Solving the Package Management Puzzle

Devin Pastoor
Vice President, Technology Solutions
Metrum Research Group
devinp@metrumrg.com
What does Metrum do?
i've seen things...

terrible things...
Strategizing and advocating for R is hard because we’re trying to solve a much larger slice of problems with a single toolchain
Clearly see consistent principles

Reproducibility, collaboration, discoverability, efficiency, accuracy are all key elements that drive and validate decisions.
It needs to work

And should keep working consistently
What do you do with R?

- Tables and Figures
- Regulatory Statistical analysis
- Voodoo AI
- Build applications that will run at longer time-scales (shiny apps/api’s etc)
What does R look like in your organization
Where does package management fit in?
No one tool/product addresses these needs or will address these needs holistically.
Many Decisions

Discovery
- New Github Packages
- Package Changes

New CRAN packages
- Usage
- Release Timelines

Decision making
- Validation Criteria
- Licensing
- Dependency Graphs

Reproducibility
- Consistent Package Versions
- Immutable Repositories

R ecosystem

Development
- Testing
- Best Practices
- SDLC

Publishing
- Documentation
What are (generally) unknown risks groups are opening themselves up to?
lockfiles are great...until they’re not
packrat
...
Package: hms
Source: CRAN
Version: 0.4.2
Hash: b4096a4f6a6736138e9a825c2baaacf0
Requires: pkgconfig, rlang

Package: htmltools
Source: CRAN
Version: 0.3.6
Hash: b24df7ea0856eab6618f6a56016d940d
Requires: Rcpp, digest

Package: htmlwidgets
Source: CRAN
Version: 1.2
Hash: de18b75f31630089b22e30d4b188cfbe
Requires: htmltools, jsonlite, yaml
...
Upgrading package versions in a project
Why lockfiles are not the (sole) solution for businesses

- Unclear how lockfile came to be constructed
- Too easy to fragment
- Unclear what packages drove the creation of lockfiles
- How to upgrade or add packages over time?
- Does not (alone) mean it will be achievable to re-construct cross-platform
CRAN is great...until it’s not
It tests against many platforms and configurations!
CRAN Package Check Results for Package dplyr

Last updated on 2018-08-15 15:46:57 CEST.

<table>
<thead>
<tr>
<th>Flavor</th>
<th>Version</th>
<th>T_install</th>
<th>T_check</th>
<th>T_total</th>
<th>Status</th>
<th>Flags</th>
</tr>
</thead>
<tbody>
<tr>
<td>r-devel-linux-x86_64-debian-clang</td>
<td>0.7.6</td>
<td>316.61</td>
<td>124.08</td>
<td>440.69</td>
<td>OK</td>
<td></td>
</tr>
<tr>
<td>r-devel-linux-x86_64-debian-gcc</td>
<td>0.7.6</td>
<td>307.18</td>
<td>96.59</td>
<td>403.77</td>
<td>OK</td>
<td></td>
</tr>
<tr>
<td>r-devel-linux-x86_64-fedora-clang</td>
<td>0.7.6</td>
<td></td>
<td>640.96</td>
<td></td>
<td>NOTE</td>
<td></td>
</tr>
<tr>
<td>r-devel-linux-x86_64-fedora-gcc</td>
<td>0.7.6</td>
<td></td>
<td>637.13</td>
<td></td>
<td>NOTE</td>
<td></td>
</tr>
<tr>
<td>r-devel-windows-ix86+x86_64</td>
<td>0.7.6</td>
<td>588.00</td>
<td>217.00</td>
<td>805.00</td>
<td>NOTE</td>
<td></td>
</tr>
<tr>
<td>r-patched-linux-x86_64</td>
<td>0.7.6</td>
<td>358.81</td>
<td>114.24</td>
<td>473.05</td>
<td>OK</td>
<td></td>
</tr>
<tr>
<td>r-patched-solaris-x86</td>
<td>0.7.6</td>
<td></td>
<td>546.80</td>
<td></td>
<td>NOTE</td>
<td></td>
</tr>
<tr>
<td>r-release-linux-x86_64</td>
<td>0.7.6</td>
<td>363.78</td>
<td>113.02</td>
<td>476.80</td>
<td>OK</td>
<td></td>
</tr>
<tr>
<td>r-release-windows-ix86+x86_64</td>
<td>0.7.6</td>
<td>619.00</td>
<td>242.00</td>
<td>861.00</td>
<td>NOTE</td>
<td></td>
</tr>
<tr>
<td>r-release-osx-x86_64</td>
<td>0.7.6</td>
<td></td>
<td></td>
<td></td>
<td>NOTE</td>
<td></td>
</tr>
<tr>
<td>r-oldrel-windows-ix86+x86_64</td>
<td>0.7.6</td>
<td>819.00</td>
<td>286.00</td>
<td>1105.00</td>
<td>NOTE</td>
<td></td>
</tr>
<tr>
<td>r-oldrel-osx-x86_64</td>
<td>0.7.6</td>
<td></td>
<td></td>
<td></td>
<td>NOTE</td>
<td></td>
</tr>
</tbody>
</table>

I mainly want to take this opportunity to flex for CRAN! 💪

The package is a port of a C++ Python module to use Rcpp.

I was originally booted from CRAN due to some memory leak issues which had gone unnoticed for years in 🙁! CRAN checks caught the issue immediately. #rstats
It contains the latest blessed code!
Replying to @GoldbergData

reprex.tidyverse.org but you should install from GitHub, the CRAN version is broken (and not authored by me, despite what DESCRIPTION says)

12:26 AM - 4 Jun 2018

2 Retweets  9 Likes
Hadley Wickham @hadleywickham · May 15

Replying to @P_Boueri @MichaelFrasco

We also have enough development resources we can rapidly roll out a fix if we need to. Currently CRAN release time is biggest risk: dplyr is broken on CRAN; dev version is ok and we submitted almost a month ago and it’s still not up
It builds binaries!
Joris Meys @JorisMeys

. @krmlr Just noticed there's a pretty serious conflict between #rstats pkg rlang 0.3.0 and tibble 2.0.0. It's related to the soft deprecation of new_list_along().

I've seen a 2.0.1 milestone for tibble on @github so I guess you're aware already. Any idea on the timeline?

7:19 AM - 12 Jan 2019

2 Retweet 2 Likes

Kirill Müller @krmlr · 6h
Replying to @JorisMeys @github
The update should appear on CRAN any time now. It's delayed by a package that accesses some API in its tests and led to a false positive with CRAN's automated checks.

lionel @lionelhenry · 33m
Replying to @JorisMeys @krmlr @github
the conflict is not serious, it's just a failing assumption in unit tests. However it does prevent the creation of CRAN binaries for tibble 2.0, which can cause updating issues...
devtools: Tools to Make Developing R Packages Easier

Collection of package development tools.

Version: 2.0.0
Depends: R (≥ 3.0.2)
Imports: callr, cli, digest, git2r (≥ 0.23.0), http (≥ 0.4), jsonlite, memoise (≥ 1.0.0), methods, pkgbuild (≥ 1.0.2), pkgload (≥ 1.0.1), rcmdrcheck (≥ 1.3.0), remotes (≥ 2.0.0), rstudioapi (≥ 0.6.0.9000), sessioninfo, stats, tools, usethis (≥ 1.4.0), utils, withr
Suggests: BioCInstaller, BioCManager, bitops, covr (≥ 3.2.0), crayon, curl (≥ 0.9), evaluate, fognhorn (≥ 1.0.1), gmailr (≥ 0.7.0), knitr, lintr (≥ 0.2.1), mockery, pingr, MASS, pkgdown, Rcpp (≥ 0.10.0), rhub, rmarkdown, roxygen2 (≥ 6.1.0), rversions, spelling (≥ 1.1), testthat (≥ 1.0.2), whisker
Published: 2018-10-19
Author: Hadley Wickham [aut], Jim Hester [aut, cre], Winston Chang [aut], RStudio [cph], R Core team [ctb] (Some namespace and vignette code extracted from base R)
Maintainer: Jim Hester <james.hester at rstudio.com>
BugReports: https://github.com/r-lib/devtools/issues
License: GPL-2 | GPL-3 [expanded from: GPL (≥ 2)]
URL: https://github.com/r-lib/devtools
NeedsCompilation: no
Materials: README NEWS
CRAN checks: devtools results

Downloads:
Reference manual: devtools.pdf
Vignettes: Devtools dependencies
Package source: devtools_2.0.0.tar.gz
OS X binaries: r-release: devtools_1.13.6.tgz, r-oldrel: devtools_1.13.6.tgz
Old sources: devtools archive
devtools: Tools to Make Developing R Packages Easier

Collection of package development tools.

Version: 2.0.2
Depends: R (≥ 3.0.2)
Imports: callr, cli, digest, git2r (≥ 0.23.0), http (≥ 0.4), jsonlite, memoise (≥ 1.0.0), pkgbuild (≥ 1.0.2), r sessioninfo, stats, tools, usethis (≥ 1.4.0), utils, withr
Suggests: BiocManager, bitops, covr (≥ 3.2.0), crayon, curl (≥ 0.9), evaluate, foghorn (≥ 1.1.0), mailr (≥ 1.0.2), markdown, roxygen2 (≥ 6.1.0), rversions, spelling (≥ 1.1), testthat (≥ 1.0.2), whisker
Published: 2019-04-08
Author: Hadley Wickham [aut], Jim Hester [aut, cre], Winston Chang [aut], RStudio [cph], R Core tea
Maintainer: Jim Hester <james.hester at rstudio.com>
BugReports: https://github.com/r-lib/devtools/issues
License: GPL-2 | GPL-3 [expanded from: GPL (≥ 2)]
URL: https://github.com/r-lib/devtools
NeedsCompilation: no
Materials: README NEWS
CRAN checks: devtools results

Downloads:

Reference manual: devtools.pdf
Vignettes: Devtools dependencies
Package source: devtools_2.0.2.tar.gz
Windows binaries: r-devel: devtools_2.0.1.zip, r-release: devtools_2.0.1.zip, r-oldrel: devtools_2.0.2.zip
OS X binaries: r-release: devtools_1.13.6.tar.gz, r-oldrel: devtools_1.13.6.tar.gz
Old sources: devtools archive
devtools: Tools to Make Developing R Packages Easier

Collection of package development tools.

Version: 2.0.2
Depends: R (≥ 3.0.2)
Imports: callr, cli, digest, git2r (≥ 0.23.0), http (≥ 0.4), jsonlite, memoise (≥ 1.0.0), pkgbuild (≥ 1.0.2), rsessioninfo, stats, tools, usethis (≥ 1.4.0), utils, withr
Suggests: BioCManager, bitops, covr (≥ 3.2.0), crayon, curl (≥ 0.9), evaluate, fohorn (≥ 1.1.0), gmapr (≥ 1.0.2), rmarkdown, roxygen2 (≥ 6.1.0), rversions, spelling (≥ 1.1), testthat (≥ 1.0.2), whisker
Published: 2019-04-08
Author: Hadley Wickham [aut], Jim Hester [aut, cre], Winston Chang [aut], RStudio [cph], R Core team
Maintainer: Jim Hester <james.hestert@rstudio.com>
BugReports: https://github.com/r-lib/devtools/issues
License: GPL-2 | GPL-3 [expanded from: GPL (≥ 2)]
URL: https://github.com/r-lib/devtools
NeedsCompilation: no
Materials: README NEWS
CRAN checks: devtools results

Downloads:

Reference manual: devtools.pdf
Vignettes: Devtools dependencies
Package source: devtools_2.0.2.tar.gz
Windows binaries: r-devel: devtools_2.0.2.zip, r-release: devtools_2.0.2.zip, r-oldrel: devtools_2.0.2.zip
OS X binaries: r-release: devtools_2.0.2.tgz, r-oldrel: devtools_1.13.6.tgz
Old sources: devtools archive

Reverse dependencies:
Version: 2.0.1
Check: tests
Result: ERROR

Running ‘test-that.R’ [191s/202s]
Running the tests in ‘tests/test-that.R’ failed.
Last 13 lines of output:
Loading required package: devtools
<simpleError: pandoc version 1.12.3 or higher is required and was not found (see the help page ?rmarkdown::pandoc_available).>  
--- 1. Error: Package readme can be built (@test-build-readme.R#8) -----------------------------  
pandoc version 1.12.3 or higher is required and was not found (see the help page ?rmarkdown::pandoc_available).  
  1: build_readme("testReadme") at testthat/test-build-readme.R:8  
     2: callr::r(build, args = list(pkg_path = pkg$path, readme_path = readme_path, ... = ..., quiet = quiet), show = TRUE, spinner = FALSE)  
     3: get_result(output = out, options)  

--- testthat results ------------------------------------------------------------------------  
OK: 42 SKIPPED: 11 FAILED: 1  
1. Error: Package readme can be built (@test-build-readme.R#8)  

Error: testthat unit tests failed
Execution halted
Flavor: r-release-osx-x86_64
Why CRAN is not the (sole) solution for businesses

- Moving target
- Platform inconsistencies
  - Binary vs src
  - non-CRAN platform
- CRAN philosophies not necessarily aligned with enterprise business
  - “We need it to work everywhere”
  - What is secure/unsecure
  - If you’re on CRAN we don’t want you to break...but good luck otherwise
- CRAN submission acceptance and timelines can be opaque
- Even professional developers are still people - everyone makes mistakes
THIS IS FINE.
Think about **First Principles**

the fundamental concepts or assumptions on which a theory, system, or method is based.
Package repositories should be treated immutably
Package installation requirements should be declarative and customizable
Changes to any system should be clearly conveyed to the user before action takes place
Individual tools should be modular and extensible
No description, website, or topics provided.

Manage topics

- 495 commits
- 11 branches
- 24 releases
- 4 contributors

Branch: master v0.5.0-beta.3

- doc: update readme with newer installer
  - release: v0.4.0
  - cmd
  - configlib
  - cran
  - desc
  - docs/images
  - gpsr
  - integration_tests
  - logger

Latest commit b667fe1f 19 hours ago
3 months ago
3 months ago
4 months ago
4 months ago
22 days ago
4 months ago
3 months ago
5 months ago

Create new file Find File

Clone or download
How do you start the conversation of what works for you and your organization?
Identify the the risks/gaps that matter to you
If you want others to join - get your approach and code in the “open” early.

Communication is Key
When possible, stand on the shoulders of others

But remember first principles
Come talk to me

Find me at devinp@metrumrg.com and @devinpastooror (I don’t tweet much)