Putting Order into CHAOSS

Metrics to Analyze Code Development

CHAOSSCon Europe - Brussels - 1st Feb. 2019

Ana Jiménez Santamaría, Daniel Izquierdo Cortázar @bitergia, https://speakerdeck.com/bitergia





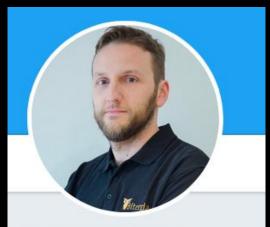
About us:



Ana Jimenez

@anajsana95

Marketing at bitergia, software development analytics 🙌 Gamer and anime lover 🙉 Self-taught illustrator 🧐 日本語の学生 🛱 España 🏧-USA 🌉



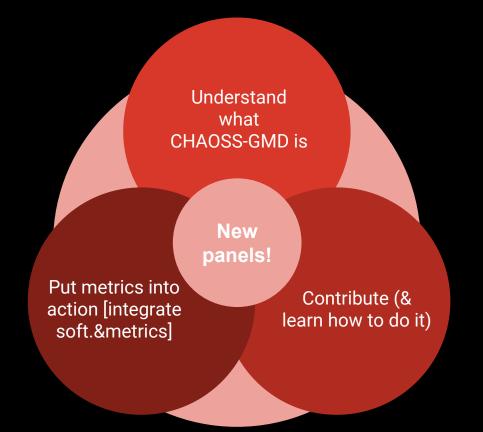
Daniel Izquierdo

@dizquierdo

@Bitergia co-founder. Open and Inner Source advocate. Love data analytics. Python developer.



Motivation







CHAOSS Growth, maturity and decline group

C.H.A.O.S.S = Community Health Analytics Open Source Software

↔ Code ① Issues 12 11	Pull requests (1) 🛛 📙 P	rojects 0 💷 Wiki	<u>네</u> Insights							
/orking Group focused on Growt	h-Maturity-Decline metr	ics and software								
117 commits	¥ 14 branches	11 7 cor	ntributors		MIT ق					
Branch: master - New pull request]		Create new file	Upload files	Find file	Clone or download -				
sgoggins Merge pull request #63 fro	m chaoss/jgbarah-patch-11			i	Latest comr	nit 8a4514a 2 days ago				
focus_areas	removed old folders from	removed old folders from my branch 3 days ag								
implementations	Rename some directories, to simplify structure 2 months ago									
metrics	Merge branch 'master' into dicortazar-gmd-open-issues 2 months ago									
use_cases	[use_cases] First try at adding metrics to a use case 23 days ago									
.gitignore	link fix test					3 months ago				
LICENSE	Initial commit 9 months									
README.md	Update README.md 3 days a									
contributing.md	read me addressing issue 35 2 months ago									
gmd_metrics.md	fix links 2 months ag									
meeting notes.md	dec 19th notes					14 days ago				



Software development projects life cycle:

Growth, maturity and decline

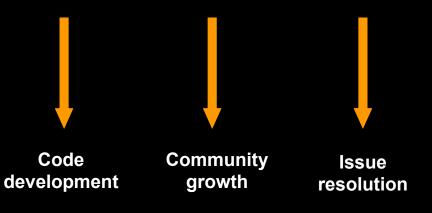




Identify project's state analyzing code development:

How can I know my community health?

Identifying my project state: Growth Maturity or Decline?





Contributing

Contributing

You can start by introducing yourself on the CHAOSS mailing list (see below) explaining your interest. Then, you can have a look at the archives of the mailing lists, at the minutes of past meetings, and at the

- issues and
- · pull requests in
- · contents of this repository.

To contribute, of course you can participate in the in online meetings, and in issues and pull requests. Currently, main lines of work are:

- Use cases. Proposal and discussion of use cases that help to understand metrics in context. You can propose your use cases, and/or contribute to refine those that have been proposed already.
- Focus areas (definitions, goals, questions). Currently, in the [definition of GMD metrics] we strucuture it in the following focus areas: Issue Resolution, Code Development, and Community Growth.

For each area of interest, we're following the goal-question-metric methodology defining questions and metrics that help to answer them.

You can contribute by proposing new goals for a focus area, or new questions for learning about those goals, or new metrics for answering those questions. Or by helping to refine goals, and questions.

See more details in the contributing file.



Contributing

We're in the process of discussing questions and metrics. Meanwhile, below is the former list of metrics, and their related questions.

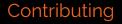
Metric	Question								
Pull Requests Merged	What is the number of code commits?								
Lines of Code Changed	What is the number of lines of code changed?								
Code Reviews	What is the number of code reviews?								
Pull Request Merge Duration	What is the duration of time between code merge request and code commit?								
Code Review Efficiency	What is the number of merged code changes/number of abandoned code change requests?								
Maintainer Response to Merge Request Duration	What is the duration of time for a maintainer to make a first response to a code merge request?								
Code Review Iteration	What is the number of iterations that occur before a merge request is accepted or declined?								
Forks	Forks are a concept in distributed version control systems like GitHub. It is a proxy for the approximate number of developers who have taken a shot at building and deploying the codebase <i>for development</i> .								
Pull Requests Open	Number of open pull requests.								
Pull Requests Closed	Number of closed pull requests.								
Pull Request Comment Duration	The difference between the timestamp of the pull request creation date and the most recent comment on the pull request.								
Pull Request Comment Diversity	Number of each people discussing each pull request.								
Pull Request Comments	Number of comments on each pull request.								



Contributing [for a newcomer]







Open Ticket & Discussion

Creation Process

Review & Merge

Open ticket at grimoirelab-sigils Have discussion if needed Panel Documentation Export panel as JSON file Export documentation as md file. Pull Request to grimoirelab-sigils The maintainer may look for improvements (There were indeed!) When ready to go, merge the panel.



Putting metrics into action:

Lines of code changed panel visualization:

Blue6												
Dashboard / Lines of Code Changed			Full screen Shar	re Edit < 🛛 Last 2 years 🗲								
Search (e.g. status:200 AND extension:PHP)	Search (e.g. status:200 AND extension:PHP)											
NOT Bos Add a filter+												
S Lines of Code Changed	Lines Added vs. Lines Removed Evolution		Repositories									
3 Organization		Lines Added		Search								
Select 👻		Lines Removed	Repository 🗘	# Commits ‡								
Project			https://github.com/cloudfoundry-incubator/kubo-ci.git	399								
Select	gi -2000000 -		https://github.com/cloudfoundry/bosh.git	364								
Apply changes Cancel changes Clear form	ਊ -400000 -		https://github.com/cloudfoundry/cloud_controller_ng.git	324								
	<u>=</u> -600000 -		https://github.com/cloudfoundry/istio.git	317								
	-8000000 2017-04-01 2017-07-01 2017-10-01 2018-01-01 2018-04-01 2018-07-01 2018-10-01		https://github.com/cloudfoundry/runtime-ci.git	301								
	Time		https://github.com/cloudfoundry/cli.git	282								
			https://github.com/cloudfoundry-incubator/cli-sapi.git	274								
Big Numbers	Commits Evolution		https://github.com/cloudfoundry/bosh-bootloader.git	267								
	200	Cloudfoundry	https://github.com/cloudfoundry-incubator/cloud_controller_ng_sapi.git	266								
	in the state of th	 cloudfoundry-incubat 	https://github.com/cloudfoundry/cf-deployment.git	244								
517 1 24099798 26337721 Authors Median Modified Files Added Lines Removed Lines	100- 100- 100- 2017/04-01 2017/07-01 2017/10-01 2016/01 2016/04 2016/04 2016/07-01 2016/10-01 Time		Export: Raw 🛦 Formatted 🛦	1 2 3 4 543 »								
Lines of Code Changed												

Lines of Code Changed

This panel focuses on Number of lines of code changed. We based the panel on Git data.

The idea behind the panel is having a look on how community contributes to the repositories by having a look at the magnitude of those changes with several metrics related with Git commits, specially modified lines and the amount of affected files. By using this panel we can identify and analyze behaviors of specific organizations, projects or repositories.

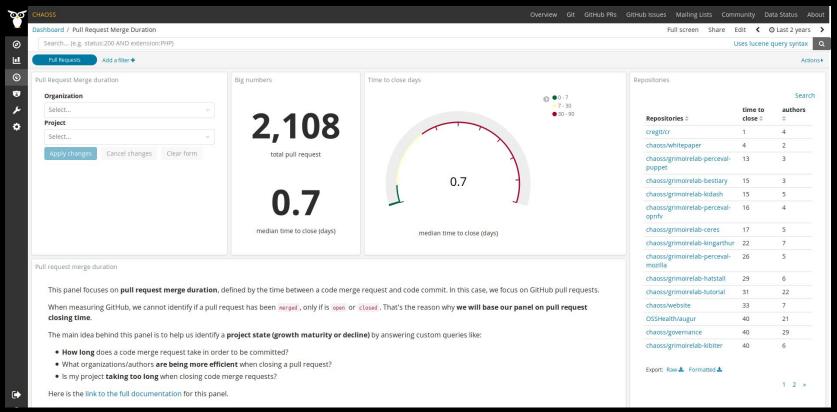
There are some filters added by default:

- NOT Merges : It ignores those commits whose number of modified files is 0 (this is, commits from merges).
- NOT Bots : It ignores those commits whose author is marked as a Bot.

Here is the link to the full documentation for this panel.



Pull Request Merge Duration panel visualization:





Putting metrics

into action:

Pull Requests Merged

Putting metrics into action:

CHAOSS		Overview Git GitHub PRs G	SitHub Issues Mailing Lists Comm	nunity Data Status	About
Dashboard / Pull Requests Merged			Full screen Share E	dit 🔇 🛛 Last 5 ye	years ゝ
*			U	lses lucene query syn	ntax Q
NOT author bot Add a filter +					Actions
GitHub Pull Requests Merged	Big Numbers	Evolution	Repositories		
Organization		40 III Pull Requests Merged		S	Search
Select	1,380	30 Trend	Repository ≑	Commits 🗘 Autho	ors ‡
Project	Commits	20	chaoss/grimoirelab-perceval	264 2	
Select			chaoss/grimoirelab-elk	206 9	
Apply changes Cancel changes Clear form			chaoss/grimoirelab-sigils	201 6	
	34		OSSHealth/augur	128 11	
	Authors	-10	chaoss/grimoirelab-sirmordred	89 7	
		2015-01-01 2016-01-01 2017-01-01 2018-01-01 2019-01-01	chaoss/grimoirelab-sortinghat	84 3	
			chaoss/wg-diversity-inclusion	70 5	
Citize Dell Deserver Manard			chaoss/website	64 7	

GitHub Pull Requests Merged

This panel focuses on merged GitHub Pull Requests. As GitHub API does not always identifies merges, we based the panel directly on Git data.

The idea behind the panel is having a look on how community performs and evolved merging Pull Requests. By using this panel we can identify trends and analyze behaviours of specific organizations, projects or repositories.

Commits: a filter is applied within visualizations to get only those that we consider as "merges". We consider a commit as a "merge" if satisfies at least one of the following conditions:

- Number of files is 0.
- Commit message starts with "Merge"

1 2 3 »

3

5

39

chaoss/grimoirelab-tutorial

Export: Raw 🛓 Formatted 🛓

chaoss/grimoirelab-manuscripts 38

bout >



Maintainer First Response

Ø												w Git Gitl	Hub PRs GitHub Issu	ies Mailing Lists		ty Data Statu	
-	Dashboard / Maintaine	er Response to Me	rge Request Dur	tion										Full screen SI	hare Edit	< O Last 5	years 🕻
Ð	Search (e.g. status	s:200 AND extensio	on:PHP)												Uses	lucene query sy	/ntax O
<u>1</u>	Add a filter 🕇																
9	Filter			Summary							Repository Data						
9	Organization															ç	Search
5	Select		÷									Ave	erage Time to First	Median Time	to First	# Pull	
,	Project										Repository 🛱	Att	ention (Days) 🕆	Response (Da	ays) 🌣	Requests 🕸	£.
1	Select		-		0.02			2 204		2 220	chaoss/grimoirela	b-tutorial 19.	824	0.42		42	
	Apply	Apply Cancel Clear changes form			0.03 Median Time to First Response (Days)	Average	2.394 Average Time to First Response (Days)		2,330 # Pull Requests	chaoss/grimoirela sortinghat	ib- 12.5	12.596	0.495		101		
			form						chaoss/grimoirela	b 7.2	03	1.795		101	_		
											chaoss/grimoirela	b-hatstall 5.5	07	0.29		58	
											chaoss/grimoirela sirmordred	b- 5.1	3	0.045		207	
	Average Time to First F	Response in Days	over Time								chaoss/wg-gmd	4.6	4	0.715		51	_
	250										chaoss/metrics	4.2	39	0.46		82	
	Maintainer Fir 200 Trend	st Response to Merge I	Request in Days	1							OSSHealth/augur	3.3	84	0.4		67	
	150										chaoss/wg-divers inclusion	ity- 2.84	4	0.67		93	
	100 50				1						chaoss/grimoirela manuscripts	b- 2.6	24	0.02		85	
	0			-	- · · · ·			-	- to		Export: Raw 📥 Fo	ormatted 🛓					÷
	-50	2014-07-01	2015-01-01	2015-07	7-01 2016-01-01	2016-07-01	2017-01-01	2017-07-01	2018-01-01	2018-07-01 2019-01	-01					173	

Help

This panel focuses on any GitHub Pull Request and measures its first time to attention.

The idea behind the panel is having a look on how community performs and evolves replying to pull requests. By using this panel we can identify trends and analyze certain behaviours of specific organizations, projects or repositories.

- Summary: This widget contains a summary of the information about the time to first response in average and in median.
- Repository Data: This table provides an overview of the same information but split by repository in GitHub.
- Average Time to First Response in Days over Time: This evolutionary chart displays the behaviour of the average metric over time.

More in depth documentation can be found at the Maintainer Merge Request Duration Page



.

What we achieved

Understand

Contribute

Put metrics into action

New panels!

Further Work

Match panels and my pet project GMD status

Produce new panels

Enter into discussions in the working group



Let's go for questions!



Ana Jiménez, Daniel Izquierdo

anajsldizquierdo@bitergia.com