Panorama des outils et plateformes pour déposer ses codes

Semaine Data-SHS Paris 2022

Timothée Giraud

Dec 6, 2022

Pourquoi déposer ses codes de recherche ?

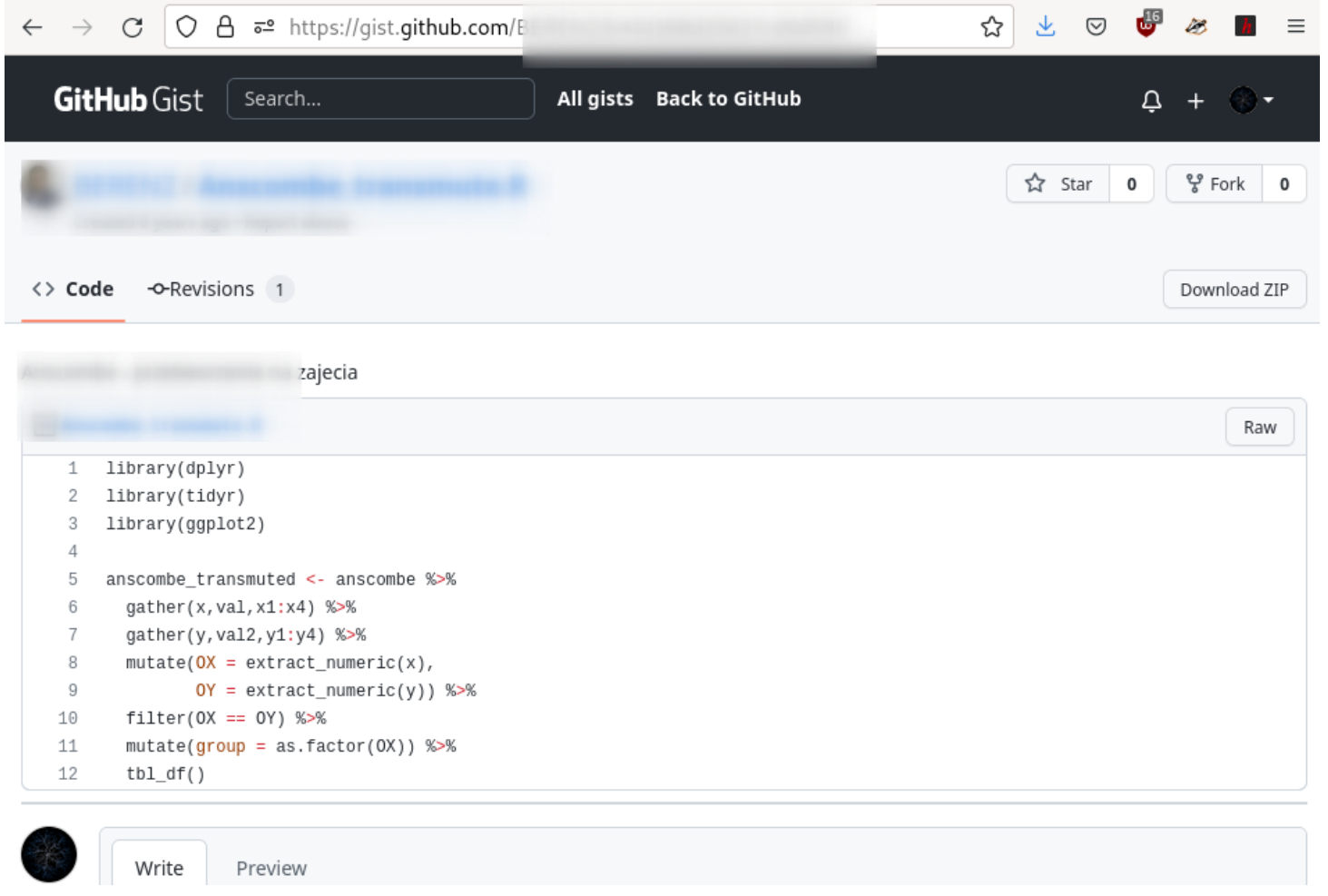
Pour soi-même, pour les autres, pour la science !

- Réutilisation
- Traçabilité
- Collaboration
- Transparence et vérifiabilité

Que déposer ?

Des fragments de code

Un pense-bête sur une opération spécifique.



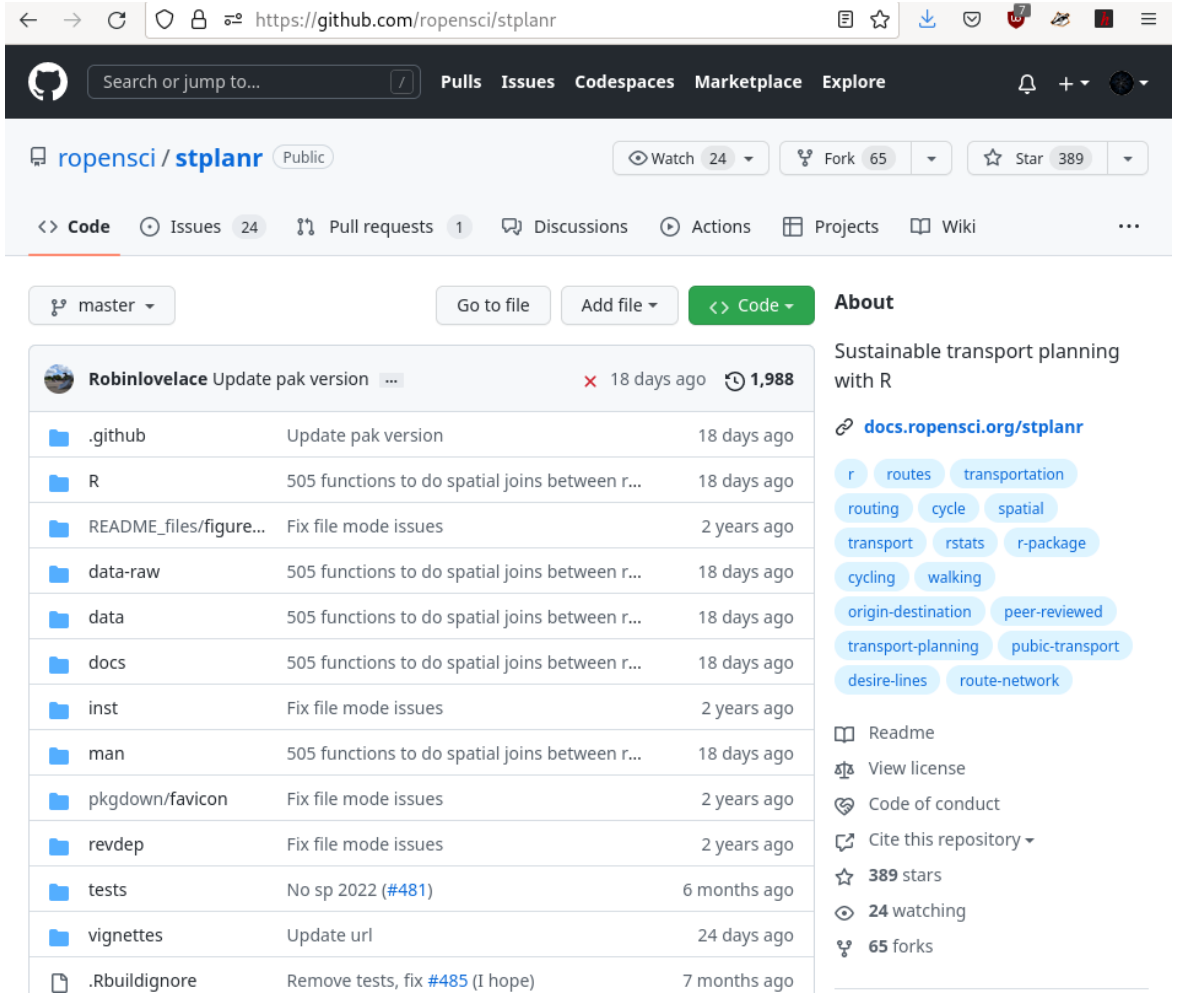
Exemple de gist

Des analyses completes

Le code utilisé pour créer les analyses, les figures d'un article.

Des logiciels

Code construit pour la réutilisation, généricisé.



Les plateformes et les outils

Les logiciels de gestion de versions

Git, GitHub, GitLab

Git est un logiciel de gestion de versions distribué pour

- suivre les changements dans les fichiers texte,
- gérer l'historique du code source,
- partager le code avec des dépots distants.

GitHub et **GitLab** hébèrgent des dépots distants + des services pour gérer des projets (issue tracking, collaboration, hébergement web, CI/CD).

GitHub, GitLab

GitHub est une entreprise de Microsoft qui produit et héberge un logiciel non libre (ou privateur) sur le site github.com.

GitLab est une entreprise qui produit un logiciel libre (GitLab CE) et un logiciel non libre (GitLab EE & gitlab.com).

GitLab

Huma-Num
la TGIR des humanités numériques

GitLab : hébergement, versionning et partage de code

L'instance GitLab d'Huma-Num permet l'hébergement sécurisé et le partage maîtrisé de fichiers de code selon le protocole *git*.

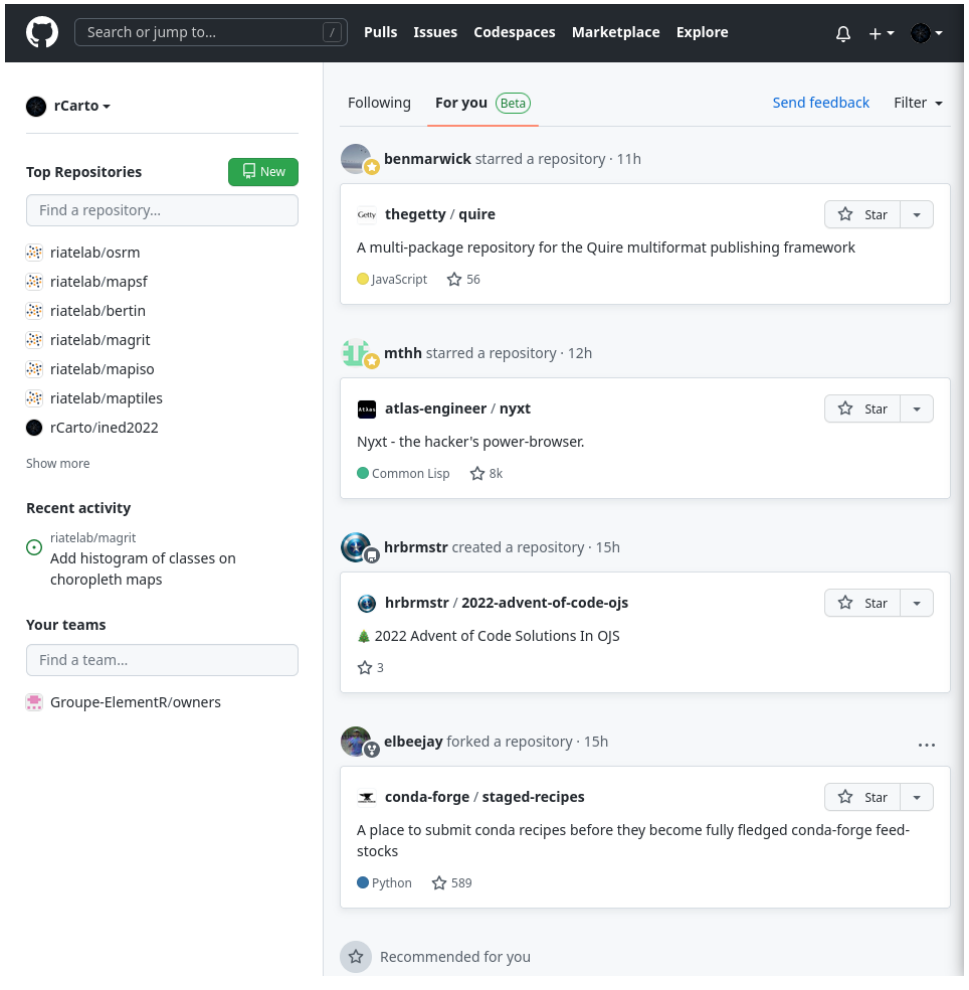
Il s'agit d'une implémentation du logiciel Gitlab.

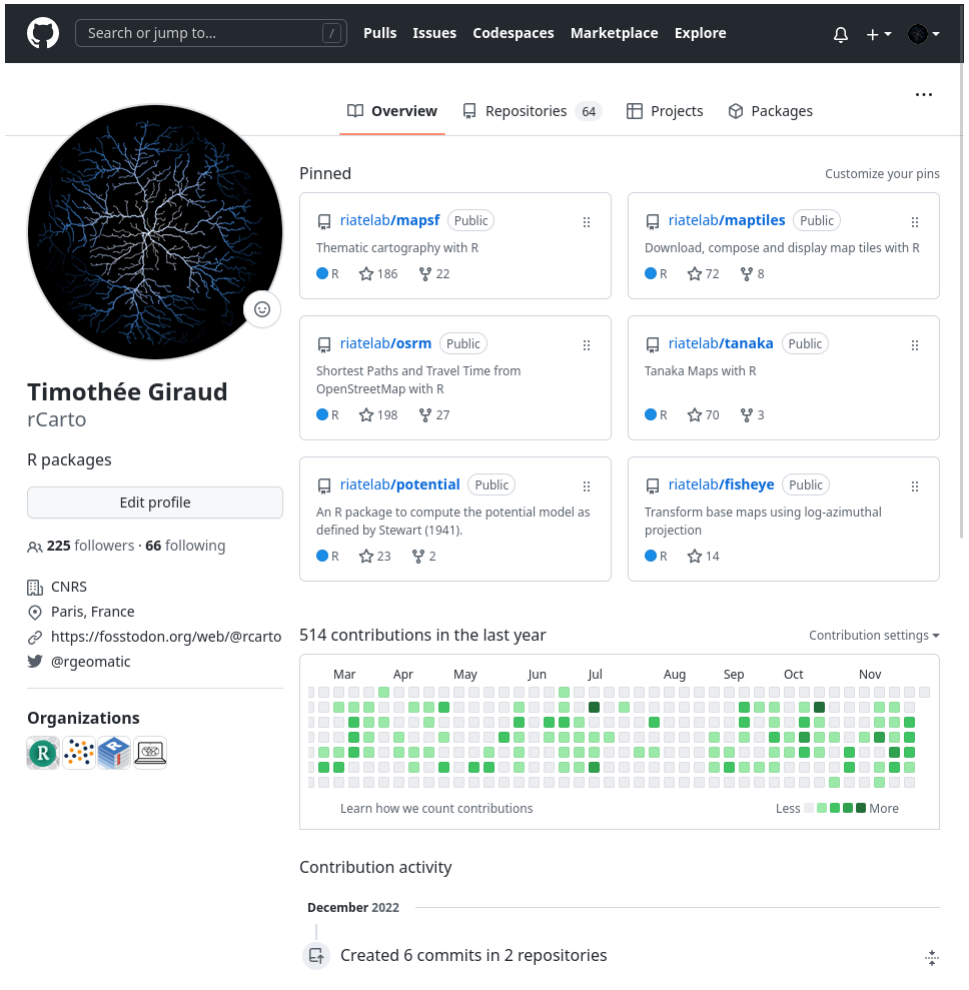
Les principales fonctionnalités sont la gestion de version et des dépôts (git), l'intégration continue, la génération de sites web (pages), la gestion de tickets (issues).

Demander l'ouverture d'un compte Gitlab : la demande d'un compte Gitlab se fait à partir de l'interface HumanID. Pour cela, le cas échéant il est nécessaire de disposer d'un compte HumanID (voir la documentation).

GitHub

GitHub est un peu le réseau social des développeurs.
Sa centralisation et la masse de ses utilisateurs sont deux éléments qui facilitent les collaborations et donnent une bonne visibilité aux projets.


Search or jump to...
Pulls
Issues
Codespaces
Marketplace
Explore
rCarto
Top Repositories
New
Find a repository...
riatelab/osrm
riatelab/mapsf
riatelab/bertin
riatelab/magrit
riatelab/mapiso
riatelab/maptiles
rCarto/ined2022
Show more
Recent activity
riatelab/magrit
Add histogram of classes on choropleth maps
Your teams
Find a team...
Groupe-ElementR/owners
Following
For you
Beta
Send feedback
Filter
benmarwick starred a repository · 11h
thegetty / quire
Star
A multi-package repository for the Quire multiformat publishing framework
JavaScript
56
mthh starred a repository · 12h
atlas-engineer / nyxt
Star
Nyxt - the hacker's power-browser.
Common Lisp
8k
hrbrmstr created a repository · 15h
hrbrmstr / 2022-advent-of-code-ojs
Star
2022 Advent of Code Solutions In OJS
3
elbeejay forked a repository · 15h
conda-forge / staged-recipes
Star
A place to submit conda recipes before they become fully fledged conda-forge feed-stocks
Python
589
Recommended for you


Search or jump to...
Pulls
Issues
Codespaces
Marketplace
Explore
Overview
Repositories 64
Projects
Packages
Timothée Giraud
rCarto
R packages
Edit profile
225 followers · 66 following
CNRS
Paris, France
https://fosstodon.org/web/@rcarto
@rgeomatic
Organizations
Pinned
Customize your pins
riatelab/mapsf Public
Thematic cartography with R
R 186 22
riatelab/maptiles Public
Download, compose and display map tiles with R
R 72 8
riatelab/osrm Public
Shortest Paths and Travel Time from OpenStreetMap with R
R 198 27
riatelab/tanaka Public
Tanaka Maps with R
R 70 3
riatelab/potential Public
An R package to compute the potential model as defined by Stewart (1941).
R 23 2
riatelab/fisheye Public
Transform base maps using log-azimuthal projection
R 14
514 contributions in the last year
Contribution settings
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Learn how we count contributions
Less
More
Contribution activity
December 2022
Created 6 commits in 2 repositories

Git, GitHub, GitLab

Ces plateformes sont spécifiquement dédiées au developpement de logiciels.

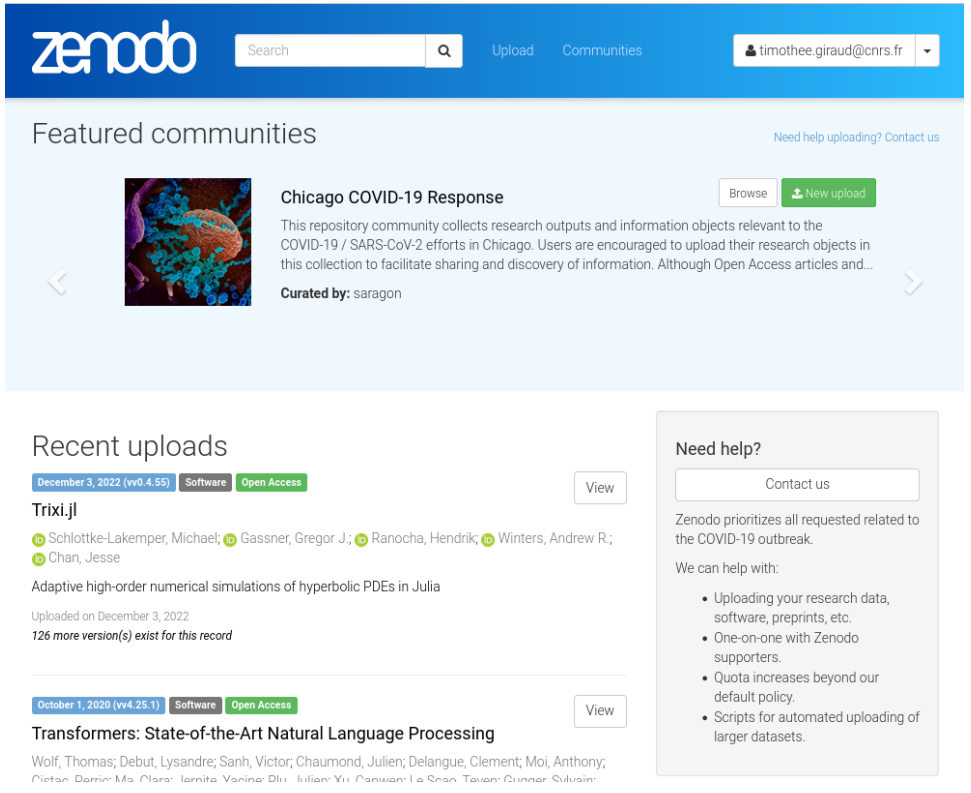
Ce sont des services, parfois commerciaux, sans aucune garantie de perénité.

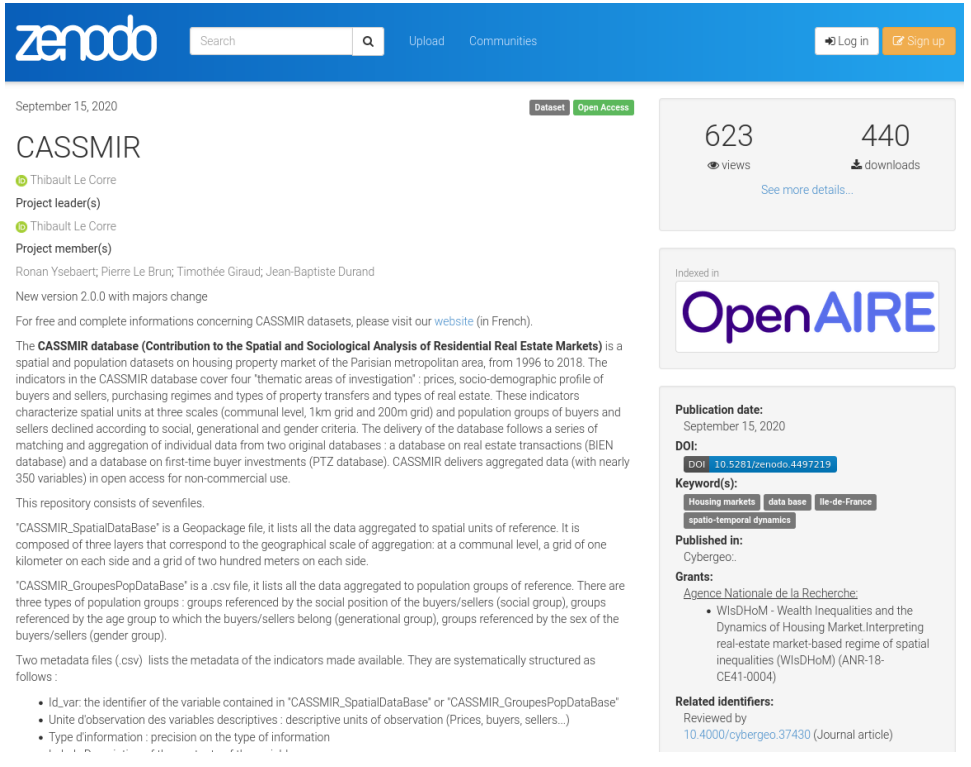
Il ne s’agit en aucun cas de services d’archivage.

Les services d’archivage

Ici l’objectif est bien d’archiver les codes, les bases de données, les figures…

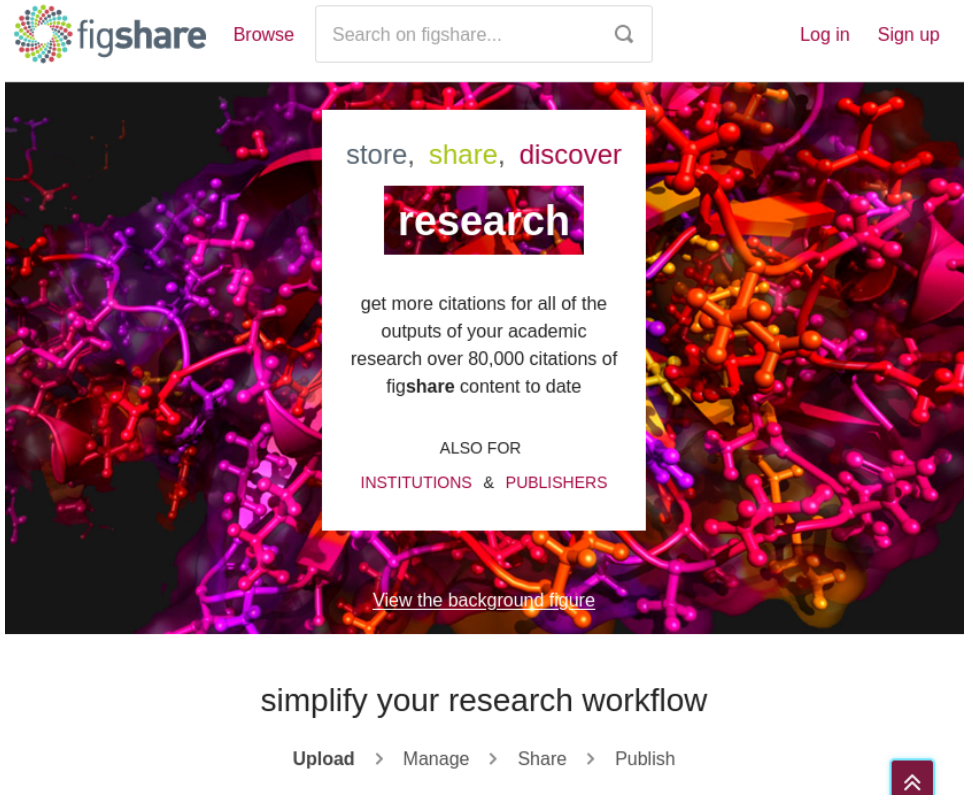
Zenodo

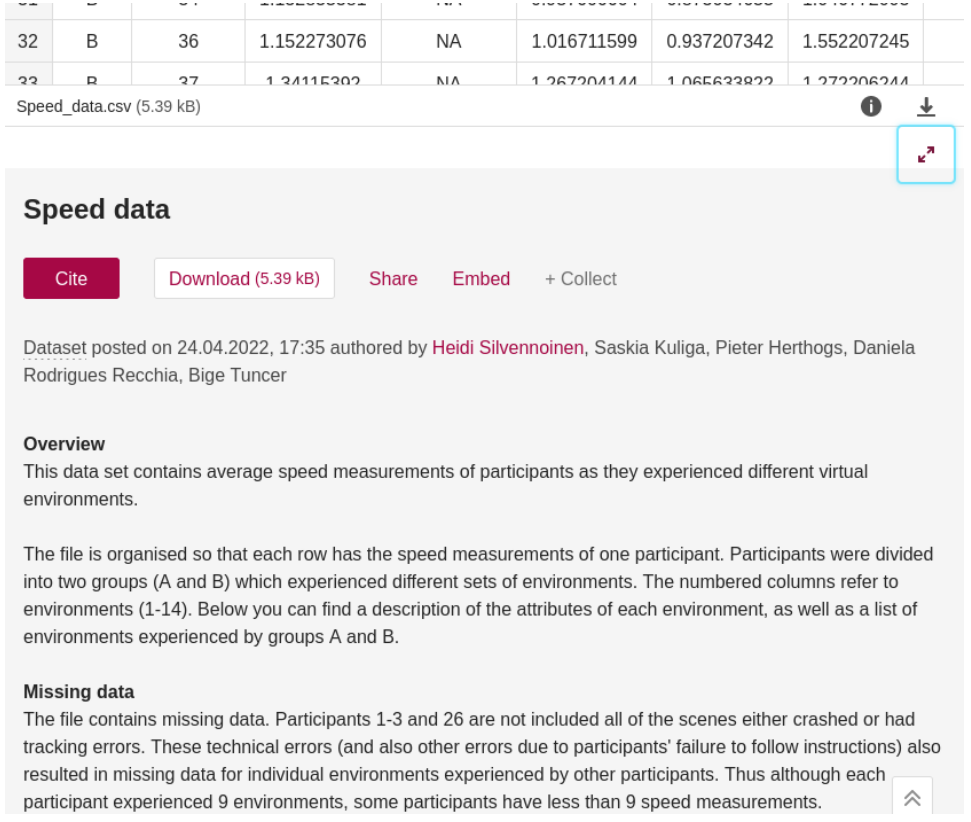




Un service développé et hébergé par le CERN.

Figshare





Un service commercial de l’entreprise qui possède le groupe Nature Publishing.

D’autres services

- DRYAD
  CONNECTING THE NATURAL WORLD
- HAL
- nakala
- Software Heritage
  THE GREAT LIBRARY OF SOURCE CODE
- …

Les répertoires d'extensions de logiciel

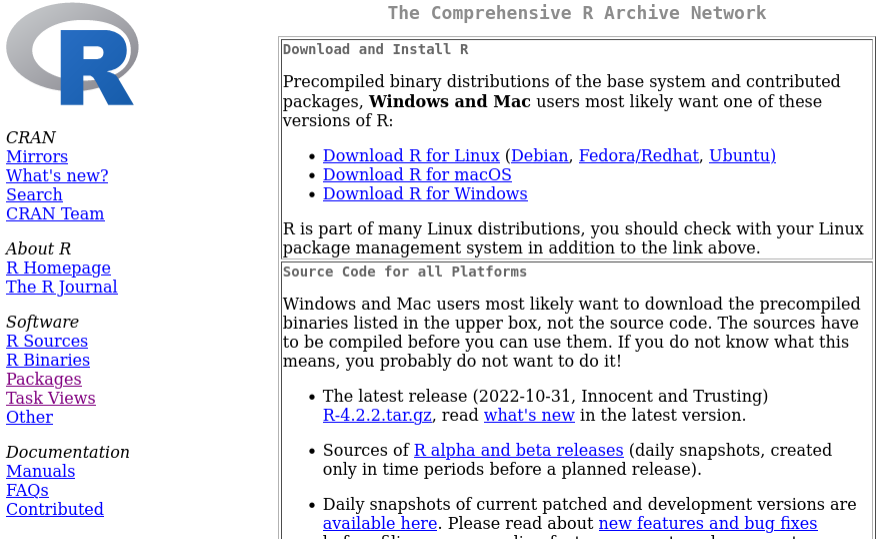
Différents languages ou logiciels proposent des répertoires pour la mise à disposition d'extensions.

Les extensions sont proposées de manière plus ou moins harmonisée.

Un accès est souvent donné aux versions anciennes.

Les répertoires d'extensions de logiciel

Pour les librairies R, les *packages* : le **CRAN** (**C**omprehensive **R A**rchive **N**etwork).



CRAN
Mirrors
What's new?
Search
CRAN Team

About R
R Homepage
The R Journal

Software
R Sources
R Binaries
Packages
Task Views
Other

Documentation
Manuals
FAQs
Contributed

The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

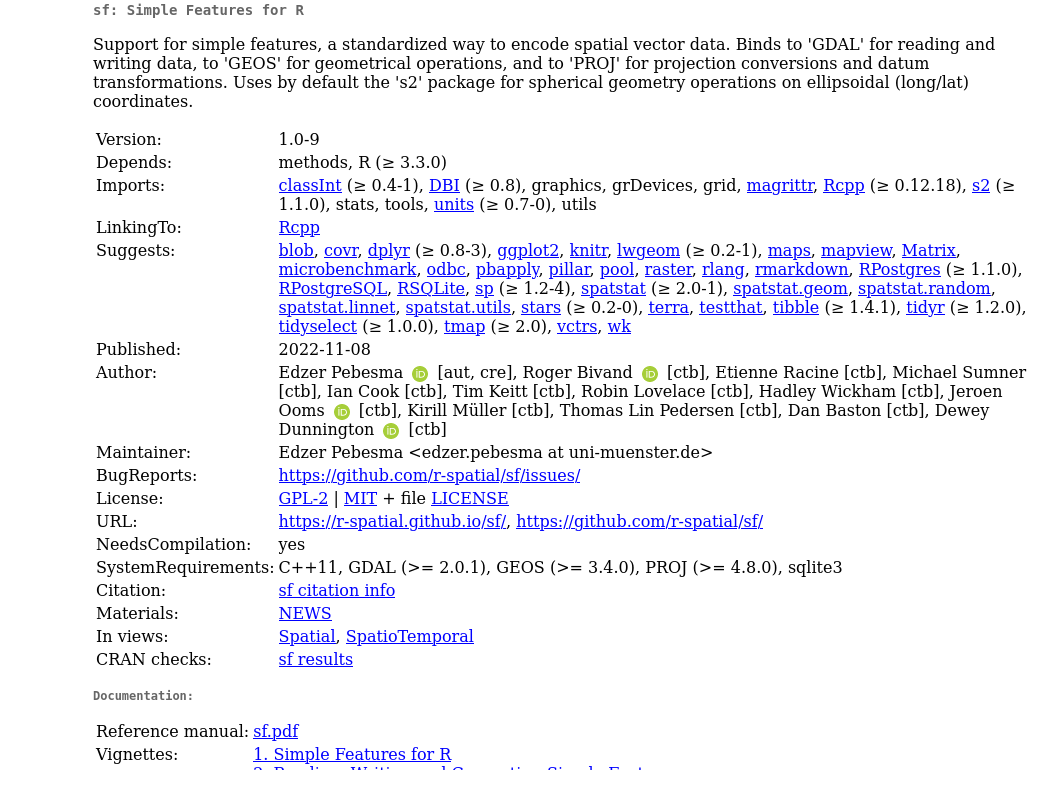
- Download R for Linux (Debian, Fedora/Redhat, Ubuntu)
- Download R for macOS
- Download R for Windows

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (2022-10-31, Innocent and Trusting) R-4.2.2.tar.gz, read what's new in the latest version.
- Sources of R alpha and beta releases (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are available here. Please read about new features and bug fixes



sf: Simple Features for R

Support for simple features, a standardized way to encode spatial vector data. Binds to 'GDAL' for reading and writing data, to 'GEOS' for geometrical operations, and to 'PROJ' for projection conversions and datum transformations. Uses by default the 's2' package for spherical geometry operations on ellipsoidal (long/lat) coordinates.

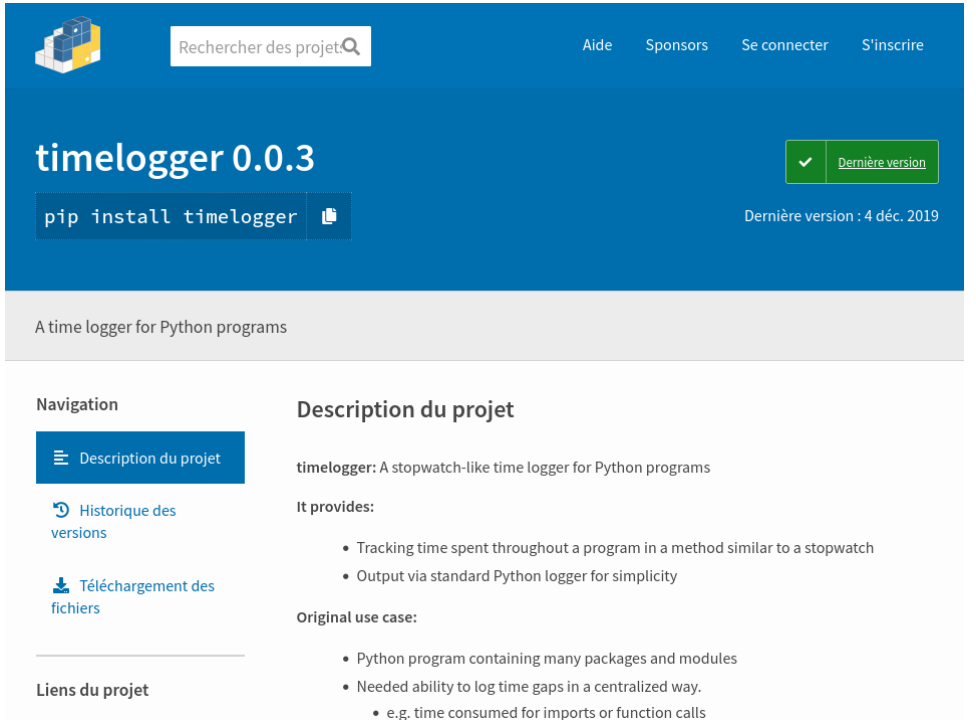
Version:	1.0-9
Depends:	methods, R (≥ 3.3.0)
Imports:	classInt (≥ 0.4-1), DBI (≥ 0.8), graphics, grDevices, grid, magrittr, Rcpp (≥ 0.12.18), s2 (≥ 1.1.0), stats, tools, units (≥ 0.7-0), utils
LinkingTo:	Rcpp
Suggests:	blob, covr, dplyr (≥ 0.8-3), ggplot2, knitr, lwgeom (≥ 0.2-1), maps, mapview, Matrix, microbenchmark, odbc, pbapply, pillar, pool, raster, rlang, rmarkdown, RPostgres (≥ 1.1.0), RPostgreSQL, RSQLite, sp (≥ 1.2-4), spatstat (≥ 2.0-1), spatstat.geom, spatstat.random, spatstat.linnet, spatstat.utils, stars (≥ 0.2-0), terra, testthat, tibble (≥ 1.4.1), tidyr (≥ 1.2.0), tidyselect (≥ 1.0.0), tmap (≥ 2.0), vctrs, wk
Published:	2022-11-08
Author:	Edzer Pebesma [aut, cre], Roger Bivand [ctb], Etienne Racine [ctb], Michael Sumner [ctb], Ian Cook [ctb], Tim Keitt [ctb], Robin Lovelace [ctb], Hadley Wickham [ctb], Jeroen Ooms [ctb], Kirill Müller [ctb], Thomas Lin Pedersen [ctb], Dan Baston [ctb], Dewey Dunnington [ctb]
Maintainer:	Edzer Pebesma <edzer.pebesma at uni-muenster.de>
BugReports:	https://github.com/r-spatial/sf/issues/
License:	GPL-2 | MIT + file LICENSE
URL:	https://r-spatial.github.io/sf/, https://github.com/r-spatial/sf/
NeedsCompilation:	yes
SystemRequirements:	C++11, GDAL (>= 2.0.1), GEOS (>= 3.4.0), PROJ (>= 4.8.0), sqlite3
Citation:	sf citation info
Materials:	NEWS
In views:	Spatial, SpatioTemporal
CRAN checks:	sf results

Documentation:

Reference manual:	sf.pdf
Vignettes:	1. Simple Features for R

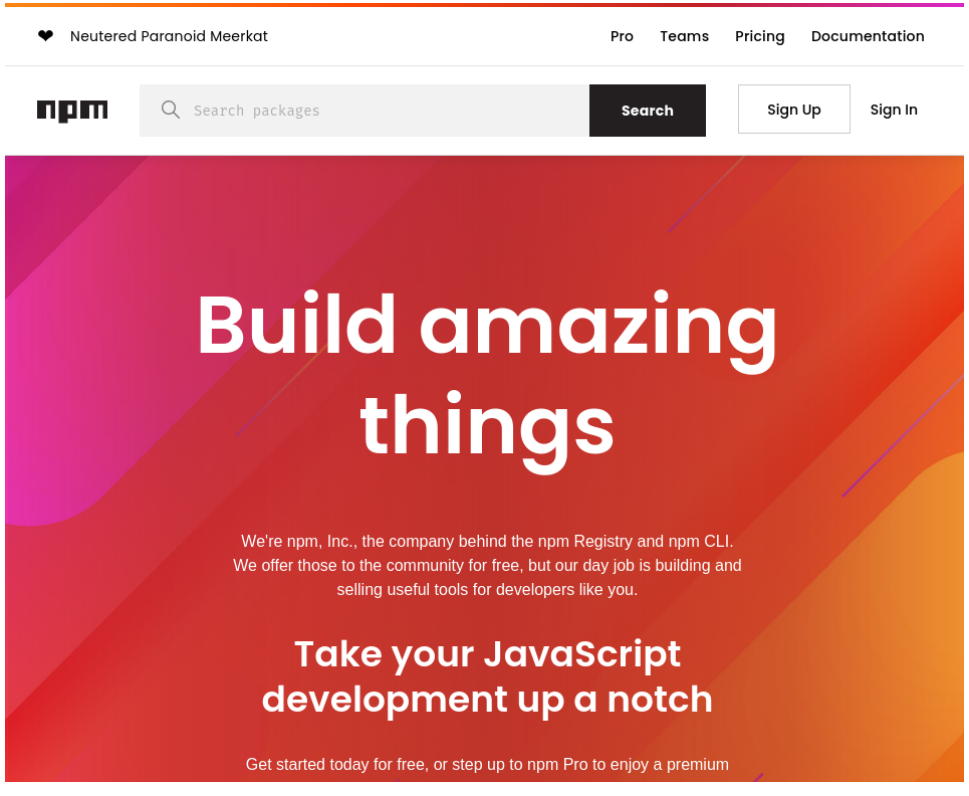
Les répertoires d'extensions de logiciel

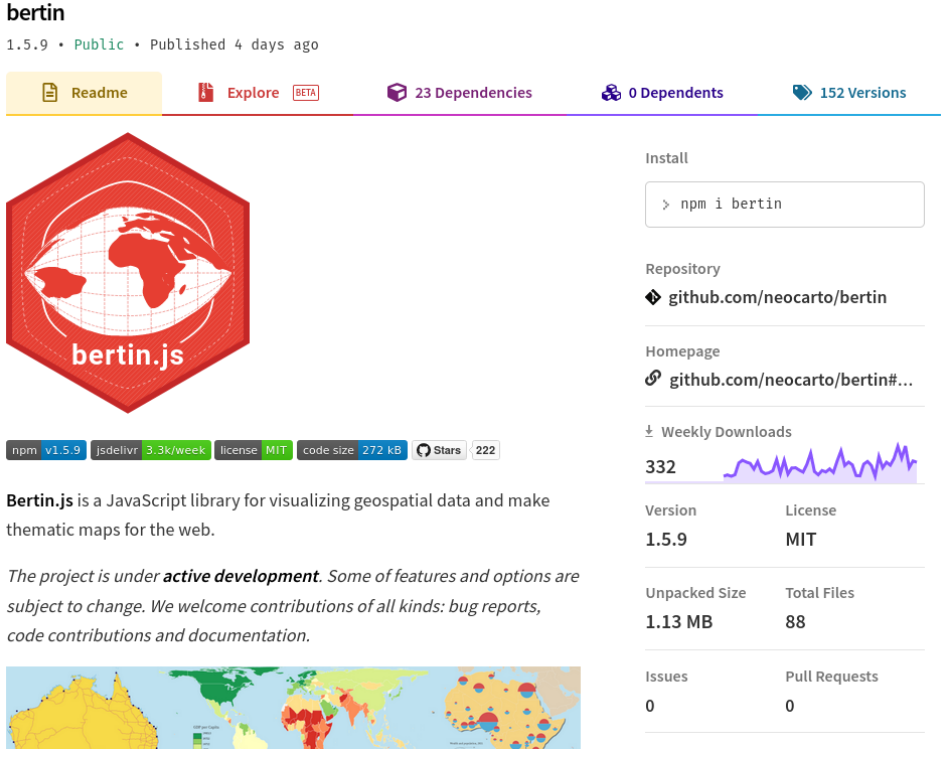
Pour les librairies Python : **PyPi** (**Py**thon **P**ackage **I**ndex).



Les répertoires d'extensions de logiciel

Pour les librairies Javascript : **npm** (**n**pm is not a **p**ackage **m**anager).





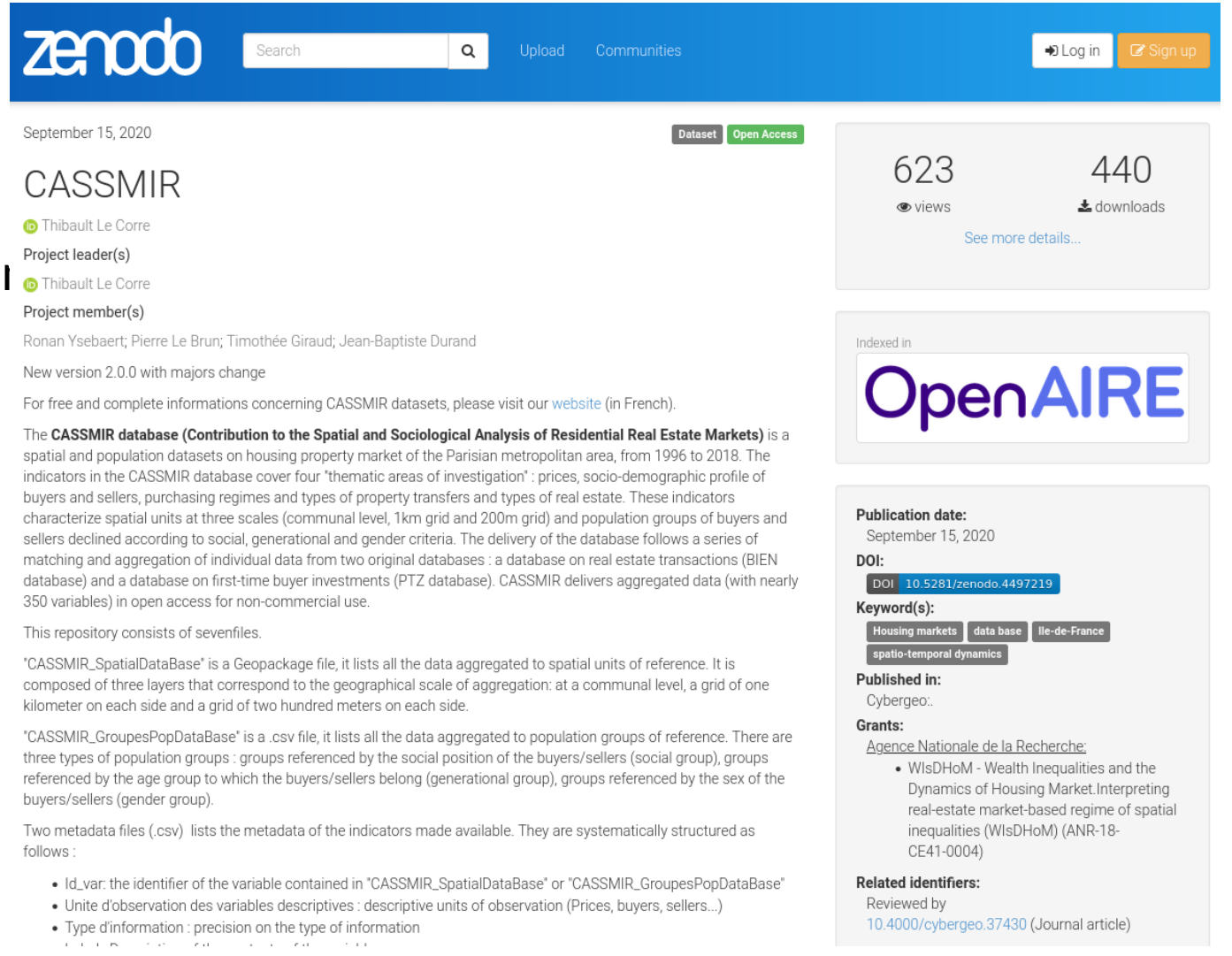
Exemple

Un (data) paper...

Thibault Le Corre. (2021). Une base de données pour étudier vingt années de dynamiques du marché immobilier résidentiel en Île-de-France, In *Cybergeo: European Journal of Geography*, Data papers, document 992. https://doi.org/10.4000/cyb

portant sur une base de données…

Thibault Le Corre. (2020). CASSMIR [Data set]. In Cybergeo (2.0.0). Zenodo. https://doi.org/10.5281/zen



dont la construction technique est décrite par un site web / notebook…

Génération d'échantillons de données individuelles

Afin de permettre à l'utilisateur de prendre connaissance des données sources, et dans un souci de reproductibilité du travail, des échantillons tests sont diffusés. La production de ces échantillons est contrainte par les restrictions d'accès aux données, pour des questions de confidentialité de la donnée individuelle mais, aussi, par l'interdiction de délivrer une donnée qui ne respecterait pas la propriété intellectuelle des services producteurs de la donnée. Ces deux contraintes peuvent être surmontées en proposant des données qui n'offrent pas la possibilité reconstituer la réalité. L'objectif est ainsi de donner accès à un échantillon des données BIEN et PTZ en passant par une anonymysation des données individuelles et la génération d'un espace fictif.

En dehors de l'import des bases de données sources non anonymisées, l'intégralité du code ci-dessous est exécutable en important les échantillons de données individuelles anonymisées

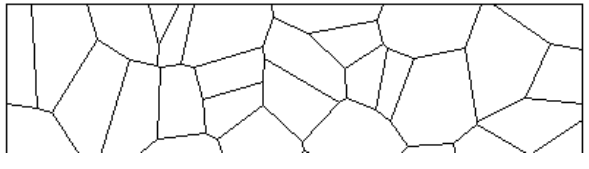
Aucune information produite à partir de ce jeu de données anonymisées serait suceptible d'être utilisée comme une expression de "l'état réel du marché". Ce jeu de données est spécialement dédié à la reproductibilité du travail de construction de la base de données CASSMIR et aucun autre usage peut en être fait.

```
library(sf)
library(tidyverse)
library(ade4)
library(cartography)
library(ggplot2)
library(SpatialPosition)
```

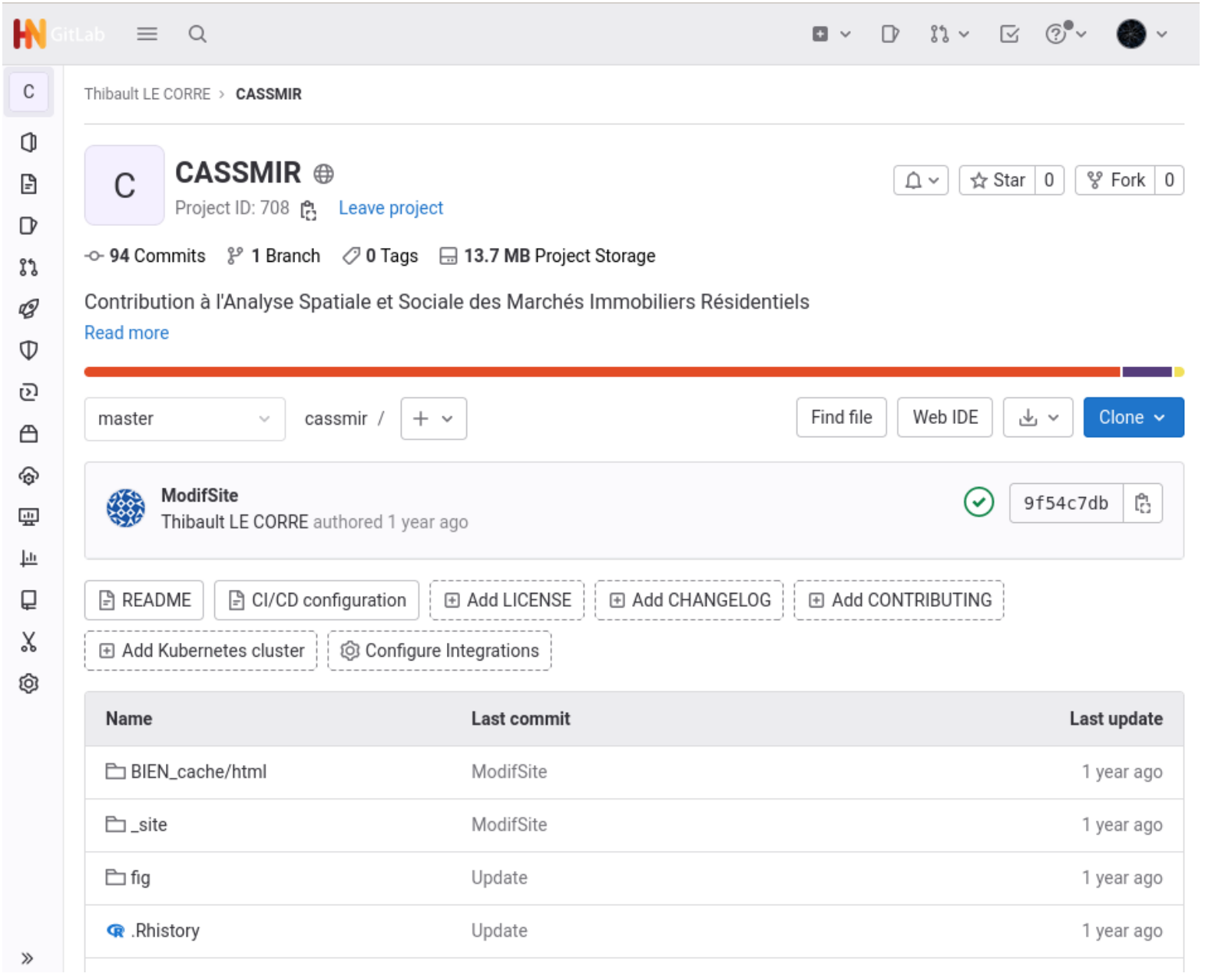
Générer un espace fictif

Cette première étape a pour objectif de créer un espace discret qui ne représente aucun autre espace dans la réalité. Cet espace est contraint par des limites en absicces et ordonnées distantes de 20 km. Cet espace est ensuite découpé par la génération de polygones de Voronoï qui symbolisent des limites administratives communales. A ce découpage s'ajoute la création de deux grilles carroyées dont la taille équivaut à celle des carreaux Insee de 200m et de 1Km de côté.

Exemple d'un espace fictif



hébergé sur une forge.


GitLab
Thibault LE CORRE › CASSMIR
C
CASSMIR
Project ID: 708
Leave project
Star 0
Fork 0
94 Commits
1 Branch
0 Tags
13.7 MB Project Storage
Contribution à l'Analyse Spatiale et Sociale des Marchés Immobiliers Résidentiels
Read more
master
cassmir /
Find file
Web IDE
Clone
ModifSite
Thibault LE CORRE authored 1 year ago
9f54c7db
README
CI/CD configuration
Add LICENSE
Add CHANGELOG
Add CONTRIBUTING
Add Kubernetes cluster
Configure Integrations
Name
Last commit
Last update
BIEN_cache/html
ModifSite
1 year ago
_site
ModifSite
1 year ago
fig
Update
1 year ago
.Rhistory
Update
1 year ago

Merci de votre attention