technologie Docker). Le partage, la valorisation et l'analyse dynamique des données seront inclus dans le panel. FAIR correspond à l'acronyme anglais "Findable, Accessible, Interoperable, & Reusable". Un panel initialement défini pour faire des analyses de séquençage peut être étendu pour faire leurs propres analyses. Le projet support est une étude "l'expression différentielle de gênes" à partir de données RNAseq d'Orsay.

**Pré-requis**
```



Welcome !

Bienvenue à FAIR_bioinfo

Vous trouverez ici des communications réalisées lors des sessions FAIR_bioinfo. Les communications sont en français. Tous le contenu présenté existe déjà en anglais sur internet. Nous proposons donc des ressources pour les francophones.

You will find here some communications made during the I2BC Bioinformatics Club. Communications will be mainly in French. All the content presented also exists in English on the Internet. Therefore, we propose here resources for Francophones.

Informations pratiques

- Quand ? : le dernier vendredi après midi de chaque mois (sauf juillet à définir), rdv 12h30
- Durée ? : 1h30 (questions incluses)
- Lieu ? : Salle de conférence A.Kalogeropoulos, b. 400, campus Orsay

Objectifs

L'objectif est de proposer et d'utiliser un panel d'outils permettant la réalisation d'un projet complet de bio-info en partant de rien et aboutissant à la création d'un conteneur (technologie Docker). Le partage, la valorisation et l'analyse dynamique des données seront inclus dans le panel. FAIR correspond à l'acronyme anglais "Findable, Accessible, Interoperable, & Reusable". Un panel initialement défini pour faire des analyses de séquençage peut être étendu pour faire leurs propres analyses. Le projet support est une étude "l'expression différentielle de gênes" à partir de données RNAseq d'Orsay.

Pré-requis

Question... Savoir taper sur un clavier ?

Contact

- Thomas DENECER (thomas.denecker@gmail.com)
- Claire Toffano-Noghe (claire.toffano-noghe@psud.fr)

Communications orales

https://thomasdenecker.github.io/FAIR_Bioinfo



Showcase your work

Also works directly from HTML

- ① Create a folder named "docs"
 - ▶ main file must be named index.html
- ② "Settings" → "Options" → "GitHub Pages"

GitHub Pages

GitHub Pages is designed to host your personal, organization, or project pages from a GitHub repository.

Your site is published at <https://thomasdenecker.github.io/bPeaks-application/>

Source
Your GitHub Pages site is currently being built from the `/docs` folder in the `master` branch. [Learn more.](#)

master branch /docs folder

Theme Chooser
Select a theme to publish your site with a Jekyll theme. [Learn more.](#)

Choose a theme

Custom domain
Custom domains allow you to serve your site from a domain other than `thomasdenecker.github.io`. [Learn more.](#)

Save

Enforce HTTPS
--- Required for your site because you are using the default domain (`thomasdenecker.github.io`)

HTTPS provides a layer of encryption that prevents others from snooping on or tampering with traffic to your site. When HTTPS is enforced, your site will only be served over HTTPS. [Learn more.](#)

<https://thomasdenecker.github.io/bPeaks-application/>



Showcase your work

Remember to choose a licence !

This will determine whether anyone can use, modify, and distribute your code / tool / software...

<https://help.github.com/en/articles/licensing-a-repository>

- ① Create a file named "LICENCE"
- ② GitHub will suggest templates



Showcase your work

Remember to choose a licence !

This will determine whether anyone can use, modify, and distribute your code / tool / software...

<https://help.github.com/en/articles/licensing-a-repository>

- ① Create a file named "LICENCE"
- ② GitHub will suggest templates



The CeCILL licence (v.2.1) is recommended by the CEA, CNRS and INRIA ("CEA CNRS INRIA Logiciel Libre"). Copy it directly.

<http://cecill.info/licences.fr.html>



Showcase your work

GitHub takes care of displaying the information on your repository.

The screenshot shows a GitHub repository page for 'thomasdenecker / Fair_Projet'. The top navigation bar includes 'Code', 'Issues', 'Pull requests', 'Projects', 'Wiki', 'Security', 'Insights', and 'Settings'. Below the navigation, there's a section titled 'Add a license to your project' with a list of available licenses:

- Apache License 2.0
- GNU General Public License v3.0
- MIT License
- BSD 2-Clause "Simplified" License
- BSD 3-Clause "New" or "Revised" License
- Eclipse Public License 2.0
- GNU Affero General Public License v3.0
- GNU General Public License v2.0
- GNU Lesser General Public License v2.1
- GNU Lesser General Public License v3.0
- Mozilla Public License 2.0
- The Unlicense

For the 'Apache License 2.0' option, detailed information is shown:

A permissive license similar to the BSD 2-Clause License, but with a 3rd clause that prohibits others from using the name of the project or its contributors to promote derived products without written consent.

Permissions	Limitations	Conditions
<input checked="" type="checkbox"/> Commercial use	<input checked="" type="checkbox"/> Liability	<input checked="" type="checkbox"/> License and copyright notice
<input checked="" type="checkbox"/> Modification	<input checked="" type="checkbox"/> Warranty	
<input checked="" type="checkbox"/> Derivatives		
<input checked="" type="checkbox"/> Private use		

This is not legal advice. Learn more about repository licenses.

For the 'BSD 3-Clause License' option, the details are as follows:

BSD 3-Clause License

Copyright © 2010. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
3. Neither the name of the copyright holder nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED

To adopt BSD 3-Clause "New" or "Revised" License, enter your details. You'll have a chance to review before committing a LICENSE file to a new branch or the root of your project.

Year: 2010
Full name: thomasdenecker

Review and submit



Showcase your work

Validate and merge with the main branch

thomasdenecker/Fair_Projet is licensed under the BSD 3-Clause "New" or "Revised" License		Permissions	Limitations	Conditions
	A permissive license similar to the BSD 2-Clause License, but with a 3rd clause that prohibits others from using the name of the project or its contributors to promote derived products without written consent.	<ul style="list-style-type: none">✓ Commercial use✓ Modification✓ Distribution✓ Private use	<ul style="list-style-type: none">✗ Liability✗ Warranty	<ul style="list-style-type: none"> ⓘ License and copyright notice
This is not legal advice. Learn more about repository licenses.				

Showcase your work

Kind reminder: consequence of not choosing a licence.

No License

When you make a creative work (which includes code), the work is under exclusive copyright by default. Unless you include a license that specifies otherwise, nobody else can copy, distribute, or modify your work without being at risk of take-downs, shake-downs, or litigation. Once the work has other contributors (each a copyright holder), "nobody" starts including you.

Even in the absence of a license file, you may grant some rights in cases where you publish your source code to a site that requires accepting terms of service. For example, if you publish your source code in a public repository on GitHub, you have accepted the [Terms of Service](#), by which you allow others to view and fork your repository. Others may not need your permission if [limitations and exceptions to copyright](#) apply to their particular situation. Neither site terms nor jurisdiction-specific copyright limitations are sufficient for the kinds of collaboration that people usually seek on a public code host, such as experimentation, modification, and sharing as fostered by an open source license.

You don't have to do anything to not offer a license. You may however wish to add a copyright notice and statement that you are not offering any license in a prominent place (e.g., your project's README) so that [users](#) don't assume you made an oversight. If you're going to accept others' contributions to your non-licensed project, you may wish to explore with your lawyer adding a contributor agreement to your project so that you maintain copyright permission from contributors, even though you're not granting the same.

Disallowing use of your code might not be what you intend by "no license." An [open-source license](#) allows reuse of your code while retaining copyright. If your goal is to completely opt-out of copyright restrictions, try a [public domain dedication](#).

For users

If you find software that doesn't have a license, that generally means you have no permission from the creators of the software to use, modify, or share the software. Although a code host such as GitHub may allow you to view and fork the code, this does not imply that you are permitted to use, modify, or share the software for any purpose.

Your options:

- Ask the maintainers nicely to add a license. Unless the software includes strong indications to the contrary, lack of a license is probably an oversight. If the software is hosted on a site like GitHub, open an issue requesting a license and include a link to this site. If you're bold and it's fairly obvious what license is most appropriate, open a pull request to add a license – see "suggest this license" in the sidebar of the page for each license on this site (e.g., [MIT](#)).
- Don't use the software. Find or create an alternative that is under an open source license.
- Negotiate a private license. Bring your lawyer.

Release

Release

Goal : provide users with a version of your code that has been fixed in time and labelled.

All the steps are detailed here:

- <https://help.github.com/en/articles/creating-releases>

Release

Make a release

The screenshot shows a GitHub repository page for `thomasdenecker / FAIR_Bioinfo`. The top navigation bar includes options for Watch (3), Star (3), Fork (1), and Settings. Below the header, there's a brief description: "Démonstration d'outils de bioinfo dans le cadre d'un projet". A red box highlights the "0 releases" count in the repository summary bar. A large black arrow points down from this bar to a modal window titled "Create a new release". The modal contains a small tag icon and the text: "There aren't any releases here. Releases are powered by [tagging specific points of history](#) in a repository. They're great for marking release points like v1.0." At the bottom of the modal is a "Create a new release" button.

thomasdenecker / FAIR_Bioinfo

Watch 3 Star 3 Fork 1

Code Issues 0 Pull requests 0 Projects 1 Wiki Security Insights Settings

Démonstration d'outils de bioinfo dans le cadre d'un projet

Edit

Manage topics

57 commits 2 branches 0 releases 1 environment 2 contributors View license

There aren't any releases here

Releases are powered by [tagging specific points of history](#) in a repository. They're great for marking release points like v1.0.

Create a new release



Release

Make a release

thomasdenecker / FAIR_Bioinfo

Watch 3 Star 3 Fork 1

Code Issues 0 Pull requests 0 Projects 1 Wiki Security Insights Settings

Releases Tags

Tag version Target: master

Choose an existing tag, or create a new tag on publish

Release title

Write Preview

Describe this release

Attach files by dragging & dropping, selecting or pasting them.

Attach binaries by dropping them here or selecting them.

This is a pre-release

We'll point out that this release is identified as non-production ready.

Publish release Save draft

Tagging suggestions

It's common practice to prefix your version names with the letter v.

Some good tag names might be v1.0 or v2.3.4.

If the tag isn't meant for production use, add a pre-release version after the version name.

Some good pre-release versions might be v0.2-alpha or v5.9-beta.3.

Semantic versioning

If you're new to releasing software, we highly recommend reading about [semantic versioning](#).

Release

Semantic of a release number

1.0.0
MAJOR.MINOR.PATCH

- MAJOR : changes not backwards-compatible
- MINOR : new/modified functionalities, backwards-compatible
- PATCH : bug fixes, backwards-compatible

More details : <https://semver.org/>

Release

First release for FAIR_Bioinfo

Releases Tags Edit release Delete

Latest release

1.0.0
thomasdenecker released this just now
1e307ed

1.0.0
Update Readme

Assets 2

Source code (zip)
Source code (tar.gz)

Obtain a DOI

Obtain a DOI



Obtain a DOI

Digital Object Identifier

Reference system to cite an object (A GitHub project in our case)



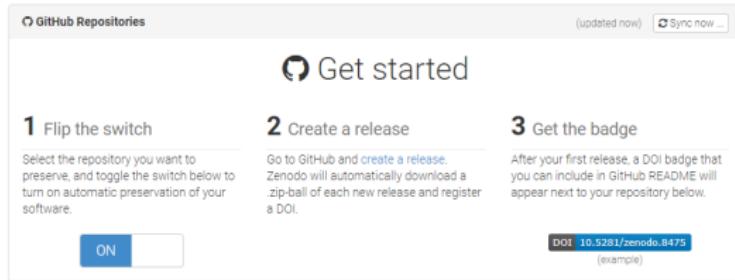
<https://guides.github.com/activities/citable-code/>

Obtain a DOI

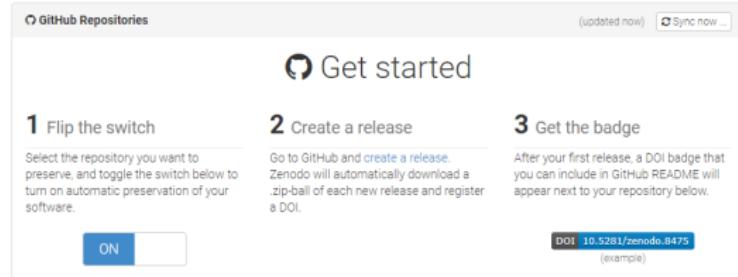
1/ Sign in to Zenodo

- With your GitHub account
- With your ORCID account (add a "Linked account" to GitHub afterwards)

2/ Go to the Settings page → GitHub tab



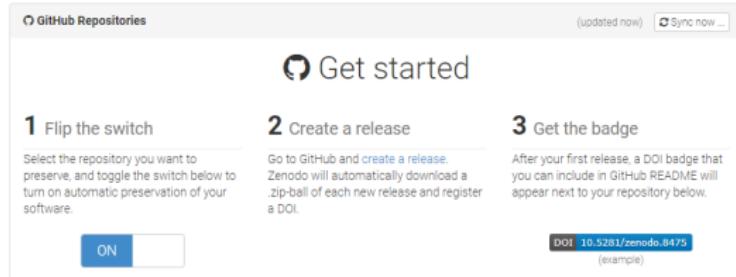
Obtain a DOI



3/ In the list below, find the project you want to link to Zenodo. Flip the switch.

 thomasdenecker/FAIR_Bioinfo 

Obtain a DOI

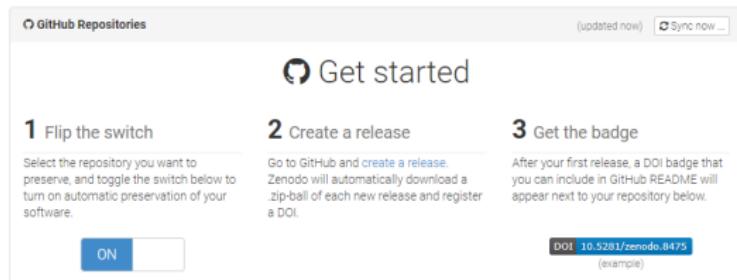


4/ On GitHub, in Settings → Webhooks, a new line has been created: Zenodo will be notified of any new release created in this project.

The screenshot shows the GitHub Webhooks settings page with one webhook configuration listed:

Action	URL	Edit	Delete
✓	https://zenodo.org/api/hooks/receivers/github/events/ (release)	Edit	Delete

Obtain a DOI



5/ Back to Zenodo. After a release, a badge will be available below the project's name, in the category Enabled repositories.



Obtain a DOI

DOI Badge

This badge points to the latest released version of your repository. If you want a DOI badge for a specific release, please follow the DOI link for one of the specific releases and grab badge from the archived record.

Markdown

```
[ ! [DOI](https://zenodo.org/badge/164655551.svg)](https://zenodo.org/badge/late
```

reStructuredText

```
.. image:: https://zenodo.org/badge/164655551.svg  
:target: https://zenodo.org/badge/latestdoi/164655551
```

HTML

```
<a href="https://zenodo.org/badge/latestdoi/164655551">  
<https://github.com/features/package-registry>

