RSMF Soft Launch

February 9, 2025

Project Overview

Vision

The Research Software Maintenance Fund supports projects that improve sustainability, reduce technical debt, and enhance community engagement.

Our project, *Enabling the Next Generation of Contributors to R*, aims to put the R project on a more sustainable footing.

Kylie Ariel Bemis

- PhD in Statistics (Purdue University)

- Teaching faculty (Northeastern University)

- R packages

  - `Cardinal` - mass spectrometry imaging

  - `matter` - out-of-memory computing

- Interests

  - R language's C API

  - R performance and extensibility

- Tribal member of Zuni Pueblo

Mikael Jagan

- Worked as research software developer at McMaster University

  - Mathematical epidemiology, inference for infectious disease models

- Co-author with Martin Mächler on Matrix package

  - Efficiency, classes with integer and complex data types

- Co-author of lme4 package with Martin Mächler and Ben Bolker

  - Generalized to support structured covariance

- Maintainer of flint: interface to Flint C library

  - Extend R with arbitrary precision

- Developed expertise in S4 and methods package as a result of maintaining above packages

Jeroen Ooms - Background

- Team member of Tidyverse

- Lead infrastructure at rOpenSci for 10+ years

- Lead developer for https://r-universe.dev

- Maintains 50+ CRAN packages: see https://jeroen.r-universe.dev

- Former maintainer of Rtools toolchains

Jeroen Ooms - Activities/interests

R-devel activities:

- CI tooling for base R, e.g. https://contributor.r-project.org/svn-dashboard/

- Maintain R SVN mirrors, e.g. https://github.com/r-devel/r-svn

- Helping Bioconductor modernize infrastructure based on r-universe

Some current ideas/interests:

- Help CRAN build for win-arm64 using free GitHub hardware

- Patches to improve install.packages() and friends

- Help R project integrate CI tooling, e.g. to test patches (aka PRs)

Mentored Contributors - Partner Supported

Michael Chirico (Google)

- Member of R Contribution Working Group

- PhD in economics

- Quantitative/statistical work in industry since

- Maintains internal infrastructure for R at Google (~2000 users)

- R package contributor: data.table and bit64 (64-bit integers in R)

- Translation: admin of Weblate, translations in data.table

- R contribution: several tens of patches to R

- Will split between translation infrastructure and bug fixes for RSMF

Mossa Reimert (A2-Ai)

- Was postdoc in veterinary epidemiology

- Now Senior Scientific Software Engineer at A2-Ai

- Developed tooling to interface Rust with R:

  - libR_sys binding library and extendr crates to interface R in Rust

  - rextendr package to scaffold extendr-powered R packages

- Interested in build process (e.g. make files), tidying C code

Kevin Ushey (Posit)

- Working on RStudio IDE (even now that Positron is being developed)

- Maintainer of renv for managing reproducible environments

Mentored Contributors - Non-RSMF

Lluís Revilla

- Working at Roche

- Bioconductor: package maintainer, Code of Conduct Committee

- R contribution:

  - Spanish and Catalan translation

  - R bugs, CRAN utilities

- Working Groups: R Contribution, R Package Repositories

R Core - RSMF funded

Simon Urbanek

- President of the R Foundation

- Associate Professor in Statistics, University of Auckland (previously AT&T)

- R Project: particularly responsible for Mac-related work
  - Mac binaries, Mac-builder, Mac GUI, Quartz graphics device

- R packages: notably rJava providing R/Java interface

- R Project infrastructure, e.g. Bugzilla, Etherpad

Peter Dalgaard

- Professor, Copenhagen Business School
- R Core Team member since August 1997
- Release Manager since December 1997
- General debugging
- Bits & bobs

Release management

- ▶ Statistical software: High emphasis on correctness
- ▶ Conservative, carefully tested releases
- ▶ Release processes (alpha/beta/RC phases, code freeze timing)
- ▶ Scheduling (and names :))
- ▶ Two-branched development – what goes where?
- ▶ Release and pre-release tarballs (scripted)

Ad-hoc Mentors

Michael Lawrence

- **Affiliation**: Freenome   R Core
- **Open-source contributions**:
  - GGobi,
  - Bioconductor sequencing infrastructure,
  - R (shallow copying and other improvements/fixes)
- **Research interests**:
  - User interfaces (LLMs)
  - Scalability
  - Particularly in the context of Bioinformatics

Paul Murrell

- Associate Professor in Statistics, University of Auckland

- Developed grid graphics system
  - Low-level plotting facilities
  - Basis for higher-level packages like lattice, ggplot2

- R Project: particularly responsible for graphics-related issues

- Mentor at R Dev Days helping novice contributors navigate C code underlying graphical functions

Deepayan Sarkar

- Professor at the Indian Statistical Institute, Delhi

- Creator of lattice graphics

- Work related to documentation

  - Converting Rd to HTML

  - Creating 1-page HTML reference manuals for packages (now shown on CRAN landing pages alongside PDF versions)

Luke Tierney: University of Iowa and R-core

- Statistical areas
    - Computational methods for Bayesian inference
    - Dynamic graphics

- Computational projects
    - Lisp-Stat
    - R-Core (since 1997)

- R-Core contributions
    - Low level infrastructure
    - Language features

Low level infrastructure

- ▶ Memory management
- ▶ Byte code compilation
- ▶ Serializaiton
- ▶ Profiling support
- ▶ ALTREP

Current and future projects

- ▶ API clarification and cleanup
- ▶ Compiler improvements
- ▶ Complex assignment improvements
- ▶ Support for larger integers

Language features

- ▶ Environments on formulas
- ▶ `with()` function
- ▶ Active bindings
- ▶ Exception handling system
- ▶ Namespaces
- ▶ Raw strings
- ▶ Native pipe operator

Other R Core

Martin Maechler - Background

- Role & Foundation

  - A founding member of the R Core Development Team (since 1997).

  - Current Secretary General of the R Foundation.

- The "Open Source" Catalyst

  - In 1995, Mächler famously persuaded R's creators (Ihaka & Gentleman) to release R under the GNU General Public License (GPL)

- Academic Background

  - Professor Emeritus and Adjunct Professor at ETH Zürich (Seminar for Statistics).

  - Personal Website: https://people.math.ethz.ch/~maechler/

Martin Maechler - Activities

- Expertise:
    - Numerical accuracy, robust statistics, high-performance computing.
    - Commit History: One of the most active contributors to base R.
- Community Support:
    - Manages dozens of R-related mailing lists, including R-help and R-devel.
- Developer Tooling
    - Project Leader for ESS (Emacs Speaks Statistics).
- Maintains Foundational "recommended" R packages:
    - Matrix: Support for sparse and dense matrix classes.
    - cluster: Methods for cluster analysis.

Contractors (Translation Infrastructure)

Fernando Corrêa

- PhD candidate and statistician at R6 in São Paulo, Brazil

- R6 is a data science consultancy firm in Brazil

  - Books, online courses, and blog posts for local community

  - Packages: including auth0 for Auth0 authentication in Shiny

- R6 staff have participated in several R Dev Days

- R6 will be the contractor for work on translation infrastructure

  - I will be the main developer

  - Other members of the Brazilian community will contribute, including Beatriz Milz

Hi, I'm Beatriz Milz (Bea)! 🇧🇷

- PhD in Environmental Sciences (University of São Paulo)
- Communities: R-Ladies São Paulo, rOpenSci, LatinR
- Co-organized the satRday 2019 in São Paulo
- Translation activities:
    - [rOpenSci dev guide](#) (Portuguese)
    - [R for Data Science](#) (2nd ed; Portuguese)
- Brazilian R community, in collaboration with R6 Consulting

Gergely Daróczi

- R package developer, e.g. for logging transforming R objects into R markdown

- Community organizer: local R User Group and local conferences.

- Set up Weblate server as experimental platform for contributing translations to base R

  - Maintaining alongside Mike Chirico

Partners

Posit

- Represented at meeting by Hadley Wickham, Chief Scientist
- Providing support especially software engineering support through
  - Tidyverse Team
  - George Stagg (AI team)
  - Kevin Ushey (RStudio team)

A2-Ai

- Science and technology solutions company, especially in pharma & life sciences

  - Open source packages for analytics, technical writing, and visualization

  - Tools for package management and testing

- Providing support via

  - Mossa Reimert (general R contribution)

  - Devin Pastoor (AI review of CRAN packages)

R Consortium

- Represented at meeting by Terry Christiani, Executive director

- R Consortium is a group of business supporting the R project

- Support different projects in the R ecosystem via

  - Working Groups

  - Infrastructure Steering Committee grants

- Providing support for developer events, especially travel support

R Foundation

- R Foundation providing travel support for R Core members to attend related events

- Also acting as legal entity through which R Core members are providing support on a voluntary basis

rOpenSci

- Represented at meeting by Noam Ross, Executive Director

    - Runs software peer review process

    - Works in epidemiology and data journalism

- rOpenSci support

    - Contract of record for some of the freelancers working on the project

    - Share expertise in community building and outreach

Next Steps

Contract

- 2nd version of main contract between Birmingham and SSI under review

- Non-academic approved contractors will need to sign a form to declare recent UK/EU funding (NumFOCUS, Jeroen Ooms, Upshift R Kft and R6 Estatistica e Treinamenos)

- Contracts with all partners will follow

Mentored Cohort

- Mentees will be paired with personal mentors

  - Guide through the program

  - Meet approx monthly

- Bug components will be distributed amongst cohort

- Projects (currently in Google Doc) will be posted on
  https://github.com/r-devel/rsmf/

  - Mentees and mentors assigned/express interest

- Mentee to prepare Q1 plans in discussion with personal and project mentees

- Chat channel initially on Slack until Zulip ready

Translations work

- R6 will create a translations patch for R-4.6.0 using existing set-up.

- From learnings, outline plans to improve processes.

Other Work in Q1

- Launch Zulip

- Drafting CoC and associated protocols

- Planning for 2026 R Dev Days and Sprints

  - Rencontres R, Nantes, Fri 19 June

  - CascadiaR, Portland, Fri 26 June

  - useR!, Warsaw, Fri 10 July

  - Birmingham, 2-4 September

- Updates to R Foundation statutes accommodating online voting

- Ongoing events: Office Hours, participation in EDIA groups

Communication

- Quarterly reports required by funder

- Monthly main team meeting

  - ~8:30 UTC October - March

  - Alternating UK/NZ friendly times April - September

  - Short report-backs, discussion and look ahead

- Monthly R Contribution Working Group meetings for discussing outreach, development infrastructure and project governance

- Ad-hoc meetings on specific projects

Discussion