

RECURSIVE PACKAGE DEPENDENCIES

UNRAVELING A TANGLED WEB

Brian Burns **UPTAKE**[®]



@bburns632



@_Brian_Burns

70%

of R packages have at least one package dependency

70%

of R packages have at least one package dependency

12%

of R packages have six or more package dependencies

An R Package Dependency

An R package dependency is another R package that must be installed in order for your package to work.

Using them **is good!**

Package dependencies typically provide:

- Non-trivial functionality
- Additional programming support
- Additional testing
- Time savings

ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics

A system for 'declaratively' creating graphics, based on "The Grammar of Graphics". You provide the data, tell 'ggplot2' how to map variables to aesthetics, what graphical primitives to use, and it takes care of the details.

Version:	3.1.1
Depends:	R (≥ 3.1)
Imports:	digest , grid , gtable (≥ 0.1.1), lazyeval , MASS , mgcv , plyr (≥ 1.7.1), reshape2 , rlang (≥ 0.2.1), scales (≥ 0.5.0), stats , tibble , viridisLite , withr (≥ 2.0.0)
Suggests:	covr , dplyr , ggplot2movies , hexbin , Hmisc , lattice , mapproj , maps , maptools , multcomp , munsell , nlme , testthat (≥ 0.11.0), vdiff , quantreg , knitr , rgeos , rpart , rmarkdown , sf (≥ 0.3-4), svglite (≥ 1.2.0.9001)
Enhances:	sp
Published:	2019-04-07
Author:	Hadley Wickham [aut, cre], Winston Chang [aut], Lionel Henry [aut], Thomas Lin Pedersen [aut], Kohske Takahashi [aut], Claus Wilke [aut], Kara Woo [aut], RStudio [cph]
Maintainer:	Hadley Wickham <hadley at rstudio.com>

If a package
dependency does
not work,
your package will
not work.

Expectations for a Package Dependency

By including a package as a dependency (or just into your regular use), you are **trusting** this package:

1. to be updated for newer versions of R,
2. to be maintained with good versioning,
3. to have no breaking changes, and
4. **to have trustworthy dependencies as well**

... and those dependencies to have trustworthy dependencies

... and those dependencies to have trustworthy dependencies

RECURSION!!!

Don't Do This

Here are some tips!



Managing Recursive Dependencies

1. **Identify** the extent of a package's recursive dependencies
2. **Evaluate** the the risk of at least the first level of package dependencies
3. **Mitigate** a risky dependency

Identify: Via CRAN & Clickin'

Seems
straightforward
enough.

Only five
dependencies.

gcForest: Deep Forest Model

R application programming interface (API) for Deep Forest which based on Zhou and Feng (2017). Deep Forest: Towards an Alternative to Deep Neural Networks. (<[arXiv:1702.08835v2](#)>) or Zhou and Feng (2017). Deep Forest. (<[arXiv:1702.08835](#)>). And for the Python module 'gcForest' (<<https://github.com/pylablanche/gcForest>>).

Version: 0.2.7

Depends: R (≥ 3.4.0)

Imports: [reticulate](#), [pkgdown](#), [crayon](#), [cli](#), utils

Suggests: [rmarkdown](#), [knitr](#)

Published: 2018-10-19

Author: Xu Jing [cre]

Maintainer: Xu Jing <274762204_at_gg.com>

Identify: Via CRAN & Clickin'


gcForest

Imports:

- **reticulate**
- pkgdown
- crayon
- cli
- utils

reticulate: Interface to 'Python'

Interface to 'Python' modules, classes, and functions. When calling into 'Python', R data types are automatically converted to their equivalent 'Python' types. When values are returned from 'Python' to R they are converted back to R types. Compatible with all versions of 'Python' ≥ 2.7 .

Version:	1.12
Depends:	R (≥ 3.0)
Imports:	utils, graphics, jsonlite , Rcpp ($\geq 0.12.7$), Matrix
LinkingTo:	Rcpp
Suggests:	testthat , knitr , callr , rmarkdown
Published:	2019-04-12
Author:	Kevin Ushey [aut, cre], JJ Allaire [aut], RStudio [cph, fnd], Yuan Tang  [aut, cph], Dirk Eddelbuettel [ctb, cph], Bryan Lewis [ctb, cph], Marcus Geelnard [ctb, cph] (TinyThread library, http://tinythreadpp.bitsnbites.eu/)
Maintainer:	Kevin Ushey <kevin at rstudio.com>

Identify: Via CRAN & Clickin'



gcForest

Imports:

- reticulate
- **pkgdown**
- crayon
- cli
- utils

pkgdown: Make Static HTML Documentation for a Package

Generate an attractive and useful website from a source package. 'pkgdown' converts your documentation, vignettes, 'README', and more to 'HTML' making it easy to share information about your package online.

Version: 1.3.0
Depends: R (≥ 3.1.0)
Imports: [callr](#) (≥ 2.0.2), [cli](#), [crayon](#), [desc](#), [digest](#), [evaluate](#), [fs](#) (≥ 1.2.0), [fanshi](#), [highlight](#), [httr](#), [magrittr](#), [MASS](#), [memoise](#), [openssl](#), [pkgload](#) (≥ 1.0.2), [purrr](#), [processx](#), [R6](#), [rematch2](#), [rlang](#) (≥ 0.2.0), [rmarkdown](#) (≥ 1.1.9007), [roxygen2](#), [rsconnect](#), [rstudioapi](#), [tibble](#), [tools](#), [whisker](#), [withr](#), [xml2](#) (≥ 1.1.1), [yaml](#)
Suggests: [covr](#), [devtools](#), [htmlwidgets](#), [jsonlite](#), [knitr](#), [leaflet](#), [testthat](#), [rticles](#)
Published: 2018-12-07
Author: Hadley Wickham  [aut, cre], Jay Hesselberth  [aut], RStudio [cph, fnd]

Identify: Via CRAN & Clickin'

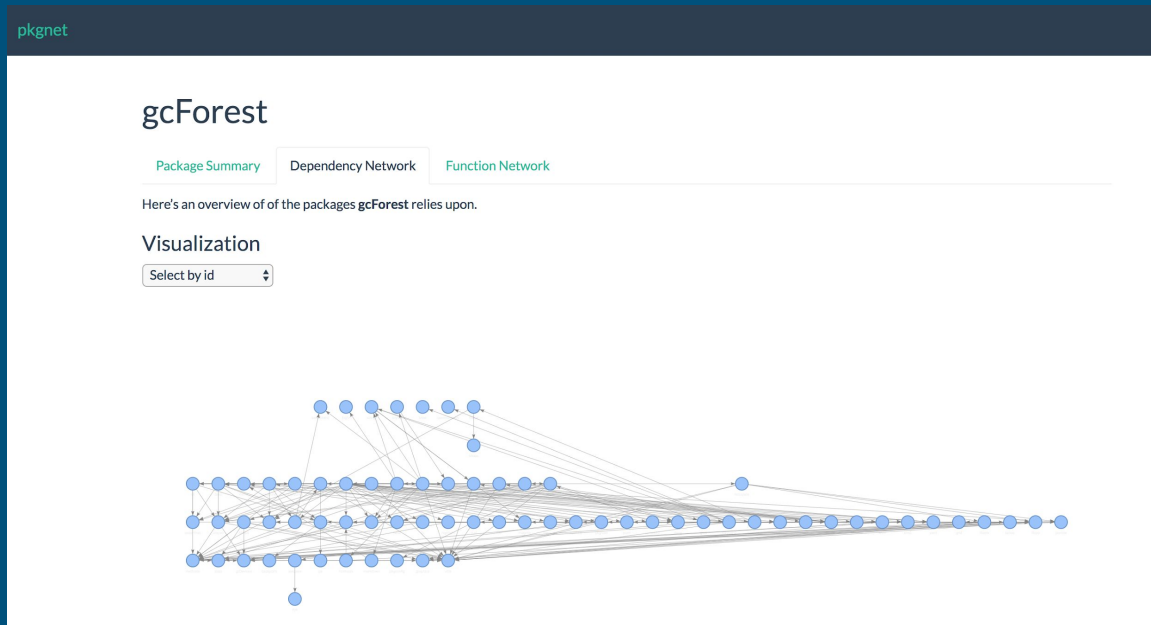
This is going to take forever!

Also, in the end, you will not have:

1. An efficient view of the dependency network as a whole
2. Interrelationships between the dependencies will not be immediately visible.

Identify: Via pkgnet

```
1 install.packages('pkgnet')
2 pkgnet::CreatePackageReport('gcForest')
```



Evaluate a Dependency

Some things to look for:

1. How **well supported** is this package?
 - Last update
 - Number of Developers
 - Github development activity
2. How **widely used** is this package?
 - Number of CRAN Downloads
 - Stack Overflow activity
3. How **widely approved** is this package?
 - Github stars
4. How **resource intensive** is this package?
 - File size
 - Installation Time

Evaluate a Dependency

it depends:
a **dialog** about **dependencies**

Jim Hester  Studio

 @jimhester  @jimhester_



An excellent resource for evaluating a dependency.

This package scrapes CRAN for stats on those mentioned measures and more.

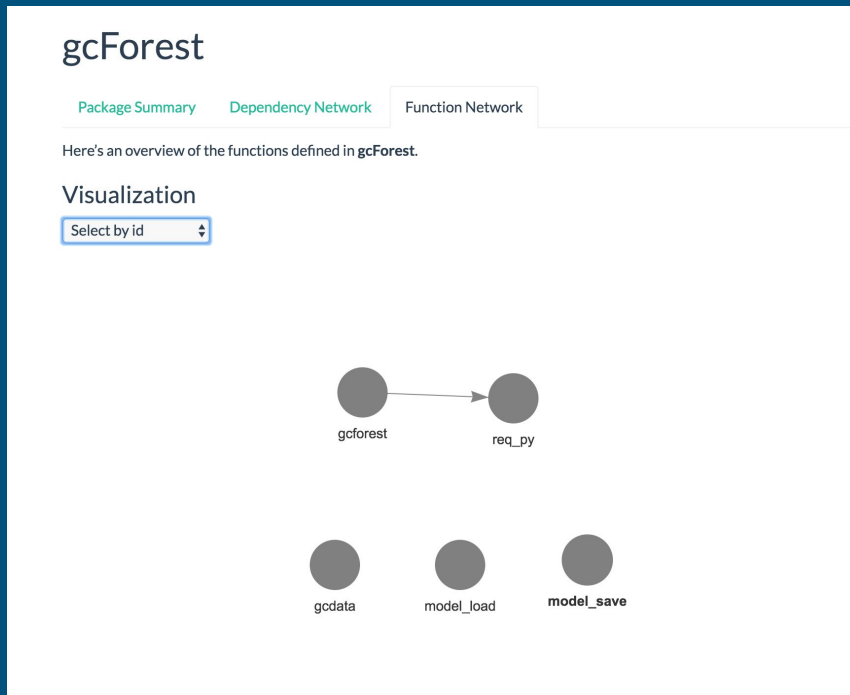
```
1 devtools::install_github("jimhester/itdepends")  
2 itdepends::dep_weight("<package name>")
```

Evaluate a Dependency

But also...

5. How **well designed** is this package?
 - Inspect the code
 - Explore function network w/ pkgnet
 - Read the docs

6. How **well tested** is this package?
 - Examine overall test coverage
 - Inspect test coverage for critical functions



Okay, I've weighed
enough things to get a
good picture of the risk.

Now what are my
options?

Ways to **Mitigate** a Risky Dependency

- **Avoid it** - if the function you plan to use is a small portion of the package, you may want to replicate just that function (use pkgnet function networks).
- **Substitute it** - at 17,000+ R packages now publically available, there's likely to be an alternative.
- **Abstract it** - wrap the code from the dependency in a thin wrapper function and reference the wrapper throughout your package. This enables use of the dependency for now but makes replacing it easier later.
- **Contribute to it** - if a package looks good but for a few things, suggest / make some edits. Open source software is a community. Contribute!

To Recap, with Recursive Dependencies:

1. **Identify** the extent of a package's recursive dependencies
 - Listed direct dependencies alone may be misleading
2. **Evaluate** the the risk of at least the first level of package dependencies
 - Utilize `pkgnet` & `itdepends` to evaluate structure, test coverage, usage, and a number of other metrics
3. **Mitigate** a risky dependency
 - Avoid, Substitute, Abstract or Contribute!

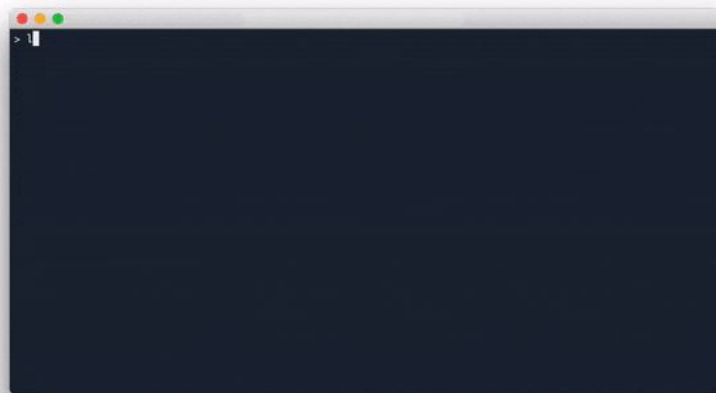
Interested in pkgnet?

Check out our website & gallery:

<https://uptakeopensource.github.io/pkgnet>

Issues, Ideas & Contributions at:

<https://github.com/uptakeopensource/pkgnet>



Thanks!

References:

- Brian Burns, James Lamb, Jay Qi and Patrick Boueri (2019). pkgnet: Get Network Representation of an R Package. R Package version 0.4.0. <https://CRAN.R-project.org/package=pkgnet>
- Cox, Russ. 2019. "Research!Rsc: Our Software Dependency Problem". Research.Swtch.Com. Accessed April 25 2019. <https://research.swtch.com/deps>.
- Hester. "itDepends: a dialog about dependencies". 2019 Rstudio Conference. <https://speakerdeck.com/jimhester/it-depends>
- "Package Basics · R Packages". 2019. R-Pkgs.Had.Co.Nz. Accessed April 25 2019. <http://r-pkgs.had.co.nz/description.html#dependencies>.
- Plakidas, Schall & Zdun. 2017. "Evolution Of The R Software Ecosystem: Metrics, Relationships, And Their Impact On Qualities". Journal Of Systems And Software 132: 119-146. Elsevier BV. doi:10.1016/j.jss.2017.06.095.
- "Rdocumentation". 2019. Rdocumentation.Org. Accessed April 25 2019. <https://www.rdocumentation.org/trends>.



Identify: Via pkgnet

Table

node	outDegree	inDegree	numRecursiveDeps	numRecursiveRevDeps	betweenness	pageRank
gcForest	5.000	0.000	70.000	0.000	0.000	0.007
pkgdown	30.000	1.000	65.000	1.000	55.400	0.008
roxygen2	13.000	1.000	33.000	2.000	8.544	0.007
pkgload	8.000	2.000	23.000	3.000	17.294	0.008
rmarkdown	12.000	1.000	23.000	2.000	16.544	0.007
pkgbuild	8.000	1.000	20.000	4.000	16.000	0.008
knitr	8.000	1.000	16.000	3.000	7.500	0.007
rematch2	1.000	1.000	15.000	2.000	0.000	0.007
rsconnect	6.000	1.000	15.000	2.000	6.333	0.007

Identify: Via pkgnet

