HOW SUCCESSFUL IS MY SOFTWARE?

Slides:

https://softwaresaved.github.io/softwaresuccess-rse2017

Notes:

http://bit.ly/softwaresuccess-rse2017

WHO I AM

Neil Chue Hong, Software Sustainability Institute

GET THESE SLIDES

https://github.com/softwaresaved/softwaresuccess-rse2017/

MEASURING SOFTWARE IS HARD

WHAT ARE WE MEASURING?

QUALITY

ENGAGEMENT

IMPACT

WHY ARE WE MEASURING IT?

IMPROVEMENT

COMPARISON

AIMS OF THE SESSION

- Wider understanding of the topic
- Identifying success indicators that work for you
- Discussion of potential generic success indicators
- Suggestions for how these indicators could be measured

INTRODUCTIONS

INTRODUCTIONS

Your name

Where you're from

One thing you want to learn from this workshop

... add it to the shared notes ...
http://bit.ly/softwaresuccess-rse2017

IDENTIFYING SUCCESSFUL SOFTWARE

PAIR ACTIVITY

1. Break into pairs
2. One person in the pair:
    - Chooses a piece of research software they think is successful
    - Describe to your partner what measures make you think it is successful
    - Your partner writes these on a Post-It note as you talk
3. Swap and do the same for the other person in the pair
4. Discuss whether you had the same ways of identifying successful software, and where they differed.

HOW DO PEOPLE IDENTIFY SUCCESS?

GROUP DISCUSSION

What was important?

What was unexpected?

Does the opposite hold (absence == failure)?

What's easier to measure?

What's more accurate?

... be ready to share ...

EXISTING APPROACHES

SOFTWARE USAGE

Downloads

Popularity (Forks/Stars)

Citations

Downstream Use

... is someone else using it ...

SOFTWARE USABILITY

Functionality

User-friendliness

Performance

Documentation Quality

Multi-platform

... often subjective ...

PROJECT GROWTH AND MATURITY

Number of Contributors

Community Activity

Contribution Acceptance

Response Rate

Release Velocity

… visible activity …

PROJECT DIVERSITY

Contribution diversity

Contributor diversity

Contributor breadth

Decision distribution

Language bias

… hidden indicators …

CODE QUALITY

Test coverage

Maintainability Indices

Cyclomatic complexity

... can mask other failures ...

PROJECT RISK

Bus factor

Known Vulnerabilities

Bug Age

Downstream Use

… reliability leads to use …

BUT ULTIMATELY...

High search ranking in Google

MEASURING SOFTWARE

GROUP EXERCISE

1. Choose three success criteria that are important to your group
2. Create a small set of metrics that would allow you to rate a piece of software against that measure.
   - Apply your metrics to two pieces of software your group is familiar with and compare, *OR*
   - Use these metrics to rank the following set of software:
     - scipy: https://github.com/scipy/scipy
     - Spyder: https://github.com/spyder-ide/spyder
     - pycse: https://github.com/jkitchin/pycse

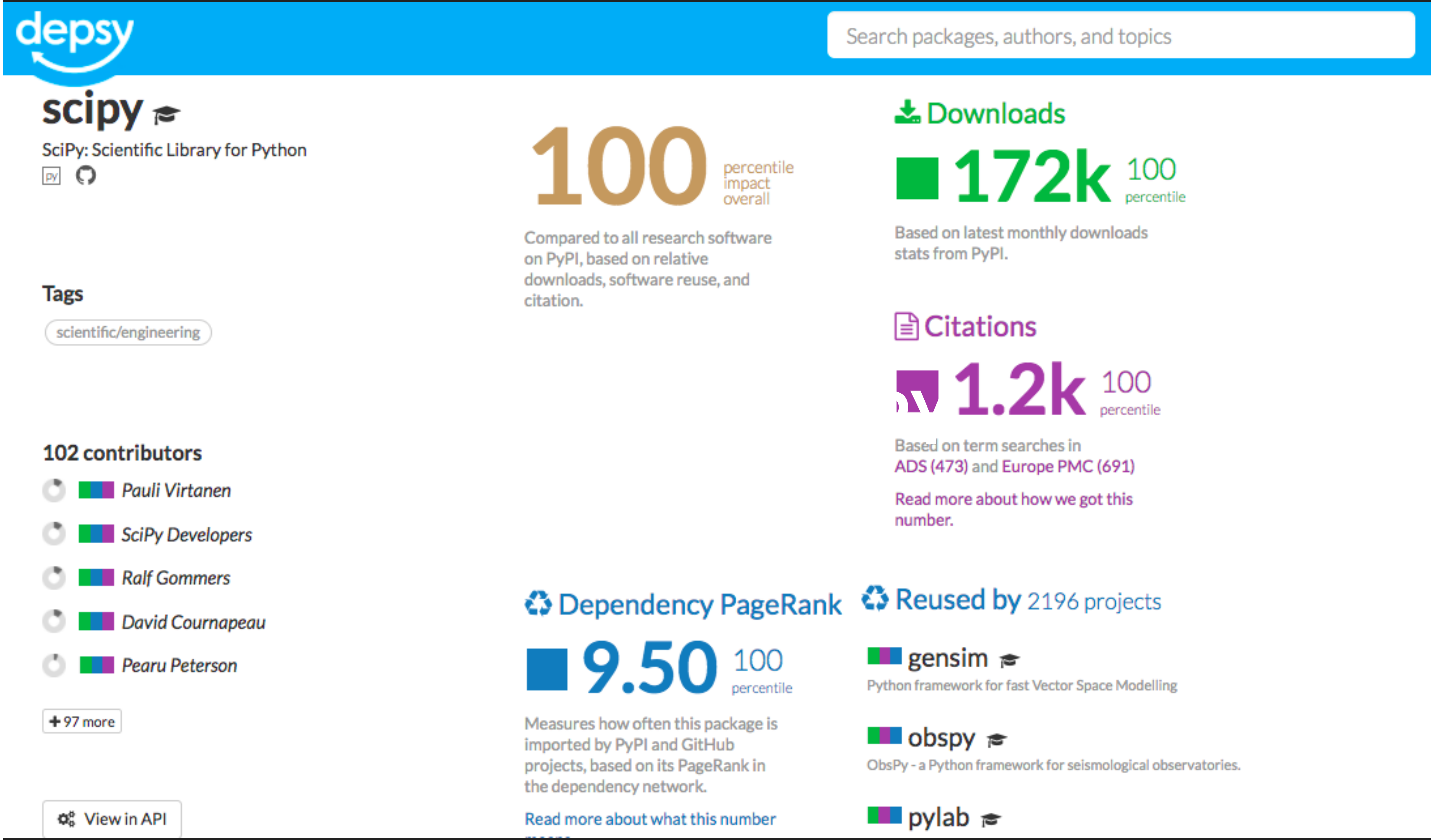
WHAT ARE EXISTING MODELS OF SUCCESS?

MANY OTHER SCHEMES

See our reference list:
https://github.com/softwaresaved/softwaresuccess-rse2017/blob/master/references.md

Contribute via pull requests

DEPSY



depsy

Search packages, authors, and topics

scipy

SciPy: Scientific Library for Python

Tags

scientific/engineering

102 contributors

- Pauli Virtanen
- SciPy Developers
- Ralf Gommers
- David Cournapeau
- Pearu Peterson

+97 more

View in API

100 percentile impact overall

Compared to all research software on PyPI, based on relative downloads, software reuse, and citation.

Downloads

172k 100 percentile

Based on latest monthly downloads stats from PyPI.

Citations

1.2k 100 percentile

Based on term searches in ADS (473) and Europe PMC (691)

Read more about how we got this number.

Dependency PageRank

9.50 100 percentile

Measures how often this package is imported by PyPI and GitHub projects, based on its PageRank in the dependency network.

Read more about what this number

Reused by 2196 projects

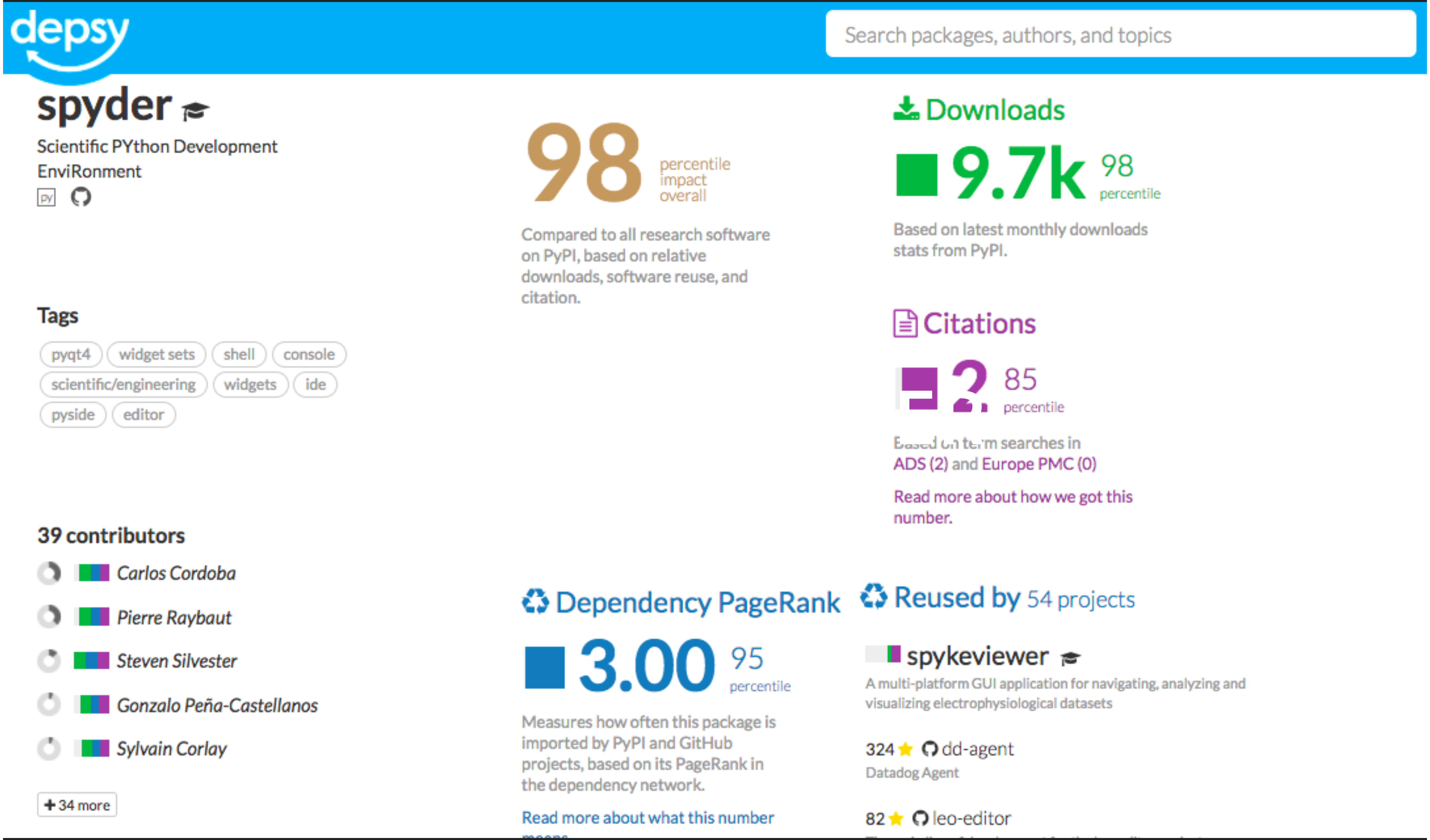
gensim

Python framework for fast Vector Space Modelling

obspy

ObsPy - a Python framework for seismological observatories.

pylab



depsy

Search packages, authors, and topics

spyder

Scientific PYthon Development EnviRonment

py

98 percentile impact overall

Compared to all research software on PyPI, based on relative downloads, software reuse, and citation.

Downloads

■ 9.7k 98 percentile

Based on latest monthly downloads stats from PyPI.

Citations

■ 2 85 percentile

Based on term searches in ADS (2) and Europe PMC (0)

Read more about how we got this number.

Tags

pyqt4 · widget sets · shell · console · scientific/engineering · widgets · ide · pyside · editor

39 contributors

Carlos Cordoba

Pierre Raybaut

Steven Silvester

Gonzalo Peña-Castellanos

Sylvain Corlay

+ 34 more

Dependency PageRank

■ 3.00 95 percentile

Measures how often this package is imported by PyPI and GitHub projects, based on its PageRank in the dependency network.

Read more about what this number means

Reused by 54 projects

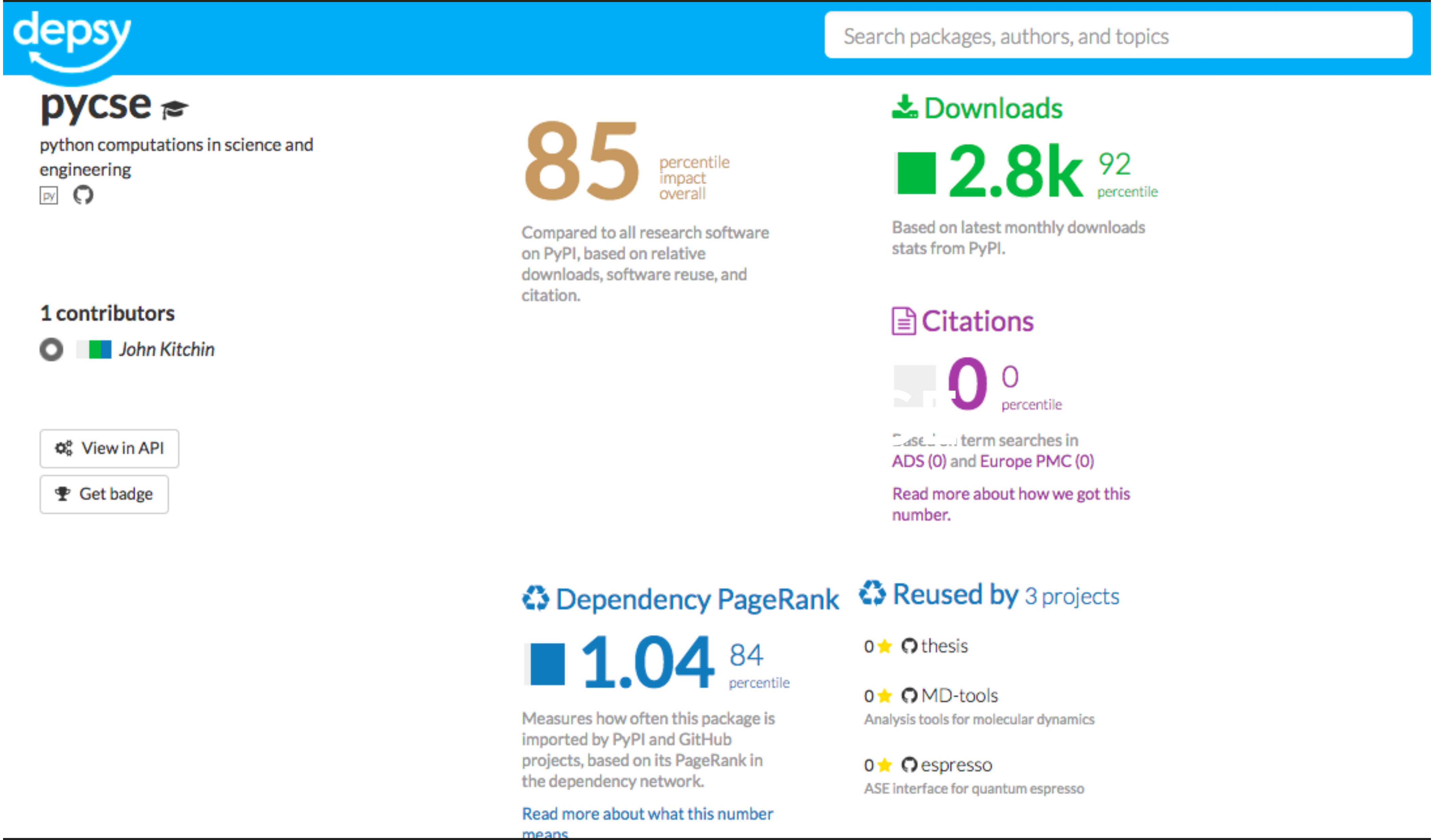
spykeviewer

A multi-platform GUI application for navigating, analyzing and visualizing electrophysiological datasets

324 ★ dd-agent

Datadog Agent

82 ★ leo-editor



pycse

python computations in science and engineering

85 percentile impact overall

Compared to all research software on PyPI, based on relative downloads, software reuse, and citation.

Downloads

2.8k 92 percentile

Based on latest monthly downloads stats from PyPI.

1 contributors

John Kitchin

Citations

0 0 percentile

Based on term searches in ADS (0) and Europe PMC (0)

Read more about how we got this number.

Dependency PageRank

1.04 84 percentile

Measures how often this package is imported by PyPI and GitHub projects, based on its PageRank in the dependency network.

Read more about what this number means

Reused by 3 projects

0 ★ thesis

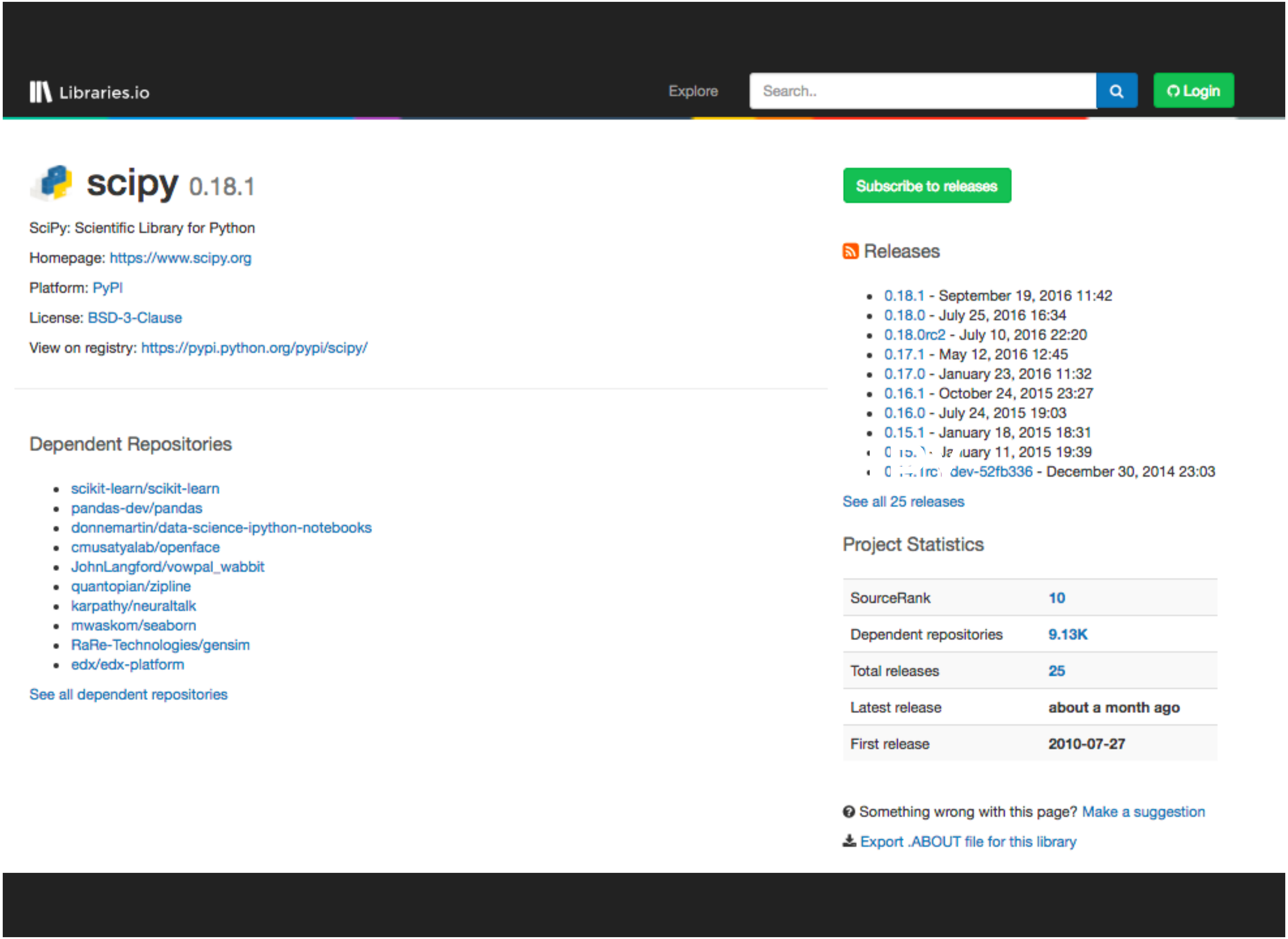
0 ★ MD-tools

Analysis tools for molecular dynamics

0 ★ espresso

ASE interface for quantum espresso

LIBRARIES.IO AND SOURCERANK



scipy 0.18.1

SciPy: Scientific Library for Python

Homepage: https://www.scipy.org

Platform: PyPI

License: BSD-3-Clause

View on registry: https://pypi.python.org/pypi/scipy/

Dependent Repositories

- scikit-learn/scikit-learn
- pandas-dev/pandas
- donnemartin/data-science-ipython-notebooks
- cmusatyalab/openface
- JohnLangford/vowpal_wabbit
- quantopian/zipline
- karpathy/neuraltalk
- mwaskom/seaborn
- RaRe-Technologies/gensim
- edx/edx-platform

See all dependent repositories

Releases

- 0.18.1 - September 19, 2016 11:42
- 0.18.0 - July 25, 2016 16:34
- 0.18.0rc2 - July 10, 2016 22:20
- 0.17.1 - May 12, 2016 12:45
- 0.17.0 - January 23, 2016 11:32
- 0.16.1 - October 24, 2015 23:27
- 0.16.0 - July 24, 2015 19:03
- 0.15.1 - January 18, 2015 18:31
- 0.15.0 - January 11, 2015 19:39
- 0.14.1rc1.dev-52fb336 - December 30, 2014 23:03

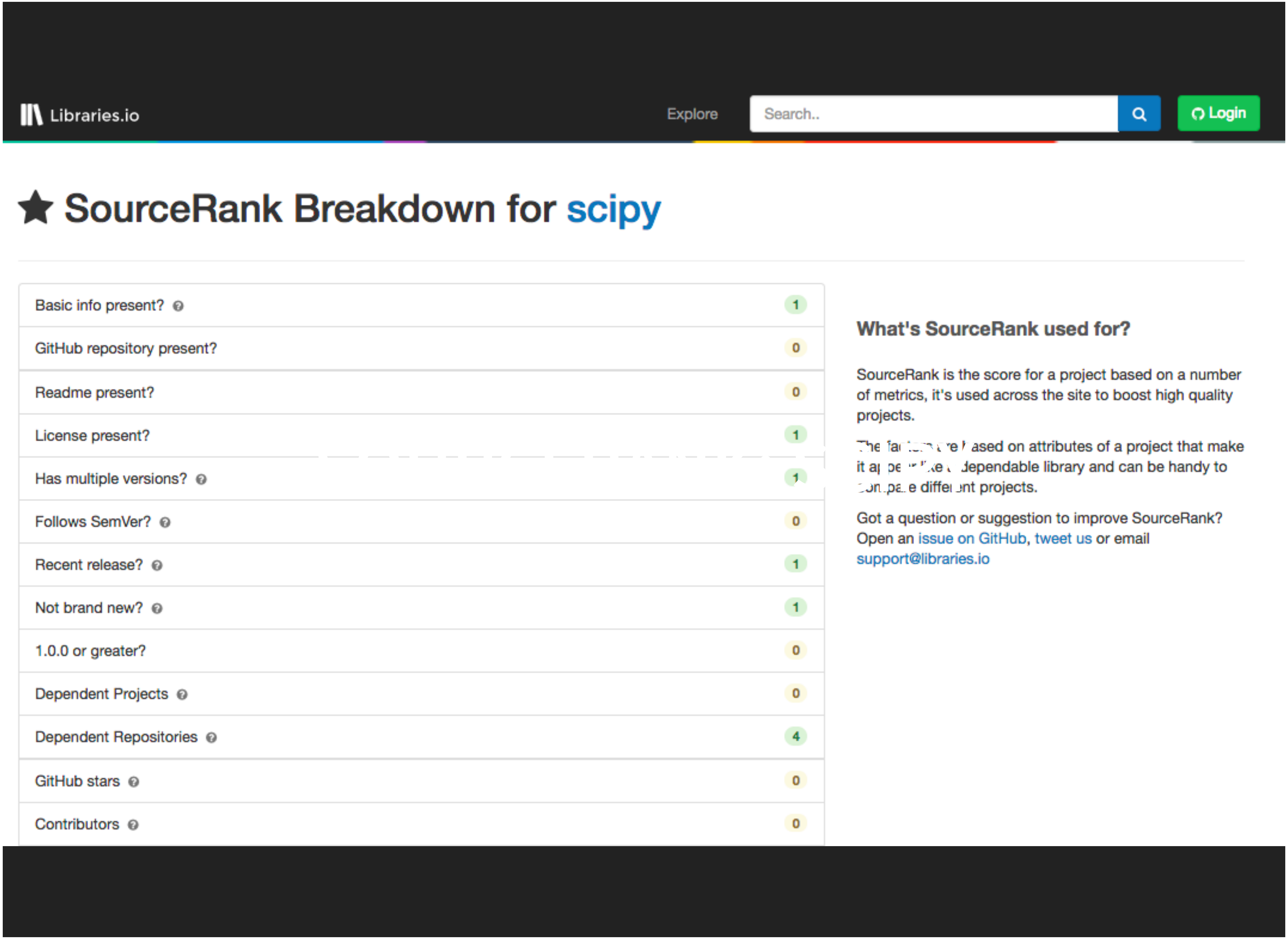
See all 25 releases

Project Statistics

SourceRank	10
Dependent repositories	9.13K
Total releases	25
Latest release	about a month ago
First release	2010-07-27

Something wrong with this page? Make a suggestion

Export .ABOUT file for this library



SourceRank Breakdown for scipy

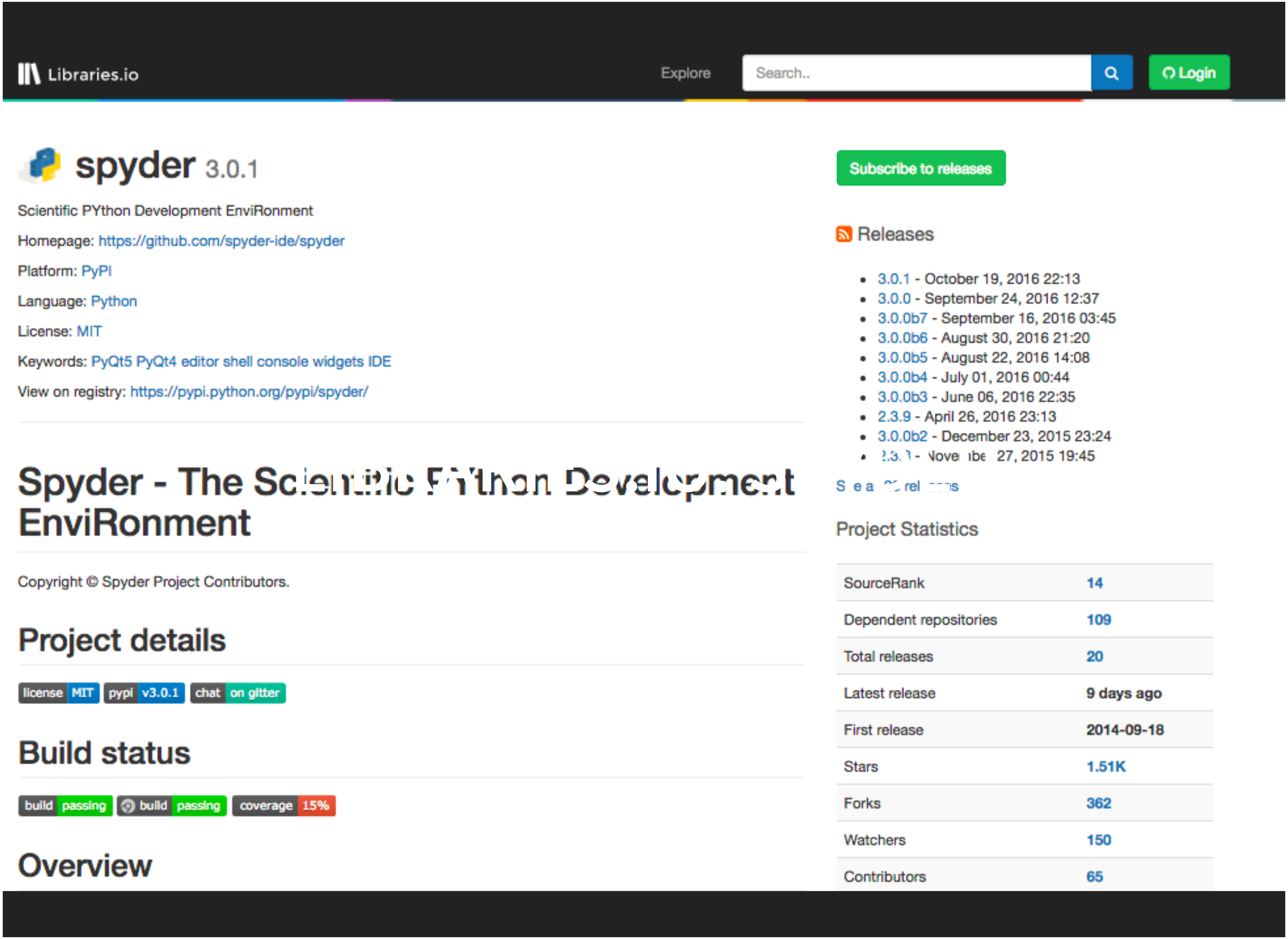
Basic info present?	1
GitHub repository present?	0
Readme present?	0
License present?	1
Has multiple versions?	1
Follows SemVer?	0
Recent release?	1
Not brand new?	1
1.0.0 or greater?	0
Dependent Projects	0
Dependent Repositories	4
GitHub stars	0
Contributors	0

What's SourceRank used for?

SourceRank is the score for a project based on a number of metrics, it's used across the site to boost high quality projects.

The factors are based on attributes of a project that make it appear like a dependable library and can be handy to compare different projects.

Got a question or suggestion to improve SourceRank? Open an issue on GitHub, tweet us or email support@libraries.io



spyder 3.0.1

Scientific PYthon Development EnviRonment

Homepage: https://github.com/spyder-ide/spyder

Platform: PyPI

Language: Python

License: MIT

Keywords: PyQt5 PyQt4 editor shell console widgets IDE

View on registry: https://pypi.python.org/pypi/spyder/

Spyder - The Scientific PYthon Development EnviRonment

Copyright © Spyder Project Contributors.

Project details

Build status

Overview

Subscribe to releases

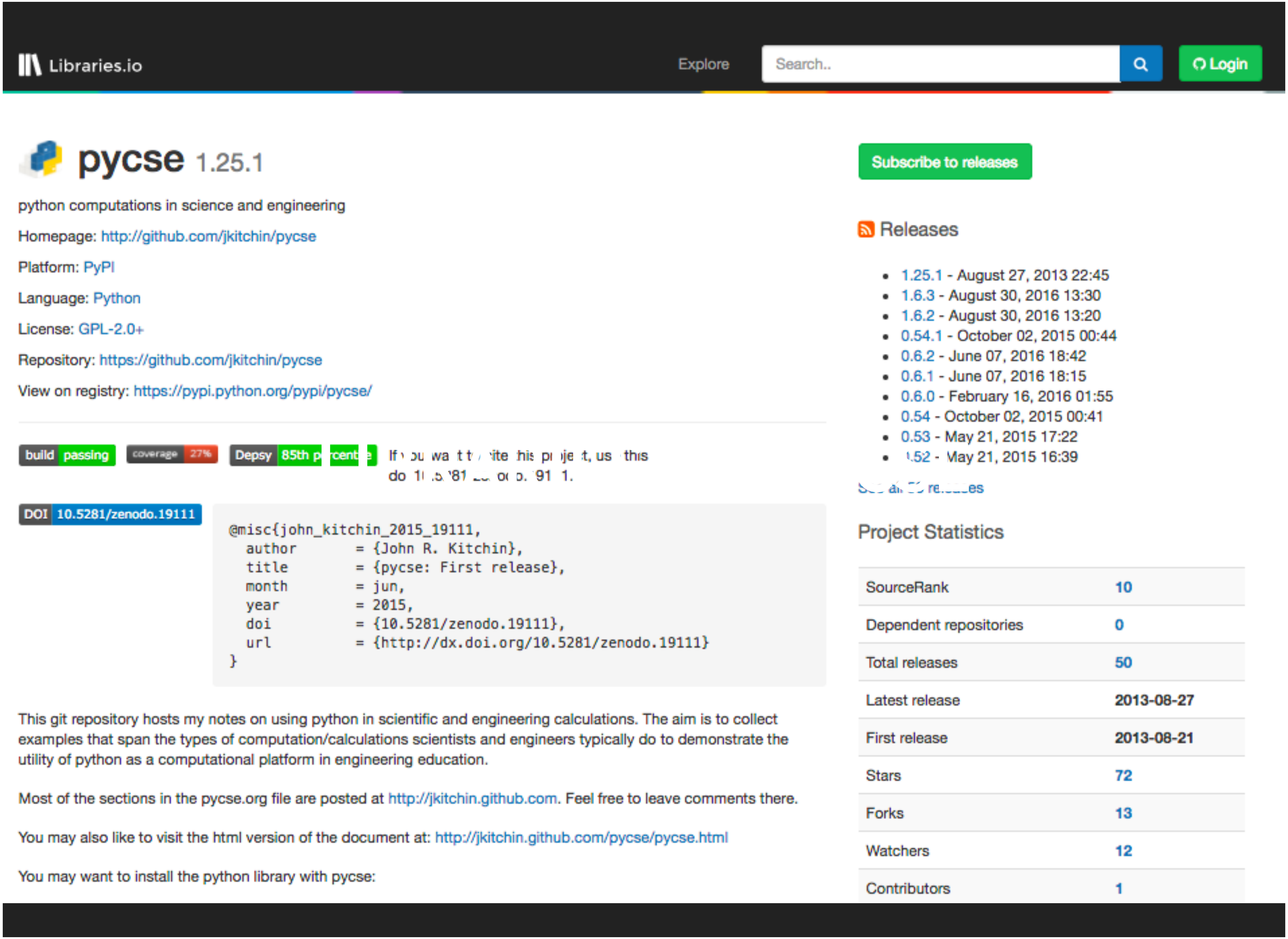
Releases

- 3.0.1 - October 19, 2016 22:13
- 3.0.0 - September 24, 2016 12:37
- 3.0.0b7 - September 16, 2016 03:45
- 3.0.0b6 - August 30, 2016 21:20
- 3.0.0b5 - August 22, 2016 14:08
- 3.0.0b4 - July 01, 2016 00:44
- 3.0.0b3 - June 06, 2016 22:35
- 2.3.9 - April 26, 2016 23:13
- 3.0.0b2 - December 23, 2015 23:24
- 2.3.8 - November 27, 2015 19:45

See all 20 releases

Project Statistics

SourceRank	14
Dependent repositories	109
Total releases	20
Latest release	9 days ago
First release	2014-09-18
Stars	1.51K
Forks	362
Watchers	150
Contributors	65



pycse 1.25.1

python computations in science and engineering

Homepage: http://github.com/jkitchin/pycse

Platform: PyPI

Language: Python

License: GPL-2.0+

Repository: https://github.com/jkitchin/pycse

View on registry: https://pypi.python.org/pypi/pycse/

build passing coverage 27% Depsy 85th percentile

If you want to cite this project, use this doi 10.5281/zenodo.19111.

DOI 10.5281/zenodo.19111

```
@misc{john_kitchin_2015_19111,
  author       = {John R. Kitchin},
  title        = {pycse: First release},
  month        = jun,
  year         = 2015,
  doi          = {10.5281/zenodo.19111},
  url          = {http://dx.doi.org/10.5281/zenodo.19111}
}
```

This git repository hosts my notes on using python in scientific and engineering calculations. The aim is to collect examples that span the types of computation/calculations scientists and engineers typically do to demonstrate the utility of python as a computational platform in engineering education.

Most of the sections in the pycse.org file are posted at http://jkitchin.github.com. Feel free to leave comments there.

You may also like to visit the html version of the document at: http://jkitchin.github.com/pycse/pycse.html

You may want to install the python library with pycse:

Releases

- 1.25.1 - August 27, 2013 22:45
- 1.6.3 - August 30, 2016 13:30
- 1.6.2 - August 30, 2016 13:20
- 0.54.1 - October 02, 2015 00:44
- 0.6.2 - June 07, 2016 18:42
- 0.6.1 - June 07, 2016 18:15
- 0.6.0 - February 16, 2016 01:55
- 0.54 - October 02, 2015 00:41
- 0.53 - May 21, 2015 17:22
- 0.52 - May 21, 2015 16:39

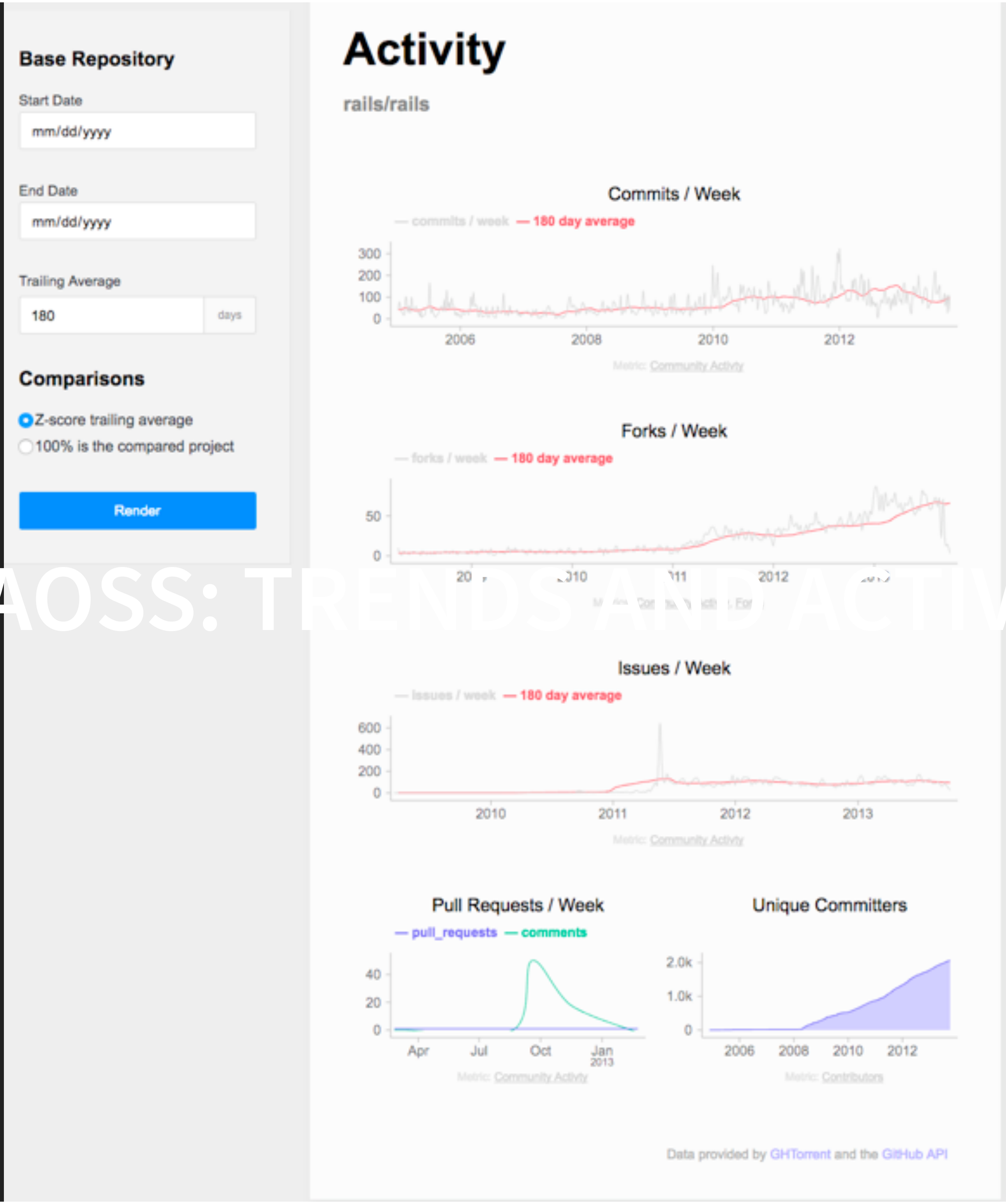
See all 50 releases

Project Statistics

SourceRank	10
Dependent repositories	0
Total releases	50
Latest release	2013-08-27
First release	2013-08-21
Stars	72
Forks	13
Watchers	12
Contributors	1

CHAOSS FRAMEWORK

CHAOSS: TRENDS AND ACTIVITY



Base Repository

Start Date

mm/dd/yyyy

End Date

mm/dd/yyyy

Trailing Average

180 days

Comparisons

● Z-score trailing average

○ 100% is the compared project

Activity

rails/rails

Commits / Week

— commits / week — 180 day average
300
200
100
0
2006
2008
2010
2012

Metric: Community Activity

Forks / Week

— forks / week — 180 day average
50
0
2012

Issues / Week

— issues / week — 180 day average
600
400
200
0
2010
2011
2012
2013

Metric: Community Activity

Pull Requests / Week

— pull_requests — comments
40
20
0
Apr
Jul
Oct
Jan 2013

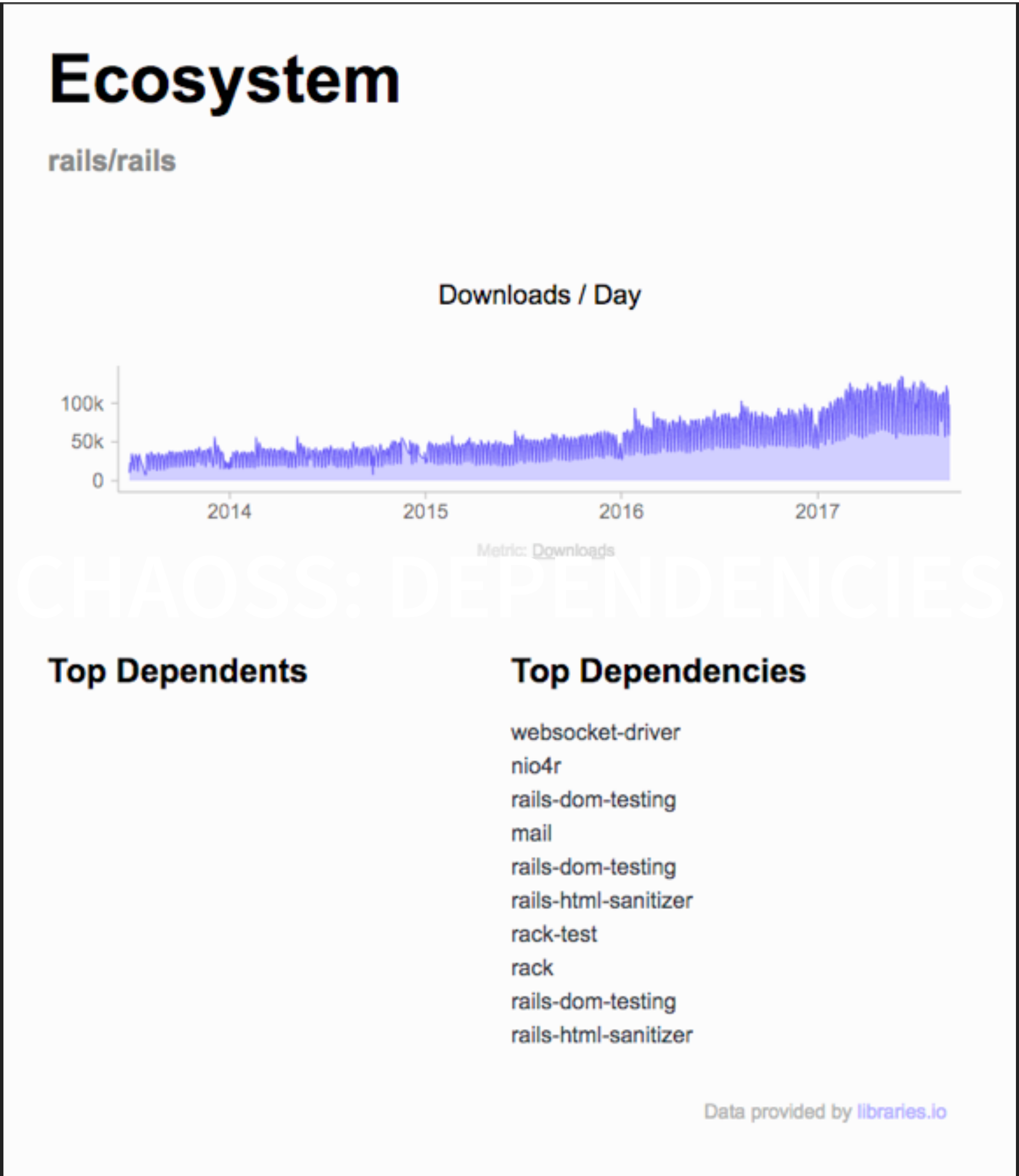
Metric: Community Activity

Unique Committers

2.0k
1.0k
0
2006
2008
2010
2012

Metric: Contributors

Data provided by GHTorrent and the GitHub API



Ecosystem

rails/rails

Downloads / Day

CHAOSS: DEPENDENCIES

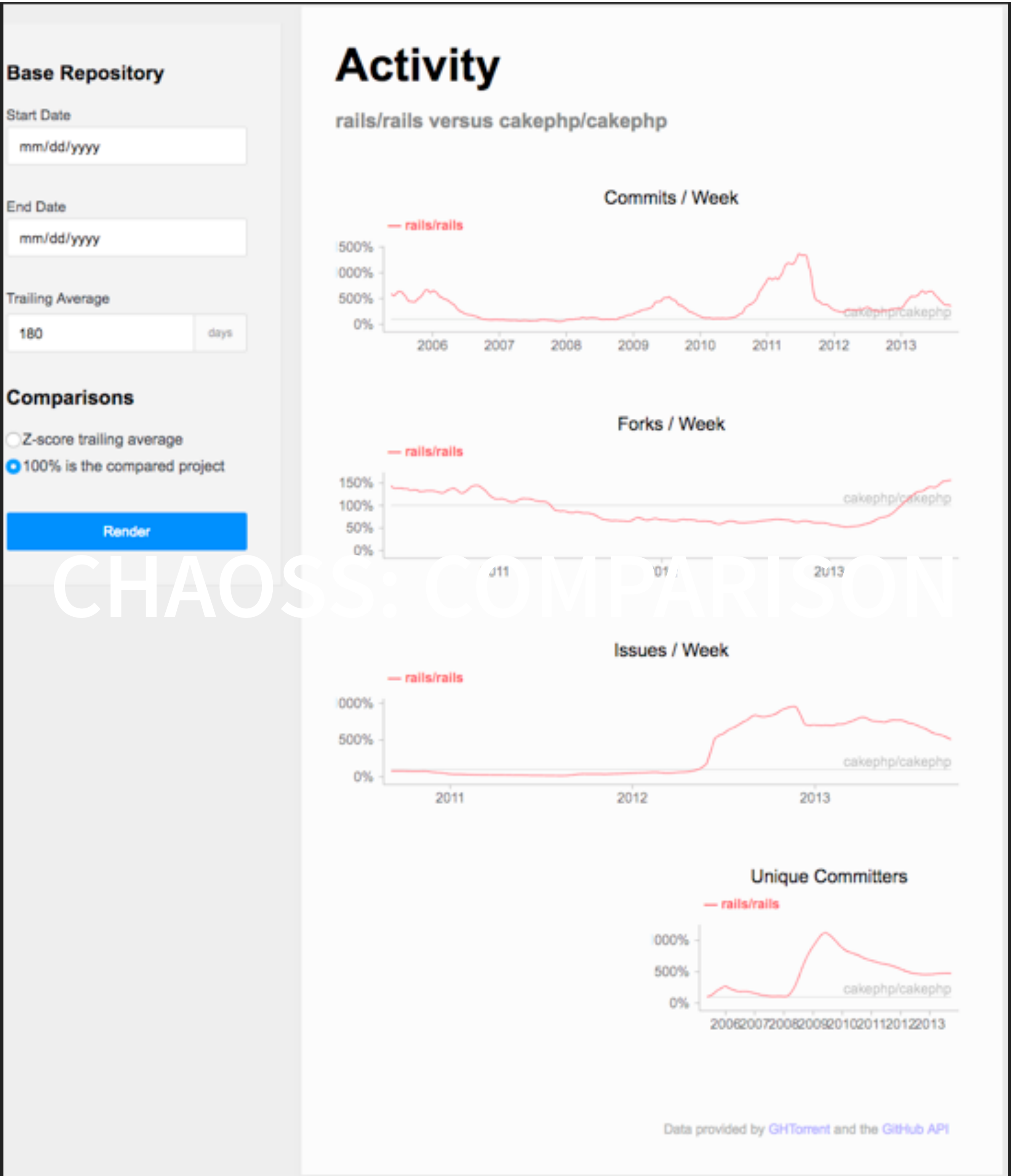
Top Dependents

Top Dependencies

websocket-driver
nio4r
rails-dom-testing
mail
rails-dom-testing
rails-html-sanitizer
rack-test
rack
rails-dom-testing
rails-html-sanitizer

Data provided by libraries.io

CHAOSS: COMPARISON



Base Repository

Start Date: mm/dd/yyyy

End Date: mm/dd/yyyy

Trailing Average: 180 days

Comparisons

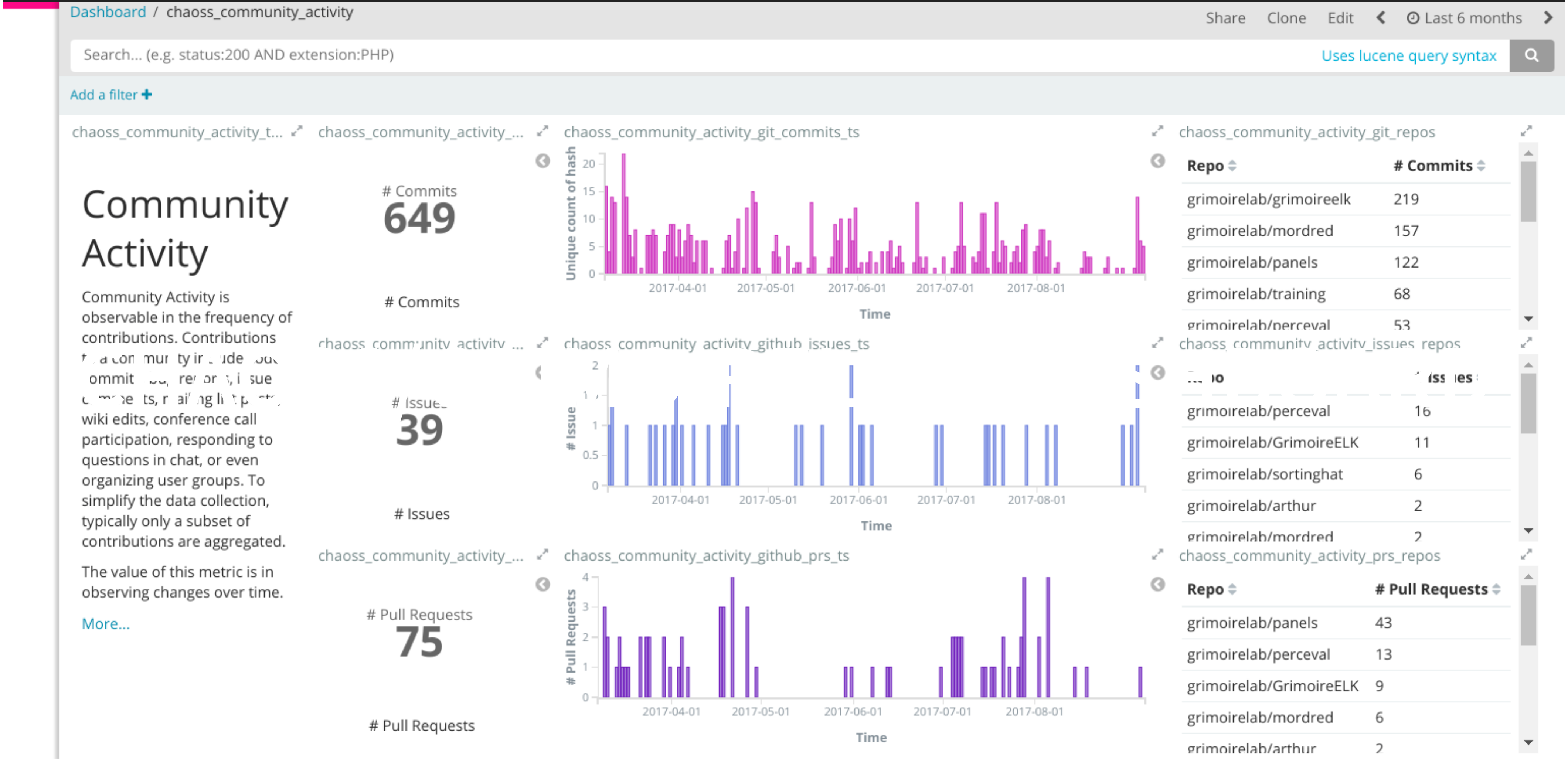
○ Z-score trailing average

● 100% is the compared project

Activity

rails/rails versus cakephp/cakephp

Data provided by GHTorrent and the GitHub API



Dashboard / chaoss_community_activity
Share
Clone
Edit
Last 6 months
Search... (e.g. status:200 AND extension:PHP)
Uses lucene query syntax
Add a filter +
chaoss_community_activity_t...
chaoss_community_activity_...
chaoss_community_activity_git_commits_ts
Unique count of hash
20
15
10
5
0
2017-04-01
2017-05-01
2017-06-01
2017-07-01
2017-08-01
Time
chaoss_community_activity_...
chaoss_community_activity_github_issues_ts
# Issue
2
1
1
0.5
0
2017-04-01
2017-05-01
2017-06-01
2017-07-01
2017-08-01
Time
chaoss_community_activity_...
chaoss_community_activity_github_prs_ts
# Pull Requests
4
3
2
1
0
2017-04-01
2017-05-01
2017-06-01
2017-07-01
2017-08-01
Time

Community Activity

Community Activity is observable in the frequency of contributions. Contributions by a community include code commits, bug reports, issue comments, mailing list posts, wiki edits, conference call participation, responding to questions in chat, or even organizing user groups. To simplify the data collection, typically only a subset of contributions are aggregated.

The value of this metric is in observing changes over time.

More...

# Commits
649
# Commits

# Issues
39
# Issues

# Pull Requests
75
# Pull Requests

chaoss_community_activity_git_repos

Repo	# Commits
grimoirelab/grimoireelk	219
grimoirelab/mordred	157
grimoirelab/panels	122
grimoirelab/training	68
grimoirelab/perceval	53

chaoss_community_activity_issues_repos

Repo	# Issues
grimoirelab/perceval	16
grimoirelab/GrimoireELK	11
grimoirelab/sortinghat	6
grimoirelab/arthur	2
grimoirelab/mordred	2

chaoss_community_activity_prs_repos

Repo	# Pull Requests
grimoirelab/panels	43
grimoirelab/perceval	13
grimoirelab/GrimoireELK	9
grimoirelab/mordred	6
grimoirelab/arthur	2

WHAT WE'VE LEARNED

WHAT WE'VE LEARNED

Many different ways to measure software

Different people will place different importance on different measures

Some things are easy to measure,
some things are useful to measure

There are already examples out there -
some may fit your requirements

Contribute your own ideas!