rOpenSci - Empowering through community, software, and Peer Review

Nicholas Tierney (rOpenSci)
발표 자료 웹 버전 보기
rOpenSci: Empowering people with community, software, and peer review

Nicholas Tierney, Monash University

R User Conference, Korea

Friday 26th October, 2018

bit.ly/ruck-njt

nj_tierney
Observe → Hypothesis → Predict → Test → Analyse → New Knowledge!
Observe → Hypothesis → Predict → Test → Analyse → New Knowledge!
Observe → Hypothesis → Predict → Test → Analyse → New Knowledge!
Observe → Hypothesis → Predict → Test → Analyse → New Knowledge!
Oh, this is bad.
Heard via Garret Grolemund's great talk
Heard via Garret Grolemund's great talk
The paper

Data

Data Cleaning

Exploratory data analysis

Code
Implicit assumptions
Implicit assumptions

Replication

Can we repeat the study and get the same results?
Implicit assumptions

Replication
Can we repeat the study and get the same results?

Review
What did a peer think?
Implicit assumptions

Replication
Can we repeat the study and get the same results?

Review
What did a peer think?

Record
Are the data and code shared?
Open Science vs. ...?
Conducting open science
<table>
<thead>
<tr>
<th>Language Rank</th>
<th>Types</th>
<th>Spectrum Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Python</td>
<td>![Icon]</td>
<td>100.0</td>
</tr>
<tr>
<td>2. C++</td>
<td>![Icon]</td>
<td>99.7</td>
</tr>
<tr>
<td>3. Java</td>
<td>![Icon]</td>
<td>97.5</td>
</tr>
<tr>
<td>4. C</td>
<td>![Icon]</td>
<td>96.7</td>
</tr>
<tr>
<td>5. C#</td>
<td>![Icon]</td>
<td>89.4</td>
</tr>
<tr>
<td>6. PHP</td>
<td>![Icon]</td>
<td>84.9</td>
</tr>
<tr>
<td>7. R</td>
<td>![Icon]</td>
<td>82.9</td>
</tr>
<tr>
<td>8. JavaScript</td>
<td>![Icon]</td>
<td>82.6</td>
</tr>
<tr>
<td>9. Go</td>
<td>![Icon]</td>
<td>76.4</td>
</tr>
<tr>
<td>10. Assembly</td>
<td>![Icon]</td>
<td>74.1</td>
</tr>
</tbody>
</table>
```r
---
title: "Viridis Demo"
output: html_document
---

```r
include = FALSE}
library(viridis)
```r

The code below demonstrates two color palettes in the [viridis](https://github.com/sjmgarnier/viridis) package. Each plot displays a contour map of the Maunga Whau volcano in Auckland, New Zealand.

```r
## Viridis colors
```r
image(volcano, col = viridis(200))
```r

```r
## Magma colors
```r
image(volcano, col = viridis(200, option = "A"))
It is not easy.
It is not easy.
Let's do it together.
Founded in 2011 by Karthik Ram, Scott Chamberlain and Carl Boettiger.
Founded in 2011 by Karthik Ram, Scott Chamberlain and Carl Boettiger.

Original motivation was to make data access easier and reproducible.
By the numbers

5 staff
1 postdoc
2 Bioconductor pkgs
192 CRAN pkgs
287 total pkgs

~ 500 code contributors citations
LOTS! of community members

(From Stefanie Butland)

bit.ly/ruck-njt · @nj_tierney
Software

Peer Review

Community
Software
Software

Data retrieval and Access
Software

Data retrieval and Access

Data visualisation
Software

Data retrieval and Access

Data visualisation

Data sharing
Software: magick
Software

FILTERS
- All
- Altmetrics
- Data Publication
- Tools
- Visualization
- Databases
- Security
- Taxonomy
- CRAN / BIOC

Type to search...

<table>
<thead>
<tr>
<th>PACKAGE</th>
<th>MAINTAINER</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>bomrang</td>
<td>Adam Sparks</td>
<td>Fetch Australian Government Bureau of Meteorology Data</td>
</tr>
<tr>
<td>charlatan</td>
<td>Scott Chamberlain</td>
<td>Make fake data</td>
</tr>
<tr>
<td>cleanEHR</td>
<td>Sinan Shi</td>
<td>The Critical Care Clinical Data Processing Tools</td>
</tr>
</tbody>
</table>

ropensci.org/packages

github.com/ropensci

github.com/ropenscilabs
Software
Software

Peer Review

Community
Peer review for software?
Onboarding
Pre-submission enquiry

Peer-Review

Acceptance
Open and Non-Adversarial

No Rejections
Example peer review thread
Example peer review thread

**batpigandme (Mara Averick) on 28 Jul 2017**
Other than a renegade missing s (just submitted PR), it all looks great. @njtierney and visdat FTW!

**njtierney on 28 Jul 2017**
woohoo! Changes merged, thanks Mara! :D

**noamross (Noam Ross) on 31 Jul 2017**
Approved! Thanks @njtierney for submitting and @batpigandme and @seaan for your reviews and all of you hanging on for the long haul!
Package promotion

[blog] Onboarding `visdat`, a tool for preliminary visualisation of whole dataframes, by @nj_tierney
ropensci.org/blog/blog/2017 ...

Onboarding visdat, a tool for preliminary visualisation of whole dataframes

Nicholas Tierney
August 22, 2017

Take a look at the data
This is a phrase that comes up when you first get a dataset.
It is also ambiguous. Does it mean to do some exploratory modelling?
Or make some histograms, scatterplots, and boxplots? Is it both?

bit.ly/ruck-njt • @nj_tierney
Onboarding → JOSS, MEE
Sign up to review

Toph Allen · Brooke Anderson · Alison Appling · Zebulun Arendsee · Taylor Arnold · Dean Attali · Mara Averick · Suzan Baert · James Balamuta · David Bapst · Joëlle Barido-Sottani · Cale Basaraba · John Baumgartner · Marcus Beck · Gabe Becker · Jason Becker · Dom Bennett · Kenneth Benoit · Aaron Berdanier · Carl Boettiger · Ben Bond-Lamberty · Alison Boyer · Jenny Bryan · Lorenzo Bussetto · Jorge Cimentada · Jon Clayden · Will Cornwell · Ildiko Czeller · Laura DeCicco · Christophe Dervieux · Amanda Dobbyn · Jasmine Dumas · Remko Duursma · Mark Edmondson · Paul Egeler · Evan Eskew · Manuel Fernandez · Rich FitzJohn · Robert Flight · Zachary Foster · Auriel Fournier · Carl Ganz · Duncan Garmonsway · Sharla Gelfand · David Gohel · Laura Graham · Charles Gray · Corinna Gries · Julia Gustavsen · W Kyle Hamilton · Ivan Hanigan · Jeff Hanson · Ted Hart · Nucharee Haswell · Verena Haunschmid · Rafael Pilliard Hellwig · Bea Hernandez · Jim Hester · Peter Hickey · Roel M. Hogervorst · Jeff Hollister · Sean Hughes · Najko Jahn · Tamora D James · Max Joseph · Soumya Kalra · Michael Kane · Andee Kaplan · Hazel Kavli · Os Keyes · Michael Koontz · Anna Krystalli · Will Landau · Erin LeDell · Thomas Leeper · Stephanie Locke · Robin Lovelace · Julia Stewart Lowndes · Tim Lucas · Andrew MacDonald · Jesse Maegan · Tristan Mahr · Ben Marwick · Miles McBain · Lucy D’Agostino McGowan · Amelia McNamara · Elaine McVey · Bryce Mecum · Francois Michonneau · Jessica Minnier · Priscilla Minotti · Paula Moraga · Ross Mounce · Lincoln Mullen · Matt Mulvahill · Dillon Niederhut · Jakub Nowosad · Paul Oldham · Jeroen Ooms · Philipp Ottolinger · Mark Padgham · Edzer Pebesma · Nistara Randhawa · Neal Richardson · tyler rinker · Emily Robinson · Bob Rudis · Kent Russell · Francisco Rodriguez Sanchez · Alicia Schep · Marco Sciaini · Julia Silge · Margaret Siple · Peter Slaughter · Mike Smith · Tuija Sonkkila · Gaurav Sood · Adam Sparks · Joseph Stachelek · Irene Steves · Michael Sumner · Sarah Supp · Filipe Teixeira · Andy Teucher · Jennifer Thompson · Joe Thorley · Tiffany Timbers · Tim Trice · Ted Underwood · Kevin Ushey · Josef Uyeda · Frans van Dunné · Remi Vergnon · Claudia Vitolo · Ben Ward · Elin Waring · Rachel Warnock · Leah Wasser · Stefan Widgren · Luke Winslow · David Winter · Kara Woo · Lauren Yamane · Taras Zakharko · Hao Zhu · Naupaka Zimmerman
Submit a package for review

ropensci/onboarding

SUBMIT A PACKAGE!

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

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

  ```
  ...
  ```

Submit new issue

bit.ly/ruck-njt · @nj_tierney
Learn more about review

rOpenSci Packages: Development, Maintenance, and Peer Review

rOpenSci onboarding editorial team: Scott Chamberlain, Anna Krystalli, Lincoln Mullen, Karthik Ram, Noam Ross, Maëlle Salmon

2018-09-28

ropensci.github.io/dev_guide/
Now sure if you should review?

So you (don't) think you can review a package

Mara Averick

August 22, 2017

Contributing to an open-source community *without* contributing code is an oft-vaunted idea that can seem nebulous. Luckily, putting vague ideas into action is one of the strengths of the [rOpenSci Community](https://ropensci.org), and their package onboarding system offers a chance to do just that.

This was my first time reviewing a package, and, as with so many things in life, I went into it worried that I’d somehow ruin the package-reviewing process— not just the package itself, but the actual onboarding infrastructure...maybe even rOpenSci on the whole.
[1] Where is the value in package peer review?

Software

Peer Review

Community
The rOpenSci unconference
The rOpenSci unconference

What it isn't
Not a conference
Not a conference

No prearranged talks
Not a hackathon
Not a hackathon

No prizes

Doesn't go all night
A place for people to:
A place for people to:

Work on ideas
A place for people to:

Work on ideas

Connect with others
A place for people to:

Work on ideas

Connect with others

Build community
A place for people to:

Work on ideas
Connect with others
Build community
Learn from others
Who comes along

Diverse people from academia, industry, government, non-profits
Who comes along

Diverse people from academia, industry, government, non-profits

Novices, nexperts, and experts
How does it work?

💻 Arrive at a venue, laptop in hand
😊 Establish a culture of acceptance
💼 Discuss issues again
👥 Form groups
💼 Work on literally anything
💬 Discuss projects at the end
Data packages

: Australian Flight information

![Graph showing the increase in Inbound Passengers from 1990 to 2010. The graph is divided into two sections: Domestic and International. The Domestic section shows a steady increase in passengers, while the International section shows a steeper increase, particularly after 2005.]
Data packages

Access information about road deaths in Australia from 1989 to 2017
Data visualisation

: plot streaming data
Data visualisation

: create and apply themes to base plots
Fun tools

: easily insert web icons into markdown documents

(used extensively throughout this talk!!!)
Thank yous
Thank yous

MONASH University

bit.ly/ruck-njt • @nj_tierney
discuss.ropensci.org
onboarding.ropensci.org
ropensci.org/community
@ropensci
References

Garret Grolemund's Keynote slides video

Karthik Ram's keynote

Amgen, 2014

Begley, 2015

Freedman, 2015

So you don't think you can review a package?
References

ropensci.org/packages
ropensci.org/onboarding
ropensci.github.io/dev_guide/

Where is the value in package peer review?

Waterslides vs Ivory Towers
References

JOSS journal

MEE Journal