

Curating a Community of Packages: Lessons from a Decade of rOpenSci Peer Review

Noam Ross, rOpenSci

 @noamross.net

UseR! 2025

Durham, NC

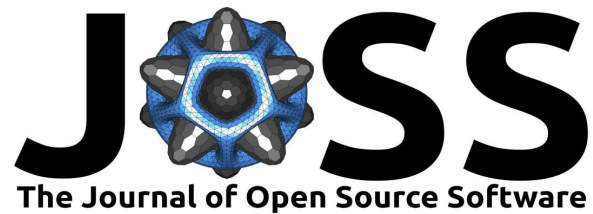


rOpenSci fosters a **culture** that values **open**,
collaborative and **reproducible** research using
shared data and **reusable** methods.

We build **technical** and **social** infrastructure for R
to enable open science in a **welcoming and**
diverse community

ropensci.org

package collections in R



CRAN Task View: Web Technologies and Services

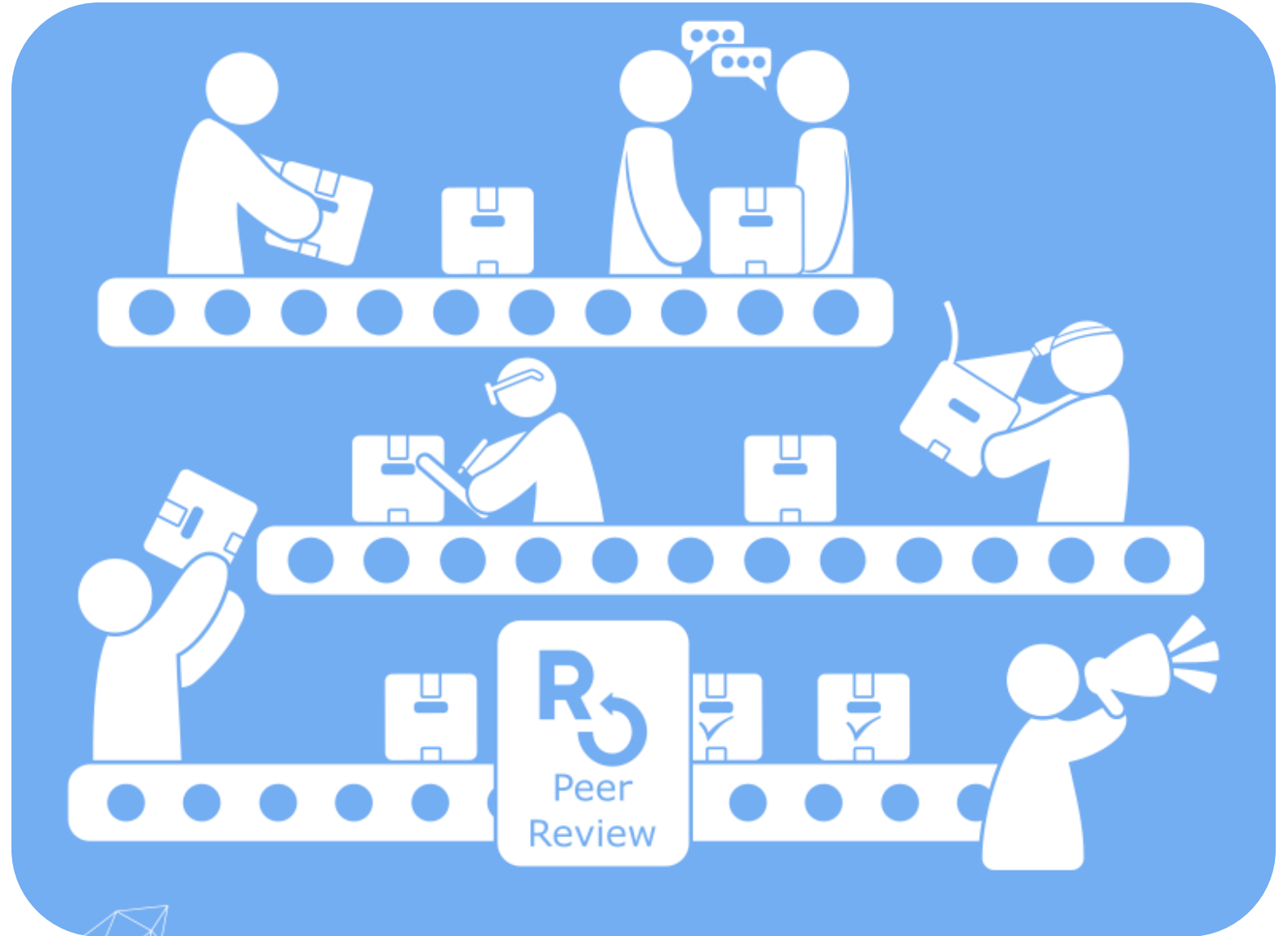
Maintainer: Mauricio Vargas Sepulveda, Will Beasley
Contact: m.sepulveda at mail.utoronto.ca
Version: 2024-10-27
URL: <https://CRAN.R-project.org/view=WebTechnologies>
Source: <https://github.com/cran-task-views/WebTechnologies/>



NOAA Fisheries Integrated
Toolbox

scope
standards
process
commitment
community

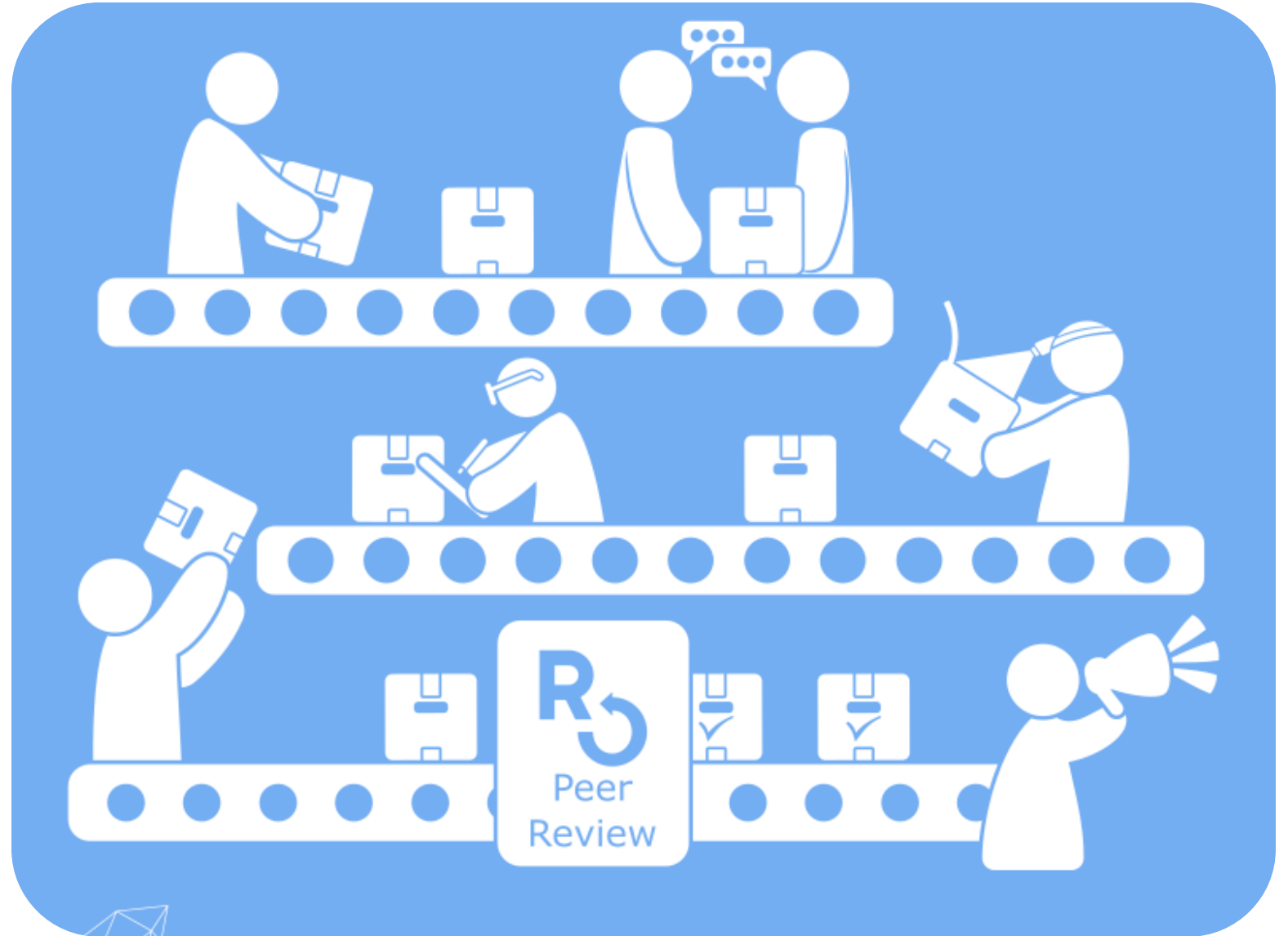
derive from **values**
and **goals**



rOpenSci software peer review

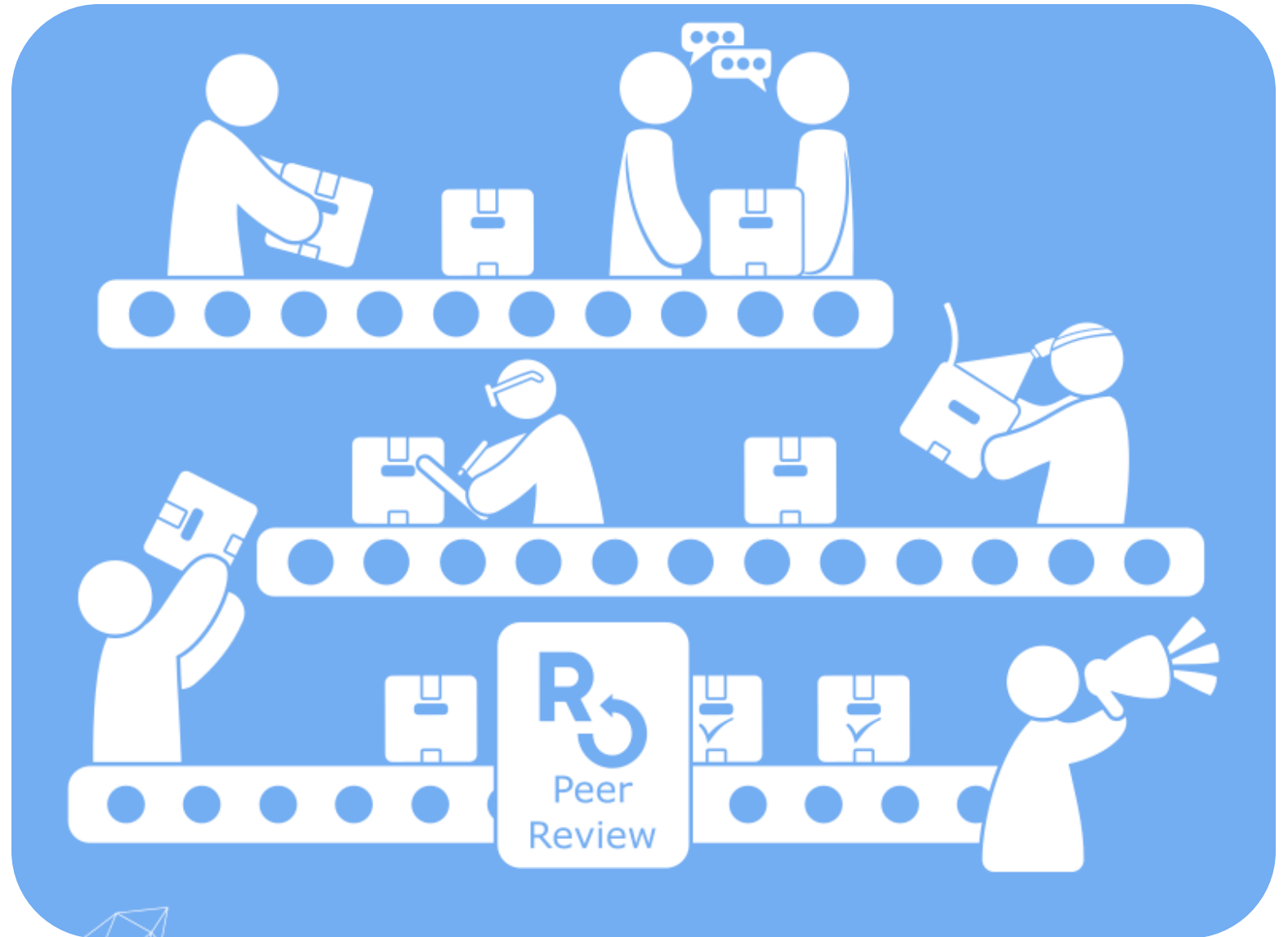
building **capacity** of
the **developer**
community and
driving adoption of
best practices

improving **quality**,
discoverability, and
sustainability of R
packages for open
science



ropensci.org/software-review

scope
standards
process
commitment
community



scope
standards
process
commitment
community

5.2 Aims and Scope

rOpenSci aims to support packages that enable reproducible research and managing the data lifecycle for scientists. Packages submitted to rOpenSci should fit into one or more of the categories outlined either below. Statistical software may also be submitted for peer review, for which we have a separate [set of guidelines and standards](#). The categories below are for general, and not statistical, software,

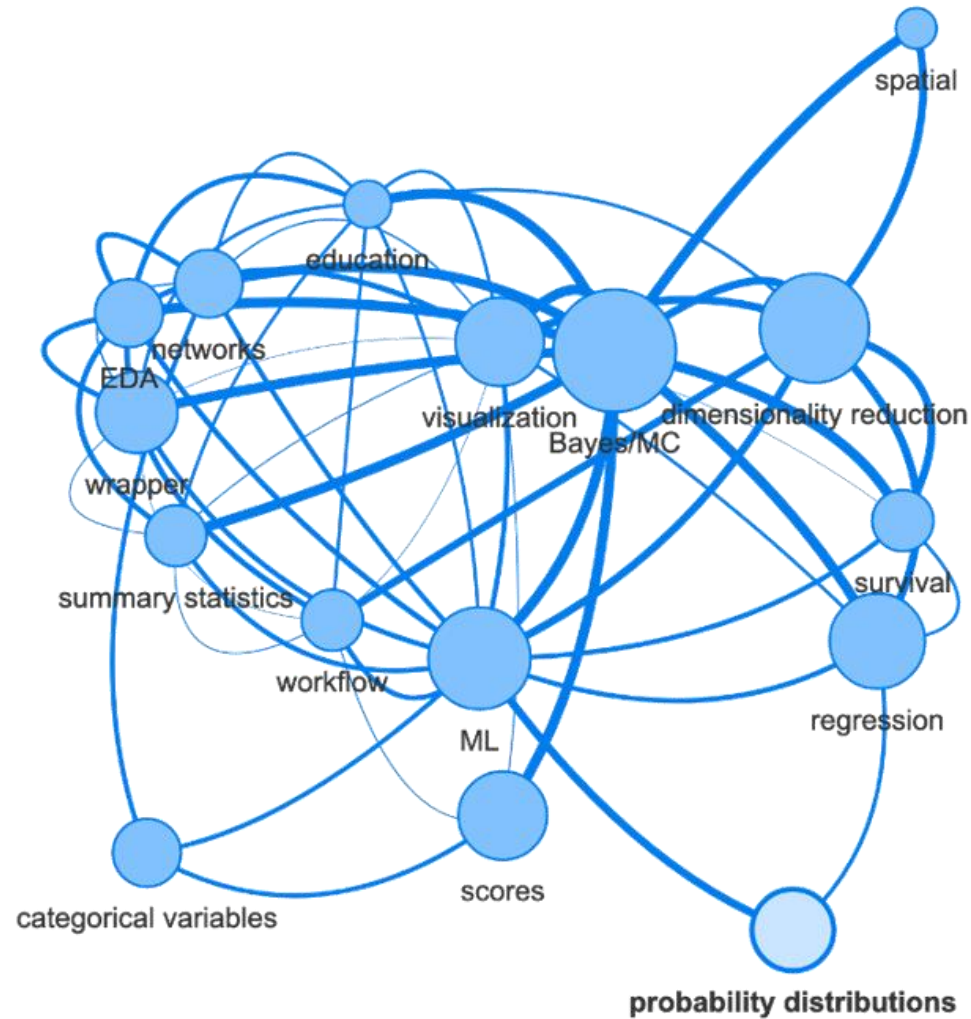
which should not be confused with the categories for statistical software.

5.2.1 Package categories

unstructured
platforms

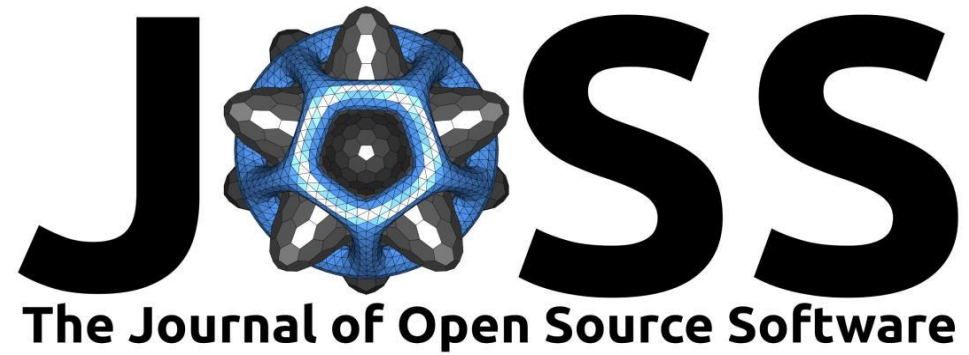
- **data retrieval:** Packages for accessing and downloading data from online sources with scientific applications. Our definition of scientific applications is broad, including data storage services, journals, and other remote servers, as many data sources may be of interest to researchers. However, retrieval packages should be focused on data *sources / topics*, rather than *services*. For example a general client for Amazon Web Services data storage would not be in-scope. (Examples: [rotl](#), [gutenbergr](#))
- **data extraction:** Packages that aid in retrieving data from unstructured sources such as text, images and PDFs, as well as parsing scientific data types and outputs from scientific equipment. Statistical/ML libraries for modeling or prediction are typically not included in this category, nor are code parsers. Trained models that

scope
standards
process
commitment



- 6.2 Bayesian and Monte Carlo Software
- 6.3 Exploratory Data Analysis and Summary Statistics
- 6.4 Machine Learning Software
- 6.5 Regression and Supervised Learning
- 6.6 Spatial Software
- 6.7 Time Series Software
- 6.8 Dimensionality Reduction, Clustering, and Unsupervised Learning
- 6.9 Probability Distributions

scope
standards
process
commitment
community



scope
standards
process
commitment
community

1.7 Documentation

1.7.1 General

- All exported package functions should be fully document
- If there is potential overlap or confusion with other packa functionality or having a similar name, add a note in the F and potentially the Description field of DESCRIPTION. Exa [rebird README](#), and the non-rOpenSci package [slurmR](#).
- The package should contain top-level documentation for package ` if there is a naming conflict). Optionally, you ca ?`foobar -package` for the package level manual file, t tag. [usethis::use_package_doc\(.\)](#) adds the template documentation.
- The package should contain at least one **HTML** vignette p coverage of package functions, illustrating realistic use ca are intended to interact. If the package is small, the vigne have very similar content.

Table of contents

1.1 Package name and metadata
1.1.1 Naming your package
1.1.2 Creating metadata for your package
1.2 Platforms
1.3 Package API
1.4 Code Style
1.5 CITATION file
1.6 README
1.7 Documentation
1.8 Documentation website
1.9 Authorship
1.10 Licence
1.11 Testing
1.12 Examples
1.13 Package dependencies
1.14 Recommended scaffolding
1.15 Version Control
1.16 Miscellaneous CRAN gotchas
1.17 Bioconductor gotchas
1.18 Further guidance

scope
standards
process
commitment
community

- **G5.4 Correctness tests** to test that statistical algorithms produce expected results to some fixed test data sets (potentially through comparisons using binding frameworks such as [RStata](#)).
 - **G5.4a** For n method from reference for simple, trivial implementation
 - **G5.4b** For n include test those implemented software, or
 - **G5.4c** Where outputs when available
- **G5.5 Correctness**
- **G5.6 Parameter** results given data algorithm should
- **G5.7 Algorithm performance tests** to test that implementation performs as expected as properties of data change. For instance, a test may show that parameters approach correct estimates within tolerance as data size increases, or that convergence times decrease for higher convergence thresholds.
- **G5.8 Edge condition tests** to test that these conditions produce expected behaviour such as clear warnings or errors when confronted with data with extreme properties including but not limited to:
 - **G5.8a** Zero-length data
 - **G5.8b** Data of unsupported types (e.g., character or complex numbers in for functions designed only for numeric data)
 - **G5.8c** Data with all- NA fields or columns or all identical fields or columns
 - **G5.8d** Data outside the scope of the algorithm (for example, data with more fields (columns) than observations (rows) for some regression algorithms)
- **G5.9 Noise susceptibility tests** Packages should test for expected stochastic behaviour, such as through the following conditions:
 - **G5.9a** Adding trivial noise (for example, at the scale of `.Machine$double.eps`) to data does not meaningfully change results
 - **G5.9b** Running under different random seeds or initial conditions does not meaningfully change results

scope
standards
process
commitment
community



scope
standards
process
commitment
community



cvitolo commented on Aug 28, 2016

Member



Thanks for the suggestion @jennybc! I have updated `rmarkdown` and `knitr` but that did not fix the problem. As a temporary workaround I manually edited the README.md file and removed the string .

I also noticed that the leaflet map is not rendered and on its place is the line , not sure what's going on there.



maelle commented on Aug 29, 2016


Member



@cvitolo great work! A few points

- There's no unit test for `ukair_get_coordinates` from what I see on codecov.io?
- In `ukair_get_hourly_data` you could output a warning if the side ID is not in the cached version of the catalogue. For instance imagine I enter a wrong ID, this is what I get:

scope
standards
process
commitment
community



ropensci-review-bot on Sep 25, 2024

Member


Checks for [fireexposuR \(v1.0.0\)](#)

git hash: [aacfb397](#)

- ✓ Package name is available
- ✓ has a 'code' field
- ✓ has a 'content-type' field
- ✓ uses 'roxygen2'
- ✓ 'DESCRIPTION' file exists
- ✓ 'DESCRIPTION' file is valid
- ✓ Package has a valid license
- ✓ All functions are documented
- ✓ Package has a valid namespace
- ✓ Package has a valid entry point
- ✓ Package has a valid R CMD SHLIB file
- ✓ R CMD SHLIB file is valid
- ✓ R CMD SHLIB file is valid
- ✓ Function names are valid

(Checks marked v)


Package License:




maurolepore on Oct 28, 2024

Member

@ropensci-review-bot assign **@ronnyhdez** as reviewer







ropensci-review-bot on Oct 28, 2024

Member

@ronnyhdez added to the reviewers list. Review due date is 2024-11-18. Thanks **@ronnyhdez** for accepting to review! Please refer to [our reviewer guide](#).

rOpenSci's community is our best asset. We aim for reviews to be open, non-adversarial, and focused on improving software quality. Be respectful and kind! See our reviewers guide and [code of conduct](#) for more.



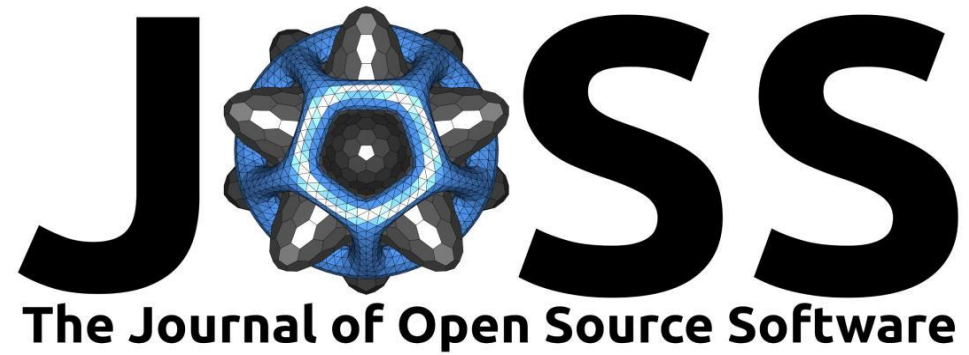


ropensci-review-bot on Oct 28, 2024

Member

@ronnyhdez: If you haven't done so, please fill [this form](#) for us to update our reviewers records.

scope
standards
process
commitment
community



scope
standards
process
commitment
community

- **G1.2** *Statistical Software should include a Life Cycle Statement describing current and anticipated future states of development.*

This package is

- In a stable state of development, with minimal subsequent development envisioned.
- In a stable state of development, with active subsequent development primarily in response to user feedback.
- In a stable state of development, with some degree of active subsequent development as envisioned by the primary authors.
- In an initially stable state of development, with a great deal of active subsequent development envisioned.

scope
standards
process
commitment
community

New Maintainer Wanted :-) #36

✓ Closed



maelle o

(or new r

If you're

info@rop

For more

- [our c](#)
- [our l](#)
- [Pack](#)
- the k

@ropens

Create su

rOpenSci 2024 Maintainer Survey

Thank you for being a package maintainer! We're grateful for your participation in rOpenSci.

We want to support package m we can. Therefore, each year we see how things are going. To se whether you would like to partic whether you are looking for a bi serving as a maintainer of your to support you as best we can, l

Would you allow us to feature you in a series of rOpenSci's social media posts telling the community about your work as a package maintainer?

☐

How is your work on your package(s) funded? *

This helps us understand potential barriers to development and maintenance. Please check all that apply.

- ☐ I work on the package as part of my job.
- ☐ I work on the package on my personal time.
- ☐ I have funding (grant, contract, etc) specifically to work on the package.
- ☐ I would rather not say.

GitHub username *

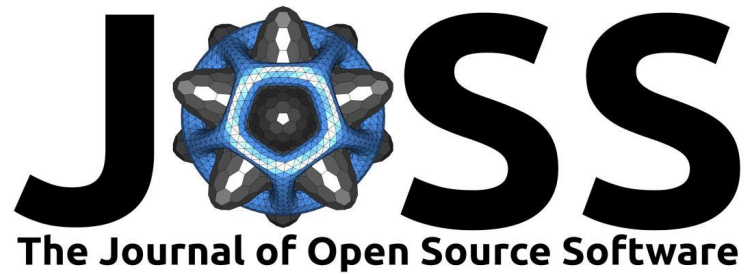
For instance "yabellini" or "jeroen"

Name *

Would you like rOpenSci to help find contributors to help with maintenance of your package?

Please briefly explain your needs. In particular, we can help polish and publicize help-wanted issues <https://ropensci.org/blog/2023/09/19/help-wanted/>

scope
standards
process
commitment
community



scope
standards
process
commitment
community



Thank you!



ropensci.org

 @noamross.net

noam@ropensci.org