

Almost There: A Study on Quasi-Contributors in Open Source Software Projects



Igor STEINMACHER
@igorsteinmacher



Gustavo PINTO
@gustavopinto

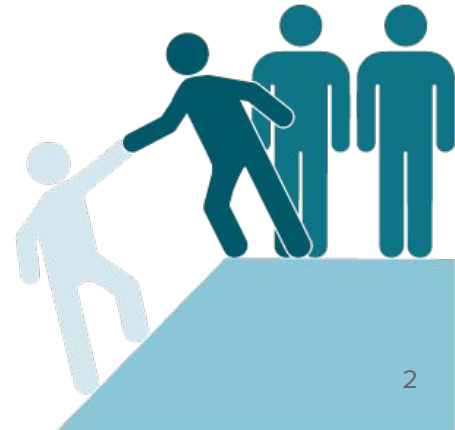


Igor WIESE
@igorwiese



Marco GEROSA
@gerosa_marco

Sustainable OSS projects



315

Views

116

CrossRef citations

0

Altmetric

Original Article

Understanding Sustained Participation in Open Source Software Projects

Yulin Fang & Derrick Neufeld

Pages 9-50 | Published online: 08 Dec 2014



ELSEVIER

Journal of Systems and Software

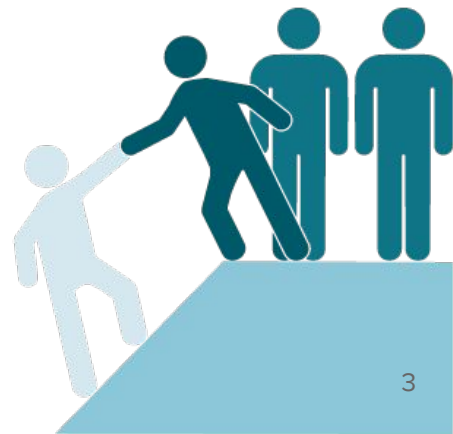
Volume 89, March 2014, Pages 128-145



Sustainability of Open Source software communities beyond a fork: How and why has the LibreOffice project evolved?

Jonas Gamalielsson  , Björn Lundell 

 [Show more](#)



315
Views
116
CrossRef citations
0
Altmetric

Original Article

Understanding Sustained Participation in Open Source Software Projects

Yulin Fang & Derrick Neufeld
Pages 9-50 | Published online: 08 Dec 2014



Journal of Systems and Software

Volume 89, March 2014, Pages 128-145



Sustainability of Open Source software communities beyond a fork: How and why has the LibreOffice project evolved?

Jonas Gamalielsson , Björn Lundell

[Show more](#)



Social Barriers Faced by Newcomers Placing Their First Contribution in Open Source Software Projects

Full Text: PDF

Authors: [Igor Stein](#)
[Tayana Co](#)
[Marco Aur](#)
[David Red](#)

An exploratory study of contribution barriers experienced by newcomers to open source software projects

Full Text: PDF

Authors: [Christoph He](#)
[Matthias Bo](#)
[Volker Gruh](#)

Published in:
• Proceeding
[CSI-SE 2014](#) Proceedi
Crowdsourcing in Soft
Pages 11-14

[Browse Conferences > Cooperative and Human Aspects...](#)

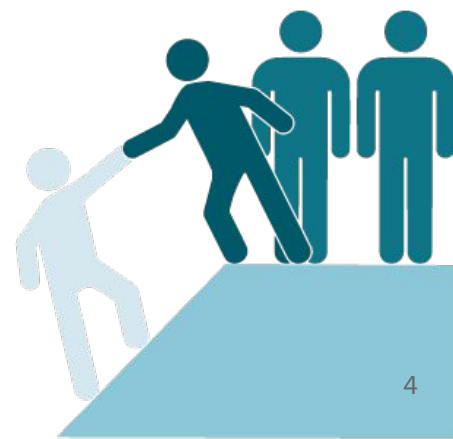
Why do newcomers abandon open source software projects?

[View Document](#)

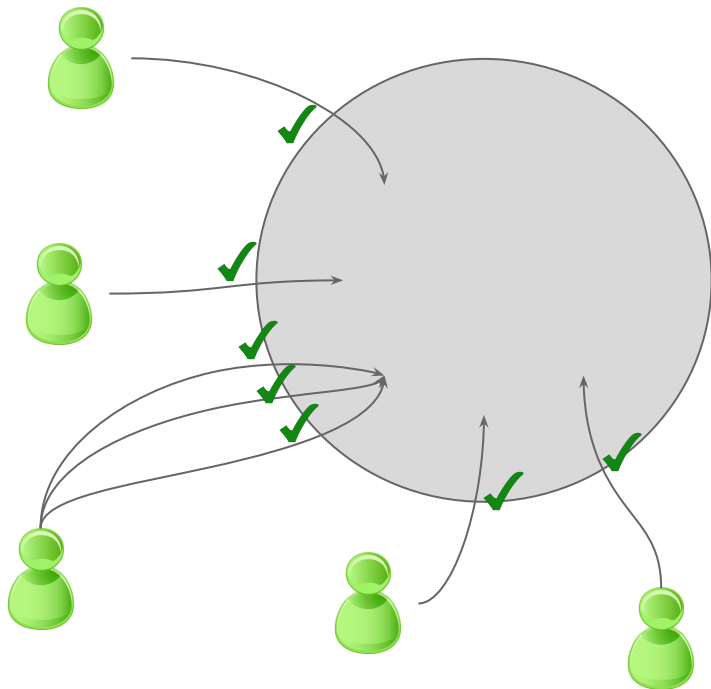
16
Paper
Citations

209
Full
Text Views

4
Author(s)
v Igor Steinmacher ; v Igor Wiese ; v Ana Paula Chaves ; v Marco Aurelio Gerosa

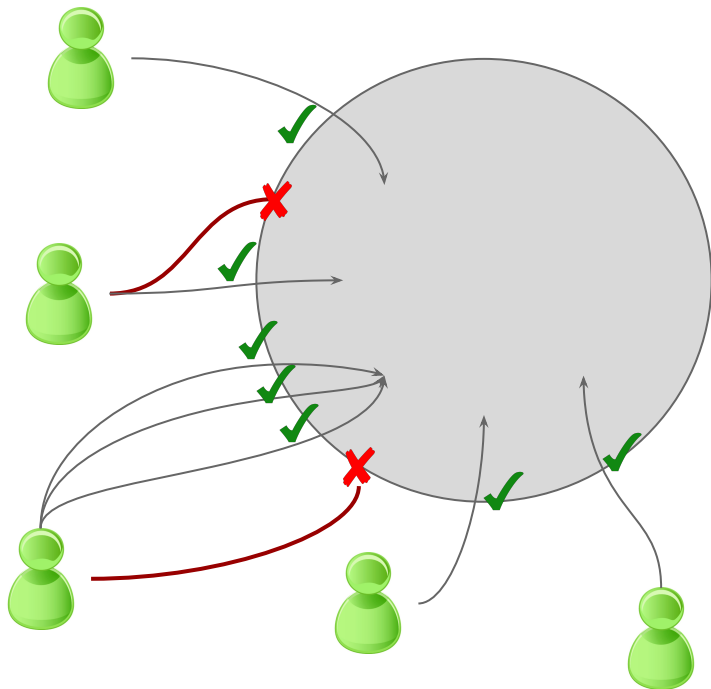


Contributing to OSS



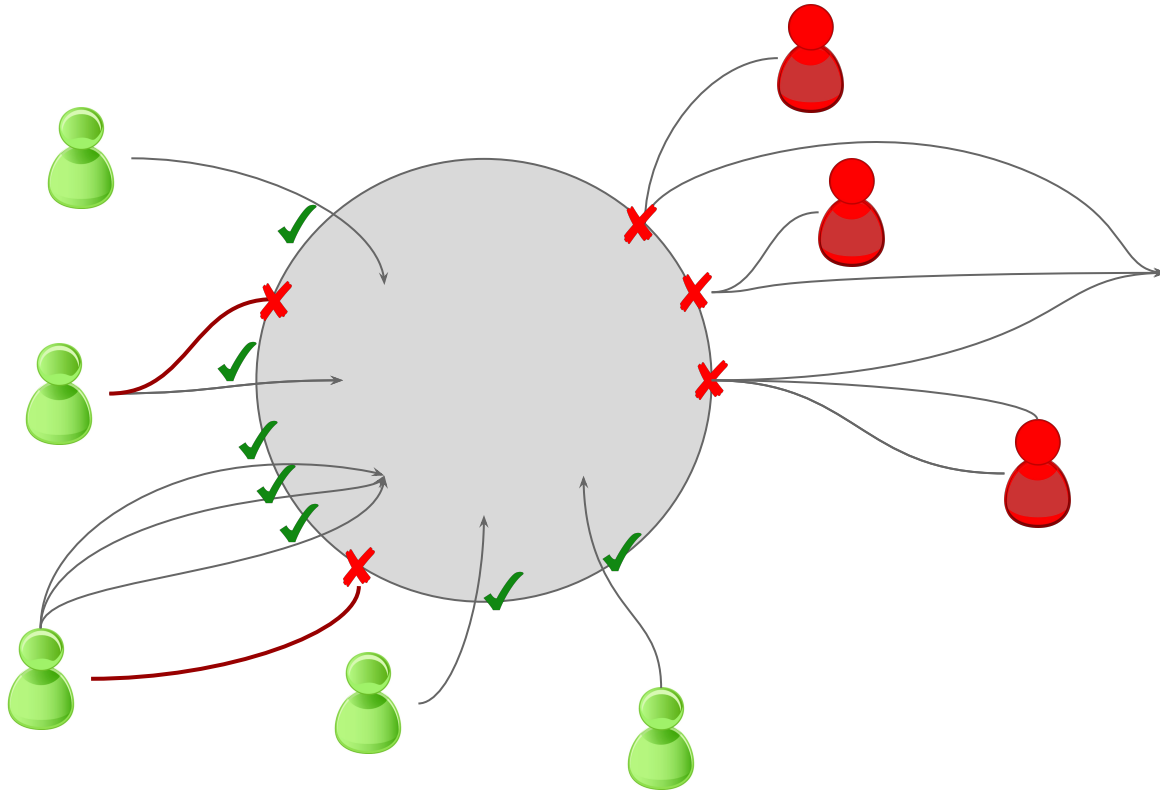
Many contributors
succeed when they
submit changes

Contributing to OSS

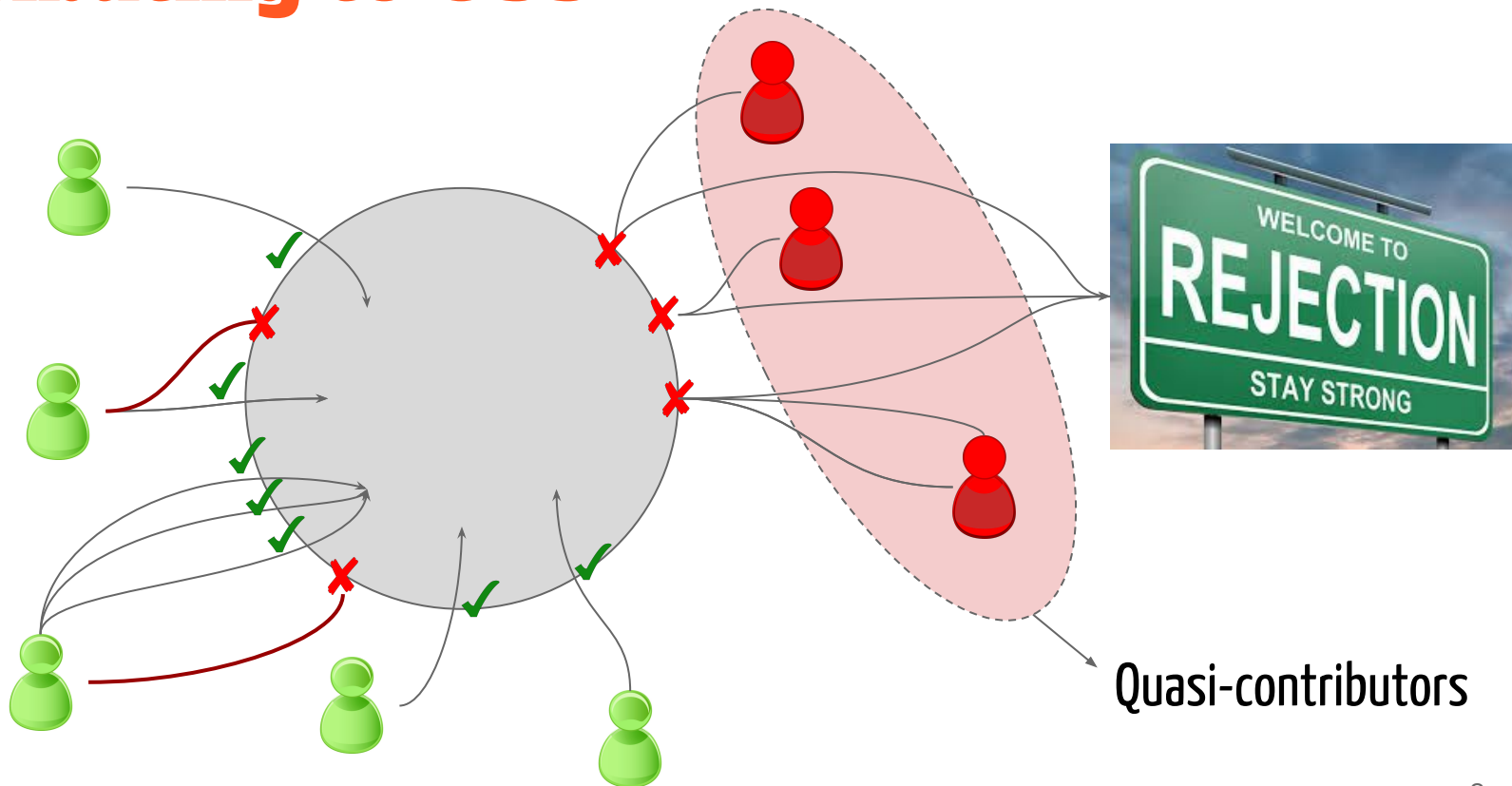


Some got some rejected patches

Contributing to OSS



Contributing to OSS



Goal

Investigate **how** and **why** quasi-contributors fail

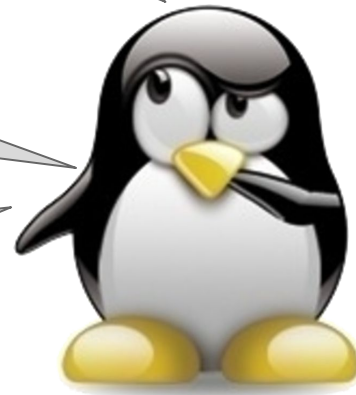


Research Questions

RQ1. How common are quasi-contributors and quasi-contributions?

RQ2. Why were the quasi-contributions not accepted?

RQ3. How do quasi-contributors perceive nonacceptance?



Sampling



Star

90,570

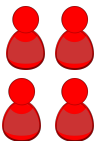


angular
bitcoin
bootstrap
caffe
d3
django
docker
flask
jenkins
joomla!
jquery
kubernetes
laravel
mongo
opencv
rails
react
redis
scikit-learn
spring
tensorflow

Mixed-Method Approach



Quasi-contributors



5,138

Quasi-contributors



234

Integrators



335 answers from
quasi-contributors



21 answers from
integrators



Manual analysis

263 PRs from
quasi-contributors

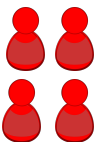
Mixed-Method Approach



Quasi-contributors



RQ1. How common are quasi-contributors and quasi-contributions?



5,138
Quasi-contributors



234
Integrators

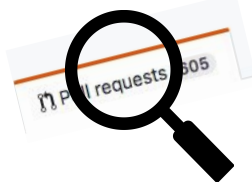


335 answers from
quasi-contributors

21 answers from
integrators

RQ2. Why were the quasi-contributions not accepted?

RQ3. How do quasi-contributors perceive nonacceptance?



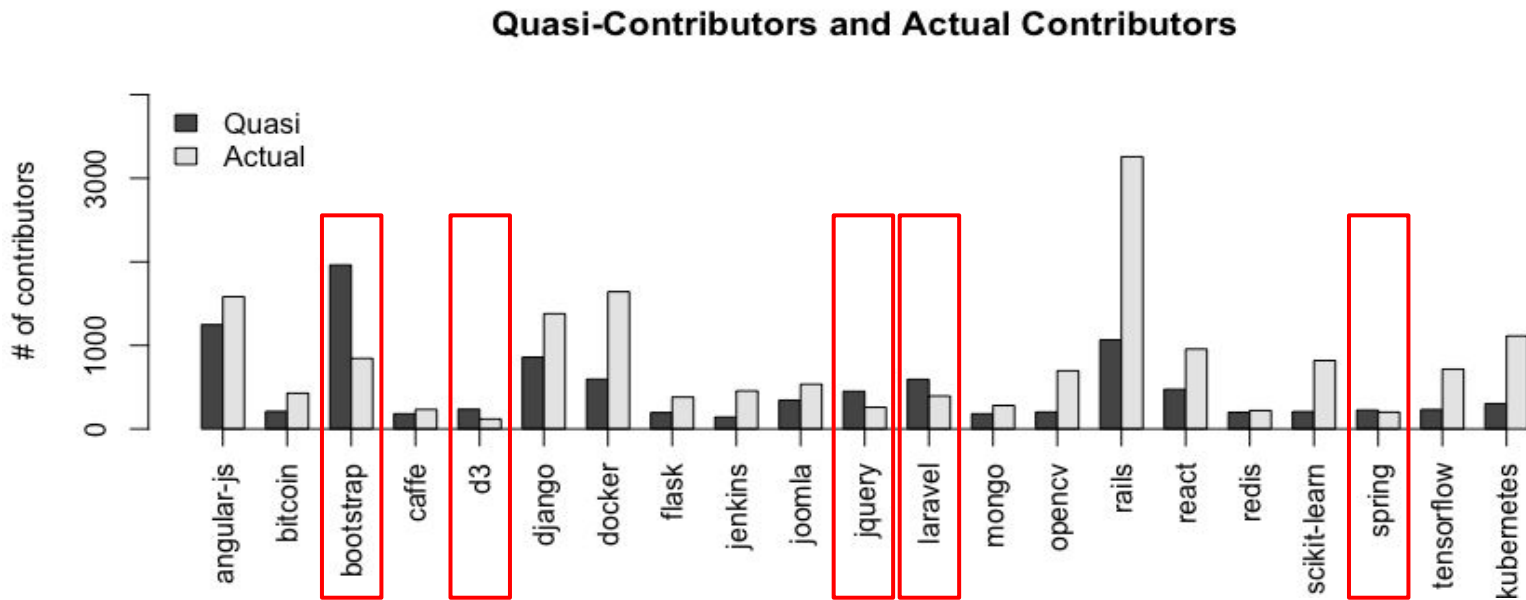
Manual analysis
263 PRs from
quasi-contributors

RQ1. How common?

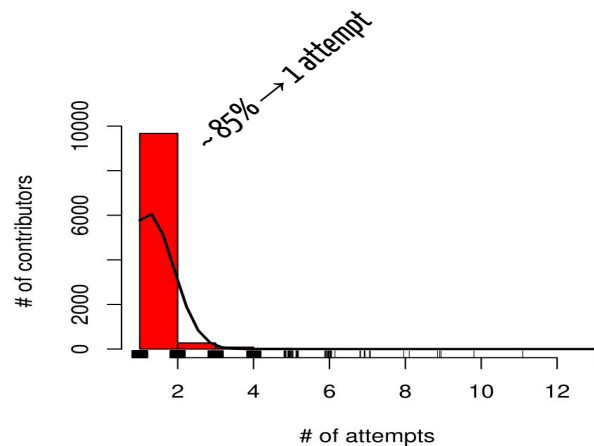
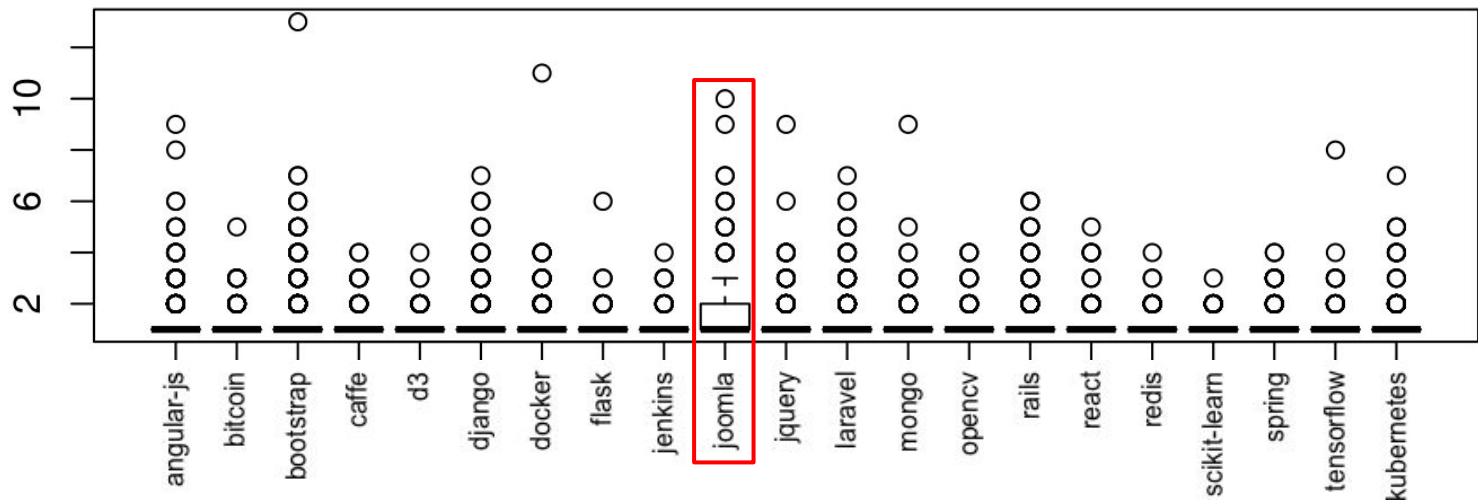
Considering all 21 projects:

→ 10,099 quasi-contributors

→ 14,623 actual contributors



RQ1. How common?



RQ2. Reasons for non-acceptance

Quasi-contributors' perspective

superseded/duplicated pull-request (52/335)

[Gousios et al., 2014; Zhou et al., 2018]



"The fix I submitted remained unmerged until someone else submitted the exact same fix ... the integrators accepted their (identical) fix and closed mine."

"Other pull-requests fixed the same issues as my pull-requests"

Gousios, G., Pinzger, M., van Deursen, A. **An Exploratory Study of the Pull-based Software Development Model.** In *ICSE 2014*

Zhou, S., Stănculescu, Ș., Leßenich, O., Xiong, Y., Wąsowski, A., Kästner, C. **Identifying Features in Forks.** In *ICSE 2018*.

RQ2. Reasons for non-acceptance

Quasi-contributors' perspective

superseded/duplicated pull-request (52/335)

[Gousios et al., 2014; Zhou et al., 2018]



"The fix I submitted remained unmerged until someone else submitted the exact same fix ... the integrators accepted their (identical) fix and closed mine."

"Other pull-requests fixed the same issues as my pull-requests"

"when you add a new feature to the project, your vision can be out of tune with the vision of the project's team, and this is natural."

"The project decided that was not a bug they wanted to provide a fix for"



mismatch between developer's and team's vision/opinion
(47/335)

RQ2. Reasons for non-acceptance

Quasi-contributors' perspective



Lack of interest from integrators
(37/335)

"I did not receive answers"

PR not needed/not relevant
(25/335)

wasn't important enough to warrant merging"

The process is too onerous, and bureaucratic."

bureaucracy
(6/335)

Quasi-contributors also offered *mea culpa*:

→ *Not an optimal solution* (20)

→ *Contributors' lack of experience/commitment* (24)

RQ2. Reasons for non-acceptance



KEEP
QUIET
BECAUSE
YOU ARE
IRRELEVANT



From integrators perspective:

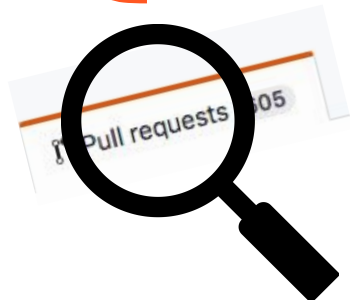
- *PR not needed/not relevant (10/21)*
- *guidelines not followed (9/21)*

"Trivial PRs that are more trouble than they are worth"

"proposed changes do not have real value"

"Does not adhere to internal style or design guidance"

RQ2. Reasons for non-acceptance



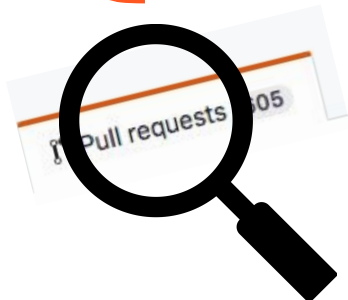
Manual analysis
263 PRs from
quasi-contributors

bbonev commented on Jun 18, 2013

I now see that similar change is already proposed by others. Sorry for the noise

Superseded/duplicated
(32/263)

RQ2. Reasons for non-acceptance



Manual analysis
263 PRs from
quasi-contributors

bbonev commented on Jun 18, 2013

Superseded/duplicated
(32/263)

timgraham commented on Feb 23, 2016

**Lack of experience/
commitment** (25/263)

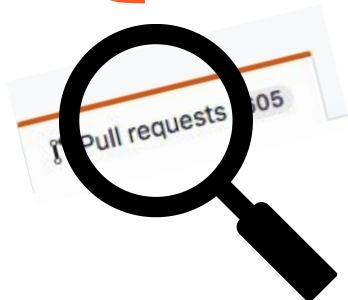
It looks to me like the test case isn't sufficient to exercise all the new code that's added (in particular, there aren't any tests for the error cases). I'm not sure this is the best solution as it seems the `to_python()` and `clean()` methods are nearly duplicated.

timgraham commented on May 7, 2016

Member

Closing due to lack of follow up. Please send a new PR if you return to it, thanks!

RQ2. Reasons for non-acceptance



Manual analysis
263 PRs from
quasi-contributors

bbonev commented on Jun 18, 2013

I now see that similar change is already proposed by others. Sorry
timgraham commented on Feb 23, 2016

Superseded/duplicated
(32/263)

**Lack of experience/
commitment** (25/263)

thaJeztah commented on Jul 31, 2015

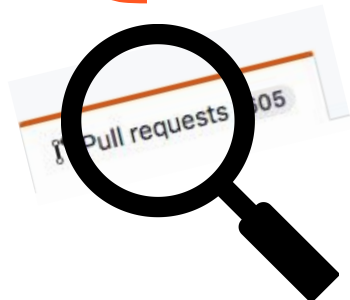
@ymzong could you test if this is still an issue? We now have new APT a
script has been changed. See <http://blog.docker.com/2015/07/new-apt-and-yum-repos/>

PR not needed/not relevant
(22/263)

ymzong commented on Jul 31, 2015

@thaJeztah Thanks for the follow-up! With the new package, the issue has been mitigated. 🐱
I'm closing the issue :)

RQ2. Reasons for non-acceptance



Manual analysis
263 PRs from
quasi-contributors

bbonev commented on Jun 18, 2013

I now see that similar change is already in the code. Sorry
timgraham commented on Feb 23, 2016

It looks to me like the change is already in the code.
@thaJeztah commented on Feb 23, 2016
there aren't any other changes in the best solution.
to_pwn commented on Feb 23, 2016
@ymlong commented on Feb 23, 2016
script? We now have new APT and yum repos.
log.docker.com/2015/07/new-apt-and-yum-repos/

timgraham commented on Jul 7, 2016

ymzong commented on Jul 31, 2015
Closing due to lack of follow up. Please send a new PR if you return to it, thanks!

@thaJeztah Thanks for the follow-up! With the new package, the issue has been mitigated. 🐱

I'm closing the issue :)

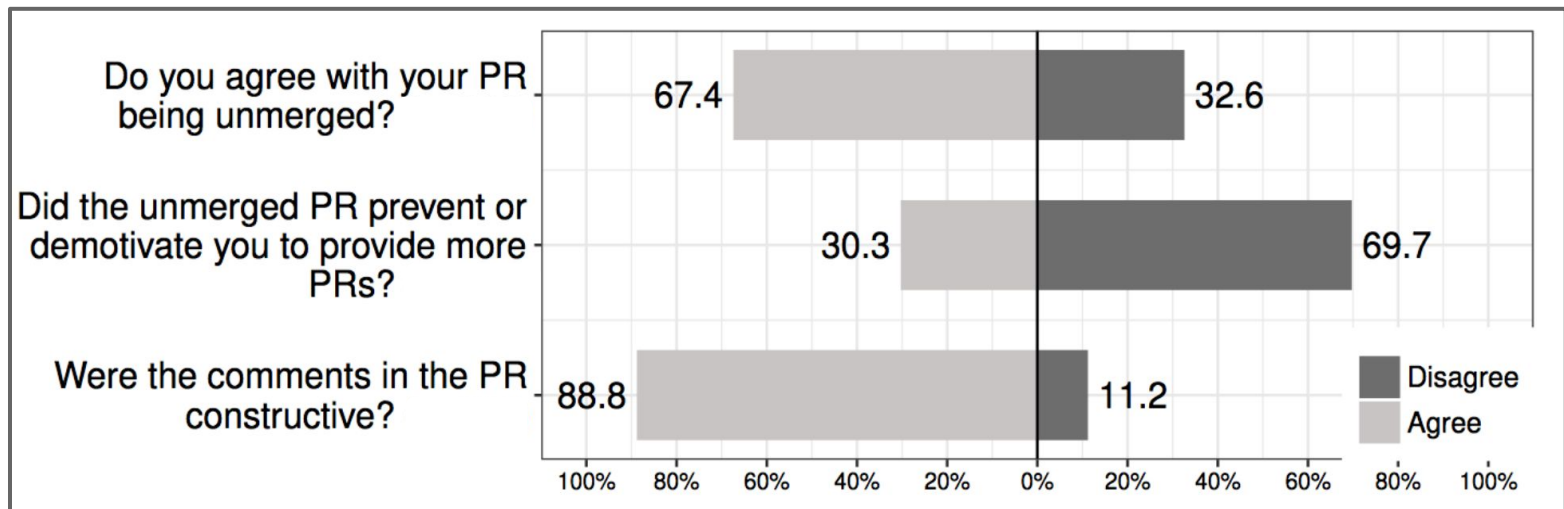
Superseded/duplicated
(32/263)

**Lack of experience/
commitment** (25/263)

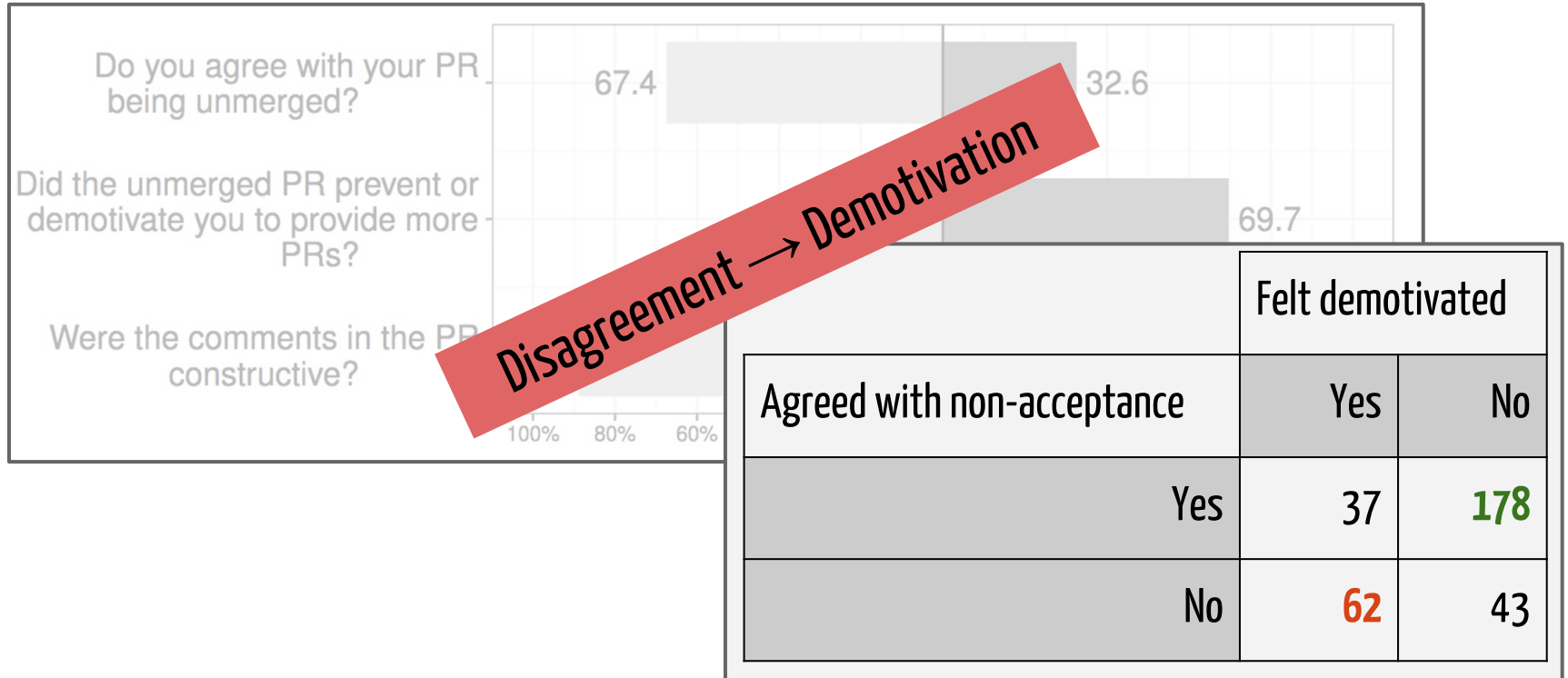
PR not needed/not relevant
(22/263)

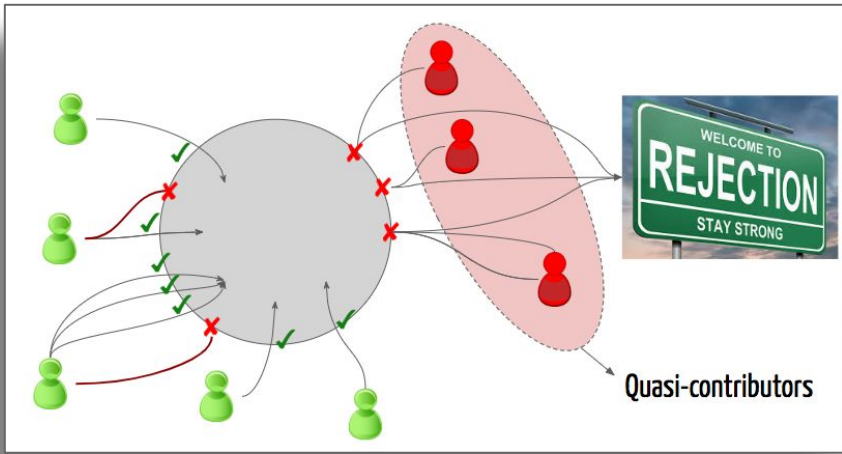
**Manual analysis results
reverberate the quasi-contributors
answers**

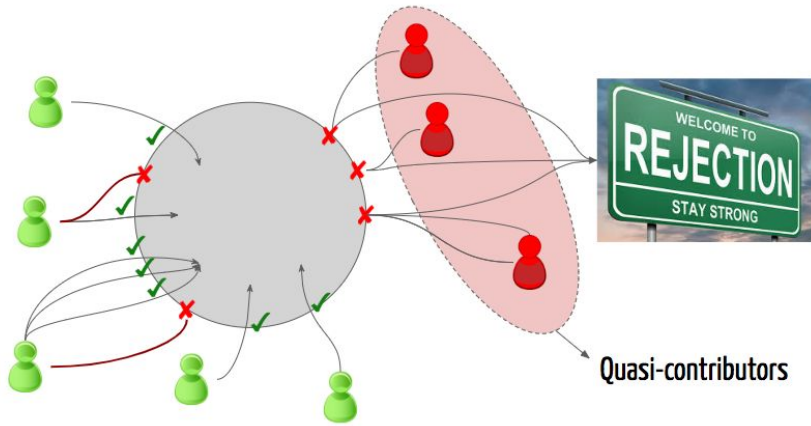
RQ3. Quasi-contributors' perception



RQ3. Quasi-contributors' perception



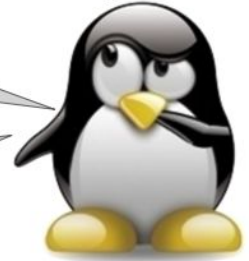


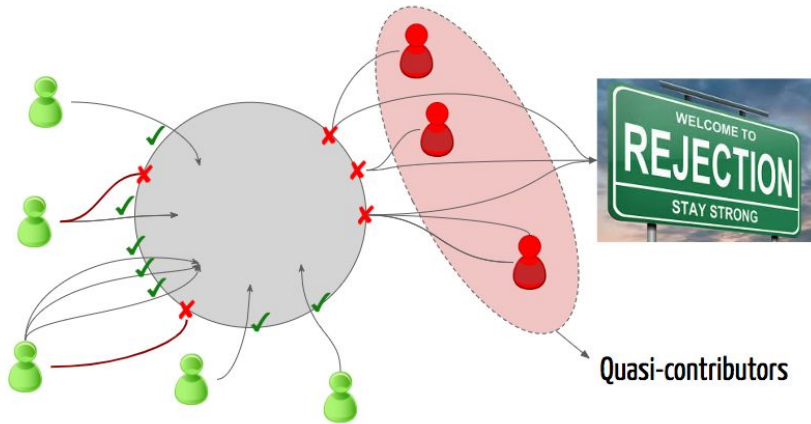


RQ1. How common are quasi-contributors and quasi-contributions?

RQ2. Why were the quasi-contributions not accepted?

RQ3. How do quasi-contributors perceive nonacceptance?

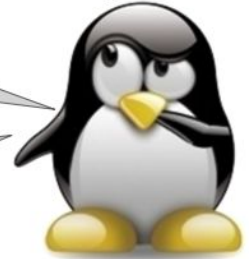




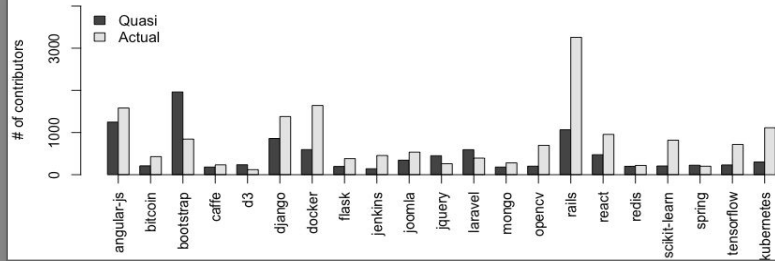
RQ1. How common are quasi-contributors and quasi-contributions?

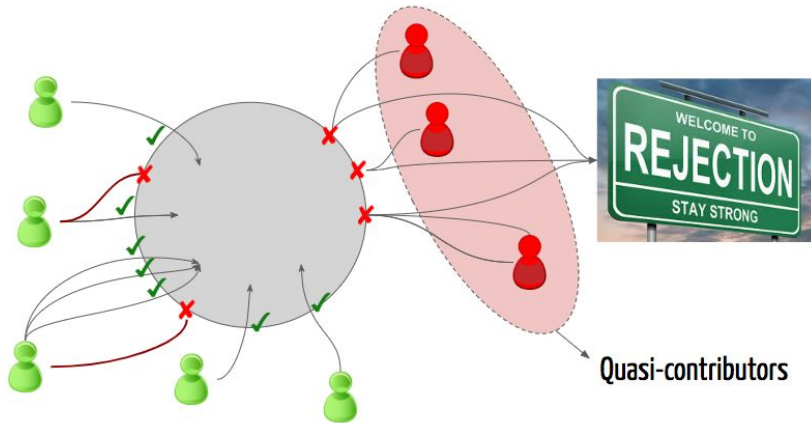
RQ2. Why were the quasi-contributions not accepted?

RQ3. How do quasi-contributors perceive nonacceptance?



Quasi-Contributors and Actual Contributors





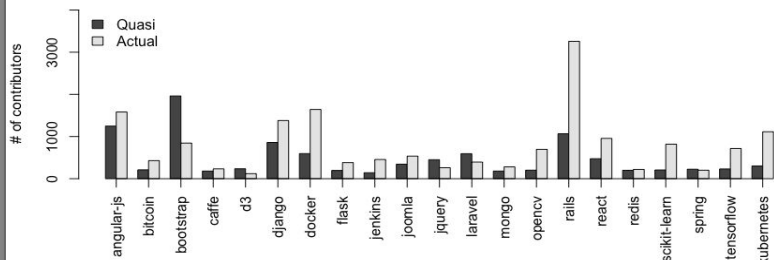
RQ1. How common are quasi-contributors and quasi-contributions?

RQ2. Why were the quasi-contributions not accepted?

RQ3. How do quasi-contributors perceive nonacceptance?



Quasi-Contributors and Actual Contributors



superseded/duplicated pull-request
(52/335)

"The fix I submitted remained unmerged until someone else submitted the exact same fix ... the integrators accepted their (identical) fix and closed mine."

"when you add a new feature to the project, your vision can be out of tune with the vision of the project's team, and this is natural."

"The project decided that was not a bug they wanted to provide a fix for"



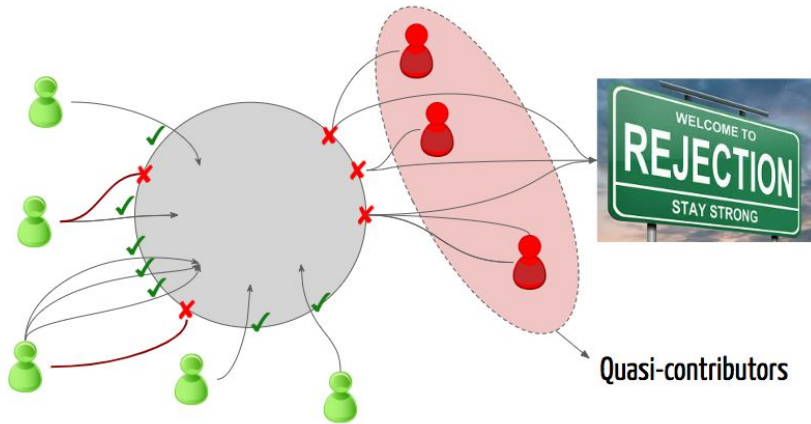
Superseded/duplicated
(32/263)

Lack of experience/
commitment (25/263)

PR not needed/not relevant
(22/263)

31, 2015

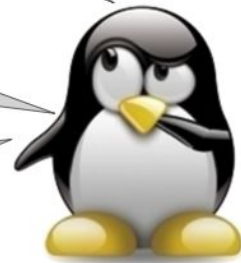
is still an issue? We now have new APT a
<http://blog.docker.com/2015/07/new-apt-and-yum-repos/>



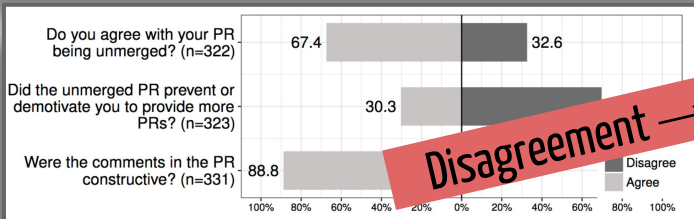
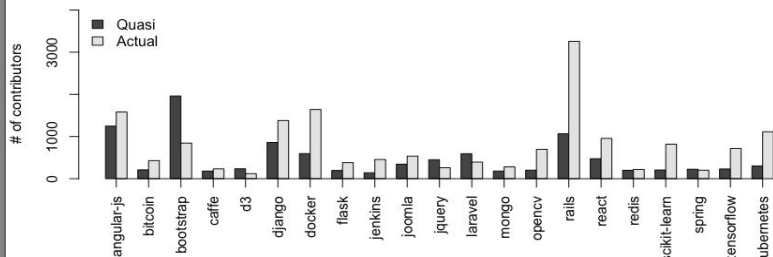
RQ1. How common are quasi-contributors and quasi-contributions?

RQ2. Why were the quasi-contributions not accepted?

RQ3. How do quasi-contributors perceive nonacceptance?



Quasi-Contributors and Actual Contributors



superseded/duplicated pull-request (52/335)

"The fix I submitted remained unmerged until someone else submitted the exact same fix ... the integrators accepted their (identical) fix and closed mine."

"when you add a new feature to the project, your vision can be out of tune with the vision of the project's team, and this is natural."

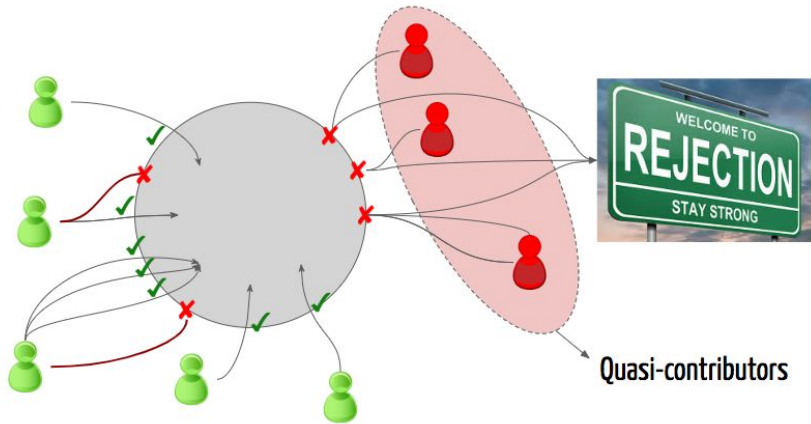
"The project decided that was not a bug they wanted to provide a fix for"



Superseded/duplicated (32/263)

Lack of experience/commitment (25/263)

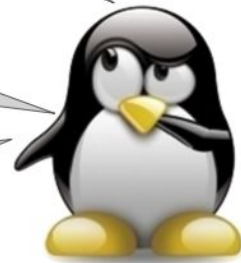
PR not needed/not relevant (22/263)



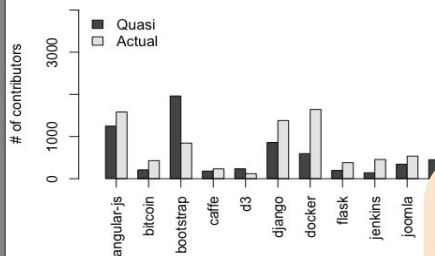
RQ1. How common are quasi-contributors and quasi-contributions?

RQ2. Why were the quasi-contributions not accepted?

RQ3. How do quasi-contributors perceive nonacceptance?



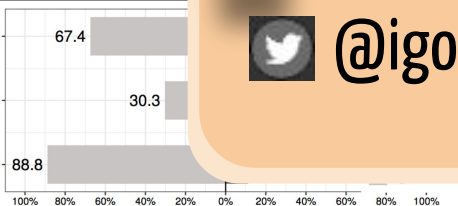
Quasi-Contributors and Actual Contributors



Do you agree with your PR being unmerged? (n=322)

Did the unmerged PR prevent or demotivate you to provide more PRs? (n=323)

Were the comments in the PR constructive? (n=331)



igor.steinmacher@nau.edu



@igorsteinmacher



superseded/duplicated pull-request (52/335)

"The fix I submitted remained unmerged until someone else submitted the exact same fix ... the integrators accepted their (identical) fix and closed mine."

Superseded/duplicated (32/263)

Lack of experience/commitment (25/263)

PR not needed/not relevant (22/263)

Conclusions

- Quasi-contributors are rather common (70% of the actual contributors)
- Different reasons lead to non-acceptance
 - Duplicated/superseded PRs
 - Mismatch between developer's and team's vision
 - Developers' fault
- Non-acceptance may incur in demotivation ($\frac{1}{3}$ felt demotivated)
- Results can be useful to OSS maintainers, newcomers, and Educators

RQ1. How common?

Projects	# Comments		# Files		# LoC Add		# LoC Del	
	ES	PV	ES	PV	ES	PV	ES	PV
angular	-0.42	0.00	0.49	0.00	0.48	0.00	0.49	0.00
bitcoin	0.11	0.00	0.26	0.00	0.14	0.00	0.26	0.00
bootstrap	-0.41	0.00	0.35	0.00	0.29	0.00	0.34	0.00
caffe	0.16	0.00	0.01	0.72	-0.01	0.39	0.13	0.12
d3	-0.35	0.00	0.01	0.38	0.01	0.39	0.01	0.12
django	-0.38	0.00	0.26	0.00	0.13	0.00	0.35	0.00
docker	0.00	0.87	0.22	0.00	0.19	0.00	0.31	0.00
flask	-0.38	0.00	0.00	0.98	-0.08	0.06	0.00	0.00
jenkins	0.00	0.00	0.26	0.00	0.21	0.00	0.32	0.00
joomla!	0.03	0.21	0.38	0.00	0.23	0.00	0.39	0.00
jquery	-0.38	0.00	0.53	0.00	0.50	0.00	0.56	0.00
kubernetes	0.00	0.00	0.22	0.00	0.22	0.00	0.33	0.00
laravel	-0.36	0.00	0.02	0.35	0.00	0.91	0.08	0.00
mongo	-0.77	0.00	0.03	0.54	-0.00	0.96	0.06	0.36
opencv	-0.24	0.00	0.18	0.00	0.17	0.00	0.31	0.00
rails	-0.28	0.00	0.13	0.00	0.04	0.05	0.27	0.00
react	-0.30	0.00	0.39	0.00	0.32	0.00	0.42	0.00
redis	0.20	0.00	0.00	0.92	0.00	0.93	0.09	0.04
scikit-learn	-0.07	0.08	0.12	0.00	-0.07	0.09	0.13	0.00
spring	-0.20	0.00	0.29	0.00	0.25	0.00	0.36	0.00
tensorflow	-0.26	0.00	0.28	0.00	0.18	0.00	0.30	0.00

RQ1. How common?

Pull-requests by quasi-contributors are more discussed

→ in line with Tsay et al. (2014) - more discussed
pull-requests are less likely to be accepted

Projects	# Comments		# Files		# LoC Add		# LoC Del	
	ES	PV	ES	PV	ES	PV	ES	PV
angular	-0.42	0.00	0.49	0.00	0.48	0.00	0.49	0.00
bitcoin	0.11	0.00	0.26	0.00	0.14	0.00	0.26	0.00
bootstrap	-0.41	0.00	0.35	0.00	0.29	0.00	0.34	0.00
caffe	0.16	0.00	0.01	0.72	-0.01	0.39	0.13	0.12
d3	-0.35	0.00	0.01	0.38	0.01	0.39	0.01	0.12
django	-0.38	0.00	0.26	0.00	0.13	0.00	0.35	0.00
docker	0.00	0.87	0.22	0.00	0.19	0.00	0.31	0.00
flask	-0.38	0.00	0.00	0.98	-0.08	0.06	0.00	0.00
jenkins	0.00	0.00	0.26	0.00	0.21	0.00	0.32	0.00
joomla!	0.03	0.21	0.38	0.00	0.23	0.00	0.39	0.00
jquery	-0.38	0.00	0.53	0.00	0.50	0.00	0.56	0.00
kubernetes	0.00	0.00	0.22	0.00	0.22	0.00	0.33	0.00
laravel	-0.36	0.00	0.02	0.35	0.00	0.91	0.08	0.00
mongo	-0.77	0.00	0.03	0.54	-0.00	0.96	0.06	0.36
opencv	-0.24	0.00	0.18	0.00	0.17	0.00	0.31	0.00
rails	-0.28	0.00	0.13	0.00	0.04	0.05	0.27	0.00
react	-0.30	0.00	0.39	0.00	0.32	0.00	0.42	0.00
redis	0.20	0.00	0.00	0.92	0.00	0.93	0.09	0.04
scikit-learn	-0.07	0.08	0.12	0.00	-0.07	0.09	0.13	0.00
spring	-0.20	0.00	0.29	0.00	0.25	0.00	0.36	0.00
tensorflow	-0.26	0.00	0.28	0.00	0.18	0.00	0.30	0.00

Tsay, J.; Dabbish M L.; Herbsleb, J. Influence of social and technical factors for evaluating contribution in GitHub. In: ICSE 2014. 356–366.

RQ1. How common?

Pull-requests by quasi-contributors are more discussed

→ in line with Tsay et al. (2014) - more discussed
pull-requests are less likely to be accepted

Quasi-contributors' pull-requests tend to be smaller

Projects	# Comments		# Files		# LoC Add		# LoC Del	
	ES	PV	ES	PV	ES	PV	ES	PV
angular	-0.42	0.00	0.49	0.00	0.48	0.00	0.49	0.00
bitcoin	0.11	0.00	0.26	0.00	0.14	0.00	0.26	0.00
bootstrap	-0.41	0.00	0.35	0.00	0.29	0.00	0.34	0.00
caffe	0.16	0.00	0.01	0.72	-0.01	0.39	0.13	0.12
d3	-0.35	0.00	0.01	0.38	0.01	0.39	0.01	0.12
django	-0.38	0.00	0.26	0.00	0.13	0.00	0.35	0.00
docker	0.00	0.87	0.22	0.00	0.19	0.00	0.31	0.00
flask	-0.38	0.00	0.00	0.98	-0.08	0.06	0.00	0.00
jenkins	0.00	0.00	0.26	0.00	0.21	0.00	0.32	0.00
joomla!	0.03	0.21	0.38	0.00	0.23	0.00	0.39	0.00
jquery	-0.38	0.00	0.53	0.00	0.50	0.00	0.56	0.00
kubernetes	0.00	0.00	0.22	0.00	0.22	0.00	0.33	0.00
laravel	-0.36	0.00	0.02	0.35	0.00	0.91	0.08	0.00
mongo	-0.77	0.00	0.03	0.54	-0.00	0.96	0.06	0.36
opencv	-0.24	0.00	0.18	0.00	0.17	0.00	0.31	0.00
rails	-0.28	0.00	0.13	0.00	0.04	0.05	0.27	0.00
react	-0.30	0.00	0.39	0.00	0.32	0.00	0.42	0.00
redis	0.20	0.00	0.00	0.92	0.00	0.93	0.09	0.04
scikit-learn	-0.07	0.08	0.12	0.00	-0.07	0.09	0.13	0.00
spring	-0.20	0.00	0.29	0.00	0.25	0.00	0.36	0.00
tensorflow	-0.26	0.00	0.28	0.00	0.18	0.00	0.30	0.00

Tsay, J.; Dabbish M L.; Herbsleb, J. Influence of social and technical factors for evaluating contribution in GitHub. In: ICSE 2014. 356–366.