



CHA OSS: Conceptualizing, Prototyping and
Evaluating Open Source Software Project
Health Metrics: Measuring Growth, Maturity,
Decline and License Risk with Software



How to track sustainable projects...



How can we know if this open source project is likely to be around in 10 years if we base a product on it?

How can we know if this open source project is ready to be used by another project?

Is there a diverse community of active contributors engaged in the project?

Are there licensing risks in using this open source project?

What is the health of the other projects that this project depends on?

... need a common understanding of open source project health.

CHAOSS Mission



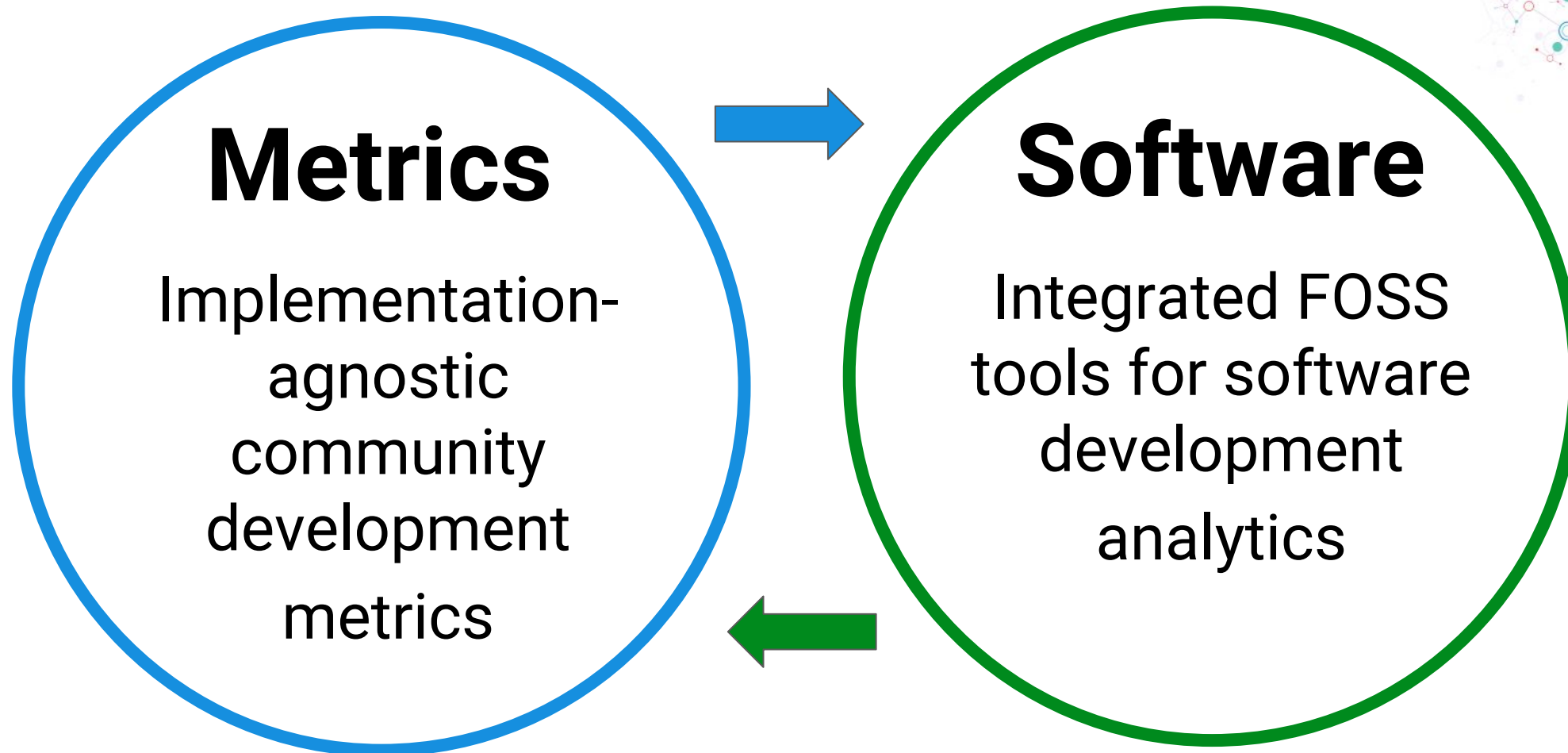
Establish implementation-agnostic metrics for measuring community activity, contributions, and health.

Produce integrated, open source software for analyzing software development in terms of these metrics.

Working in an Open Community...



Structure: Focus Around Interests



Metrics Committee



Diversity-Inclusion	Growth-Maturity-Decline
Risk	Value

wiki.linuxfoundation.org/chaoss/metrics



Diversity and Inclusion are known to challenge unchecked assumptions and lead to more open and fair collaboration practices.

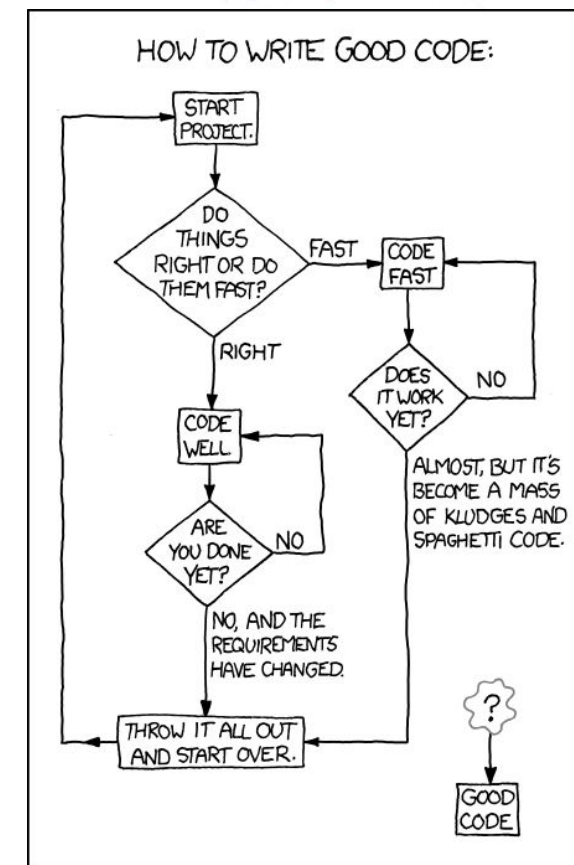
An OSS community has states: ***Growth, Maturity, and Decline***. The state that a community is in may prove important when evaluating both across and within community concerns.

The ***Risk*** metric informs how much risk an OSS community might pose. The evaluation of risk depends on situation and purpose.

Developers and organizations capture ***Value*** from engaging in OSS communities. This set of metrics can inform what this value is.

Why?

- No single health determination can be made across all open source projects, however:
 - We can start to understand what composite metrics signal and how they can be related to actions
- We aim to provide insight as local interpretations are done on the metrics
 - Provide guideposts for what others have done in similar contexts and how peer communities compare
- Develop an understanding of development processes based on facts



Software Committee

Implement Reference in Open Source

- Develop a FLOSS reference implementation of defined metrics.
- Integrate GrimoireLab, GHData, Prospector, and cregit into an Open Source Collaborative Framework
- Develop a better understanding of how contributions happen to large projects over time.



GHData

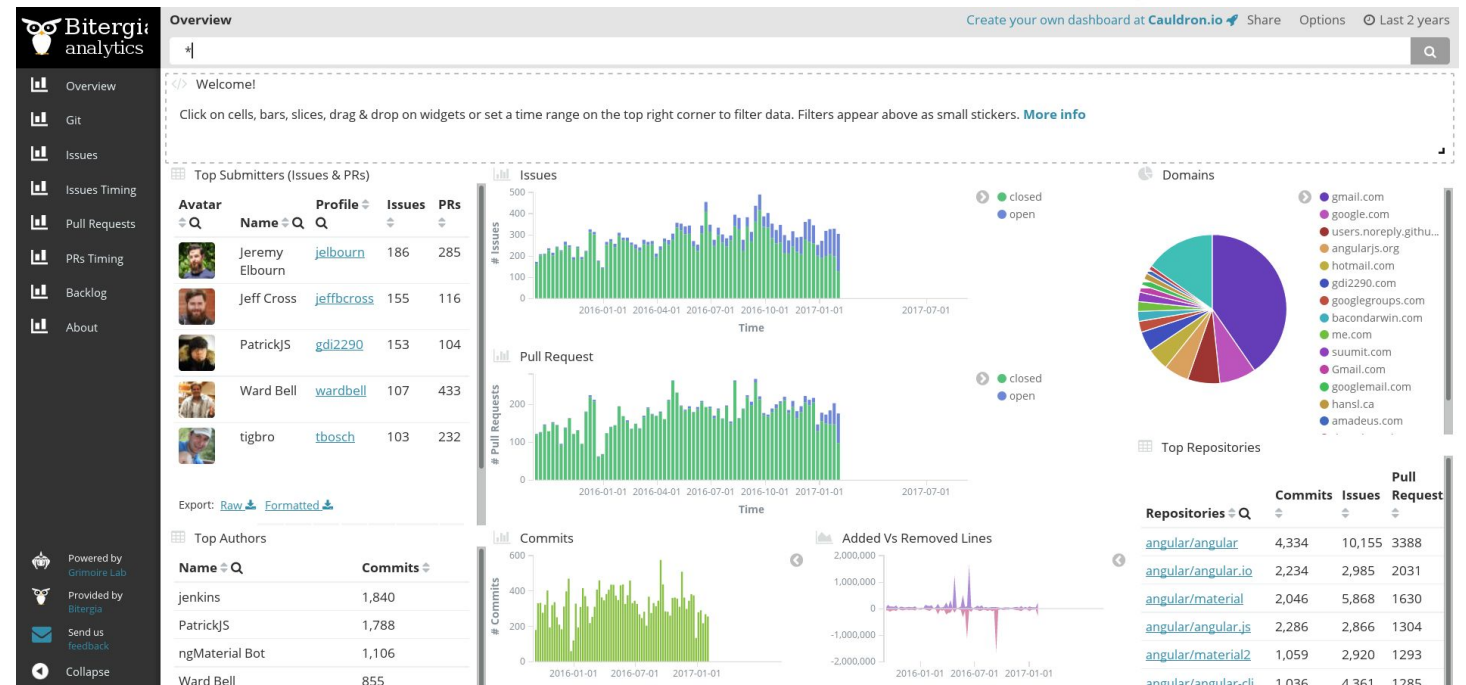
Prospector

cregit

GrimoireLab



- Retrieval from +30 data sources
- Storage of all metadata (ElasticSearch)
- Computing of interesting metrics
- Visualization
- Reports



github.com/chaoss/grimoirelab

Augur

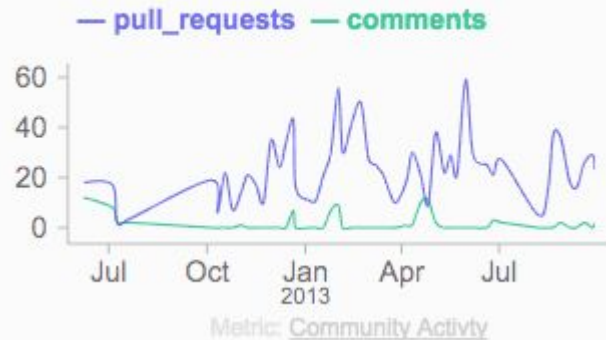
Prototyping Human Centered Metrics:
Enable Comparisons, Make Trends Central to the Use Experience



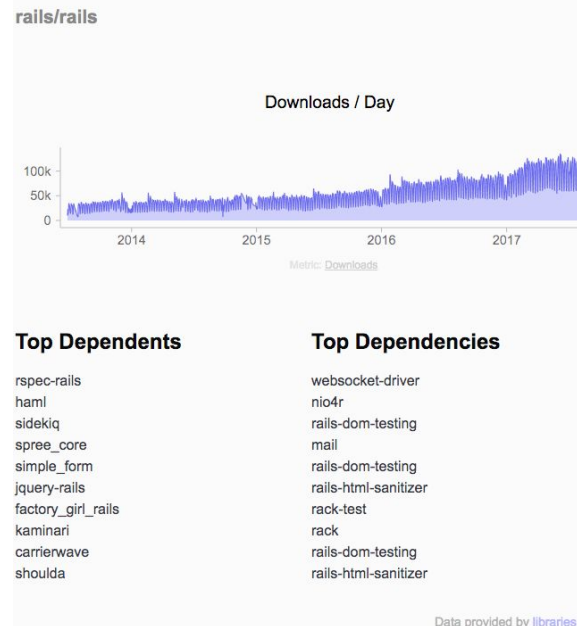
Comparisons

- Z-score trailing average
- 100% is the compared project

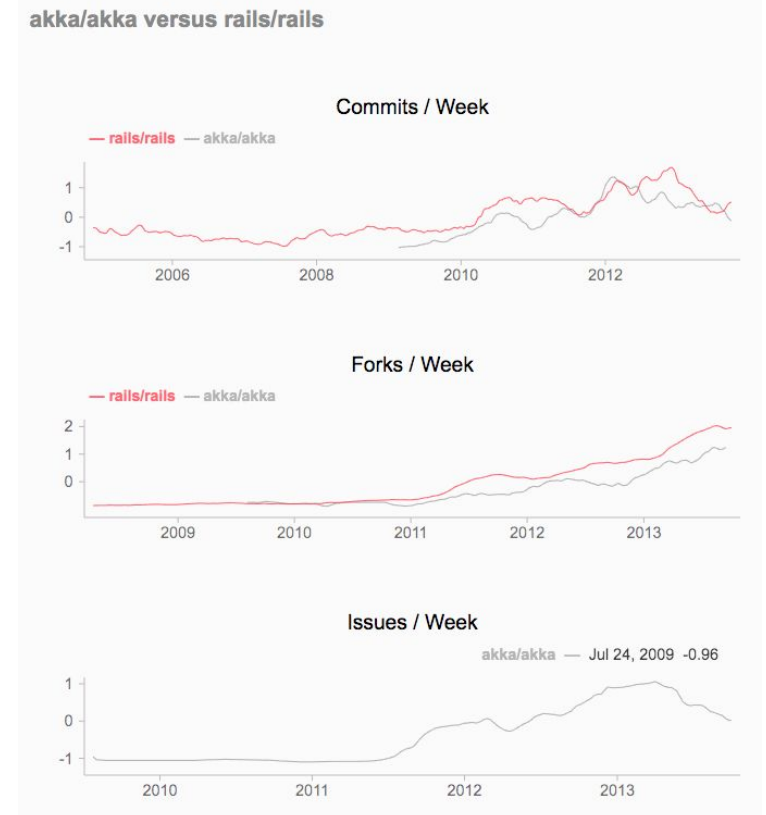
Pull Requests / Week



Ecosystem



Activity



<https://github.com/CHAOSS/augur>

Augur

Integrating DoSOCSv2 for License Risk Assessment and Management
Integrating Facade for Consistent Git Repository Mining



<https://github.com/DoSOCSv2/DoSOCSv2>

<https://github.com/brianwarner/facade>

Prospector



Metric datapoints:

- **Percentage of committers by dominant domain name: 37.01%**
Rationale: If more than 50% of COMMITTERS are from one domain (via email ID) it is dominated by one set of people. Suggested target is to have it less than 35%.
- **Percentage of commits by dominant domain name: 57.24%**
Rule: If more than 50% of CODE COMMITS are from one domain (via email ID) it is dominated by one set of people. Suggested target is to have it less than 35%.
- **Unique email address domains: 345 domains**
How many unique domains are represented?
- **Unique committers: 846 committers**
How many UNIQUE COMMITTERS are represented?

- Provides an objective, consistent and repeatable set of metrics of projects for success, sustainability and vibrancy.
- These can then coherently help assess and track continuously open source projects, which in turn would help drive the evolution of projects

<https://github.com/chaoss/prospector>



- Git-blame tracks changed lines, not tokens
 - Last person who modified part of a line, becomes “contributor” of the entire line
 - Cregit is capable of tracking the contributor of each token in a line
- In Linux Kernel (cregit.linuxsources.org):
 - blame per line is accurate in 75%
 - blame per token (using cregit) is accurate 95%
 - Results based on statistical sampling and manual analysis, with 95% reliability with +/-5% of error

github.com/cregit

Live Examples to Explore



Augur:

twitter.augurlabs.io

GrimoireLab:

opnfv.biterg.io

cauldron.io

Prospector:

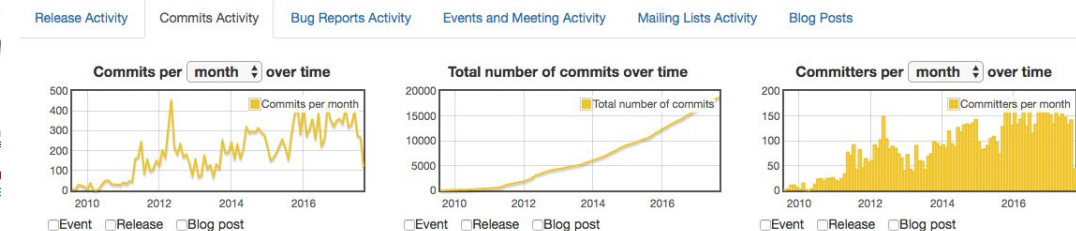
prospector.bitergia.net

Cregit:

cregit.linuxsources.org

Group	Meetings	Median RSVPs	First Meeting	Most Recent Meeting
Standard DevOps Meetup	35	2	May 20th 2017, 15:00:00.000	August 20th 2017, 15:00:00.000
Wares Online User Group	41	42	February 24th 2017, 18:30:30.000	August 23rd 2017, 01:00:00.000
Google Cloud Platform (GCP) Online Meetup	26	75.5	November 8th 2016, 19:00:00.000	August 22nd 2017, 20:00:00.000
Wares User Group - London	26	17.5	February 24th 2017, 18:30:30.000	August 24th 2017, 14:00:00.000
Trangex Kubernetes and OpenShift Meetup	21	80	April 8th 2016, 09:30:00.000	August 25th 2017, 09:00:00.000
Prague Container Meetup	20	41.5	October 20th 2016, 19:00:00.000	July 27th 2017, 18:30:00.000
Wares User Group - New York	20	7	February 24th 2017, 18:30:30.000	July 25th 2017, 19:00:00.000
South Melbourne	19	3	February 24th 2017, 18:30:30.000	June 28th 2017, 20:00:00.000
Wares User Group - Berlin	18	8	February 24th 2017, 18:30:30.000	August 24th 2017, 14:00:00.000
Google Cloud Platform User Group Singapore	17	89	June 22nd 2016, 12:30:00.000	July 4th 2017, 02:30:00.000
Kubernetes/Cloud Native Online Meetup	16	82.5	July 24th 2016, 18:45:00.000	June 20th 2017, 17:30:00.000
DevOps CT	15	14	May 27th 2016, 22:30:00.000	August 24th 2017, 22:00:00.000
Kubernetes and Cloud Native New York	13	115	March 14th 2016, 23:00:00.000	July 19th 2017, 17:30:00.000
Cloud Computing & DevOps Meetup Bangalore	12	101.5	July 9th 2016, 06:30:00.000	July 15th 2017, 06:30:00.000
Cloud Native PDX	12	30	May 5th 2016, 02:00:00.000	August 3rd 2017, 18:00:00.000
Kubernetes Los Vegas	12	7	October 7th 2016, 03:30:00.000	August 4th 2017, 03:30:00.000
Melbourne Kubernetes User Group	12	34.5	April 24th 2016, 10:00:00.000	August 15th 2017, 10:00:00.000
Cloud Native Consulting Paris	11	148	June 20th 2016, 19:00:00.000	July 5th 2017, 19:00:00.000
OpenShift and Kubernetes San Francisco	11	34	May 17th 2016, 18:00:00.000	June 14th 2017, 17:00:00.000
	11	4		

3. Activity



```
cpu_notifier_regist
if (create_hash_tabl
err = -ENOMEM
goto out;
}
entry = proc_create
if (!entry)
goto out;
proc_set_size(entry,
_hotcpu_notifier(pi
out:
cpu_notifier_register_done();
return err;
}
```

Contributors

Contributor	Count	Percentage
william lee irwin iii	58	58.00%
srivatsa s. bhat	26	26.00%
paolo ciarrocchi	5	5.00%
david howells	4	4.00%
denis v. lunev	4	4.00%
al viro	2	2.00%
dave hansen	1	1.00%

- Percentage of committers by dominant domain name: 37.10%**
Rationale: If more than 50% of COMMITTERS are from one domain (via email ID) it is dominated by one set of people. Suggested target is to have it less than 35%.
- Percentage of commits by dominant domain name: 57.27%**
Rule: If more than 50% of CODE COMMITS are from one domain (via email ID) it is dominated by one set of people. Suggested target is to have it less than 35%.



Getting Engaged with the CHAOSS Community



Mailing lists, IRC Channels and Meetings:

1. <https://chaoss.community/participate/>
2. IRC Channel: #chaoss-community on freenode
3. Wiki: <https://wiki.linuxfoundation.org/chaoss>

Participating in CHAOSS



Code

- <https://github.com/chaoss>

Meetings

- Metrics committee
 - [Weekly hangouts](#)
 - [Monthly calls](#)
- Software committee
 -

Events

- <https://wiki.linuxfoundation.org/chaoss/events>

Community starting to form...



CHAOSS

Projects News About Wiki Contact | Q

Community Health Analytics Open Source Software

LEARN MORE

Recent Posts

Come help define the metric of community growth/maturity/decline in BOF
Alex Contini | News

Learn more about Community Health Analytics Open Source Software (CHAOSS)
Alex Contini | News

Discussion: Mailing Lists, Periodic Meetings, Weekly Hangouts, IRC Channels

Code: Issues, Pull requests welcome

chaoss.community

CHAOSS

How to Get Started with CHAOSS



How to get started

- Join our [mail list](#) and introduce yourself.
- Look through the [mail list archive](#) for past discussions and self-introductions of people involved.
- Join our [monthly phone call](#) on your calendar.
- Join our [weekly hangout](#) on your calendar.

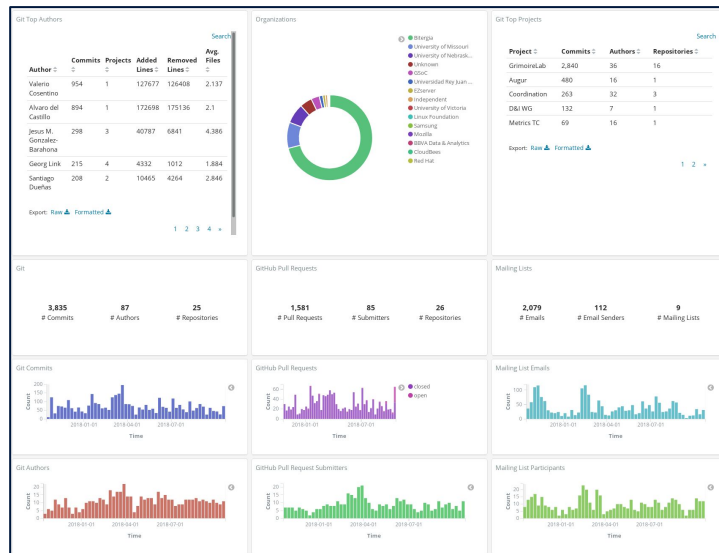
How to contribute

- Join our weekly hangouts and monthly calls.
- Share your experience with metrics.
- Let us know what you would like CHAOSS to do for you.
- Just participate in the conversation.

Project Updates



CHAOSS



Project metrics about CHAOSS available at: <https://chaoss.biterg.io>

- › Inaugural [CHAOSScon NA](#) in August was sold out, [videos of the sessions](#) were available September, and [blog post about the event](#) contributed by VMware.
- › Metrics from Growth-Maturity-Divide groups have started being adopted into Software Projects ([Augur](#), [Grimoire Lab](#) - both of which growing in functionality, user base, and new contributors :-))
- › Working Group for Diversity & Inclusion has two talks planned for OSS-EU ([Metrics that Matter for Diversity & Inclusion in Open Source](#), [Tutorial: How to Prepare a Diversity and Inclusion Report for your Community](#)) and [new mail list](#) has been set up for refining the metrics.
- › Working Group for Risk is in initial stages of organization and will incorporate License Scanning into the CHAOSS Toolkit.
- › CHAOSS community meetings being posted on [CHAOSS YouTube](#) channel
- › Participated in the 2018 GSoC - 2 students and presented *Ecosystem and Community Health* at the GSoC Googler Track at the Mentor Summit
- › CFP for [CHAOSSCon EU](#) is open until Nov 26 (will be on Feb 1, prior to FOSDEM).

Join us to extract knowledge from



CHA^OSS

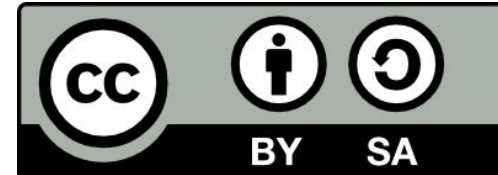
chaoss.community

CHA^OSS

License & Credits



(c) 2018 CHAOSS. Some rights reserved. This presentation is distributed under the “Attribution-ShareAlike 4.0” license, by Creative Commons, available at creativecommons.org/licenses/by-sa/4.0/



Credits:

- “Good code”, Comic by Randall Munroe, XKCD 844
License: Creative Commons Attribution-NonCommercial 2.5
xkcd.com/844/