

# TSearch

## **TSearch**: An Experimental Search System for the Julia Community

Caleb Allen

# What is “important” information?

A brief discussion of web ranking

# PageRank and the Web Corpus

## 2.5 Random Surfer Model

The definition of PageRank above has another intuitive basis in random walks on graphs. The simplified version corresponds to the standing probability distribution of a random walk on the graph of the Web. Intuitively, this can be thought of as modeling the behavior of a “random surfer”. The “random surfer” simply keeps clicking on successive links at random.

Page et al., “The PageRank Citation Ranking,” 1999.

# PageRank and the Web Corpus



The benefits of PageRank are the greatest for underspecified queries. For example, a query for “Stanford University” may return any number of web pages which mention Stanford (such as publication lists) on a conventional search engine, but using PageRank, the university home page is listed first.

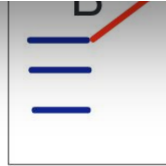


Figure 1: A and B are Backlinks of C

Page et al., “The PageRank Citation Ranking,” 1999.

# Julia Corpus

julia construct Val symbol literal

All Images Videos News Books Maps Flights More Tools

Julia Documentation  
https://docs.julialang.org › manual › metaprogramming

Metaprogramming - Julia Documentation  
Expressions provided by the parser generally only have **symbols**, other expressions, and **literal** values as their args, whereas expressions **constructed** by **Julia** ...  
Program representation · Expressions and evaluation · Macros · Code Generation

Stack Overflow  
https://stackoverflow.com › questions › what-is-a-symb...

What is a "symbol" in Julia?  
This is the essence of a **symbol**, a **symbol** is used to represent a variable in metaprogramming. Once you have **symbols** as a data type, it becomes ...  
2 answers · Top answer: Symbols in Julia are the same as in Lisp, Scheme or Ruby. However, ...  
Julia: How to create a **symbol** inside a function that matches ... 25 Aug 2016  
Literal Types in Julia - Stack Overflow 1 Jul 2022  
How to create a macro for creating a **symbol** in Julia 28 Oct 2014  
assign values to arrays with eval and **symbols** in Julia 5 Jan 2018  
More results from stackoverflow.com

Julia Documentation  
https://docs.julialang.org › manual › strings

Strings - Julia Documentation  
Characters. A Char **value** represents a single **character**: it is just a 32-bit primitive type with a special **literal** representation and appropriate arithmetic ...

JuliaHub  
https://docs.juliahub.com › HypertextLiteral.jl › yIQz

Reference · HypertextLiteral.jl  
This library provides for a `@hit()` macro and a hit string **literal**, both implementing interpolation that is aware of hypertext escape context. The `@hit` macro has ...

Read the Docs  
http://julia-doc-pe-br.readthedocs.io › latest › manual

Metaprogramming — Julia Language 0.3-dev documentation  
Expressions provided by the parser generally only have **symbols**, other expressions, and **literal** values as their args, whereas expressions **constructed** by **Julia** ...

Stack Overflow  
https://stackoverflow.com › questions › what-is-a-symb...

## What is a "symbol" in Julia?

This is the essence of a **symbol**: a **symbol** is used to represent a variable in metaprogramming. Once you have **symbols** as a data type, it becomes ...

2 answers · Top answer: Symbols in Julia are the same as in Lisp, Scheme or Ruby. However, ...

Julia: How to create a **symbol** inside a function that matches ... 25 Aug 2016

Literal Types in Julia - Stack Overflow 1 Jul 2022

How to create a macro for creating a **symbol** in Julia 28 Oct 2014

assign values to arrays with eval and **symbols** in Julia 5 Jan 2018

More results from stackoverflow.com

# Web Search: *julia linear algebra*

Julia 0.6!

julia linear algebra



All Images Videos News Books More

About 1,000,000 results (0.22 seconds)



Massachusetts Institute of Technology

<http://web.mit.edu> > doc > julia > html > stdlib > linalg

[Linear Algebra · The Julia Language](#)

In **Julia** (as in much of scientific computation), dense **linear-algebra** operations are based on the LAPACK library, which in turn is built on top of basic linear- ...



GitHub

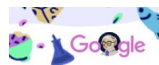
<https://github.com> > master > stdlib > LinearAlgebra > src

[LinearAlgebra.jl - JuliaLang/julia](#)

This file is a part of **Julia**. License is MIT: <https://julialang.org/license> Linear algebra module.

No official docs!

# Web Search



Search results for "julia dates" showing a snippet for JuliaDates.

Search results for "julia dates" (0.34 seconds)

About 75,400,000 results (0.34 seconds)

**JD** JuliaDates  
<https://juliadates.com> > ...

**JuliaDates**  
JuliaDates is a dating site for wealthy men and ambitious girls. People know exactly what they expect from a partner and what they ...  
[Log In](#) · [Join free](#) · [About Us](#) · [FAQ](#)

**MIT** Massachusetts Institute of Technology  
<http://web.mit.edu/share/doc/julia/html/stdlib> > ...

**Dates and Time · The Julia Language**  
All `Dates` functions are defined in the `Dates` module; note that only the `now` functions are exported; to use all other `Dates` functions, you' ...

**Wikipedia**  
[https://en.wikibooks.org/wiki/Working\\_with\\_dates...](https://en.wikibooks.org/wiki/Working_with_dates...) > ...

**Introducing Julia/Working with dates and times**  
Working with `dates` and `times` edit. Functions for working with `dates` and `times` in the standard package `Dates`. To use any of the time and date ...

**MIT** Massachusetts Institute of Technology  
<https://web.mit.edu/doc/julia/html/manual/dates> > ...

**Date and DateTime · The Julia Language**  
The `Dates` module provides two types for working with `dates`: `Date` and `DateTime`. `Date` has day and millisecond precision, respectively; both are subtypes of ...

Search results for "julia dates" showing a snippet for JuliaDates.

Search results for "julia dates" (0.34 seconds)

About 75,400,000 results (0.34 seconds)

**JD** JuliaDates  
<https://juliadates.com> > ...

**JuliaDates**  
JuliaDates is a dating site for wealthy men and ambitious girls. People know exactly what they expect from a partner and what they ...  
[Log In](#) · [Join free](#) · [About Us](#) · [FAQ](#)

←  r/Julia • 3 mo. ago  
 SquareZebra610

## Documentation

Hi,

... how do I find documentation for packages to know how



dudgybudgie · 3 mo. ago



Use the "?" REPL mode and type a module or a function to view the doc strings. Otherwise look for the git repo for the package and there's usually documentation there or a link to the docs site for that package. Remember most packages are community made so won't appear on the official Julia language docs.

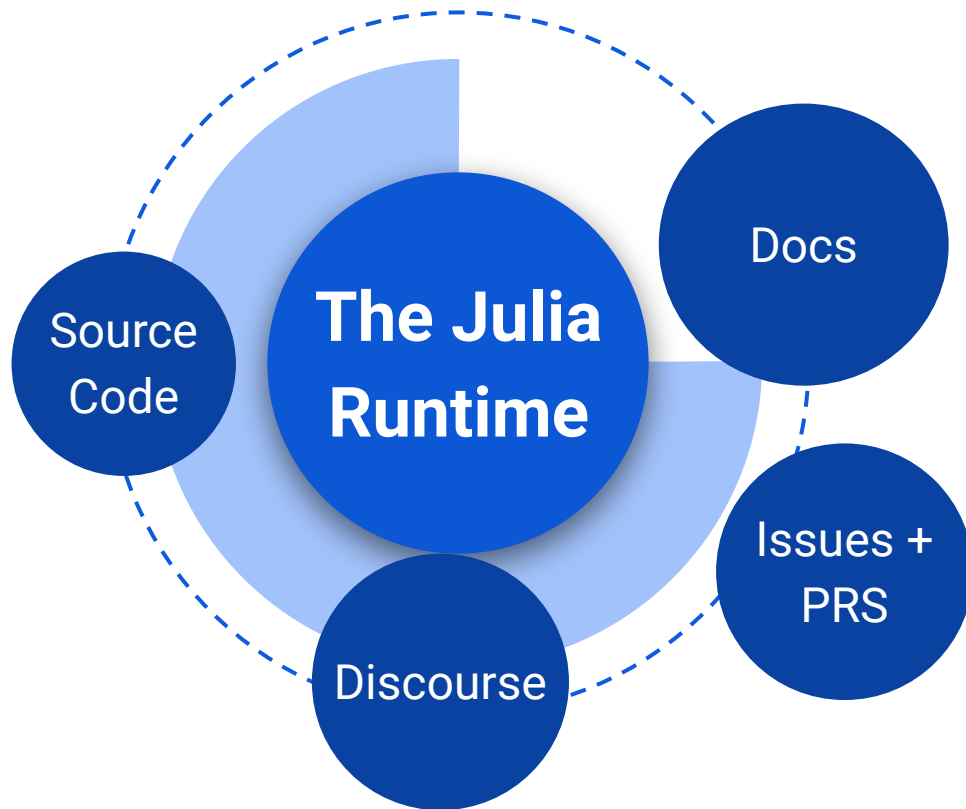
 9   Reply  Share ...

Their GitHub READMEs usually have links to docs at the top in these little boxes with dark background.

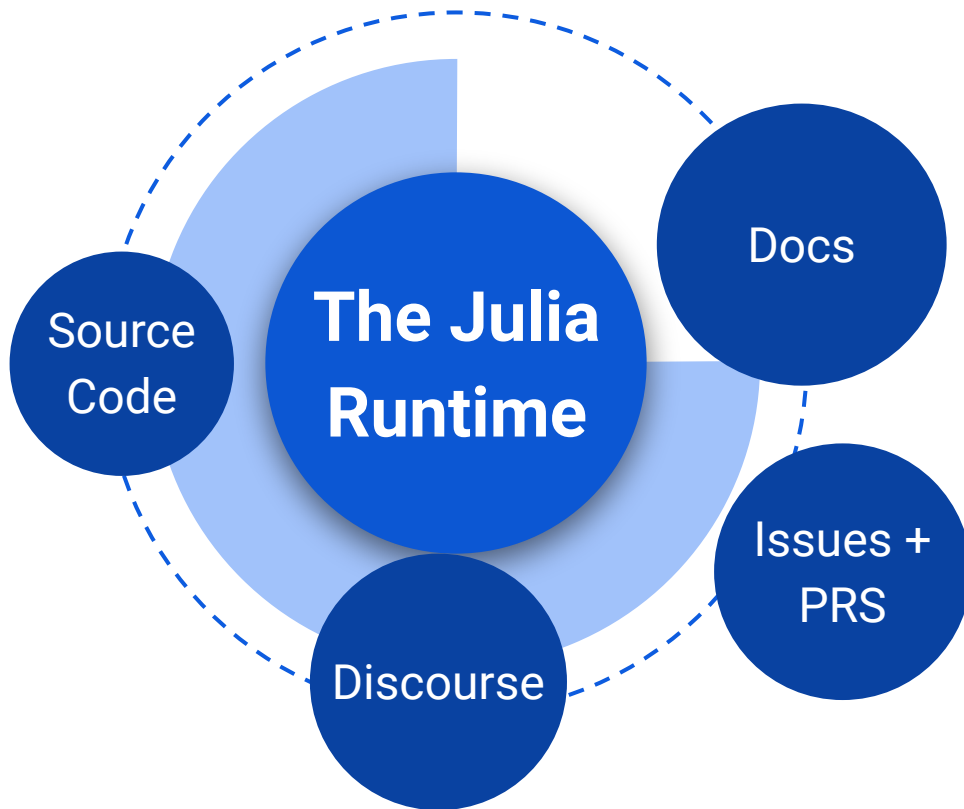


# The Julia Corpus is *Hypertext-Weak*

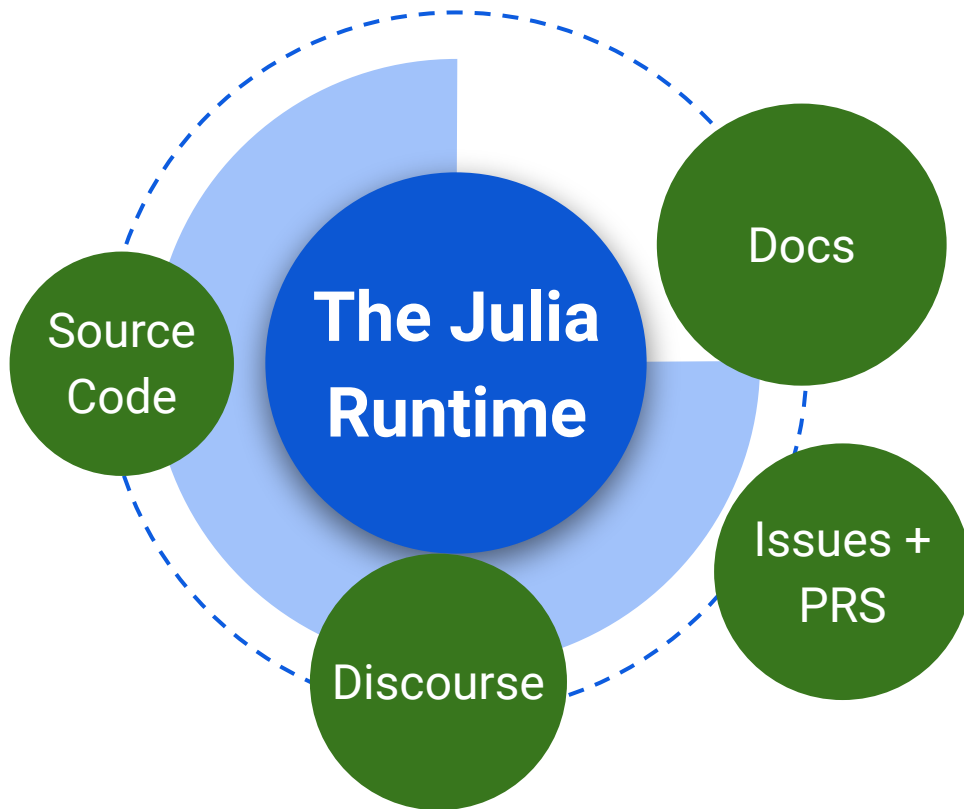
# What makes up the Julia Corpus?



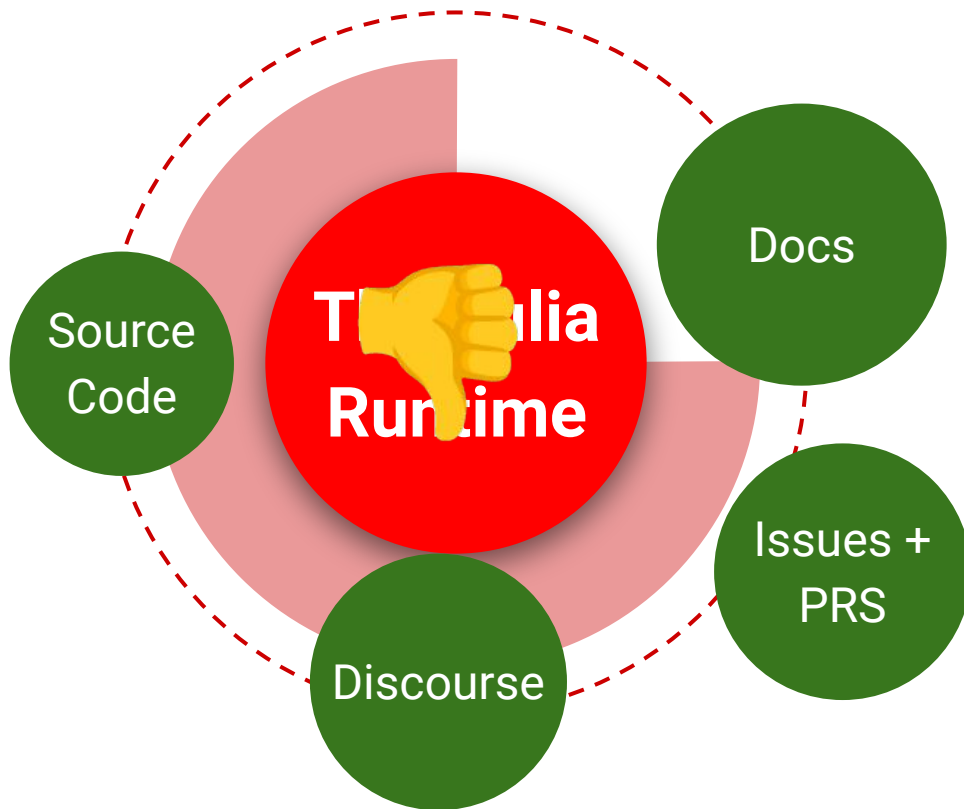
# What Julia information is on the web?



# What Julia information is on the web?



# What makes it to the web?



# TSearch: A Precision Search Tool

- A structured index built around *key tokens* from the Julia Runtime
  - type hierarchy
  - methods and method signatures
  - packages
- Like `methodswith` on steroids

# Index Structure: Precise and Dynamic

- Forward index, not inverse index
- Similar in principle to Microsoft's BitFunnel, with techniques taken from static flow analysis
- Hierarchical index accurately represents Julia structures
- Composable, incremental, performant
- Indexes of different data structures

# Crawling, tokenizing and indexing progress:

	Crawling	Tokenizing	Indexing
Base	✓	✓	✓
Stdlibs	✓	✓	✓
Packages	✓	✓	⚠
Discourse	✓	⚠	⚠
JuliaLang/julia PRs & Issues	✓	⚠	⚠



# Query structure

```
t"[Function]print*" # get all the Functions called "print*"
```

```
print_wrapped(f::Function, io::IO, args...; kws... @ Markdown
```

```
print_wrapped(io::IO, s...; width, pre, i @ Markdown
```

# Query structure

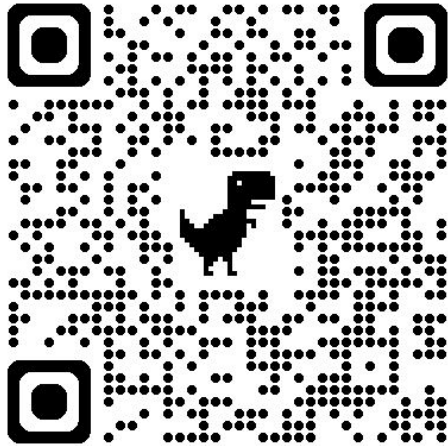
```
t"[Package]**/[Function]print*"
```

```
# find every function in every public package called "print*"
```

# Thank You!

Register your interest

<https://tsearch.org>



Thoughts?

Catch me at the hackathon, let's chat!

Also reach me online: Caleb Allen on slack, zulip, and discourse

- Inspired by “paper importance” algorithm
- Garfield, E., and I. H. Sher. “New Factors in the Evaluation of Scientific Literature through Citation Indexing.” *American Documentation* 14, no. 3 (July 1963): 195–201.  
<https://doi.org/10.1002/asi.5090140304>.

# BitFunnel

- Forward index based on the bloom-filter