

From summarizing projects to setting tags, uses of parsing R files

`{funspotr}` : R Function Spotter



Bryan Shalloway

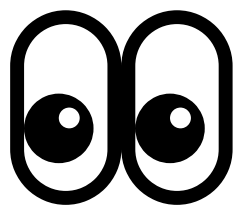
Data Science @NetApp

<https://github.com/brshallo/funspotr>

 @brshallo

 Bryan Shalloway

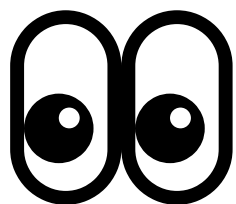
 bryanshalloway.com



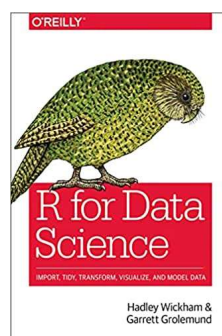
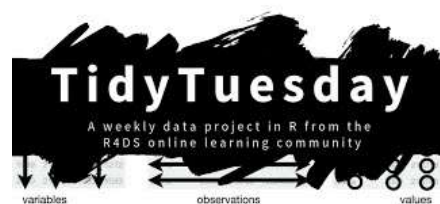
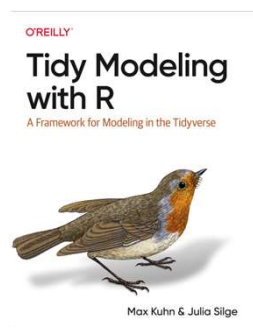
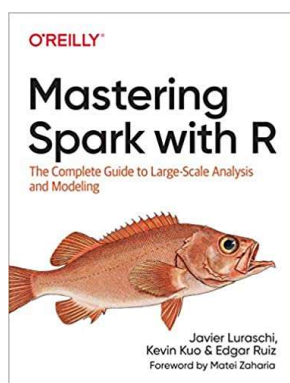
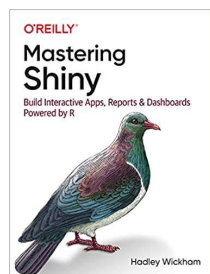
rstudio::conf
WASHINGTON D.C. 2022

<https://github.com/brshallo/funspotr>





rstudio::conf
WASHINGTON D.C. 2022

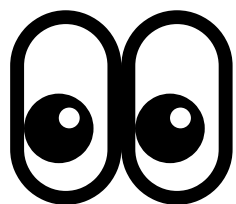


<https://github.com/brshallo/funspotr>

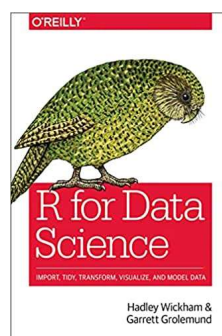
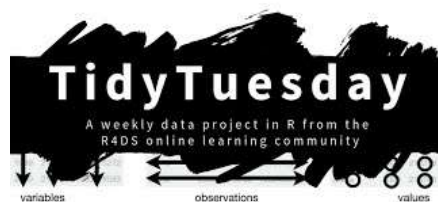
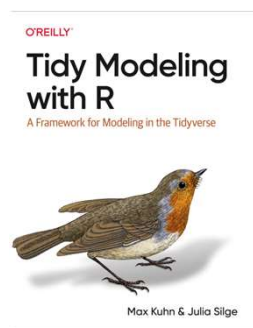
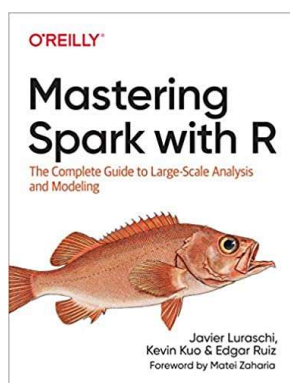
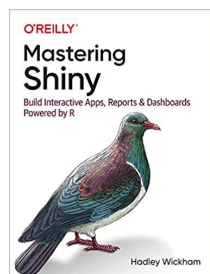
  @brshallo

 Bryan Shalloway

 bryanshalloway.com



rstudio::conf
WASHINGTON D.C. 2022



<https://github.com/brshallo/funspotr>

  @brshallo

 Bryan Shalloway

 bryanshalloway.com



Functions and notes:

- `geom_point` : Add points to plot, key args: `x` , `y` ,
- `geom_smooth` : Add line and confidence intervals to
can use `method` to change algorithm to make line.
- `geom_bar` : Stack values on top of each to make bar
`"identity"` May want to make `y = prop` to s



Alex Cookson
@alexcookson

...



Jeff Rothschild
@EatsleepfitJeff

...

Anyone other [#rstats](#) people find [@drob](#)'s [#TidyTuesday](#) screencasts useful?

I made a spreadsheet with timestamps for hundreds of specific tasks he does: bit.ly/2NnMeXH

Useful if, like me, you keep going back and ask, "Where in the video did he do [this thing] again?"

To try and learn [#tidymodels](#) I made a list of examples of using each engine, etc, by [@juliasilge](#), [@topepos](#), and [@Emil_Hvitfeldt](#). It's far from exhaustive, and in progress, but has helped a ton. [#rstats](#)

{funspotr} : R Function Spotter

<https://github.com/brshallo/funspotr>



@brshallo



Bryan Shalloway



bryanshalloway.com


```

library(funspotr)

list_files_github_repo("juliasilge/juliasilge.com") |>
  spot_funs_files() |>
  unnest_results()
#> # A tibble: 3,531 x 4
#>   funs          pkgs          relative_paths          absolute_paths
#>   <chr>         <chr>         <chr>              <chr>
#> 1 library      base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 2 options      base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 3 austen_books janeaustenr    content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 4 unnest_tokens tidytext     content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 5 count        dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 6 ungroup      dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 7 group_by     dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 8 summarize    dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 9 sum          base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 10 left_join    dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> # ... with 3,521 more rows

```

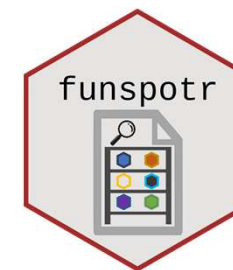
```
library(funspotr)

list_files_github_repo("juliasilge/juliasilge.com") |>
  spot_funs_files() |>
  unnest_results() |>
  DT::datatable()
```

Show 10 entries

Search: rand_forest

funcs	pkgs	contents	
All	All	All	All
rand_forest	parsonip	content/blog/2020-02-05-intro-tidymodels.Rmarkdown	https://raw.githubusercontent.com/juliasilge/juliasilge.com/main/content/blog/2020-02-05-intro-tidymodels.Rmarkdown
rand_forest	parsonip	content/blog/2020-02-18-food-hyperparameter-tune.Rmarkdown	https://raw.githubusercontent.com/juliasilge/juliasilge.com/main/content/blog/2020-02-18-food-hyperparameter-tune.Rmarkdown
rand_forest	parsonip	content/blog/datasaurus-multiclass/index.Rmarkdown	https://raw.githubusercontent.com/juliasilge/juliasilge.com/main/content/blog/datasaurus-multiclass/index.Rmarkdown



<https://github.com/brshallo/funspotr>

 @brshallo

 Bryan Shalloway

 bryanshalloway.com


```
library(funspotr)

list_files_github_repo("juliasilge/juliasilge.com") |>
  spot_funs_files() |>
  unnest_results()
```

`list_files_*()`

`spot_*()`

`{helpers}()`

```
library(funspotr)

list_files_github_repo("juliasilge/juliasilge.com") |>
  spot_funs_files() |>
  unnest_results()
```

```
list_files_*()
```

```
spot_*()
```

```
{helpers}()
```

```
library(funspotr)

list_files_github_repo("juliasilge/juliasilge.com") |>
  spot_funs_files() |>
  unnest_results()
```

`list_files_*` ()

`spot_*` ()

`{helpers}` ()

```
library(funspotr)

list_files_github_repo("juliasilge/juliasilge.com") |>
  spot_funs_files() |>
  unnest_results()
```

`list_files_*()`

`spot_*()`

`{helpers}()`

```

library(funspotr)

list_files_github_repo("juliasilge/juliasilge.com")

#> # A tibble: 103 x 2
#>   relative_paths absolute_paths
#>   <chr>         <chr>
#> 1 content/blog/2016-01-19-Water-World.Rmd https://raw.g~
#> 2 content/blog/2016-06-27-Term-Frequency-tf-idf.Rmd https://raw.g~
#> 3 content/blog/2016-08-05-Something-Strange.Rmd https://raw.g~
#> 4 content/blog/2016-09-30-Non-Academic-Careers.Rmd https://raw.g~
#> 5 content/blog/2016-10-28-Tidy-Text-Mining.Rmd https://raw.g~
#> 6 content/blog/2016-11-25-Gobbledygook.Rmd https://raw.g~
#> 7 content/blog/2017-04-15-Gender-Pronouns.Rmd https://raw.g~
#> 8 content/blog/2017-06-18-tidytext-0-1-3.Rmd https://raw.g~
#> 9 content/blog/2017-06-24-using-tidycensus-to-fetch-population-~ https://raw.g~
#> 10 content/blog/2017-07-26-navigating-r-package-universe.Rmd https://raw.g~
#> # ... with 93 more rows

```

```

library(funspotr)

list_files_github_repo("juliasilge/juliasilge.com") |>
  spot_funs_files()

#> # A tibble: 103 x 3
#>   relative_paths absolute_paths spotted
#>   <chr>          <chr>          <list>
#> 1 content/blog/2016-01-19-Water-World.Rmd https://raw.g~ <named list>
#> 2 content/blog/2016-06-27-Term-Frequency-tf-idf.Rmd https://raw.g~ <named list>
#> 3 content/blog/2016-08-05-Something-Strange.Rmd https://raw.g~ <named list>
#> 4 content/blog/2016-09-30-Non-Academic-Careers.Rmd https://raw.g~ <named list>
#> 5 content/blog/2016-10-28-Tidy-Text-Mining.Rmd https://raw.g~ <named list>
#> 6 content/blog/2016-11-25-Gobbledygook.Rmd https://raw.g~ <named list>
#> 7 content/blog/2017-04-15-Gender-Pronouns.Rmd https://raw.g~ <named list>
#> 8 content/blog/2017-06-18-tidytext-0-1-3.Rmd https://raw.g~ <named list>
#> 9 content/blog/2017-06-24-using-tidycensus-to-fetc~ https://raw.g~ <named list>
#> 10 content/blog/2017-07-26-navigating-r-package-uni~ https://raw.g~ <named list>
#> # ... with 93 more rows

```



```
library(funspotr)

list_files_github_repo("juliasilge/juliasilge.com") |>
  spot_funs_files()

#> # A tibble: 103 x 3
#>   relative_paths absolute_paths spotted
#>   <chr>          <chr>          <list>
#> 1 content/blog/2016-01-19-Water-World.Rmd https://raw.g~ <named list>
#> 2 content/blog/2016-06-27-Term-Frequency-tf-idf.Rmd https://raw.g~ <named list>
#> 3 content/blog/2016-08-05-Something-Strange.Rmd https://raw.g~ <named list>
#> 4 content/blog/2016-09-30-Non-Academic-Careers.Rmd https://raw.g~ <named list>
#> 5 content/blog/2016-10-28-Tidy-Text-Mining.Rmd https://raw.g~ <named list>
#> 6 content/blog/2016-11-25-Gobbledygook.Rmd https://raw.g~ <named list>
#> 7 content/blog/2017-04-15-Gender-Pronouns.Rmd https://raw.g~ <named list>
#> 8 content/blog/2017-06-18-tidytext-0-1-3.Rmd https://raw.g~ <named list>
#> 9 content/blog/2017-06-24-using-tidycensus-to-fetc~ https://raw.g~ <named list>
#> 10 content/blog/2017-07-26-navigating-r-package-uni~ https://raw.g~ <named list>
#> # ... with 93 more rows
```

```

library(funspotr)

list_files_github_repo("juliaspot_funs_files()

#> # A tibble: 103 x 3
#>   relative_paths
#>   <chr>
#> 1 content/blog/2016-01-19
#> 2 content/blog/2016-06-27
#> 3 content/blog/2016-08-05
#> 4 content/blog/2016-09-30
#> 5 content/blog/2016-10-28
#> 6 content/blog/2016-11-25
#> 7 content/blog/2017-04-15
#> 8 content/blog/2017-06-18
#> 9 content/blog/2017-06-24
#> 10 content/blog/2017-07-26
#> # ... with 93 more rows

... |>
  with(spotted[[7]])
#> $result
#> # A tibble: 34 x 2
#>   funs          pkgs
#>   <chr>        <chr>
#> 1 library      base
#> 2 options      base
#> 3 theme_set    ggplot
#> 4 theme_roboto (unknown)
#> 5 austen_books janeaustenr
#> 6 unnest_tokens tidytext
#> 7 c            base
#> 8 count        dplyr
#> 9 separate     tidyr
#> 10 filter       dplyr
#> # ... with 24 more rows
#>
#> $error
#> NULL

```

	paths	spotted
		<list>
raw.g~		<named list>
raw.g~		<named list>
raw.g~		<named list>
raw.g~		<named list>
raw.g~		<named list>
raw.g~		<named list>
raw.g~		<named list>
raw.g~		<named list>
raw.g~		<named list>
raw.g~		<named list>

```

library(funspotr)

list_files_github_repo("juliasilge/juliasilge.com") |>
  spot_funs_files() |>
  unnest_results()
#> # A tibble: 3,531 x 4
#>   funs          pkgs          relative_paths          absolute_paths
#>   <chr>         <chr>         <chr>              <chr>
#> 1 library      base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 2 options      base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 3 austen_books janeaustenr  content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 4 unnest_tokens tidytext    content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 5 count        dplyr        content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 6 ungroup      dplyr        content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 7 group_by     dplyr        content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 8 summarize    dplyr        content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 9 sum          base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 10 left_join    dplyr        content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> # ... with 3,521 more rows

```

```

library(funspotr)

list_files_github_repo("juliasilge/juliasilge.com") |>
  spot_funs_files() |>
  unnest_results()
#> # A tibble: 3,531 x 4
#>   funs          pkgs          relative_paths          absolute_paths
#>   <chr>         <chr>         <chr>              <chr>
#> 1 library      base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 2 options      base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 3 austen_books janeaustenr    content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 4 unnest_tokens tidytext     content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 5 count        dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 6 ungroup      dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 7 group_by     dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 8 summarize    dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 9 sum          base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 10 left_join    dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> # ... with 3,521 more rows

```

spot_funs()
...because R search spaces...

rstudio::conf
WASHINGTON D.C. 2022

```
library(pkgs2)  
require(pkg3)  
pkg::fun()
```

{callr}



{import}

 : $f(x)$

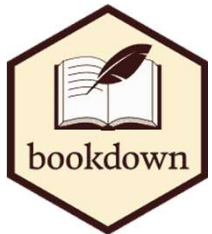
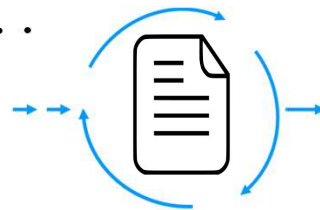
apply(utils::find)

$f(x)$

utils::getParseData()

{formatr},

...



```

library(funspotr)

list_files_github_repo("juliasilge/juliasilge.com") |>
  spot_funs_files() |>
  unnest_results()
#> # A tibble: 3,531 x 4
#>   funs          pkgs          relative_paths          absolute_paths
#>   <chr>         <chr>         <chr>              <chr>
#> 1 library      base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 2 options      base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 3 austen_books janeaustenr  content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 4 unnest_tokens tidytext    content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 5 count        dplyr        content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 6 ungroup      dplyr        content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 7 group_by     dplyr        content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 8 summarize    dplyr        content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 9 sum          base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 10 left_join    dplyr        content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> # ... with 3,521 more rows

```




Chelsea Parlett-Pelleriti
@ChelseaParlett

...

rstudio::conf
WASHINGTON D.C. 2022

I have so many snippets of `#rstats` code in `Untitled*.R` files, that have useful bits of code that I want to catalog somewhere, but I don't know where would be best, and so instead I have many open R files at all times.



<https://github.com/brshallo/funspotr>



@brshallo



Bryan Shalloway



bryanshalloway.com



Chelsea Parlett-Pelleriti
@ChelseaParlett

...

rstudio::conf
WASHINGTON D.C. 2022

I have so many snippets of `#rstats` code in `Untitled*.R` files, that have useful bits of code that I want to catalog somewhere, but I don't know where would be best, and so instead I have many open R files at all times.



GitHub Gist

```
list_files_github_gists("brshallo") |>  
spot_funs_files() |>  
unnest_results()
```

```
funspotr:::clean()
```

<https://github.com/brshallo/funspotr>



@brshallo



Bryan Shalloway



bryanshalloway.com

```
---  
title: Identifying R Functions & Packages...  
author: Bryan Shalloway  
date: '2022-01-18'  
tags:  
  - "ML"  
  - "Machine Learning"  
  - "R"  
  - "Rstats"  
...
```



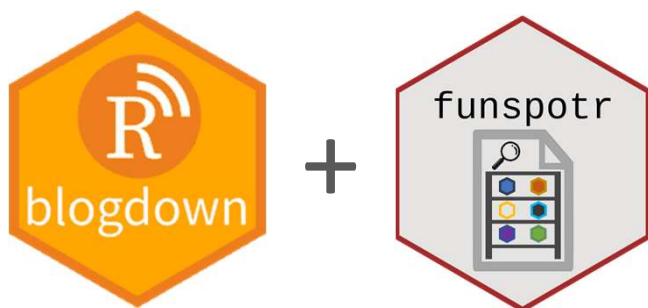
<https://github.com/brshallo/funspotr>

  @brshallo

 Bryan Shalloway

 bryanshalloway.com

```
---  
title: Identifying R Functions & Packages...  
author: Bryan Shalloway  
date: '2022-01-18'  
tags:  
  - "`r funspotr::spot_tags()`"  
...
```



<https://github.com/brshallo/funspotr>

 @brshallo

 Bryan Shalloway

 bryanshalloway.com

Identifying R Functions & Packages Used in GitHub Repos (funspotr part 1)

➔ Tags: [dplyr](#) [funspotr](#) [yaml](#) [purrr](#) [fs](#) [readr](#) [here](#)

2022-01-18

Tags

- [ameshousing](#) (3)
- [animatrixr](#) (3)
- [broom](#) (3)
- [chrome-api](#) (1)
- [corr](#) (1)
- [devtools](#) (6)
- [digest](#) (1)
- [dplyr](#) (23)

```

library(funspotr)

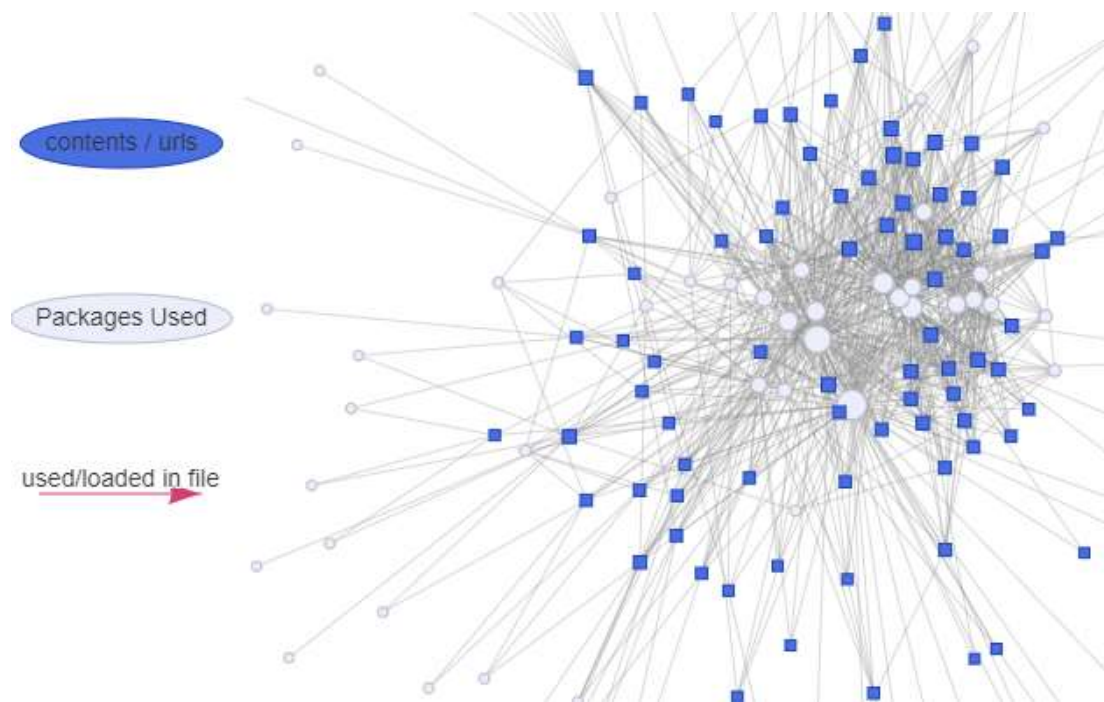
list_files_github_repo("juliasilge/juliasilge.com") |>
  spot_funs_files() |>
  unnest_results()
#> # A tibble: 3,531 x 4
#>   funs          pkgs          relative_paths          absolute_paths
#>   <chr>         <chr>         <chr>              <chr>
#> 1 library      base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 2 options      base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 3 austen_books janeaustenr    content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 4 unnest_tokens tidytext    content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 5 count        dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 6 ungroup      dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 7 group_by     dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 8 summarize    dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 9 sum          base          content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> 10 left_join    dplyr         content/blog/2016-06-27-Term-Freque~ https://raw.g~
#> # ... with 3,521 more rows

```

funspotr::network_plot()

```
library(funspotr)

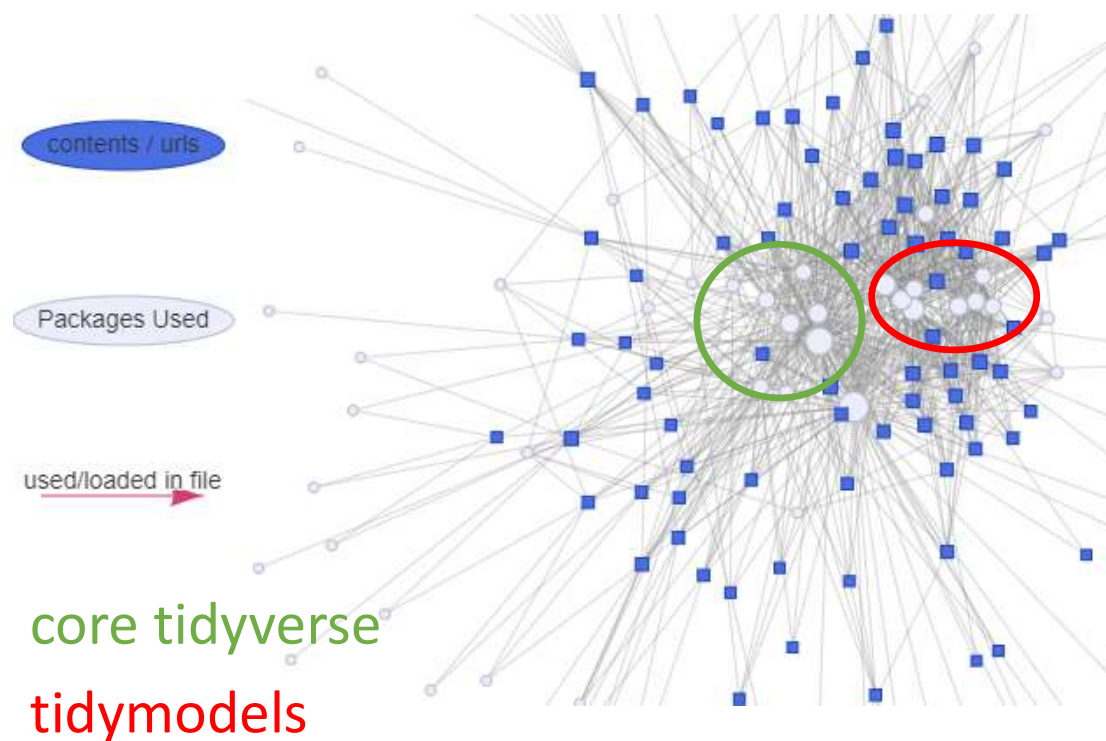
list_files_github_repo(
  "juliasilge/juliasilge.com"
) |>
spot_pkgs_files() |>
unnest_results() |>
network_plot()
```



funspotr::network_plot()

```
library(funspotr)

list_files_github_repo(
  "juliasilge/juliasilge.com"
) |>
spot_pkgs_files() |>
unnest_results() |>
network_plot()
```



funspotr::network_plot()

```
library(funspotr)

list_files_github_repo(
  "juliasilge/juliasilge.com"
) |>
spot_pkgs_files() |>
unnest_results() |>
network_plot()
```

...related

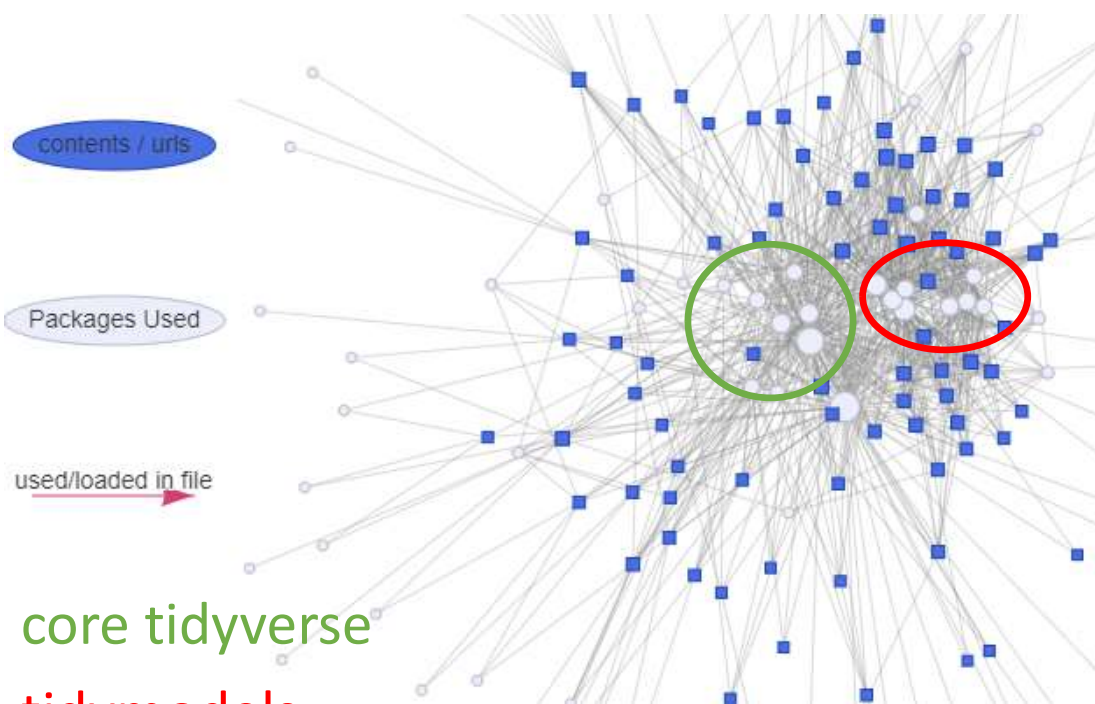
GITHUB NEXT: Visualizing a codebase

<https://githubnext.com/projects/repo-visualization>

OBSIDIAN

<https://obsidian.md/>

<https://github.com/brshallo/funspotr>

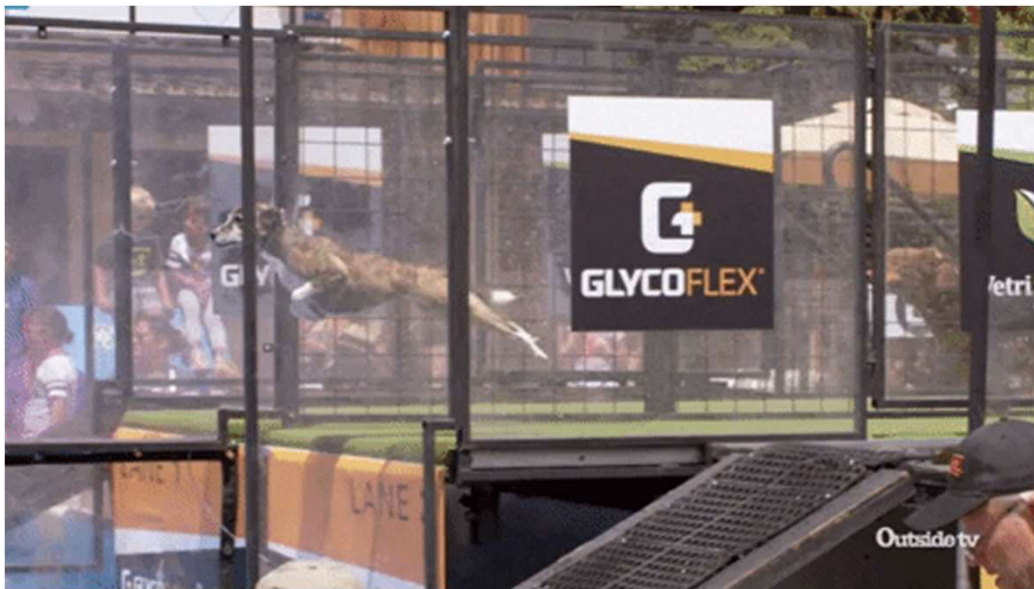


@brshallo

Bryan Shalloway

bryanshalloway.com

```
remotes::install_github("brshallo/funspotr")
```



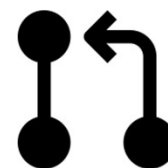
<https://github.com/brshallo/funspotr>

  @brshallo

 Bryan Shalloway

 bryanshalloway.com

```
remotes::install_github("brshallo/funspotr")
```



<https://github.com/brshallo/funspotr>

  @brshallo

 Bryan Shalloway

 bryanshalloway.com

Thank you!

- {funspotr}: <https://brshallo.github.io/funspotr/>
- Posts (old API):
 - <https://www.bryanshalloway.com/2022/01/18/identifying-r-functions-packages-used-in-github-repos/>
 - <https://www.bryanshalloway.com/2022/02/07/identifying-r-functions-packages-in-your-github-gists/>
 - <https://www.bryanshalloway.com/2022/03/17/network-plots-of-code-collections-funspotr-part-3/>
- Slides: <https://github.com/brshallo/funspotr-rstudioconf2022>