

Contributing to data science communities - My experiences with rOpenSci

Najko Jahn (State and University Library Göttingen)

najko.jahn@sub.uni-goettingen.de

@najkoja

About me

Hi, I am Najko and work as a Data Analyst at the SUB Göttingen

My Role:

- Design and implement new services for analysing scholarly communication to support researchers, and to inform decision-making
- Particular focus on open data and open tools to access, discover and analyse metadata about scholarly works

How I work

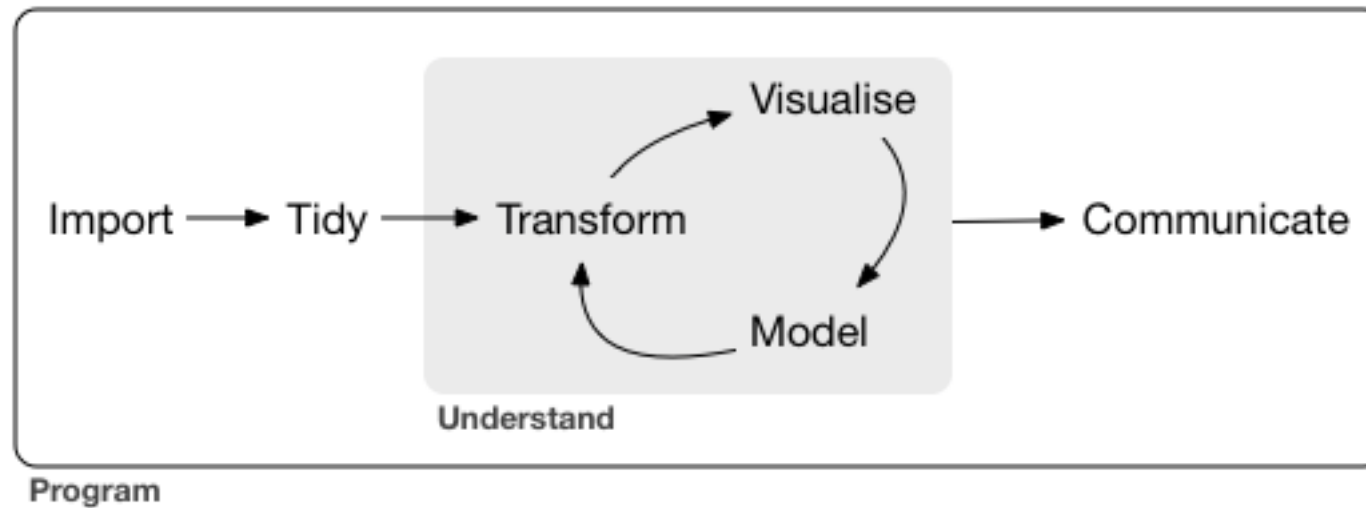


Figure from H. Wickham & G. Grolemund, 2017. *R for Data Science*. Boston.

Learning routes: challenges

- Career paths vary
- Lack of peers within the library
- Heterogeneous training landscape
- Sustainability and quality assurance of code



A photograph of three women sitting at a table with laptops, smiling and engaged in conversation. The image is overlaid with a semi-transparent blue filter. The woman on the left is wearing glasses and a name tag that says 'SAM'. The woman in the middle is wearing a name tag that says 'ERIN'.

Transforming science through open data and software

We build software with a community of users and developers, and educate scientists about transparent research practices.

Interact & Share

Use our packages to acquire your data from both your own and from various data sources, analyze it, add your narrative and generate a final document in any of widely used formats such as Word, Markdown, PDF or LaTeX. Combine our tools with the rich ecosystem of R packages.



**Access a Vast Trove
of Scientific Data**



**Interactive Data
Analysis and DataViz**



**Share Documents
and Your Data**

[BROWSE PACKAGES](#)

[Contribute your package to our suite](#)

Our community is our best asset

We are building a welcoming and diverse global community of software users and developers from a range of research domains. We aim to build capacity of novices, experts, and the “nexperts” in between. We welcome participation and civil conversations that adhere to our [code of conduct](#). There are many ways to get involved.

Follow us on [Twitter](#). **Read our blog** to learn about our software packages, best practices, events and the people who make up our community. **Subscribe** to [rOpenSci News](#) to get [semi-monthly updates](#) on our activities sent to your inbox.

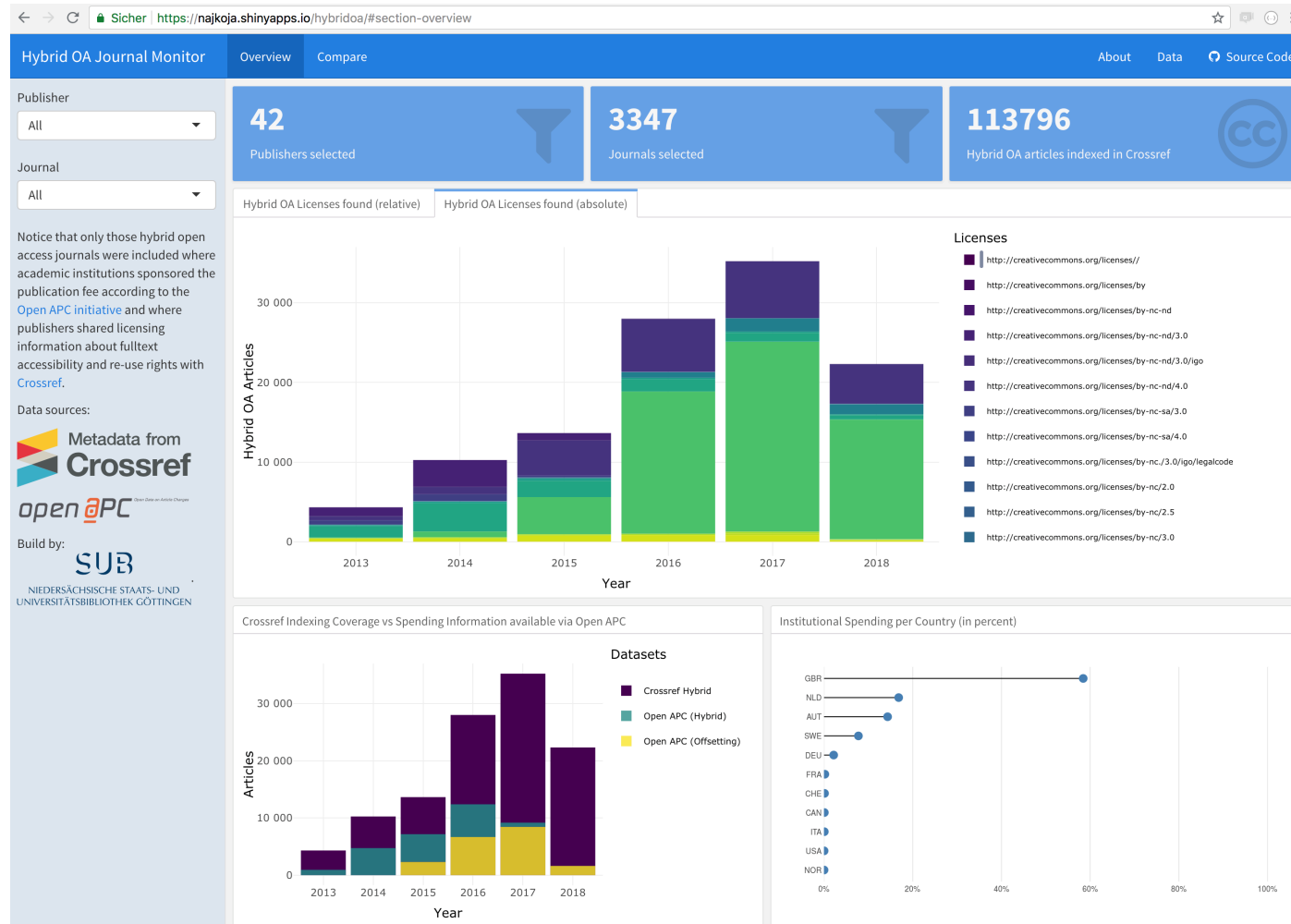
Ask and answer #rstats questions on our [discussion forum](#), on [Stackoverflow](#), or [Twitter](#). **Attend** our quarterly [Community Calls](#) to hear about best practices, Q & As with well known developers, and to learn more about rOpenSci developments. These are free and open for anyone to attend and provide an opportunity to connect with rOpenSci community members around the world.

Use our software. **Improve** package [documentation or code](#). **Review a package**. **Submit** your package for [open peer review](#) and onboarding. **Apply** to participate in our [annual unconference](#).

How do I use and participate in rOpenSci activities?

- Use rOpenSci as **resource to learn** about new data science developements and tools
- Use rOpenSci packages in my **work** and **provide feedback**
- **Develop and maintain** clients to access and discover literature databases with R
- **Review** packages

Hybrid OA Dashboard



rcrossref for data import

plotly for interactive visualisation

europemc - R Interface to Europe PMC RESTful Web Service

build passing

build passing

codecov 88%

CRAN 0.3

downloads 284/month

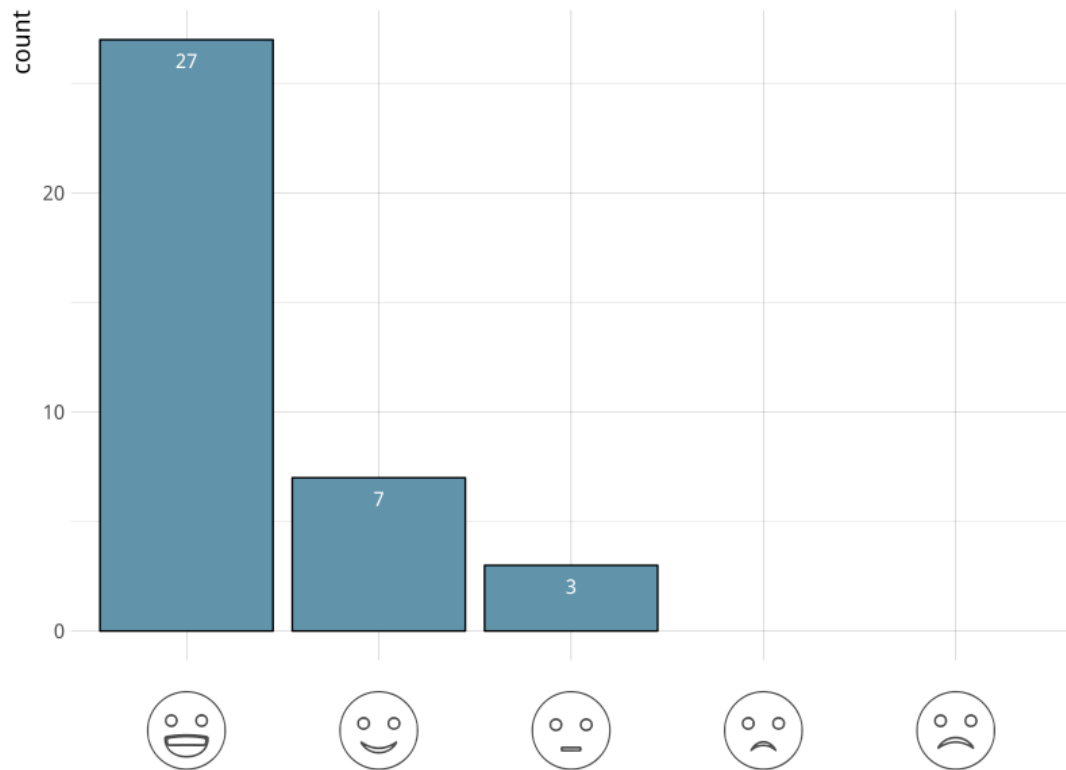
Peer Reviewed

europemc facilitates access to the [Europe PMC RESTful Web Service](#).

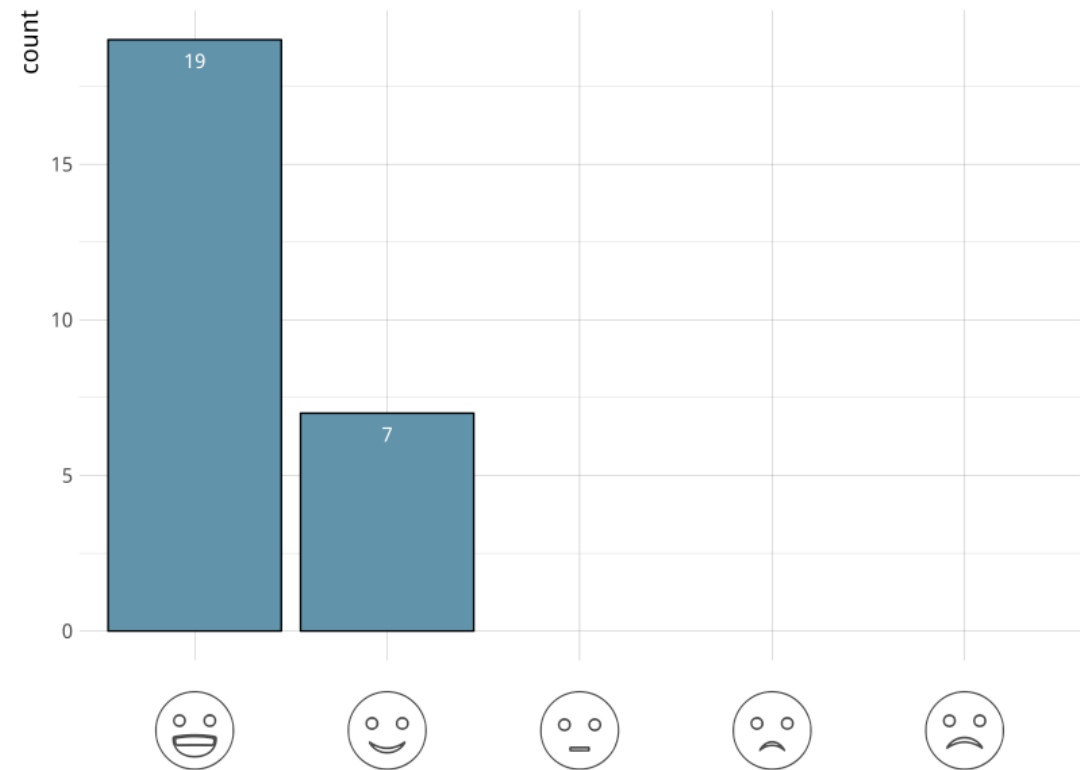
```
europemc::epmc_search("Lagotto Romagnolo")
#> # A tibble: 42 x 27
#>   id      source pmid  doi  title      authorString      journalTitle issue
#>   <chr> <chr> <chr> <chr> <chr>      <chr>              <chr>          <chr>
#> 1 28583... MED    28583... 10.1... Basal A... Syrjä P, Anwar ... Vet Pathol    6
#> 2 25945... MED    25945... 10.1... Behavio... Jokinen TS, Tii... J Vet Inter... 4
#> 3 24354... MED    24354... 10.1... FDG-PET... Jokinen TS, Haa... Vet Radiol ... 3
#> 4 17552... MED    17552... 10.1... Benign ... Jokinen TS, Met... J Vet Inter... 3
#> 5 17490... MED    17490... 10.1... Cerebel... Jokinen TS, Rus... J Small Ani... 8
#> 6 29056... MED    29056... 10.1... Relatio... Byosiere SE, Fe... Behav Proce... <NA>
#> 7 27525... MED    27525... 10.1... Genetic... Donner J, Kauko... PLoS One      8
#> 8 29166... MED    29166... 10.1... Frequen... Zierath S, Hugh... PLoS One     11
#> 9 29237... MED    29237... 10.1... Molecul... Yu Y, Hasegawa ... BMC Vet Res    1
#> 10 25875... MED    25875... 10.1... A misse... Kyöstilä K, Syr... PLoS Genet    4
#> # ... with 32 more rows, and 19 more variables: journalVolume <chr>,
#> #   pubYear <chr>, journalIssn <chr>, pageInfo <chr>, pubType <chr>,
#> #   isOpenAccess <chr>, inEPMC <chr>, inPMC <chr>, hasPDF <chr>,
#> #   hasBook <chr>, citedByCount <int>, hasReferences <chr>,
#> #   hasTextMinedTerms <chr>, hasDbCrossReferences <chr>,
#> #   hasLabsLinks <chr>, hasTMAccessionNumbers <chr>,
#> #   firstPublicationDate <chr>, pmcid <chr>, hasSuppl <chr>
```

Onboarding

As a reviewer, how satisfied are you overall with the review process?



As an author, how satisfied are you overall with the review process?



<https://ropensci.org/blog/2018/04/17/author-survey/>



tts commented on 28 Jun 2017 • edited ▾



Quite happy with fixes from my part.

One thing I didn't check earlier was the behaviour of the `Done` button in the RStudio addin. At the moment, it doesn't respond to a click. AFAIK you need to add a listener to `input$done` Something like

```
shiny::observeEvent(input$done, {  
  stopApp()  
})
```

as [explained here](#)

From my current (default) addins, [Add Crossref Citations](#) wraps `stopApp` with `invisible` but I'm not sure if that's necessary.



njahn82 referenced this issue in [ropensci/roadoi](#) on 28 Jun 2017

bugfix stopping shiny addin #15

Merged



njahn82 commented on 28 Jun 2017

Member



Very helpful @tts, I fixed the bug as suggested.



In summary

- Creating opportunities for skilling up is critical!
- Fostering community engagement of library staff in existing open science communities is mutual beneficial:
 - Engaging in data science communities educates librarians about open and reproducible research practices
 - It promotes re-use of open tools and services
 - Community participation helps to assure quality of coding