

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

# Les paquets de rOpenSci

<https://ropensci.org/packages/>







# Les paquets de rOpenSci

<https://ropensci.org/packages/>

## FILTERS

- ☐ All
- ☐ Data Publication
- ☒ Visualization
- ☐ Geospatial
- ☐ Images Processing
- ☐ Computing Infrastructure
- ☐ Taxonomy
- ☐ Altmetrics
- ☐ Tools
- ☐ Databases
- ☐ Web
- ☐ Literature
- ☐ Security
- ☐ CRAN / BIOC available

 Type to search...


PACKAGE	MAINTAINER	DESCRIPTION	DETAILS
<a href="#">iheatmapr</a>	<a href="#">Alicia Schep</a>	Interactive, Complex Heatmaps	 
<a href="#">plotdap</a>	<a href="#">Carson Sievert</a>	Easily visualize data from 'ERDDAP' Servers via the rerddap package	 
<a href="#">plotly</a>	<a href="#">Carson Sievert</a>	Programmatic interface to the Plot.ly API -construct plots	 







# Les paquets de rOpenSci

<https://ropensci.org/packages/>

## FILTERS

- ☐ All
- ☐ Data Publication
- ☐ Visualization
- ☐ Geospatial
- ☒ Images Processing
- ☐ Computing Infrastructure
- ☐ Taxonomy
- ☐ Altmetrics
- ☐ Tools
- ☐ Databases
- ☐ Web
- ☐ Literature
- ☐ Security
- ☐ CRAN / BIOC available

 Type to search...

PACKAGE	MAINTAINER	DESCRIPTION	DETAILS
<a href="#">ijtiff</a>	<a href="#">Rory Nolan</a>	TIFF I/O for ImageJ Users	 
<a href="#">magick</a>	<a href="#">Jeroen Ooms</a>	Advanced Image-Processing in R, using the imagemagick library	 
<a href="#">tesseract</a>	<a href="#">Jeroen Ooms</a>	Bindings to Tesseract OCR engine	 







# Les paquets de rOpenSci

<https://ropensci.org/packages/>

## FILTERS

- ☐ All
- ☐ Data Publication
- ☐ Visualization
- ☐ Geospatial
- ☐ Images Processing
- ☐ Computing Infrastructure
- ☒ Taxonomy
- ☐ Altmetrics
- ☐ Tools
- ☐ Databases
- ☐ Web
- ☐ Literature
- ☐ Security
- ☐ CRAN / BIOC available

 Type to search...

PACKAGE	MAINTAINER	DESCRIPTION	DETAILS
<a href="#">binomen</a>	<a href="#">Scott Chamberlain</a>	Taxonomic name classes and parsing methods	 
<a href="#">ritis</a>	<a href="#">Scott Chamberlain</a>	Wrapper to the Integrated Taxonomic Information Service (ITIS) API	 
<a href="#">tava</a>	<a href="#">Scott</a>	Taxonomic Classes	 



# rOpenSci onboarding !

Comment assurer la qualité de tous les paquets ?

# rOpenSci onboarding !

Comment assurer la qualité de tous les paquets ?

Revue de logiciels ouverts. *Onboarding* = accueil et intégration.

# rOpenSci onboarding !

Comment assurer la qualité de tous les paquets ?

Revue de logiciels ouverts. *Onboarding* = accueil et intégration.

- Stimule l'adoption des meilleures pratiques et standards.

# rOpenSci onboarding !

Comment assurer la qualité de tous les paquets ?

Revue de logiciels ouverts. *Onboarding* = accueil et intégration.

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

# rOpenSci onboarding !

Comment assurer la qualité de tous les paquets ?

Revue de logiciels ouverts. *Onboarding* = accueil et intégration.

- Stimule l'adoption des meilleures 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](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](https://ropensci.github.io/dev_guide/building.html)

Tout un livre ! [https://ropensci.github.io/dev\\_guide](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

# rOpenSci onboarding

ropensci / onboarding

Watch ▾

25

★ Star

123

🍴 Fork

63

<> Code

🔔 Issues 30

🔗 Pull requests 0

📁 Projects 0

📖 Wiki

📊 Insights

rOpenSci Onboarding: peer-review of community-contributed packages

r

rstats

r-packages

peer-reviews

peer-review

software-quality

📦 244 commits

🌿 9 branches

📦 0 releases

👤 Fetching contributors

Branch: master ▾

New pull request

Create new file

Upload files

Find file

Clone or download ▾



**Maëlle Salmon** Update names

Latest commit f3c381d Jun 25, 2018

📄 CNAME	added a CNAME	Aug 21, 2017
📄 README.Rmd	Update names	Jun 25, 2018
📄 README.md	Update names	Jun 25, 2018
📄 editors_guide.md	Update editors_guide.md	Jun 14, 2018

# Onglet de suivi des "problèmes"

ropensci / onboarding

Watch

25

Star

123

Fork

63

Code

Issues 30

Pull requests 0

Projects 0

Wiki

Insights

is:issue is:open

Labels

Milestones

New issue

30 Open ✓ 153 Closed

Author

Labels

Projects

Milestones

Assignee

Sort

Submission: smapr 1/editor-checks package topic:data-access topic:geospatial

#231 opened Jun 20, 2018 by mbjoseph 0 of 8



1

DataPackageR submission 1/editor-checks package

#230 opened Jun 18, 2018 by gfinak 11 of 19



3

Pre-submission inquiry: repvar package

#229 opened Jun 18, 2018 by zkamvar

2

Submission: tiler 2/seeking-reviewer(s) package

#226 opened Jun 12, 2018 by leonawicz 11 of 19



5

submission: arkdb 3/reviewer(s)-assigned package topic:databases

#224 opened Jun 8, 2018 by cboettig 11 of 19



4

Submission: smapr 1/editor-checks package



1



# Soumettre un paquet

ropensci / bowerbird

Watch ▾

4

★ Star

14

🍴 Fork

2

<> Code

🔔 Issues 5

🔗 Pull requests 0

📁 Projects 0

📖 Wiki

📊 Insights

Keep a collection of sparkly data resources <https://ropensci.github.io/bowerbird/>

r

rstats

r-package

📁 566 commits

🌿 8 branches

📦 0 releases

👤 Fetching contributors

Branch: master ▾

New pull request

Create new file

Upload files

Find file

Clone or download ▾

🔄 Fetching latest commit...

📁 R	add bb_modify_source function	Feb 24, 2018
📁 docs		
📁 inst/extdata		
📁 man	add bb_modify_source function	Feb 24, 2018
📁 tests		

# Soumettre un paquet

ropensci / onboarding

Watch

25

Star

123

Fork

63

Code

Issues 30

Pull requests 0

Projects 0

Wiki

Insights

is:issue is:open

Labels

Milestones

New issue

30 Open ✓ 153 Closed

Author

Labels

Projects

Milestones

Assignee

Sort

Submission: smapr 1/editor-checks package topic:data-access topic:geospatial

#231 opened Jun 20, 2018 by mbjoseph 0 of 8



1

DataPackageR submission 1/editor-checks package

#230 opened Jun 18, 2018 by gfinak 11 of 19



3

Pre-submission inquiry: repvar package

#229 opened Jun 18, 2018 by zkamvar

2

Submission: tiler 2/seeking-reviewer(s) package

#226 opened Jun 12, 2018 by leonawicz 11 of 19



5

submission: arkdb 3/reviewer(s)-assigned package topic:databases

#224 opened Jun 8, 2018 by cboettig 11 of 19



4

Submission: smapr 1/editor-checks package



1

# Soumettre un paquet

ropensci / onboarding

Watch

25

Star

123

Fork

63

Code

Issues 30

Pull requests 0

Projects 0

Wiki

Insights



Title

Write

Preview

AA B i

“ ” < > ↺

☰ ☷ ☶

@

📌

↶

### Summary

- What does this package do? (explain in 50 words or less):

- Paste the full DESCRIPTION file inside a code block below:

...

Attach files by dragging & dropping, [selecting them](#), or pasting from the clipboard.

Styling with Markdown is supported

Submit new issue

# Pas de rejets mais...

Paquets hors du domaine visé non accueillis.

[https://ropensci.github.io/dev\\_guide/policies.html](https://ropensci.github.io/dev_guide/policies.html)

# Pas de rejets mais...

Paquets hors du domaine visé non accueillis.

[https://ropensci.github.io/dev\\_guide/policies.html](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, reproductibilité, données géospatiales, analyse de texte.

# Pas de rejets mais...

Paquets hors du domaine visé non accueillis.

[https://ropensci.github.io/dev\\_guide/policies.html](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, reproductibilité, données géospatiales, analyse de texte.
- application *scientifique*

# Pas de rejets mais...

Paquets hors du domaine visé non accueillis.

[https://ropensci.github.io/dev\\_guide/policies.html](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, reproductibilité, données géospatiales, analyse de texte.
- application *scientifique*
- meilleur que les paquets similaires

# Pas de rejets mais...

Paquets hors du domaine visé non accueillis.

[https://ropensci.github.io/dev\\_guide/policies.html](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, reproductibilité, données géospatiales, analyse de texte.
- application *scientifique*
- meilleur que les paquets similaires

En cas de doute, demande de pré-soumission !



# Demande de pré-soumission ?

ropensci / onboarding

Watch

25

Star

123

Fork

63

Code

Issues 30

Pull requests 0

Projects 0

Wiki

Insights

## Doctr: Pre-Submission Inquiry #110

New issue

Closed

ctlente opened this issue Apr 13, 2017 · 2 comments



ctlente commented Apr 13, 2017



This is a pre-submission of the `doctr` package. I'd just like to know if I should consider going for the formal submission or whether the package doesn't fit rOpenSci's guidelines.

### Summary

- What does this package do? (explain in 50 words or less):  
`doctr` fits in rOpenSci's "reproducibility" category and its main goal is to automate EDA workflows while at the same time making them reproducible, predictable and consistent. Like a real doctor, it has functions for examining, diagnosing and assessing the progress of its "patients".
- URL for the package (the development repository, not a stylized html page):  
<https://github.com/ctlente/doctr>

### Assignees

No one assigned

### Labels

0/presubmission

out-of-scope

### Projects

None yet

### Milestone

No milestone

# Demande de pré-soumission ?

ropensci / onboarding

Watch

25

Star

123

Fork

63

Code

Issues 30

Pull requests 0

Projects 0

Wiki

Insights

## Presub enquiry: bowerbird #129

New issue

Closed

raymondben opened this issue Jul 2, 2017 · 7 comments



raymondben commented Jul 2, 2017 · edited

Member



See <https://github.com/AustralianAntarcticDivision/bowerbird>

This is a data retrieval package, but I'm enquiring as to whether this fits within your desired scope for such packages. Specifically, it facilitates downloading and maintaining a local data repository from one or more third-party data providers. It comes with (at latest count) 50-odd predefined data sources, generally themed around marine and Antarctic environmental data. The basic usage is to define a configuration of one or more sources that you want mirrored locally, then run the synchronisation function on daily or other regular intervals to keep your local copy in sync with the remote masters. Most of the download functionality is achieved by passing calls to `wget` and asking it to mirror the remote source, although a few of the predefined sources have more elaborate handlers.

It seems to me that this provides a nice complement to existing ropensci packages that provide on-demand retrieval of remote data. I'm not sure if this is a good fit for the ropensci ecosystem, but I'm enquiring as to whether this fits within your desired scope for such packages.

Assignees

No one assigned

Labels

0/presubmission

package

Projects

None yet

Milestone

No milestone

# Procédé de revue

# Vérifications par l'éditeur·rice

maelle commented Aug 4, 2017 • edited by skott ▾Member

**Editor checks:**

- [x] **Fit**: The package meets criteria for [fit](#) and [overlap](#)
- [x] **Automated tests**: Package has a testing suite and is tested via Travis-CI or another CI service.
- [x] **License**: The package has a CRAN or OSI accepted license
- [x] **Repository**: The repository link resolves correctly
- ☐ **Archive** (JOSS only, may be post-review): The repository DOI resolves correctly
- ☐ **Version** (JOSS only, may be post-review): Does the release version given match the GitHub release (v1.0.0)?

**Editor comments**

Thanks for your submission @raymondben! I'm currently seeking one reviewer (already found one)!

Here is the output of `goodpractice::gp()` for you and the reviewers to consider (note: we don't expect you to have 100% unit test code coverage!)

It is good practice to

- ✱ write unit tests **for** all functions, and all package code **in** general. **79%** of code lines are covered by test cases.

```
R/aadc_eds_handler.R:14:NA
R/aadc_eds_handler.R:15:NA
R/aadc_eds_handler.R:16:NA
R/aadc_eds_handler.R:17:NA
R/aadc_eds_handler.R:27:NA
... and 385 more lines
```

- ✱ write short and simple functions. These functions have high cyclomatic complexity: `bb_source` (56).
- ✱ avoid long code lines, it is bad **for** readability. Also, many people prefer editor windows that are about **80** characters wide. Try make your lines shorter than **80** characters

```
R/aadc_eds_handler.R:28:1
R/aadc_eds_handler.R:32:1
R/aadc_eds_handler.R:36:1
R/aadc_eds_handler.R:38:1
R/aadc_eds_handler.R:40:1
... and 480 more lines
```

- ✱ avoid calling `setwd()`, it changes the global environment. If you need it, consider using `on.exit()` to restore the working directory.

```
R/aadc_eds_handler.R:44:5
R/amps_handler.R:44:9
R/provenance.R:50:5
R/sync.R:63:5
R/utils.R:77:5
... and 2 more lines
```

Reviewers: @MilesMcBain @lwasser  
Due date: 2017-11-07

# Revue



# 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

ropensci / onboarding

Watch

25

Star

123

Fork

63

Code

Issues 30

Pull requests 0

Projects 0

Wiki

Insights

is:issue is:open

Labels

Milestones

New issue

30 Open ✓ 153 Closed

Author

Labels

Projects

Milestones

Assignee

Sort

Submission: smapr 1/editor-checks package topic:data-access topic:geospatial

#231 opened Jun 20, 2018 by mbjoseph 0 of 8



1

DataPackageR submission 1/editor-checks package

#230 opened Jun 18, 2018 by gfinak 11 of 19



3

Pre-submission inquiry: repvar package

#229 opened Jun 18, 2018 by zkamvar

2

Submission: tiler 2/seeking-reviewer(s) package

#226 opened Jun 12, 2018 by leonawicz 11 of 19



5

submission: arkdb 3/reviewer(s)-assigned package topic:databases

#224 opened Jun 8, 2018 by cboettig 11 of 19



4

Submission: smapr 1/editor-checks package



1

# Tisser les fils de "problèmes" GitHub

# Tisser les fils de "problèmes" GitHub

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

# Tisser les fils de "problèmes" GitHub

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

Mon expérience

# Tisser les fils de "problèmes" GitHub

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

Mon expérience

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

# Résultat : fils de GitHub tissés

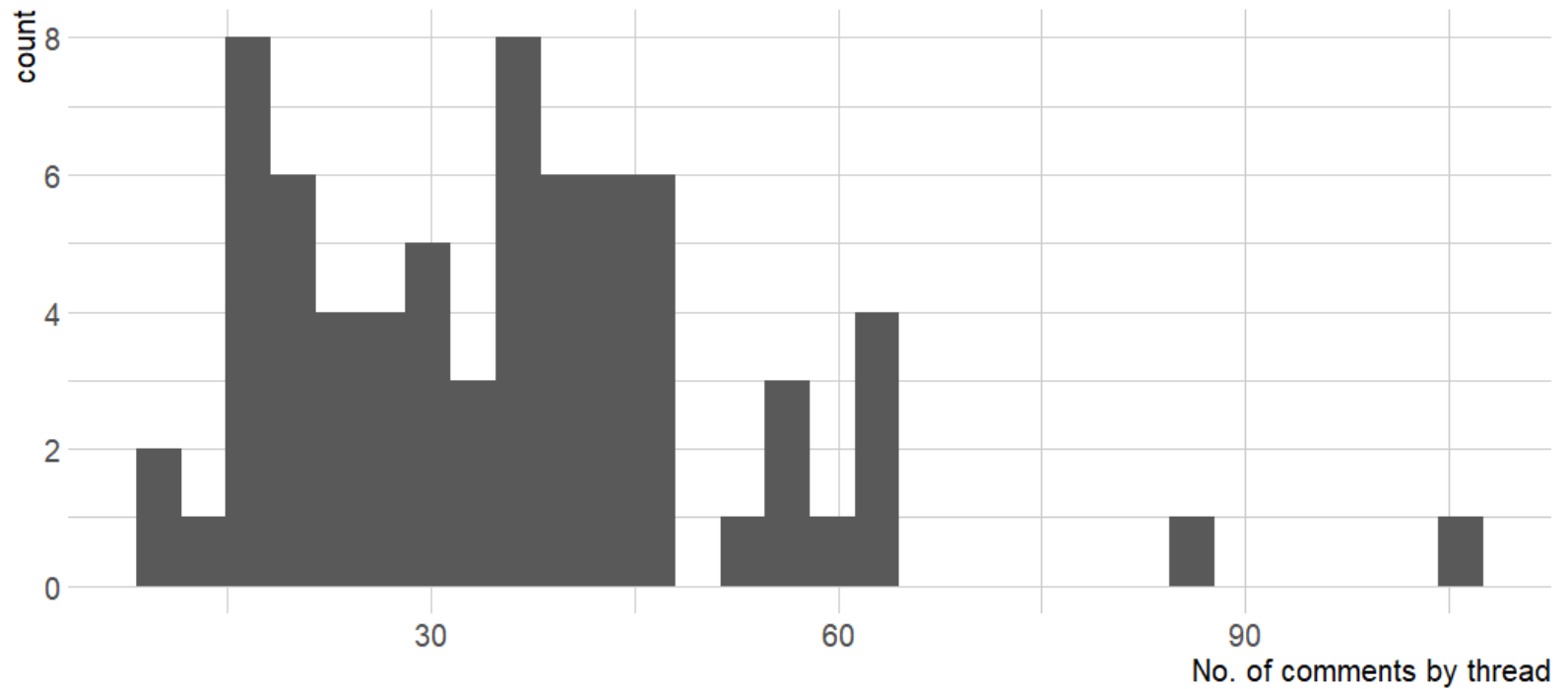
Magie de GitHub V4 + un peu de domptage de données...



# Résultat : fils de GitHub tissés

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

# Résultat : fils de GitHub tissés



# Données dans les dépôts accueillis

# Données dans les dépôts accueillis

ropensci / bowerbird

Watch

4

Star

14

Fork

2

Code

Issues 5

Pull requests 0

Projects 0

Wiki

Insights

Branch: master

Commits on Jun 7, 2018

try apt update for trusty travis build



0f277c6



raymondben committed Jun 7, 2018



Merge pull request #20 from maelle/master

Verified



fe2f8fb



raymondben committed Jun 7, 2018



Update GitHub links

Verified



68aeb56



maelle committed Jun 7, 2018



Commits on Mar 5, 2018






# Données dans les dépôts accueillis

- Clone via  <https://github.com/ropensci/git2r>






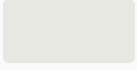
# Données dans les dépôts accueillis

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

# Données dans les dépôts accueillis






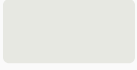
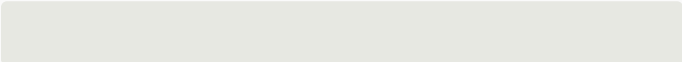
- Clone via  <https://github.com/ropensci/git2r>
- Remise à leur état avant la soumission grâce à .   
 ... Machine à remonter le temps !
- Obtension de l'historique des modifications via .  
<https://github.com/lorenzwalther/gitsum>

# Données dans les dépôts accueillis

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



# Données dans les dépôts accueillis

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


# Regardons les données

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

# 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

Travaux pour les réviseur·se·s



# Temps de revue vs. nombre d'exports

# Notes sur la quantification du travail

# Notes sur la quantification du travail

- Difficile de définir des mesures.

# Notes sur la quantification du travail

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

# Notes sur la quantification du travail

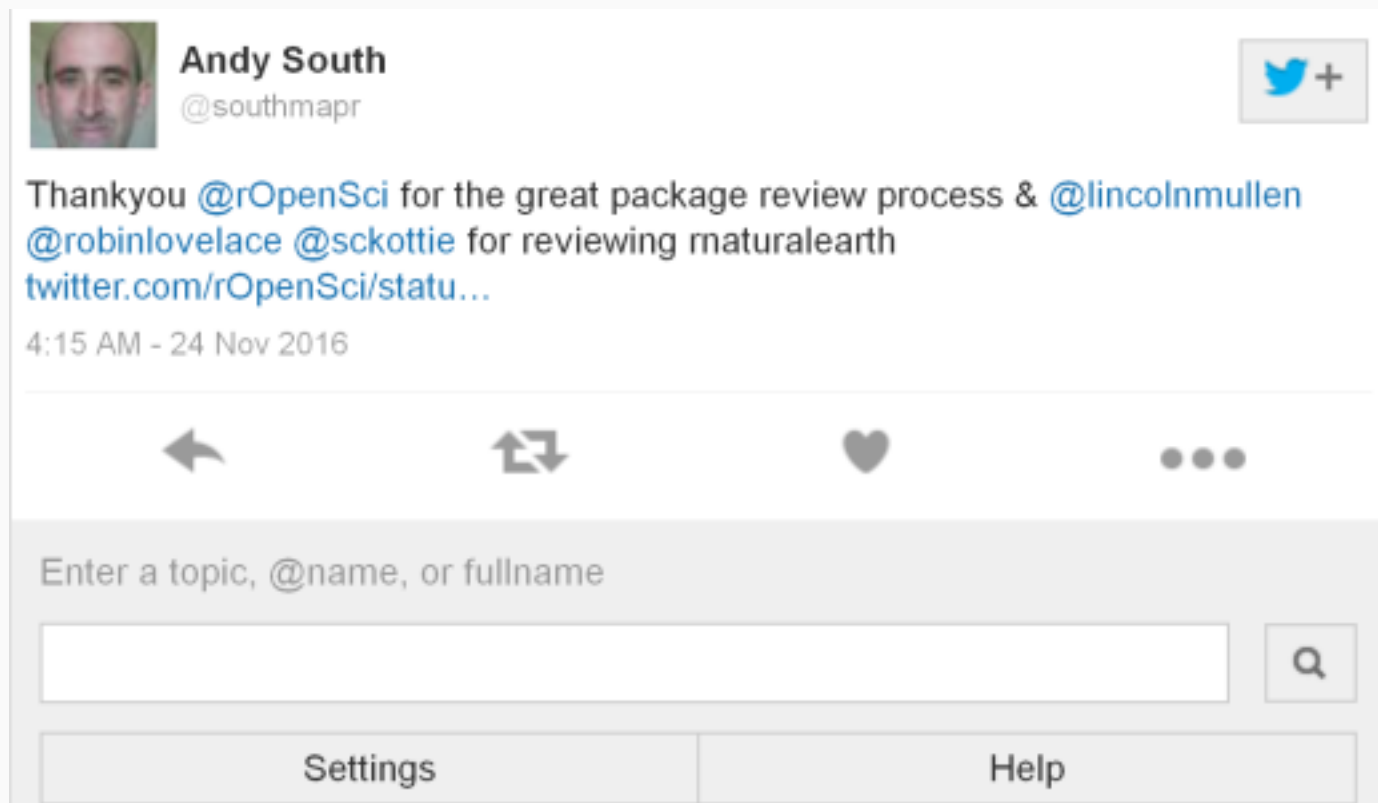
- Difficile de définir des mesures.
- Travail fait par des *volontaires* !
- Diminution du temps par automatisation.

# Notes sur la quantification du travail

- 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 gardiens 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 !

- [redacted] et [redacted]
- <https://www.tidyttextmining.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

# Lignes les plus positives !

# Engagez-vous, qu'il·elle·s disaient !

 rOpenSci

 Dépôt officiel <https://github.com/ropensci/onboarding>

 Blogue <https://ropensci.org/tags/onboarding/> en particulier <https://ropensci.org/blog/2018/04/26/a-saturday-series/>

 Lisez notre livre [https://ropensci.github.io/dev\\_guide/](https://ropensci.github.io/dev_guide/)

 Portez-vous volontaire pour être réviseur·r·se  
<https://ropensci.org/onboarding/>

# Éditeur·rice·s

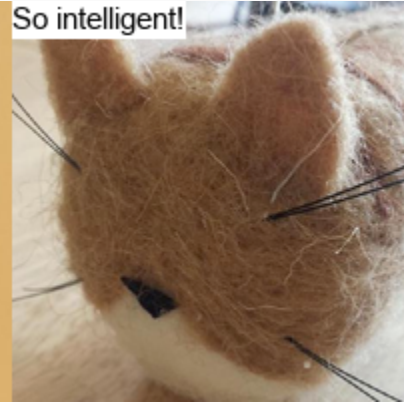
So ingenious!



So clever!



So intelligent!



So clever!



So wise!



So talented!



# Merci !



