Comparing package releases
with diffify.com

Dr. Myles Mitchell

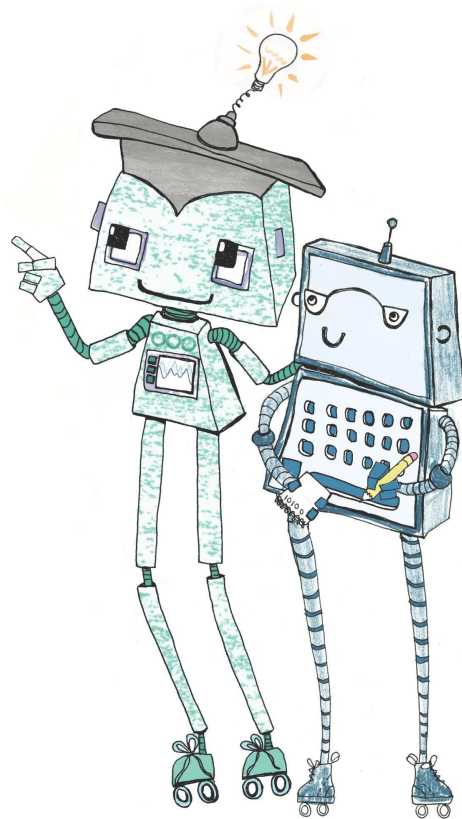Data Scientist @
Jumping Rivers

diffify

jumping rivers

Jumping Rivers

Bespoke data science solutions in support & development

✉ hello@jumpingrivers.com 🐦 @jumping_uk

- Machine learning
- Dashboard development
- R packages
- APIs
- Data pipelines
- Code review

Me and diffify

- **3-month internship at JR in 2020**

  - Started Diffify with a few interns

- **Finished my PhD in 2021**

- **Joined full time in January 2022**

  - Larger team now working on Diffify

  - Official launch **29th April**!

Motivation

- **A painfully familiar scenario**…

  - You're working on a big coding project 😀

  - You log in one day to find that nothing works! 😱

  - One of the project dependencies has been updated… 😩

- **If you're lucky**…

  - NEWS file

  - GitHub release notes

Motivation

- **Diffify to the rescue! 😎**

  - Compare R package versions at the click of a button

  - Understand how the changes will affect your code

  - Discover new package features

- **Currently works for R packages on CRAN**

- **Extending to other package repositories and Python!**

CRAN

- **Comprehensive R Archive Network**

  - ~20k packages

  - Every day: ~8 new packages; ~35 package updates

- **That's a lot to keep track of...**

Inspiration - CRANberries

CRAN:New   CRAN:Updated   CRAN:Removed   Misc   About

Package chameleon updated to version 0.2-2 with previous version 0.2-1 dated 2022-07-13

Title: Automatic Colors for Multi-Dimensional Data
Description: Assign distinct colors to arbitrary multi-dimensional data, considering its structure.
Author: Oren Ben-Kiki [aut, cre], Weizmann Institute of Science [cph]
Maintainer: Oren Ben-Kiki <oren@ben-kiki.org>

Diff between chameleon versions 0.2-1 dated 2022-07-13 and 0.2-2 dated 2022-07-13

```
DESCRIPTION            |   6 +++---
MD5                    |   8 ++++----
NEWS                   |   2 +-
R/chameleon.R          |  16 ++++++++--------
inst/doc/examples.html |   6 +++---
5 files changed, 19 insertions(+), 19 deletions(-)
```

More information about chameleon at CRAN
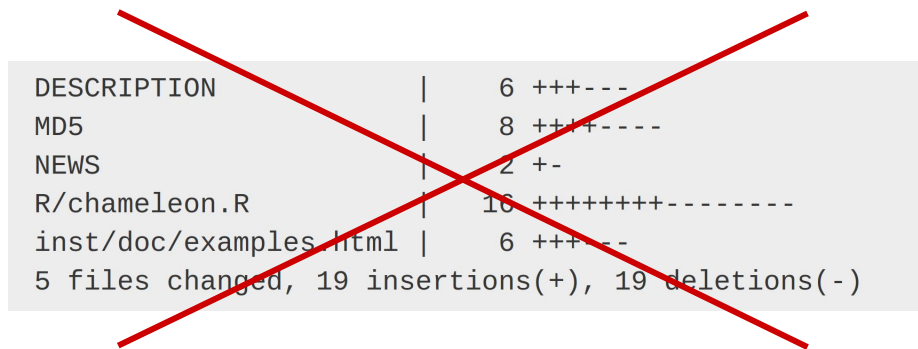Permanent link

How diffify differs…

- **Focus on the most important changes**

  - Dependencies

  - Exports

  - Functions

- **Readability + accessibility**

  - Neatly organised, interactive summaries

  - Avoid illegible git diffs

```
DESCRIPTION              |    6 +++---
MD5                      |    8 ++++----
NEWS                     |    2 +-
R/chameleon.R            |   16 +++++++--------
inst/doc/examples.html   |    6 +++--
5 files changed, 19 insertions(+), 19 deletions(-)
```

May make sense to package devs, not to users!

Live demo...

- **Worked example**

    - You've been developing a Shiny app which depends on {dplyr}

    - Updated {dplyr} from version 0.8.2 to 1.0.2 - now nothing works!

    - Let's use Diffify to find out what has changed!

- **Homepage:** https://diffify.com/

R package structure

- **Really strict package requirements:**

    - DESCRIPTION - package metadata

    - NAMESPACE - package exports / imports

    - man/*.Rd - package documentation

- **Sometimes:**

    - NEWS - self-explanatory

```
<package>
|    DESCRIPTION
|    MD5
|    NAMESPACE
|    NEWS.md
|    README.md
|
└──man
|    |    *.Rd
|    |
|    |
└──R
|    |    *.R
|    |
```
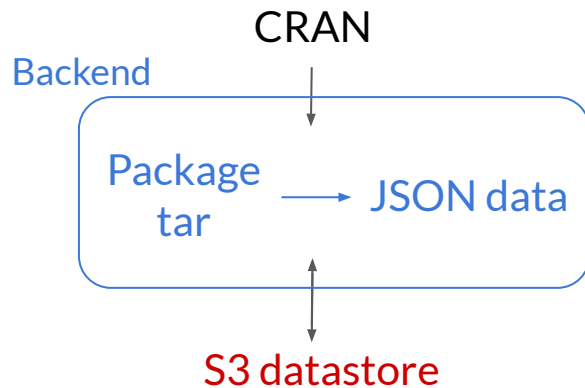
Workflow

- **Backend**

  - Built in Python (some R)

  - Every 2 hrs - looks for updates on CRAN

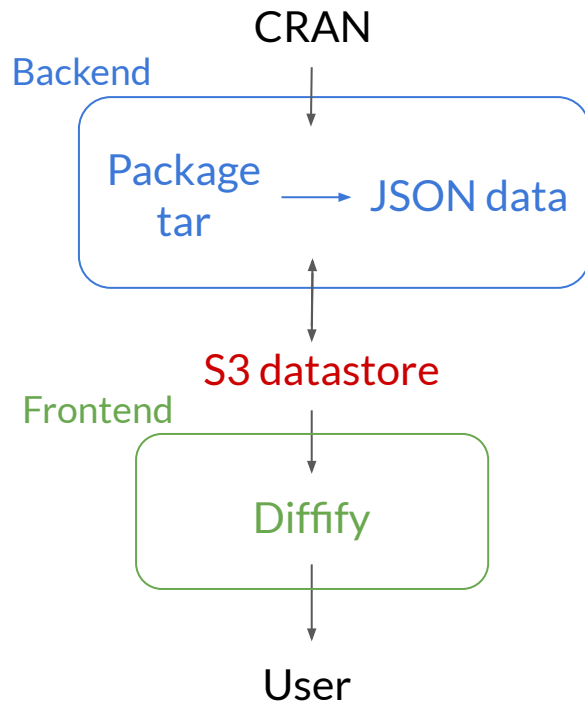 - Pushes package data to remote datastore

CRAN

Backend

Package tar → JSON data

S3 datastore

Workflow

- **Frontend**

  - Built using JavaScript + CSS

  - Queries datastore

  - Presents diff to the user

CRAN

Backend

Package tar → JSON data
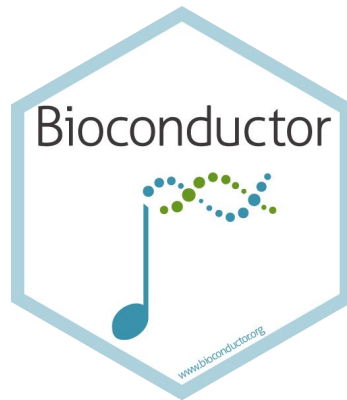
S3 datastore

Frontend

Diffify

User

Future of diffify...

- **Extend beyond CRAN**

  - Other repositories - Bioconductor

  - Dev versions - GitHub

- **Extend beyond R**

  - Python packages from PyPI

  - **Hopefully** going live September

Acknowledgements

- **Diffify is a team effort**

- **Backend devs**
  - Jamie Owen
  - Russ Hyde
  - Parisa Gregg

- **Frontend devs**
  - Tim Brock
  - Liam Kalita

- **General guidance / opinions** - Colin Gillespie

Shiny conference

- Please come to our in-person **Shiny conference** in October!

- Registration and event details here: https://shiny-in-production.jumpingrivers.com

Thanks for listening!

Useful links:

- App: https://diffify.com/

- Github: https://github.com/jumpingrivers/diffify

- Email: myles@jumpingrivers.com

- Shiny conference: https://shiny-in-production.jumpingrivers.com

Diffify - Homepage

Welcome to diffify!

Diffify provides you with a comparison between different versions of any R package stored on CRAN.

Say you were using a particular version of a package in a project and now a new version is available. With Diffify you are easily able to check what has been changed in the new release. In particular, Diffify will provide you with information from the News file as well as changes in the dependencies, namespace and functions of the package.

Get started!

Report an issue

Diffify - News

Package

dplyr

Later version

1.0.9

Earlier version

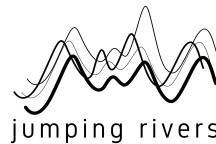1.0.3

News

1.0.9   1.0.8   **1.0.7**   1.0.6   1.0.5   1.0.4

- across() uses the formula environment when inlining them (#5886).

- summarise.rowwise_df() is quiet when the result is ungrouped (#5875).

- c_across() and across() key deparsing not confused by long calls (#5883).

- across() handles named selections (#5207).

Dependencies

Diffify - Dependencies

Dependencies

Depends | **Imports** | Suggests | Enhances

[filter..]

Added ➕ 1   **Removed ➖ 1**   **Changed ⇄ 4**   6 results

➕ **pillar**   1.5.1

➖ **ellipsis**

⇄ **lifecycle**   0.2.0 changes to 1.0.1

⇄ **rlang**   0.4.9 changes to 1.0.2

Diffify - Namespace

📑 **Namespace** ▲

| filter.. |

Added ➕ **22**　　**Removed** ➖ **5**　　**export** **6**　**S3method** **21**

27 results

➕ **add_count.data.frame**　　　　　　　　　　　　**S3method**

➕ **add_count.default**　　　　　　　　　　　　　**S3method**

➕ **filter_bullets.default**　　　　　　　　　　　**S3method**

➕ **filter_bullets.dplyr:::filter_incompatible_size**　**S3method**

➕ **filter_bullets.dplyr:::filter_incompatible_type**　**S3method**

⚒ Functions ▲

filter..

| Added ➕ 5 | Removed ➖ 1 | Changed ⇄ 7 | 13 results |

➕ if_all   Arguments ⌃

	Version 1.0.9	
	Name	**Default value**
1	.cols	everything()
2	.fns	NULL
3	...	
4	names	NULL

Diffify - Functions
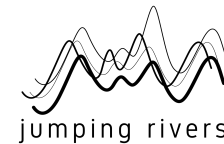
Added + 5   Removed — 1   Changed ⇄ 7   13 results

⇄ compute   Arguments ⌄

⇄ distinct_prepare   Arguments ⌃

	Version 1.0.3		Version 1.0.9	
	Name	**Default value**	**Name**	**Default value**
1	.data		.data	
2	vars		vars	
3	group_vars	character()	group_vars	character()
4	.keep_all	FALSE	.keep_all	FALSE
5			caller_env	caller_env(2)