Package and collaboration networks in CRAN

Ioannis Kosmidis

Twitter IKosmidis_
Email ioannis.kosmidis@warwick.ac.uk
Website http://www.ikosmidis.com

Reader in Data Science
University of Warwick & The Alan Turing Institute

05 November 2018
R User Group Oxford
Oxford
Outline

1. CRAN today
2. Exploring CRAN
3. cranly networks
4. cranly summaries
5. Dependence trees
As of today 05 Nov 2018 CRAN has

- 18546 authors
- contributing in 12755 packages
Percentage of CRAN packages by some authors

```r
R> library("dplyr")
R> authors <- c("R Core", "Brian Ripley", "Achim Zeileis", "Hadley Wickham", + "Dirk Eddelbuettel", "Kurt Hornik", + "Yihui Xie", "Ioannis Kosmidis")
R> author_summaries %>% group_by(author) %>%
+ summarize(`%CRAN` = 100 * n_packages / nrow(package_summaries)) %>%
+ filter(author %in% authors) %>%
+ arrange(desc(`%CRAN`))

<table>
<thead>
<tr>
<th>author</th>
<th>%CRAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hadley Wickham</td>
<td>1.05</td>
</tr>
<tr>
<td>R Core</td>
<td>0.64</td>
</tr>
<tr>
<td>Dirk Eddelbuettel</td>
<td>0.59</td>
</tr>
<tr>
<td>Kurt Hornik</td>
<td>0.56</td>
</tr>
<tr>
<td>Achim Zeileis</td>
<td>0.42</td>
</tr>
<tr>
<td>Brian Ripley</td>
<td>0.32</td>
</tr>
<tr>
<td>Yihui Xie</td>
<td>0.32</td>
</tr>
<tr>
<td>Ioannis Kosmidis</td>
<td>0.06</td>
</tr>
</tbody>
</table>
```
In the course of a couple of decades CRAN became
- a rich and diverse software ecosystem
- a large database of authors, tools and knowledge which are naturally linked to each other
- hard to explore and keep track of
Outline

1. CRAN today
2. Exploring CRAN
3. cranly networks
4. cranly summaries
5. Dependence trees
cranberries by Dirk Eddelbuettel

🌐 http://dirk.eddelbuettel.com/cranberries/
🐦 CRANberriesFeed

Aggregates information about new, updated and removed packages from CRAN available and organises it in an RSS feed and a clean interface
cranly: Package and collaboration networks in CRAN

Exploring CRAN II

cranlogs R package by Gabor Csardi

👀 https://cran.r-project.org/package=cranlogs
👀 https://cranlogs.r-pkg.org

API for package downloads from RStudio CRAN mirror, and badges

R> cranlogs::cran_top_downloads("last-week")

<table>
<thead>
<tr>
<th>rank</th>
<th>package</th>
<th>count</th>
<th>from</th>
<th>to</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>rlang</td>
<td>212802</td>
<td>2018-10-29</td>
<td>2018-11-04</td>
</tr>
<tr>
<td>2</td>
<td>Rcpp</td>
<td>194573</td>
<td>2018-10-29</td>
<td>2018-11-04</td>
</tr>
<tr>
<td>3</td>
<td>ggplot2</td>
<td>187144</td>
<td>2018-10-29</td>
<td>2018-11-04</td>
</tr>
<tr>
<td>4</td>
<td>glue</td>
<td>150919</td>
<td>2018-10-29</td>
<td>2018-11-04</td>
</tr>
<tr>
<td>5</td>
<td>stringi</td>
<td>148461</td>
<td>2018-10-29</td>
<td>2018-11-04</td>
</tr>
<tr>
<td>6</td>
<td>dplyr</td>
<td>146945</td>
<td>2018-10-29</td>
<td>2018-11-04</td>
</tr>
<tr>
<td>7</td>
<td>digest</td>
<td>143414</td>
<td>2018-10-29</td>
<td>2018-11-04</td>
</tr>
<tr>
<td>8</td>
<td>WGCNA</td>
<td>140912</td>
<td>2018-10-29</td>
<td>2018-11-04</td>
</tr>
<tr>
<td>9</td>
<td>stringr</td>
<td>137661</td>
<td>2018-10-29</td>
<td>2018-11-04</td>
</tr>
<tr>
<td>10</td>
<td>fansi</td>
<td>133979</td>
<td>2018-10-29</td>
<td>2018-11-04</td>
</tr>
</tbody>
</table>

https://cranlogs.r-pkg.org/badges/tidyverse ⇒

 downloads 1M/month
**CRANsearcher** R package by Becca Krouse and Agustin Calatroni

https://CRAN.R-project.org/package=CRANsearcher

Shiny interface and RStudio plugin for searching packages by topic

R> CRANsearcher::CRANsearcher()
available.packages from *utils*

Returns a matrix of details corresponding to packages currently available at one or more repositories

```r
R> library("tibble")
R> available.packages() %>% as.tibble
```

# A tibble: 13,333 x 17

<table>
<thead>
<tr>
<th>Package</th>
<th>Version</th>
<th>Priority</th>
<th>Depends</th>
<th>Imports</th>
<th>LinkingTo</th>
<th>Suggests</th>
<th>Enhances</th>
</tr>
</thead>
<tbody>
<tr>
<td>A3</td>
<td>1.0.0</td>
<td>&lt;NA&gt;</td>
<td>R (&gt;= ~ &lt;NA&gt; &lt;NA&gt; randomF~ &lt;NA&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>abbyyR</td>
<td>0.5.4</td>
<td>&lt;NA&gt;</td>
<td>R (&gt;= ~ httr, ~ &lt;NA&gt; testtha~ &lt;NA&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>abc</td>
<td>2.1</td>
<td>&lt;NA&gt;</td>
<td>R (&gt;= ~ &lt;NA&gt; &lt;NA&gt; &lt;NA&gt; &lt;NA&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>abc.dev~</td>
<td>1.0</td>
<td>&lt;NA&gt;</td>
<td>R (&gt;= ~ &lt;NA&gt; &lt;NA&gt; &lt;NA&gt; &lt;NA&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABC.RAP</td>
<td>0.9.0</td>
<td>&lt;NA&gt;</td>
<td>R (&gt;= ~ graphi~ &lt;NA&gt; knitr, ~ &lt;NA&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCana~</td>
<td>1.2.1</td>
<td>&lt;NA&gt;</td>
<td>R (&gt;= ~ plotrix &lt;NA&gt; &lt;NA&gt; &lt;NA&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>abcdeF~</td>
<td>0.4</td>
<td>&lt;NA&gt;</td>
<td>Rglpk,~ &lt;NA&gt; &lt;NA&gt; LIM,syb~ &lt;NA&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCopt~</td>
<td>0.15.0</td>
<td>&lt;NA&gt;</td>
<td>&lt;NA&gt; Rcpp, ~ Rcpp testtha~ &lt;NA&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCp2</td>
<td>1.2</td>
<td>&lt;NA&gt;</td>
<td>MASS &lt;NA&gt; &lt;NA&gt; &lt;NA&gt; &lt;NA&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>abcrf</td>
<td>1.7.1</td>
<td>&lt;NA&gt;</td>
<td>R(&gt;= 3~ &quot;readr~ Rcpp, Rc~ &lt;NA&gt; &lt;NA&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

# ... with 13,323 more rows, and 9 more variables: License <chr>, License_is_FOSS <chr>, License_restricts_use <chr>, OS_type <chr>, Archs <chr>, MD5sum <chr>, NeedsCompilation <chr>, File <chr>, Repository <chr>
### CRAN_package_db from **tools**

Returns a character data frame with most DESCRIPTION metadata for the current packages in the CRAN package repository

```r
R> p_db <- tools::CRAN_package_db()
R> names(p_db)
```

```
[1] "Package"         "Version"         "Priority"         "Depends"
[2] "Imports"         "LinkingTo"        "Suggests"         "Enhances"
[3] "License"         "License_is_FOSS"  "License_restricts_use" "OS_type"
[4] "MD5sum"          "Authors@R"        "NeedsCompilation"  "Additional_repositories"
[5] "Archs"           "BuildManual"      "Biarch"            "BugReports"
[6] "Built"           "ByteCompile"      "BuildResaveData"   "BuildVignettes"
[7] "Classification/JEL" "Classification/MSC" "Classification/MSC-2010" "Classification/ACM"
[8] "Classification/MSC-2010" "Classification/ACM-2012" "Collate" "Copyright"
[10] "Collate.windows" "Contact"        "Contact"          "Contact"
[11] "Description"    "Encoding"        "Encoding"         "Encoding"
[12] "LazyData"        "LazyDataCompression" "LazyDataCompression" "LazyDataCompression"
[13] "Maintainer"      "Note"            "Note"             "Note"
[14] "SysDataCompression" "SystemRequirements"  "SystemRequirements" "SystemRequirements"
[15] "URL"             "VignetteBuilder"  "VignetteBuilder"   "VignetteBuilder"
[16] "Path"            "X-CRAN-Comment"   "X-CRAN-Comment"    "X-CRAN-Comment"
[17] "Reverse linking to" "Reverse depends"  "Reverse depends"  "Reverse depends"
[18] "Reverse linking to" "Reverse imports" "Reverse imports"  "Reverse imports"
[19] "Reverse linking to" "Reverse linking to" "Reverse linking to"  "Reverse linking to"
[20] "Reverse linking to" "Reverse linking to" "Reverse linking to"  "Reverse linking to"
```

```r
[34] "Collate/unix"
```
Why bother more and make **cranly**? I

**Altruistic reasons**

Set of tools for discovery of interconnections in CRAN

Data objects for modelling software networks

Interactive shiny app for package and collaborations discovery (not there yet but should be easy now)
Why bother more and make cranly? II

Less altruistic reasons

Wanted to do what is at
🌐 ikosmidis.com/software
and keep track of who/what/why links to my R packages

Wanted a tool that helps me find referees for my editorial work (seriously!)
I am a big fan of “literate programming” principles

clean_CRAN_db: “clean up” the CRAN database information
build_network: build networks out of it
build_dependence_tree: build package dependence trees
subset: subset cranly_network objects
summary: summarize cranly_network objects
plot: plot cranly_network objects or summaries of those
various extractor functions
Outline

1. CRAN today
2. Exploring CRAN
3. cranly networks
4. cranly summaries
5. Dependence trees
The DESCRIPTION file

R> (lubridate_desc <- packageDescription("lubridate"))

Package: lubridate
Type: Package
Version: 1.7.4
Title: Make Dealing with Dates a Little Easier
Description: Functions to work with date-times and time-spans: fast and user friendly parsing of date-time data, extraction and updating of components of a date-time (years, months, days, hours, minutes, and seconds), algebraic manipulation on date-time and time-span objects. The 'lubridate' package has a consistent and memorable syntax that makes working with dates easy and fun. Parts of the 'CCTZ' source code, released under the Apache 2.0 License, are included in this package. See <https://github.com/google/cctz> for more details.
Authors@R: c(person("Vitalie", "Spinu", email = "spinuvit@gmail.com", role = c("aut","cre")), person("Garrett", "Groleumund", role = "aut"), person("Hadley", "Wickham", role = "aut"), person("Ian", "Lyttle", role="ctb"), person("Immanuel", "Constigan", role = "ctb"), person("Jason", "Law", role="ctb"), person("Doug", "Mitarotonda", role="ctb"), person("Joseph", "Larmarange", role="ctb"), person("Jonathan", "Boiser", role="ctb"), person("Chel Hee", "Lee", role = "ctb")
Maintainer: Vitalie Spinu <spinuvit@gmail.com>
License: GPL (>= 2)
Depends: methods, R (>= 3.0.0)
Imports: stringr, Rcpp (>= 0.12.13),
LinkingTo: Rcpp,
Suggests: testthat, knitr, covr
Enhances: chron, fts, timeSeries, timeDate, tis, tseries, xts, zoo
SystemRequirements: A system with zoneinfo data (e.g. /usr/share/zoneinfo) as well as a recent-enough C++11 compiler (such as g++-4.8 or later). On Windows the zoneinfo included with R is used.
VignetteBuilder: knitr
LazyData: true
Collate: 'Dates.r' 'POSIXt.r' 'RcppExports.R' 'util.r' 'parse.r' 'timeSpans.r' 'intervals.r' 'diffTimes.r'
'durations.r' ..... 
RoxygenNote: 6.0.1
BugReports: https://github.com/tidyverse/lubridate/issues
NeedsCompilation: yes
Packaged: 2018-04-10 15:18:02 UTC; vspinu
Author: Vitalie Spinu [aut, cre], Garrett Groleumund [aut], Hadley Wickham [aut], Ian Lyttle [ctb], Imanuel Constigan [ctb], Jason Law [ctb], Doug Mitarotonda [ctb], Joseph Larmarange [ctb], Jonathan Boiser [ctb], Chel Hee Lee [ctb]
Repository: CRAN
Date/Publication: 2018-04-11 10:08:43 UTC
Package directives

R> lubridate_desc$Suggests

[1] "testthat, knitr, covr"

R> lubridate_desc$Imports

[1] "stringr, Rcpp (>= 0.12.13),"

R> lubridate_desc$Depends

[1] "methods, R (>= 3.0.0)"

R> lubridate_desc$Enhances

[1] "chron, fts, timeSeries, timeDate, tis, tseries, xts, zoo"

R> lubridate_desc$LinkingTo

[1] "Rcpp,"
Package directive networks

Package directives “link” packages to each other, in a sense a package citation network by mere definition of what “package dependence” is in R.

Package directives define a **package directives network**
cranly: Package and collaboration networks in CRAN

Package directive networks

```r
R> library("cranly")
R> library("igraph")
R> directives <- with(lubridate_desc, {
+   clean_up_directives(c(Suggests, Imports, Depends, Enhances, LinkingTo))
+ })
R> pdir_net <- data.frame(from = unique(unlist(directives)), to = "lubridate")
R> plot(graph_from_edgelist(as.matrix(pdir_net), directed = FALSE))
```
Author collaboration network

R> lubridate_desc$Author

[1] "Vitalie Spinu [aut, cre],
 Garrett Grolemund [aut],
 Hadley Wickham [aut],
 Ian Lyons [ctb],
 Ian Zhu [ctb],
 Artur Wilczynski [ctb],
 Guillaume Rousset [ctb],
 David Firth [ctb],
 Joe Sill [ctb],
 Hong Zhou [ctb],
 Nima Hejazi [ctb],
 Mine Cetinkaya-Rundel [ctb],
 Doug Mitarotonda [ctb],
 Joseph Larmarange [ctb],
 Jonathan Boiser [ctb],
 Chel Hee Lee [ctb],
 Sadrnezha Abdolvahab [ctb],
 Jie Sun [ctb],
 Ananda, H. [ctb],
 Victor Ng [ctb],
 Jeremy Yen [ctb],
 Edzer Pebesma [ctb],
 Bart Emsley [ctb],
 Eric Brunner [ctb],
 Joseph Lee [ctb],
 Marco Avendano [ctb],
 Joon Yun [ctb],
 Chris Wood [ctb],
 Moshe Franklin [ctb],
 Alex B. G M [ctb],
 Simon Wood [ctb],
 John Chennupati [ctb],
 Rong Lin [ctb],
 Mike Hudson [ctb],
 Peter J. Rousseeuw [ctb],
 Davorin Lesku [ctb],
 Paul Northrup [ctb],
 Kevin Blackwell [ctb],
 Peter Langfelder [ctb],
 Pleun Brons [ctb],
 Michael B. Gastner [ctb],
 Christophe Dardenne [ctb],
 David Weinberg [ctb],
 Andrew C. Lerch [ctb],
 Sabin Shibahara [ctb],
 Gregory R. Dworkin [ctb],
 Brian Zhang [ctb],
 Giedrius Markauskas [ctb],
 Jared Sengupta [ctb],
 Chris Wrenn [ctb],
 Robert M. Harb [ctb],
 Michael P. Kellam [ctb],
 Brian R. Neyman [ctb],
 Mitesh A. Trivedi [ctb],
 Mark Leisch [ctb],
 Patrick M. Burns [ctb],
 Theodore R. Jupin [ctb],
 Thibaut Parent [ctb],
 Vincent Patry [ctb],
 Yves Acton [ctb],
 Peter Flach [ctb],
 James H. Fisher [ctb],
 Pieter J. Heskes [ctb],
 Richard B. Sigmund [ctb],
 Glenn Hurlbert [ctb],
 Michael G. Anelli [ctb],
 John A. Peel [ctb],
 David M. Smith [ctb],
 Michael J. Picheny [ctb],
 Marcus D. Hutter [ctb],
 John F. McGee [ctb],
 Andrew D. Heiberger [ctb],
 Joseph T. Park [ctb],
 Ryan J. Drude [ctb],
 Justin J. Han [ctb],
 Koen Christiaensen [ctb],
 Eunyoung Shin [ctb],
 Mark Wainberg [ctb],
 Mark A. Field [ctb],
 John C. W. Bowers [ctb],
 Adam J. Brown [ctb],
 David F. Martin [ctb],
 David J. Miller [ctb],
 ...
CRAN_package_db returns a data frame with current information in most fields in the DESCRIPTION file (and a bit more) for all packages

```r
R> p_db$Author[grep("guidance", p_db$Author)][1] %>% strwrap(70)
```

[1] "Ravi Varadhan [aut, cph, trl], Paul Gilbert [aut, cre], Marcos Raydan"
[2] "[ctb] (with co-authors, wrote original algorithms in fortran. These"
[3] "provided some guidance for implementing R code in the BB package.),"
[4] "JM Martinez [ctb] (with co-authors, wrote original algorithms in"
[5] "fortran. These provided some guidance for implementing R code in the"
[6] "BB package.), EG Birgin [ctb] (with co-authors, wrote original"
[7] "algorithms in fortran. These provided some guidance for implementing"
[8] "R code in the BB package.), W LaCruz [ctb] (with co-authors, wrote"
[9] "original algorithms in fortran. These provided some guidance for"
[10] "implementing R code in the BB package.)"

```r
R> p_db$Author[grep("Queen", p_db$Author)][1] %>% strwrap(70)
```

[1] "Alex M Chubaty [aut, cre], Her Majesty the Queen in Right of Canada,"
[2] "as represented by the Minister of Natural Resources Canada [cph]"

For package and collaboration networks, need a reliable systematic way of extracting package and author names from the fields

cranly R package provides clean_up_directives and clean_up_author


```r
R> p_db$Author[grep("guidance", p_db$Author)][1]

[1] "Ravi Varadhan [aut, cph, trl],
Paul Gilbert [aut, cre],
Marcos Raydan [ctb] (with co-authors, wrote original algorithms in fortran. These provided some guidance for implementing R code in the BB package.)"

R> p_db$Author[grep("guidance", p_db$Author)][1] %>% clean_up_author

[[1]]
[1] "Ravi Varadhan" "Paul Gilbert" "Marcos Raydan" "JM Martinez" 
[5] "EG Birgin" "W LaCruz"

R> p_db$Author[grep("Queen", p_db$Author)][1]

[1] "Alex M Chubaty [aut, cre],
Her Majesty the Queen in Right of Canada, as represented by the Minister of Natural Resources Canada [cph]"

R> p_db$Author[grep("Queen", p_db$Author)][1] %>% clean_up_author

[[1]]
[1] "Alex M Chubaty"
```
`packageDescription("tidyverse")$Imports`

```
[1] "broom (>= 0.4.2), cli (>= 1.0.0), crayon (>= 1.3.4), dplyr (>=\n0.7.4), dbplyr (>= 1.1.0), forcats (>= 0.2.0), ... rstudioapi (>=
0.7), rvest (>= 0.3.2), stringr (>= 1.2.0), tibble (>= 1.3.4),
tidyr (>= 0.7.2), xml2 (>= 1.1.1)"
```

`packageDescription("tidyverse")$Imports %>% clean_up_directives`

```
[[1]]
 [1] "broom"  "cli"    "crayon"  "dplyr"  "dbplyr"
 [6] "forcats" "ggplot2" "haven"   "hms"    "httr"
[11] "jsonlite" "lubridate" "magrittr" "modelr"  "purrr"
[16] "readr"   "readxl"  "reprex"  "rlang"  "rstudioapi"
[21] "rvest"   "stringr" "tibble"  "tidyr"   "xml2"
```
CRAN\_package\_db() to cranly\_db objects

```r
R> library("cranly")
R> package_db <- clean_CRAN_db()
R> class(package_db)

[1] "cranly_db" "data.frame"
```

```r
R> lubridate_fields <- package_db %>% filter(package == "lubridate")
R> lubridate_fields$imports

[[1]]
[1] "stringr" "Rcpp"

R> lubridate_fields$suggests

[[1]]
[1] "testthat" "knitr" "covr"

R> lubridate_fields$author

[[1]]
[1] "Vitalie Spinu" "Garrett Grolemund" "Hadley Wickham"
[4] "Ian Lyttle" "Imanuel Constigan" "Jason Law"
[7] "Doug Mitarotonda" "Joseph Larmarange" "Jonathan Boiser"
[10] "Chel Hee Lee"
```
Organise the information in `cranly_db` objects in networks

There are two obvious network **perspectives** for this information:

<table>
<thead>
<tr>
<th>Perspective</th>
<th>Nodes</th>
<th>Edge formation</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>author</td>
<td>authors</td>
<td>same author field in</td>
<td>collaboration network</td>
</tr>
<tr>
<td>package</td>
<td>packages</td>
<td>directives fields</td>
<td>package directives network</td>
</tr>
</tbody>
</table>
cranly_network objects

R> package_net <- build_network(package_db, perspective = "package")
R> str(package_net, 1)

List of 2
  $ edges:'data.frame': 81845 obs. of 3 variables:
  $ nodes:'data.frame': 12755 obs. of 64 variables:
    - attr(*, "class")= chr [1:2] "cranly_network" "list"
    - attr(*, "timestamp")= POSIXct[1:1], format: "2018-11-05 13:16:38"
    - attr(*, "perspective")= chr "package"

R> author_net <- build_network(package_db, perspective = "author")
R> str(author_net, 1)

List of 2
  $ edges:'data.frame': 95490 obs. of 10 variables:
  $ nodes:'data.frame': 18546 obs. of 2 variables:
    - attr(*, "class")= chr [1:2] "cranly_network" "list"
    - attr(*, "timestamp")= POSIXct[1:1], format: "2018-11-05 13:16:38"
    - attr(*, "perspective")= chr "author"
cranly: Package and collaboration networks in CRAN

cranly_network objects

R> head(package_net$nodes[, 1:4])

<table>
<thead>
<tr>
<th></th>
<th>package</th>
<th>version</th>
<th>priority</th>
<th>depends</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A3</td>
<td>1.0.0</td>
<td>&lt;NA&gt;</td>
<td>xtable, pbapply</td>
</tr>
<tr>
<td>2</td>
<td>a4Base</td>
<td>&lt;NA&gt;</td>
<td>&lt;NA&gt;</td>
<td>NA</td>
</tr>
<tr>
<td>3</td>
<td>a4Core</td>
<td>&lt;NA&gt;</td>
<td>&lt;NA&gt;</td>
<td>NA</td>
</tr>
<tr>
<td>4</td>
<td>abbyyR</td>
<td>0.5.4</td>
<td>&lt;NA&gt;</td>
<td>abc.data, nnet, quantreg, MASS, locfit</td>
</tr>
<tr>
<td>5</td>
<td>abc</td>
<td>2.1</td>
<td>&lt;NA&gt;</td>
<td>abc.data, nnet, quantreg, MASS, locfit</td>
</tr>
<tr>
<td>6</td>
<td>abc.data</td>
<td>1.0</td>
<td>&lt;NA&gt;</td>
<td></td>
</tr>
</tbody>
</table>

R> head(package_net$edges)

<table>
<thead>
<tr>
<th></th>
<th>from</th>
<th>to</th>
<th>type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>httr</td>
<td>abbyyR</td>
<td>imports</td>
</tr>
<tr>
<td>2</td>
<td>XML</td>
<td>abbyyR</td>
<td>imports</td>
</tr>
<tr>
<td>3</td>
<td>curl</td>
<td>abbyyR</td>
<td>imports</td>
</tr>
<tr>
<td>4</td>
<td>readr</td>
<td>abbyyR</td>
<td>imports</td>
</tr>
<tr>
<td>5</td>
<td>plyr</td>
<td>abbyyR</td>
<td>imports</td>
</tr>
<tr>
<td>6</td>
<td>progress</td>
<td>abbyyR</td>
<td>imports</td>
</tr>
</tbody>
</table>
### cranly: Package and collaboration networks in CRAN

#### cranly_network objects

```r
R> head(author_net$nodes[, 1:2])

<table>
<thead>
<tr>
<th>author</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Scott Fortmann-Roe</td>
<td></td>
</tr>
<tr>
<td>Gaurav Sood</td>
<td></td>
</tr>
<tr>
<td>Csillery Katalin</td>
<td></td>
</tr>
<tr>
<td>Lemaire Louisiane</td>
<td></td>
</tr>
<tr>
<td>Francois Olivier</td>
<td></td>
</tr>
<tr>
<td>Blum Michael</td>
<td></td>
</tr>
</tbody>
</table>

#### Package objects

```r
R> head(author_net$edges[, 1:3])

<table>
<thead>
<tr>
<th>from</th>
<th>to</th>
<th>package</th>
</tr>
</thead>
<tbody>
<tr>
<td>Csillery Katalin</td>
<td>Lemaire Louisiane</td>
<td>abc</td>
</tr>
<tr>
<td>Csillery Katalin</td>
<td>Francois Olivier</td>
<td>abc, abc.data</td>
</tr>
<tr>
<td>Csillery Katalin</td>
<td>Blum Michael</td>
<td>abc</td>
</tr>
<tr>
<td>Lemaire Louisiane</td>
<td>Francois Olivier</td>
<td>abc</td>
</tr>
<tr>
<td>Lemaire Louisiane</td>
<td>Blum Michael</td>
<td>abc</td>
</tr>
<tr>
<td>Francois Olivier</td>
<td>Blum Michael</td>
<td>abc</td>
</tr>
</tbody>
</table>
```
Interrogating `cranly_network` objects

Intuitive, “literate programming”, extractors:

- `package_by(x, author = NULL, exact = FALSE)`
- `package_with(x, name = NULL, exact = FALSE)`
- `author_with(x, name = NULL, exact = FALSE)`
- `author_of(x, package = NULL, exact = FALSE)`
- `suggests(x, package = NULL, exact = FALSE)`
- `imports(x, package = NULL, exact = FALSE)`
- `depends(x, package = NULL, exact = FALSE)`
- `linking_to(x, package = NULL, exact = FALSE)`
- `enhances(x, package = NULL, exact = FALSE)`
Interrogating `cranly_network` objects II

R> package_net %>% author_of("trackeR")

[1] "Yang Liu"  "Brian C Battaile"
[3] "Andrea Rodriguez-Martinez" "Rafael Ayala"
[5] "Yacine Debbabi"  "Lara Selles Vidal"
[7] "Ioannis Kosmidis" "Hannah Frick"
[9] "Robin Hornak"

R> package_net %>% author_of("trackeR", exact = TRUE)

[1] "Ioannis Kosmidis" "Hannah Frick" "Robin Hornak"

R> author_net %>% author_of("brglm")

[1] "Ioannis Kosmidis"  "Kjell Konis"
[3] "Euloge Clovis Kenne Pagui" "Nicola Sartori"
[5] "Alessandra Salvan"
Interrogating `cranly_network` objects III

```r
R> package_net %>% package_by("Ioannis Kosmidis")
[1] "betareg"  "brglm"   "brglm2"   "cranly"
[5] "enrichwith" "PlackettLuce" "profileModel" "trackeR"

R> author_net %>% package_by("Ioannis Kosmidis")
[1] "betareg"  "brglm"   "brglm2"   "cranly"
[5] "enrichwith" "PlackettLuce" "profileModel" "trackeR"
```
R> package_net %>% suggests("trackeR", exact = TRUE)

[1] "testthat" "knitr" "rmarkdown" "covr"

R> package_net %>% depends("trackeR", exact = TRUE)

[1] "zoo"

R> package_net %>% imports("trackeR", exact = TRUE)

[1] "ggplot2" "ggiranges" "xml2" "RSQILite" "jsonlite"
[6] "raster" "scam" "foreach" "fda" "sp"
[11] "leaflet" "ggmap" "gridExtra" "gtable"

R> package_net %>% linking_to("trackeR", exact = TRUE)

NULL
Interrogating cranly_network objects V

All packages that have "bayes" in their name

R> package_net %>% package_with("bayes") %>% head(45)

<table>
<thead>
<tr>
<th></th>
<th>Package Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>acebayes</td>
</tr>
<tr>
<td>2</td>
<td>bayesammi</td>
</tr>
<tr>
<td>3</td>
<td>BayesBinMix</td>
</tr>
<tr>
<td>4</td>
<td>bayesboot</td>
</tr>
<tr>
<td>5</td>
<td>BayesComm</td>
</tr>
<tr>
<td>6</td>
<td>BayesCR</td>
</tr>
<tr>
<td>7</td>
<td>BayesDA</td>
</tr>
<tr>
<td>8</td>
<td>BAYESDEF</td>
</tr>
<tr>
<td>9</td>
<td>bayesdfa</td>
</tr>
<tr>
<td>10</td>
<td>BayesESS</td>
</tr>
<tr>
<td>11</td>
<td>BayesFM</td>
</tr>
<tr>
<td>12</td>
<td>bayesGDS</td>
</tr>
<tr>
<td>13</td>
<td>BayesGOF</td>
</tr>
<tr>
<td>14</td>
<td>Bayesianbetareg</td>
</tr>
<tr>
<td>15</td>
<td>BayesianNetwork</td>
</tr>
<tr>
<td>16</td>
<td>bayesImageS</td>
</tr>
<tr>
<td>17</td>
<td>bayesLife</td>
</tr>
<tr>
<td>18</td>
<td>bayesloglin</td>
</tr>
<tr>
<td>19</td>
<td>bayesLopod</td>
</tr>
<tr>
<td>20</td>
<td>BayesMallows</td>
</tr>
<tr>
<td>21</td>
<td>BayesMed</td>
</tr>
<tr>
<td>22</td>
<td>bayesmix</td>
</tr>
<tr>
<td>23</td>
<td>BayesNetBP</td>
</tr>
<tr>
<td>24</td>
<td>BayesianAnimalTracker</td>
</tr>
<tr>
<td>25</td>
<td>BayesianGLasso</td>
</tr>
<tr>
<td>26</td>
<td>BayesianTools</td>
</tr>
<tr>
<td>27</td>
<td>BayesLCA</td>
</tr>
<tr>
<td>28</td>
<td>bayeslm</td>
</tr>
<tr>
<td>29</td>
<td>bayeslongitudinal</td>
</tr>
<tr>
<td>30</td>
<td>bayesm</td>
</tr>
<tr>
<td>31</td>
<td>BayesMAMS</td>
</tr>
<tr>
<td>32</td>
<td>bayesmeta</td>
</tr>
<tr>
<td>33</td>
<td>BayesMixSurv</td>
</tr>
</tbody>
</table>
Interrogating cranly_network objects VI

All packages that have “tidy” in their name

R> package_net %>% package_with("tidy")

[1] "htmltidy"    "tidybayes"    "tidyboot"    "tidycensus"
[5] "tidygenomics" "tidygraph"    "tidyhydat"    "tidyimpute"
[9] "tidyLPA"     "tidymodels"   "tidyposterior" "tidypredict"
[13] "tidyquant"   "tidyr"        "tidyRSS"     "tidyselect"
[17] "tidystats"   "tidystringdist" "tidytext"    "tidytidbits"
[21] "tidytransit" "tidytree"     "tidyverse"   "tidyxl"
Visualising cranly network objects

Static network visualisations are not very useful for exploring cranly networks.

Need dynamic and interactive network visualisations, with ability to display extra information about nodes and edges on demand.

- **node specific**
  - package networks: directive summaries, maintainer, downloads
  - collaboration networks: number of collaborators, packages

- **edge specific**
  - package networks: how are the nodes linked?
  - collaboration networks: what do the nodes collaborate on?

**visNetwork** package: an R interface to vis.js JavaScript library
Visualising `cranly_network` objects

```r
R> ## Directives network for `brglm` package
R> plot(package_net, package = "brglm")
R>
R> ## Directives network for packages by me
R> plot(package_net, author = "Ioannis Kosmidis")
R>
R> ## Collaboration network of all authors with Ioannis in their name
R> plot(author_net, author = "Ioannis", exact = FALSE)
R>
R> ## Collaboration network of all authors that have written a package with "glm" in its name
R> plot(author_net, package = "glm", exact = FALSE)
R>
R> ## Directives network of all packages with "glm" in their names
R> plot(package_net, package = "glm", exact = FALSE, title = FALSE, legend = FALSE, +
    width = 1000, height = 800)
```

The magic happens in `subset.cranly_network`
Outline

1. CRAN today
2. Exploring CRAN
3. cranly networks
4. cranly summaries
5. Dependence trees
cranly: Package and collaboration networks in CRAN

Talking to `igraph`

`cranly` has the ability to coerce `cranly_network` objects to `igraph` ones.

```
R> as.igraph(package_net)

IGRAPH 3ad7314 DN-- 12755 81845 --
+ attr: name (v/c), version (v/c), author (v/x), date (v/c), url (v/c), license (v/c), maintainer (v/c), type (e/c)
+ edges from 3ad7314 (vertex names):
[9] utils ->ABC.RAP plotrix ->ABCoptim Rcpp ->ABCoptim graphics ->ABCoptim
[13] stats ->ABCoptim utils ->ABCoptim readr ->abcrf MASS ->abcrf
[17] matrixStats->abcrf ranger ->abcrf parallel ->abcrf stringr ->abcrf
[21] Rcpp ->abcrf ggplot2 ->ABHgenotypeR reshape2 ->ABHgenotypeR utils ->ABHgenotypeR
+ ... omitted several edges
```

```
R> as.igraph(author_net)

IGRAPH b0b10ae UN-- 18546 95490 --
+ attr: name (v/c), package (v/x), package (e/c), imports (e/x), suggests (e/x), enhances (e/x),
| depends (e/x), linkingto (e/x), version (e/c), maintainer (e/c)
+ edges from b0b10ae (vertex names):
[1] Csillery Katalin --Lemaire Louisiane Csillery Katalin --Francois Olivier
[3] Csillery Katalin --Blum Michael Lemaire Louisiane --Francois Olivier
[5] Lemaire Louisiane --Blum Michael Francois Olivier --Blum Michael
[7] Csillery Katalin --Lemaire Louisiane Csillery Katalin --Francois Olivier
[9] Csillery Katalin --Blum Michael Lemaire Louisiane --Francois Olivier
[11] Lemaire Louisiane --Blum Michael Francois Olivier --Blum Michael
+ ... omitted several edges
```

This gives access to all advanced network summaries provided by igraph, including betweenness, closeness, page rank, degree, eigen centrality.
cranly: Package and collaboration networks in CRAN

Summarising cranly_network objects

R> package_summaries <- summary(package_net)
R> package_summaries %>% as.tibble

# A tibble: 12,755 x 17
  package n_authors n_imports n_imported_by n_suggests n_suggested_by n_depends n_depended_by n_enhances
  <chr>    <int>    <dbl>       <dbl>     <dbl>       <dbl>     <dbl>       <dbl>     <dbl>
1 A3        1       0           0           2           0         2           0           0
2 a4Base    1       0           0           0           1         0           0           0
3 a4Core    1       0           1           0           0         0           1           0
4 abbyyR    1       6           0           4           0         0           0           0
5 abc       4       0           1           0           1         5           2           0
6 abc.da~   4       0           0           0           1         0           1           0
7 ABC.RAP   4       3           0           2           0         0           0           0
8 ABCana~   3       1           0           3           0         0           0           0
9 abcdeF~   2       0           0           2           0         4           0           0
10 ABCopt~  2       4           0           2           0         0           0           0
# ... with 12,745 more rows, and 8 more variables: n_enhanced_by <dbl>, n_linking_to <dbl>, n_linked_by <dbl>, betweenness <dbl>, closeness <dbl>, page_rank <dbl>, degree <dbl>, eigen_centrality <dbl>
Summarising `cranly_network` objects II

```r
R> author_summaries <- summary(author_net)
R> author_summaries %>% as.tibble

# A tibble: 18,546 x 8
author     n_packages n_collaborators betweenness closeness page_rank  degree eigen_centrality
  *       <int>           <dbl>       <dbl>      <dbl>     <dbl>   <dbl>          <dbl>
1 Scott Fortmann-Roe  2             1          0    0.00000000551 0.0000170  1 0.00000000567
2 Gaurav Sood        8             4       8949.  0.00000000551 0.0000355  4 0.00000537
3 Csillery Katalin   2             6          0    0.00000000291 0.0000603  6 0
4 Lemaire Louisiane  2             6          0    0.00000000291 0.0000603  6 0
5 Francois Olivier   2             6          0    0.00000000291 0.0000603  6 0
6 Blum Michael       2             6          0    0.00000000291 0.0000603  6 0
7 Abdulmonem Alsaleh 1             3          0    0.00000000551 0.0000264  3 0.0000149
8 Robert Weeks       1             3          0    0.00000000551 0.0000264  3 0.0000149
9 Ian Morison        1             3          0    0.00000000551 0.0000264  3 0.0000149
10 RStudio           122           539        1669125. 0.00000000551 0.00152 539 0.00291
# ... with 18,536 more rows
```
Plotting package summaries = Vanity I

R> plot(author_summaries, according_to = "n_packages")

cranly top-20 according to n_packages
Package database as of 2018-11-05 13:16:38
Plotting package summaries = Vanity II

R> plot(author_summaries, according_to = "n_collaborators")

cranly top–20 according to n_collaborators
Package database as of 2018–11–05 13:16:38

- Hadley Wickham
- Brian Ripley
- RStudio
- R Core
- Martin Maechler
- Yihui Xie
- Ben Bolker
- JJ Allaire
- Michael Friendly
- Achim Zeileis
- Adrian Baddeley
- Dirk Eddelbuettel
- Rolf Turner
- Kurt Hornik
- Romain Francois
- Duncan Murdoch
- Torsten Hothorn
- Joe Cheng
- Bill Venables
- Kevin Ushey

n_collaborators
Plotting package summaries = Vanity III

R> plot(author_summaries, according_to = "page_rank", top = 20)
Outline

1. CRAN today
2. Exploring CRAN
3. cranly networks
4. cranly summaries
5. Dependence trees
A package’s dependence tree shows what else needs to be installed with the package in an empty package library with the package.

The tree can be constructed neatly using a recursion (see `compute_dependence_tree`), leveraging the advantages of functional programming.
`tibble` dependence tree

```r
R> tibble_tree <- build_dependence_tree(package_net, "tibble")
R> plot(tibble_tree, title = FALSE, legend = FALSE)
```
The package dependence index is a rough measure of how much “baggage” an R package carries.

Average across the generation index of the packages in the tree, with weights that are inversely proportional to the popularity of each package in terms of how many other packages depend on, link to or import it.

\[
\frac{\sum_{i \in C_p; i \neq p} \frac{1}{N_i} g_i}{\sum_{i \in C_p; i \neq p} \frac{1}{N_i}}
\]

- $C_p$ is the dependence tree for the package(s) $p$.
- $N_i$ is the total number of packages that depend, link or import package $i$.
- $g_i$ is the generation that package $i$ appears in the dependence tree of package(s) $p$. 
cranly: Package and collaboration networks in CRAN

### Package dependence index II

```r
R> cranly_tree <- build_dependence_tree(package_net, "cranly")
R> cranly_dep_index <- sapply(cranly_tree$nodes$package, function(package) {
  +   tree <- build_dependence_tree(package_net, package = package)
  +   s <- summary(tree)
  +   s$dependence_index
  + })
R> sort(cranly_dep_index)
```

<table>
<thead>
<tr>
<th>Package</th>
<th>Dependence Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>assertthat</td>
<td>0.000000000</td>
</tr>
<tr>
<td>cli</td>
<td>0.000000000</td>
</tr>
<tr>
<td>colorspace</td>
<td>0.000000000</td>
</tr>
<tr>
<td>countryCode</td>
<td>0.000000000</td>
</tr>
<tr>
<td>crayon</td>
<td>0.000000000</td>
</tr>
<tr>
<td>digest</td>
<td>0.000000000</td>
</tr>
<tr>
<td>fansi</td>
<td>0.000000000</td>
</tr>
<tr>
<td>glue</td>
<td>0.000000000</td>
</tr>
<tr>
<td>gtable</td>
<td>0.000000000</td>
</tr>
<tr>
<td>htmltools</td>
<td>0.000000000</td>
</tr>
<tr>
<td>jsonlite</td>
<td>0.000000000</td>
</tr>
<tr>
<td>labeling</td>
<td>0.000000000</td>
</tr>
<tr>
<td>lattice</td>
<td>0.000000000</td>
</tr>
<tr>
<td>lazyeval</td>
<td>0.000000000</td>
</tr>
<tr>
<td>magrittr</td>
<td>0.000000000</td>
</tr>
<tr>
<td>MASS</td>
<td>0.000000000</td>
</tr>
<tr>
<td>Matrix</td>
<td>0.000000000</td>
</tr>
<tr>
<td>munsell</td>
<td>0.000000000</td>
</tr>
<tr>
<td>nlme</td>
<td>0.000000000</td>
</tr>
<tr>
<td>pkgconfig</td>
<td>0.000000000</td>
</tr>
<tr>
<td>plyr</td>
<td>0.000000000</td>
</tr>
<tr>
<td>R6</td>
<td>0.000000000</td>
</tr>
<tr>
<td>RColorBrewer</td>
<td>0.000000000</td>
</tr>
<tr>
<td>Rcpp</td>
<td>0.000000000</td>
</tr>
<tr>
<td>rlang</td>
<td>0.000000000</td>
</tr>
<tr>
<td>stringi</td>
<td>0.000000000</td>
</tr>
<tr>
<td>stringr</td>
<td>0.000000000</td>
</tr>
<tr>
<td>utf8</td>
<td>0.000000000</td>
</tr>
<tr>
<td>viridisLite</td>
<td>0.000000000</td>
</tr>
<tr>
<td>withr</td>
<td>0.000000000</td>
</tr>
<tr>
<td>yaml</td>
<td>0.000000000</td>
</tr>
<tr>
<td>pillar</td>
<td>0.000000000</td>
</tr>
<tr>
<td>igraph</td>
<td>0.000000000</td>
</tr>
<tr>
<td>scales</td>
<td>0.000000000</td>
</tr>
<tr>
<td>htmlwidgets</td>
<td>0.000000000</td>
</tr>
<tr>
<td>mgcv</td>
<td>0.0004696439</td>
</tr>
<tr>
<td>visNetwork</td>
<td>0.012230725</td>
</tr>
<tr>
<td>tibble</td>
<td>0.018003822</td>
</tr>
<tr>
<td>reshape2</td>
<td>0.193331573</td>
</tr>
<tr>
<td>ggplot2</td>
<td>0.281780864</td>
</tr>
<tr>
<td>cranly</td>
<td>0.502184168</td>
</tr>
<tr>
<td>2.221914406</td>
<td></td>
</tr>
</tbody>
</table>
```
What’s next? I

Licence compatibility checks (cudos to Mark Hornick for table)

<table>
<thead>
<tr>
<th></th>
<th>MIT</th>
<th>BSD_3_clause</th>
<th>Apache 2.0</th>
<th>LGPL-2.1</th>
<th>LGPL\n(&gt;= 2.1)</th>
<th>LGPL-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIT</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>BSD_3_clause</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Apache 2.0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>LGPL-2.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LGPL\n(&gt;= 2.1)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>LGPL-3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>LGPL (&gt;= 3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>GPL-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>GPL (&gt;= 2)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GPL-3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GPL (&gt;=\n3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Affero GPL-3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>LGPL (&gt;= 3)</th>
<th>GPL-2</th>
<th>GPL (&gt;= 2)</th>
<th>GPL-3</th>
<th>GPL (&gt;=\n3)</th>
<th>Affero</th>
<th>GPL-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIT</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>BSD_3_clause</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Apache 2.0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>LGPL-2.1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>LGPL\n(&gt;= 2.1)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>LGPL-3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>LGPL (&gt;= 3)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GPL-2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GPL (&gt;= 2)</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GPL-3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GPL (&gt;=\n3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Affero GPL-3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
What’s next? II

Extractors and topic search in package descriptions

Shiny app

RStudio plugin (anyone?)
Thank you!