Notre système de revue en revue!

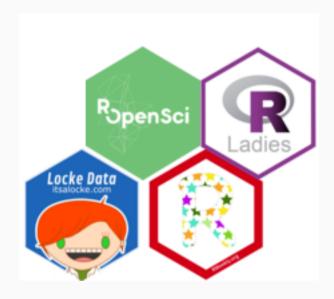
Présenter & analyser rOpenSci onboarding

Maëlle Salmon, éditrice pour rOpensci onboarding

🏏 ma_salmon 🜍 maelle 🌐 masalmon.eu

2018/06/23

Ma place dans la communauté R



- Travail pour rOpenSci et Locke Data
- R-Ladies Global social master
- Membre de l'équipe de R Weekly
- Éditrice bénévole pour rOpenSci

Qu'est-ce rOpenSci?

Qu'est-ce rOpenSci?

- Communautés de chercheur·se·s et de développeur·se·s de logiciels
- Paquets R pour la science ouverte et reproductible
- Contributions par le personnel et la communauté

https://ropensci.org/packages/

https://ropensci.org/packages/

FILTERS All All Altmetrics Tools Visualization Databases Geospatial Web Images Processing Computing Infrastructure Taxonomy Altmetrics O Altmetrics Web C tools Calculate tools CRAN / BIOC available

□ Type to search					
PACKAGE	MAINTAINER	DESCRIPTION	DETAILS		
iheatmapr	Alicia Schep	Interactive, Complex Heatmaps	CRAN (
plotdap	Carson Sievert	Easily visualize data from 'ERDDAP' Servers via the rerddap package	CRAN (
plotly	Carson	Programmatic interface to the Plot.ly API -construct plots	CRAN ()		

https://ropensci.org/packages/

Taxonomy

FILTERS All Data Publication Visualization Geospatial Images Processing Computing Infrastructure

- O Altmetrics
 O Tools
 O Databases
 O Web
 O Literature
 O Security
 □ CRAN / BIOC available
- Type to search... **PACKAGE** DESCRIPTION MAINTAINER **DETAILS** Rory Nolan ijtiff TIFF I/O for ImageJ Users CRAN (Advanced Image-Processing in R, using the imagemagick magick Jeroen Ooms library CRAN (Bindings to Tesseract OCR engine tesseract Jeroen

https://ropensci.org/packages/

FILTERS

- O All
- Data Publication
- Visualization
- Geospatial
- Images Processing
- Computing Infrastructure
- Taxonomy

- Altmetrics
- O Tools
- Databases
- O Web
- Literature
- Security
- CRAN / BIOC available

Type to search...

,			
PACKAGE	MAINTAINER	DESCRIPTION	DETAILS
binomen	Scott Chamberlain	Taxonomic name classes and parsing methods	CRAN (
ritis	Scott Chamberlain	Wrapper to the Integrated Taxonomic Information Service (ITIS) API	CRAN (7)
tava	Scott	Tavonomic Classes	CRAN 🕝 6

Comment assurer la qualité de tous les paquets ?

Comment assurer la qualité de tous les paquets ?

Revues de logiciels ouvertes. *Onboarding* = accueil et intégration.

Comment assurer la qualité de tous les paquets ?

Revues de logiciels ouvertes. *Onboarding* = accueil et intégration.

• Stimule l'adoption des meilleurs pratiques et standards.

Comment assurer la qualité de tous les paquets ?

Revues de logiciels ouvertes. *Onboarding* = accueil et intégration.

- Stimule l'adoption des meilleurs pratiques et standards.
- Construit une communauté de pratique.

Comment assurer la qualité de tous les paquets ?

Revues de logiciels ouvertes. *Onboarding* = accueil et intégration.

- Stimule l'adoption des meilleurs pratiques et standards.
- Construit une communauté de pratique.
- Partenariats avec le Journal of Open Source Software et Methods in Ecology and Evolution

Passer quoi en revue?

Passer quoi en revue?

Jeff Leek https://github.com/jtleek/rpackages

Documentation > Usability > Speed > Statistical superiority



Critères de revue

- Licence compatible avec l'open-source initiative (OSI)
- Documentation complète
- Haut taux de couverture de test
- Code lisible
- Usabilité

Critères de revue

- Licence compatible avec l'open-source initiative (OSI)
- Documentation complète
- Haut taux de couverture de test
- Code lisible
- Usabilité

https://ropensci.github.io/dev_guide/building.html

Critères de revue

- Licence compatible avec l'open-source initiative (OSI)
- Documentation complète
- Haut taux de couverture de test
- Code lisible
- Usabilité

https://ropensci.github.io/dev_guide/building.html

Tout un livre! https://ropensci.github.io/dev_guide

Comment faire la revue ?

• Ouverte & non-accusatoire

Comment faire la revue?

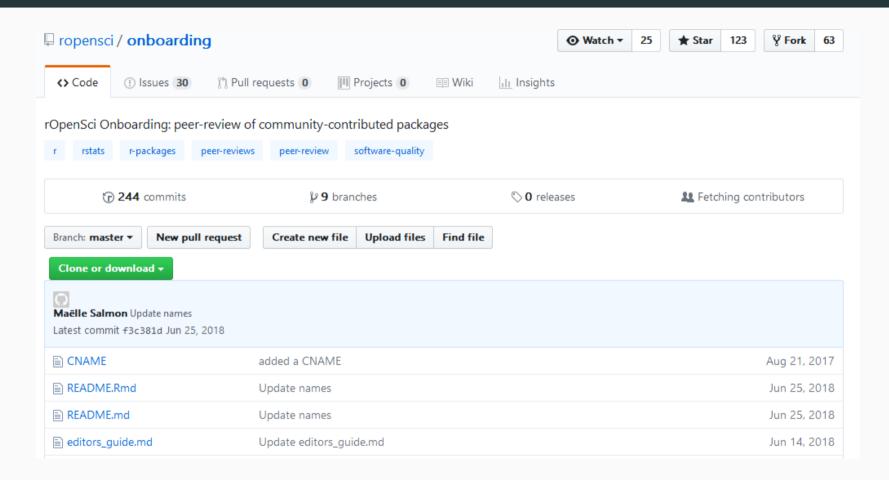
- Ouverte & non-accusatoire
- Pas de rejet

Comment faire la revue ?

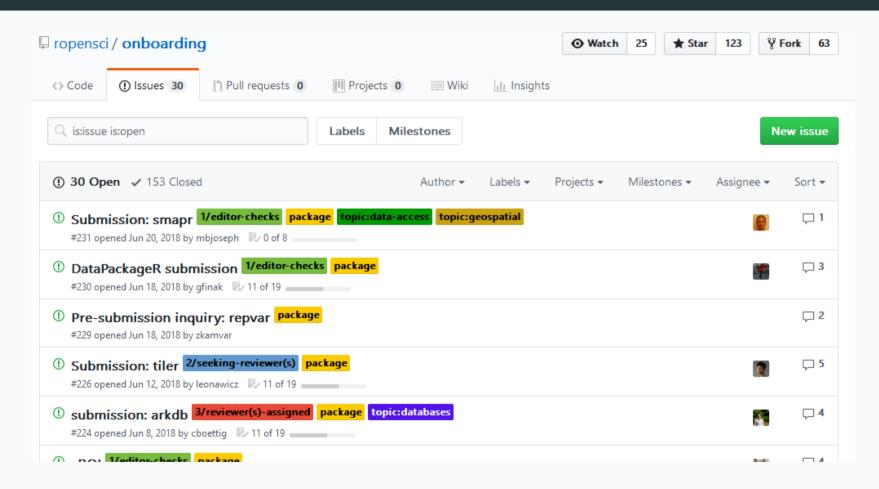
- Ouverte & non-accusatoire
- Pas de rejet
- Un processus constructif pour tous les acteurs

Comment faire la revue ?

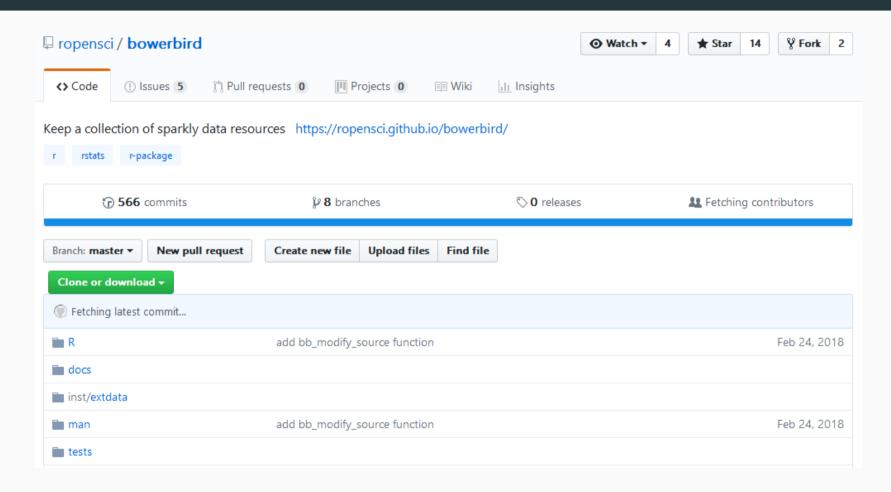
- Ouverte & non-accusatoire
- Pas de rejet
- Un processus constructif pour tous les acteurs
- En pratique, utilisation de l'infrastructure de GitHub



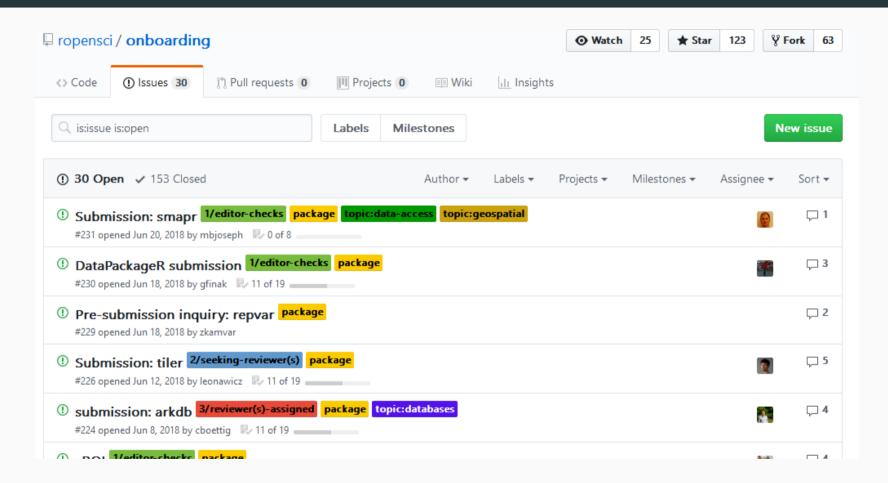
Onglet de suivi des "problèmes"



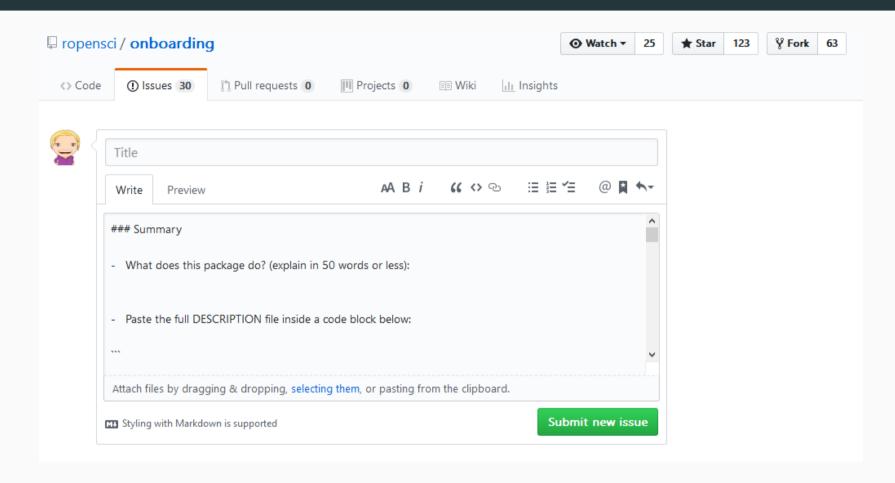
Soumettre un paquet



Soumettre un paquet



Soumettre un paquet



Paquets hors du domaine visé non accueillis.

https://ropensci.github.io/dev_guide/policies.html

Paquets hors du domaine visé non accueillis.

https://ropensci.github.io/dev_guide/policies.html

 nos catégories: accès à des données, extraction de données, accès à des formats de base de données, manipulation de données, dépôts de données,reproducibilité, données géospatialies, analyse de texte.

Paquets hors du domaine visé non accueillis.

https://ropensci.github.io/dev_guide/policies.html

- nos catégories: accès à des données, extraction de données, accès à des formats de base de données, manipulation de données, dépôts de données,reproducibilité, données géospatialies, analyse de texte.
- application scientifique

Paquets hors du domaine visé non accueillis.

https://ropensci.github.io/dev_guide/policies.html

- nos catégories: accès à des données, extraction de données, accès à des formats de base de données, manipulation de données, dépôts de données,reproducibilité, données géospatialies, analyse de texte.
- application scientifique
- meilleur que les paquets similaires

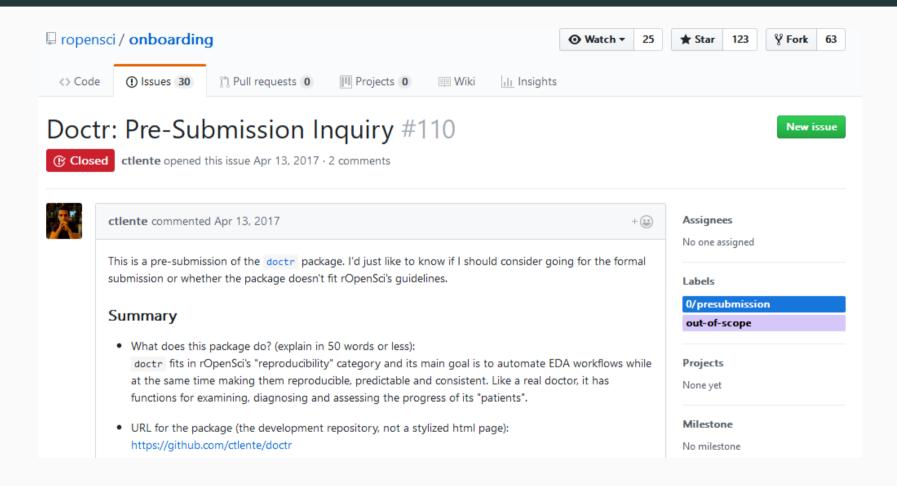
Paquets hors du domaine visé non accueillis.

https://ropensci.github.io/dev_guide/policies.html

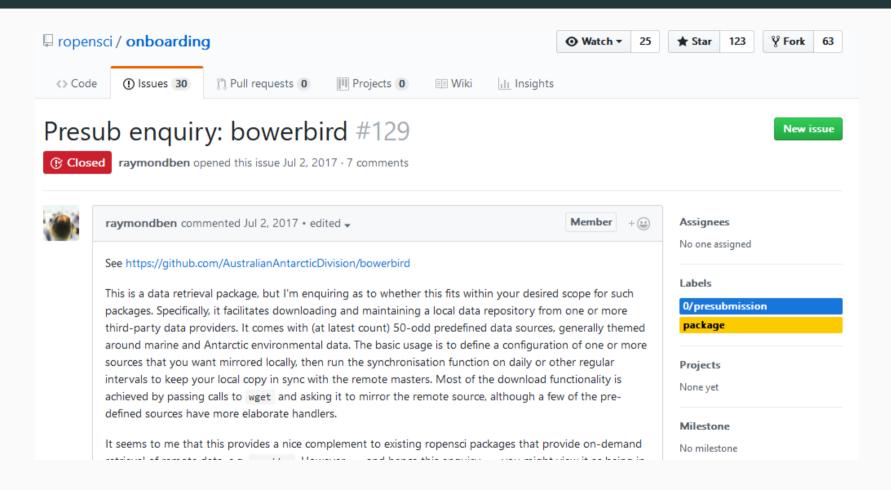
- nos catégories: accès à des données, extraction de données, accès à des formats de base de données, manipulation de données, dépôts de données,reproducibilité, données géospatialies, analyse de texte.
- application scientifique
- meilleur que les paquets similaires

En cas de doute, demande de pré-soumission!

Demande de pré-soumission?

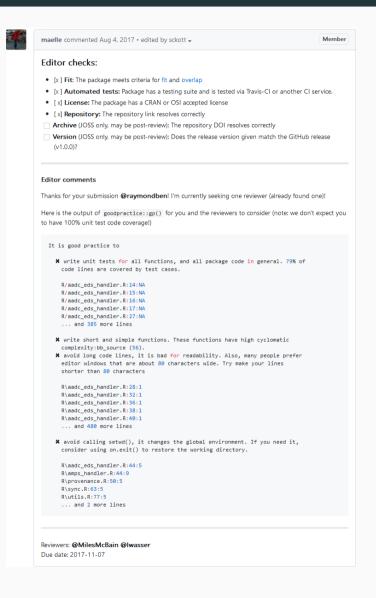


Demande de pré-soumission?



Procédé de revue

Vérifications par l'éditeur·rice



Revues

Revues

Parkage Beview Na policy price of the Monthly Service of London Service of the Service of London Service of Lon Improve function documentation Eaplein coeffig first argument Vests (aggre) or operator to oppress (aggre) is used very little within 15th package, for I question it worth as a dependency given the amount of installation time and number of invasion dependencies of adult. (An inagine issuescies) being used with automated reproductive pipelines where this may be an insur-Unusual unage of guette Section for the contract instance of the contr The function for those args to live in is probably statumeter_upper(). This might be a bit select Fyou move every from the one of laster, and isospect above.

Procédé de revue

Discussion jusqu'à acceptation et transfert du dépôt

Procédé de revue

Discussion jusqu'à acceptation et transfert du dépôt

Souvent un article de blogue

https://ropensci.org/tags/review/

Un portrait par les données?

Comment faire une analyse de données de rOpenSci onboarding?

Un portrait par les données?

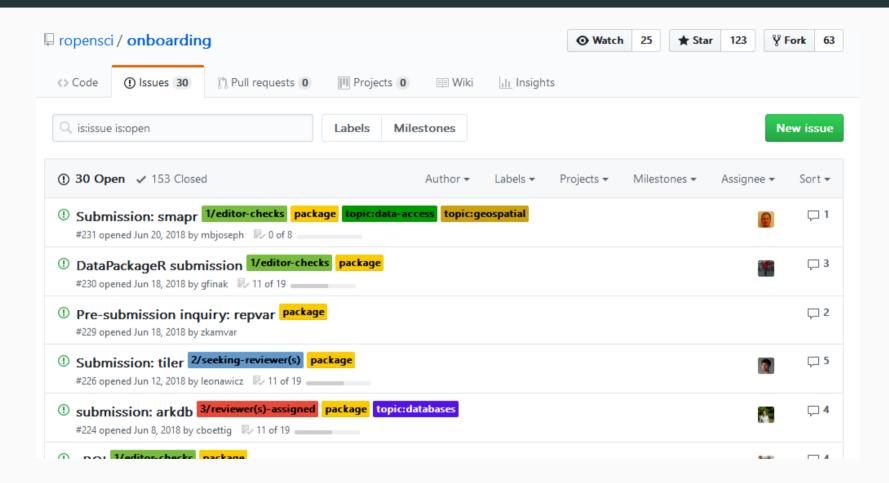
Comment faire une analyse de données de rOpenSci onboarding?

Rectanglons onboarding! Rectangler, mot inventé par Jenny Bryan

Image par Allie Brosh



Données dans le suivi des problèmes



GitHub GraphQL API v4. Mieux que v3? Obtenez seulement les données voulues.

GitHub GraphQL API v4. Mieux que v3? Obtenez seulement les données voulues.

Mon expérience

GitHub GraphQL API v4. Mieux que v3? Obtenez seulement les données voulues.

Mon expérience

- Explorateur en ligne de l'API
- ghql https://github.com/ropensci/ghql
- jq play pour écrire des commandes jq pour dompter le JSON
- jqr https://github.com/ropensci/jqr

Résultat : fils de GitHub tissés

Magie de GitHub V4 + un peu de domptage de données...

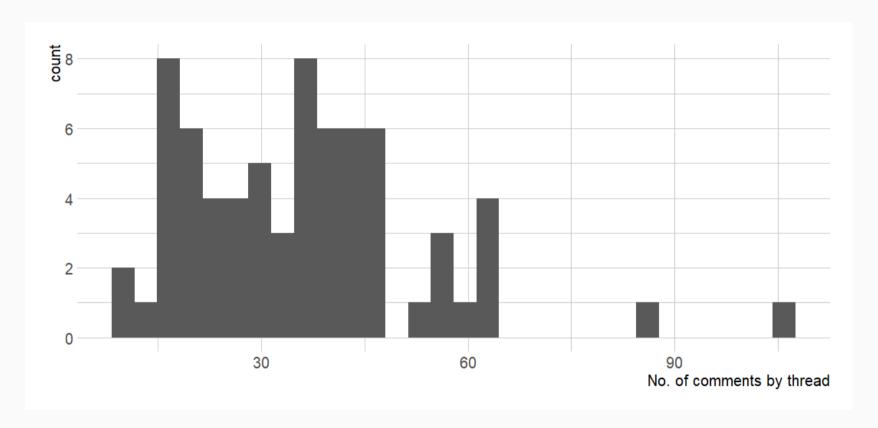
Résultat : fils de GitHub tissés

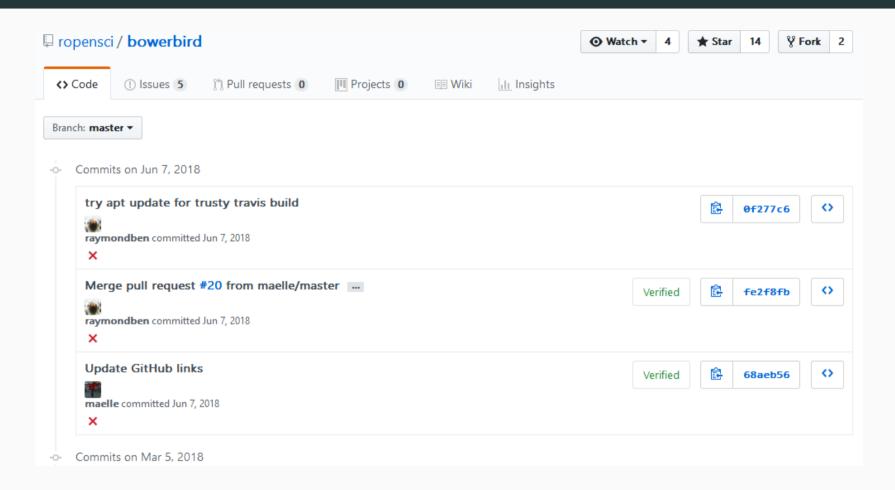
Magie de GitHub V4 + un peu de domptage de données...

```
## # A tibble: 2,521 x 10
   title created at
                     closed at
                                                body user issue package
###
   <chr> <dttm>
                             <dttm> <chr> <chr> <int> <chr>
###
###
   1 rrli~ 2015-03-31 00:25:14 2015-04-13 23:26:38 "- 1~ rich~
                                                               6 rrlite
   2 rrli~ 2015-04-01 17:30:51 2015-04-13 23:26:38 "hev~ scko~ 6 rrlite
###
   3 rrli~ 2015-04-01 17:36:03 2015-04-13 23:26:38 "@sc~ kart~ 6 rrlite
###
   4 rrli~ 2015-04-02 03:36:09 2015-04-13 23:26:38 "Sur~ jero~ 6 rrlite
###
   5 rrli~ 2015-04-02 03:50:43 2015-04-13 23:26:38 "IMO~ gabo~ 6 rrlite
###
   6 rrli~ 2015-04-02 03:53:57 2015-04-13 23:26:38 "Ide~ rich~
                                                               6 rrlite
###
   7 rrli~ 2015-04-02 18:58:53 2015-04-13 23:26:38 "> H~ kart~ 6 rrlite
###
   8 stpl~ 2015-04-08 23:56:17 2015-10-29 14:14:35 "- 1~ Robi~
                                                               10 stplanr
###
   9 rrli~ 2015-04-10 21:52:39 2015-04-13 23:26:38 "@ri~ stew~
                                                               6 rrlite
###
                                                                6 rrlite
## 10 rrli~ 2015-04-10 22:10:48 2015-04-13 23:26:38 "Tha~ rich~
## # ... with 2,511 more rows, and 3 more variables: is review <lgl>,
## #
     commenter <chr>, role <chr>
```

Résultat : fils de GitHub tissés

[1] 70





Clone via git2r https://github.com/ropensci/git2r

- Clone via git2r https://github.com/ropensci/git2r
- Remise à leur état avant la soumission grâce à git2r. git reset hard... Machine à remonter le temps!

- Clone via git2r https://github.com/ropensci/git2r
- Remise à leur état avant la soumission grâce à git2r. git reset hard... Machine à remonter le temps!
- Obtension de l'historique des modifications via gitsum.
 https://github.com/lorenzwalthert/gitsum

- Clone via git2r https://github.com/ropensci/git2r
- Remise à leur état avant la soumission grâce à git2r. git reset hard... Machine à remonter le temps!
- Obtension de l'historique des modifications via gitsum.
 https://github.com/lorenzwalthert/gitsum
- Comptage des lignes de code via cloc.
 https://github.com/hrbrmstr/cloc

- Clone via git2r https://github.com/ropensci/git2r
- Remise à leur état avant la soumission grâce à git2r. git reset hard... Machine à remonter le temps!
- Obtension de l'historique des modifications via gitsum.
 https://github.com/lorenzwalthert/gitsum
- Comptage des lignes de code via cloc.
 https://github.com/hrbrmstr/cloc
- Comptage des fonctions et classes exportées par devtools::parse_ns_file

Regardons les données

- Le code est dispo dans https://ropensci.org/blog/2018/04/26/a-satrday-ct-series/
- Paquets utilisés pour les visualisations : ggplot2, hrbrthemes, viridis.

Combien de travail représente rOpenSci

Travail dans les dépôts

Âge à la soumission

Âge à la soumission

Temps de revue

Quelle taille font les paquets?

Travail des auteurs

Travails pour les réviseur·se·s

Temps de revue vs. nombre d'exports

• Difficile de définir des mesures.

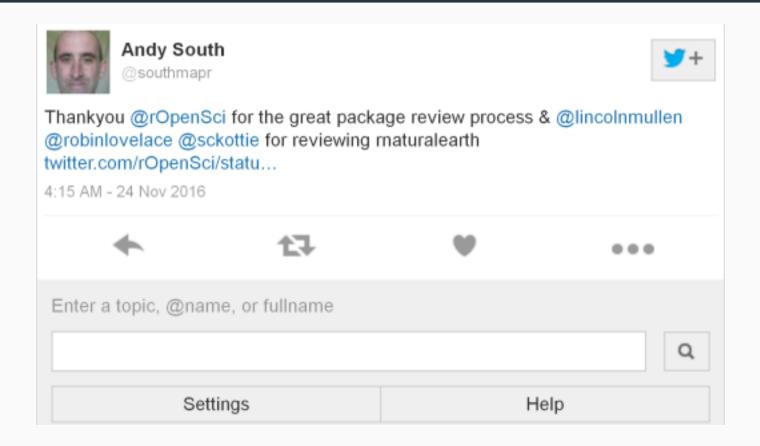
- Difficile de définir des mesures.
- Travail fait par des volontaires!

- Difficile de définir des mesures.
- Travail fait par des volontaires!
- Diminution du temps par automatisation.

- Difficile de définir des mesures.
- Travail fait par des volontaires!
- Diminution du temps par automatisation.
- Et les éditeur·rice·s alors ? Selon Tim Trice (traduit), "anges guardiens du début à la fin du procédé d'accueil". Un paquet tous les un ou deux mois.

Procédé de qualité... et sympa?

Procédé de qualité... et sympa?



Météo sociale de rOpenSci onboarding

Essai d'Ann Gentle dans http://open-advice.org/

Impressions, exemples... approche plus générale?

Météo sociale de rOpenSci onboarding

Essai d'Ann Gentle dans http://open-advice.org/

Impressions, exemples... approche plus générale?

Analyse de texte!

Météo sociale de rOpenSci onboarding

Essai d'Ann Gentle dans http://open-advice.org/

Impressions, exemples... approche plus générale?

Analyse de texte!

- tidytext et sentimentr
- https://www.tidytextmining.com/ de Julia Silge et David Robinson

Mots les plus communs

Bigrammes les plus communs

Corrélations par paire

Sentiment

Sentiment

Mots des lignes négatives

Lignes les plus négatives

```
## # A tibble: 15 x 2
      line
                                                                      sentiment
###
      <chr>
                                                                          <dbl>
##
    1 Multinomics no more things, although do make sure to add mor~
                                                                         -1.63
##
    2 not sure what you mean, but i'll use different object names ~
                                                                         -1.20
###
    3 error in .local(.object, ...):
###
                                                                         -1
    4 error:
                                                                          -1
###
   5 #### miscellaneous
                                                                         -1
###
    6 error: command failed (1)
                                                                         -0.866
###
   7 - get plate size from number of columns: maybe throwing an e~
###
                                                                         -0.786
   8 "this code returns an error, which is good, but it would be ~
                                                                         -0.744
###
    9 0 errors | 0 warnings | 0 notes
                                                                         -0.722
## 10 once i get to use this package more, i'm sure i'll have more~
                                                                         -0.721
  11 - i now realize i've pasted the spelling mistakes without th~
                                                                         -0.707
  12 minor issues:
                                                                         -0.707
## 13 ## minor issues
                                                                         -0.707
## 14 replicates issue
                                                                         -0.707
## 15 visualization issue
                                                                         -0.707
```

Lignes les plus positives !

```
## # A tibble: 15 x 2
      line
                                                                      sentiment
###
      <chr>
                                                                          <dbl>
###
    1 absolutely - it's really important to ensure it really has b~
                                                                           1.84
##
   2 overall, really easy to use and really nicely done.
                                                                           1.73
###
   3 this package is a great and lightweight addition to working ~
###
                                                                           1.46
    4 i am very grateful for your approval and i very much look fo~
##
                                                                           1.26
   5 thank you very much for the constructive thoughts.
                                                                           1.24
###
   6 thanks for the approval, all in all a very helpful and educa~
                                                                           1.22
###
   7 - really good use of helper functions
###
                                                                           1.14
   8 - i believe the utf note is handled correctly and this is ju~
                                                                           1.13
###
   9 seem more unified and consistent.
                                                                           1.13
## 10 very much appreciated!
                                                                           1.13
  11 - well organized, readable code
                                                                           1.1
## 12 - wow very extensive testing! well done, very thorough
                                                                           1.1
## 13 - i'm delighted that you find my work interesting and i'm ve~
                                                                           1.08
## 14 thank you very much for your thorough and thoughtful review,~
                                                                           1.08
## 15 great, thank you very much for accepting this package. i am ~
                                                                           1.07
```

Engagez-vous, qu'il·elle·s disaient!

- **y** rOpenSci
- O Dépôt officiel https://github.com/ropensci/onboarding
- Blogue https://ropensci.org/tags/onboarding/ en particulier https://ropensci.org/blog/2018/04/26/a-satrday-ct-series/
- Lisez notre livre https://ropensci.github.io/dev_guide/
- ✔ Portez-vous volontaire pour être réviseu·r·se https://ropensci.org/onboarding/

Éditeur·rice·s

thank(c("karthik", "noamross", "sckott"))



thank(c("lmullen", "annakrystalli", "ropensci"))



Merci!

