

StackOverflow and GitHub: Associations Between Software Development and Crowdsourced Knowledge

Bogdan Vasilescu
TU Eindhoven
@b_vasilescu

Vladimir Filkov
UC Davis

Alexander Serebrenik
TU Eindhoven
@aserebrenik

Standing on the shoulders of others

Developers:

- reuse components and libraries
- forage on the Web for information



Standing on the shoulders of others

Developers:

- reuse components and libraries
- forage on the Web for information



Standing on the shoulders of others

Developers:

- reuse components and libraries
- forage on the Web for information



How do I edit an incorrect commit message in Git?



I stupidly did a Git commit while half asleep, and wrote the totally wrong thing in the commit message.

3554

How do I change the commit message?



I have not yet pushed the commit to anyone.

1020

[git](#) [version-control](#) [commit](#) [rebase](#) [git-rewrite-history](#)

[share](#) | [edit](#) | [flag](#)

edited Apr 25 at 20:04



Roberto Tyley

2,296 • 1 • 14 • 33

asked Oct 7 '08 at 15:44



Laurie Young

24.5k • 7 • 26 • 41

532

For those somewhat new to git: Laurie's point about having not yet pushed is important. Like rebasing, this is changing the history. If someone has cloned/pulled from your repo between the original and

20 Answers

active

oldest

[votes](#)



`git commit --amend -m "New commit message"`

3858

Used to amend the tip of the current branch. Prepare the tree object you would want to replace the latest commit as usual (this includes the usual `-i/-o` and explicit paths), and the commit log editor is seeded with the commit message from the tip of the current branch. The commit you create replaces the current tip -- if it was a merge, it will have the parents of the current tip as parents -- so the current top commit is discarded.

+50

It is a rough equivalent for:

```
$ git reset --soft HEAD^  
$ ... do something else to come up with the right tree ...  
$ git commit -c ORIG_HEAD
```

but can be used to amend a merge commit.

[share](#) | [edit](#) | [flag](#)

edited Mar 17 '12 at 1:25



Alex Lockwood

22.7k • 9 • 55 • 87

answered Oct 7 '08 at 15:50



EfForEffort

42.6k • 3 • 15 • 30



`git commit --amend -m "your new message"`

1048

[share](#) | [edit](#) | [flag](#)

edited Jan 11 '11 at 10:33



dbr

55.3k • 23 • 147 • 228

answered Feb 8 '10 at 4:26



lfx_cool

10.5k • 1 • 5 • 2

tagged

[git](#) × 30882

[version-control](#) × 8369

[commit](#) × 1198

[rebase](#) × 436

[git-rewrite-history](#) × 29

asked 4 years ago

viewed 504432 times

active today

Linked

0 [How do I edit an incorrect commit message in git](#)

0 [changing git comment of a pushed commit](#)

0 [changing the commit message in git](#)

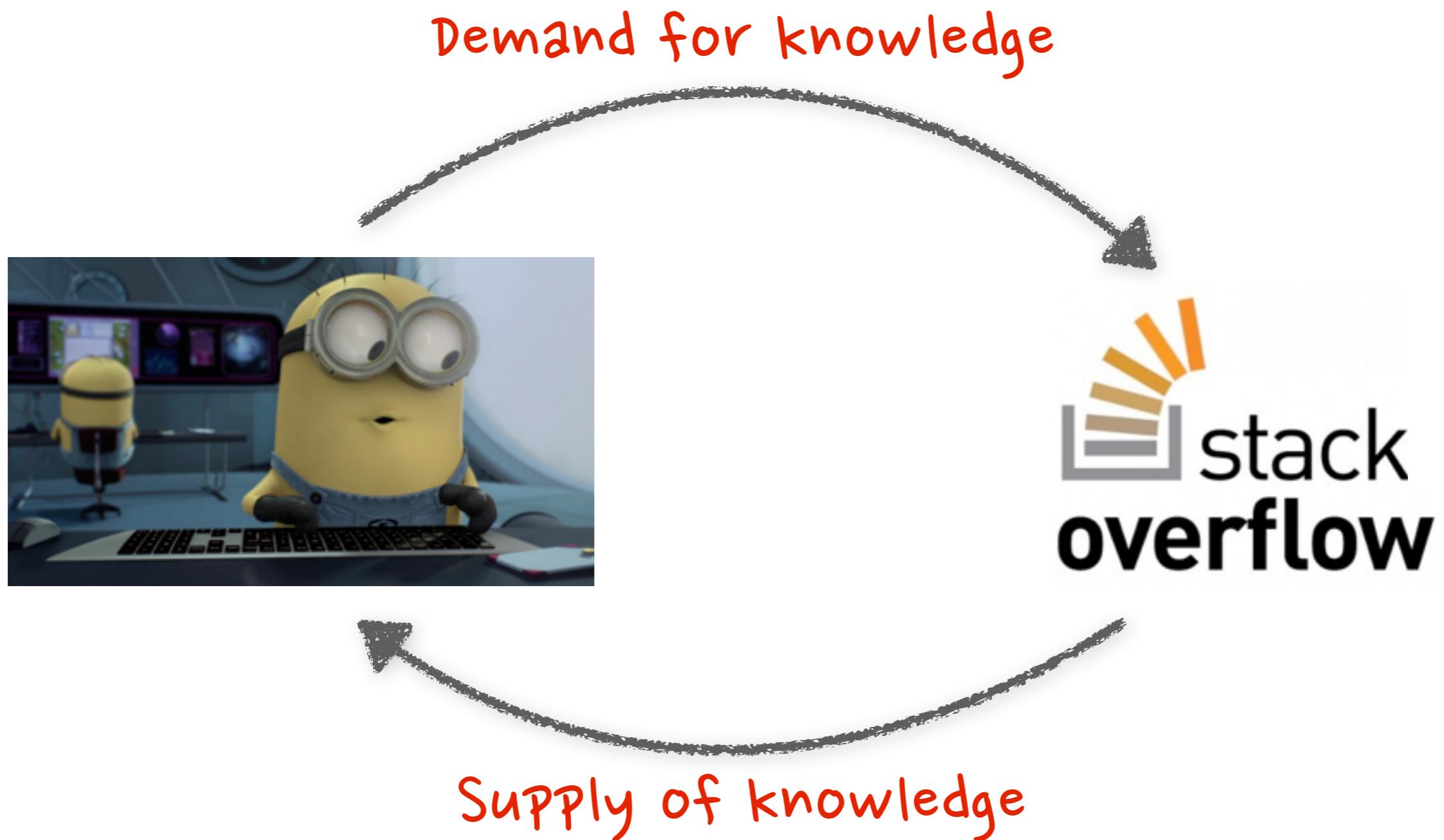
-1 [git: How to change the commit message of an already committed file, non interactively?](#)

1 [Is there any way to change commit message for a specific commit in git?](#)

3 [Change the wrong commit message in Git](#)

0 [Making a Class](#)

Writing code vs. seeking and sharing knowledge



Is participation in SO related to productivity of developers?

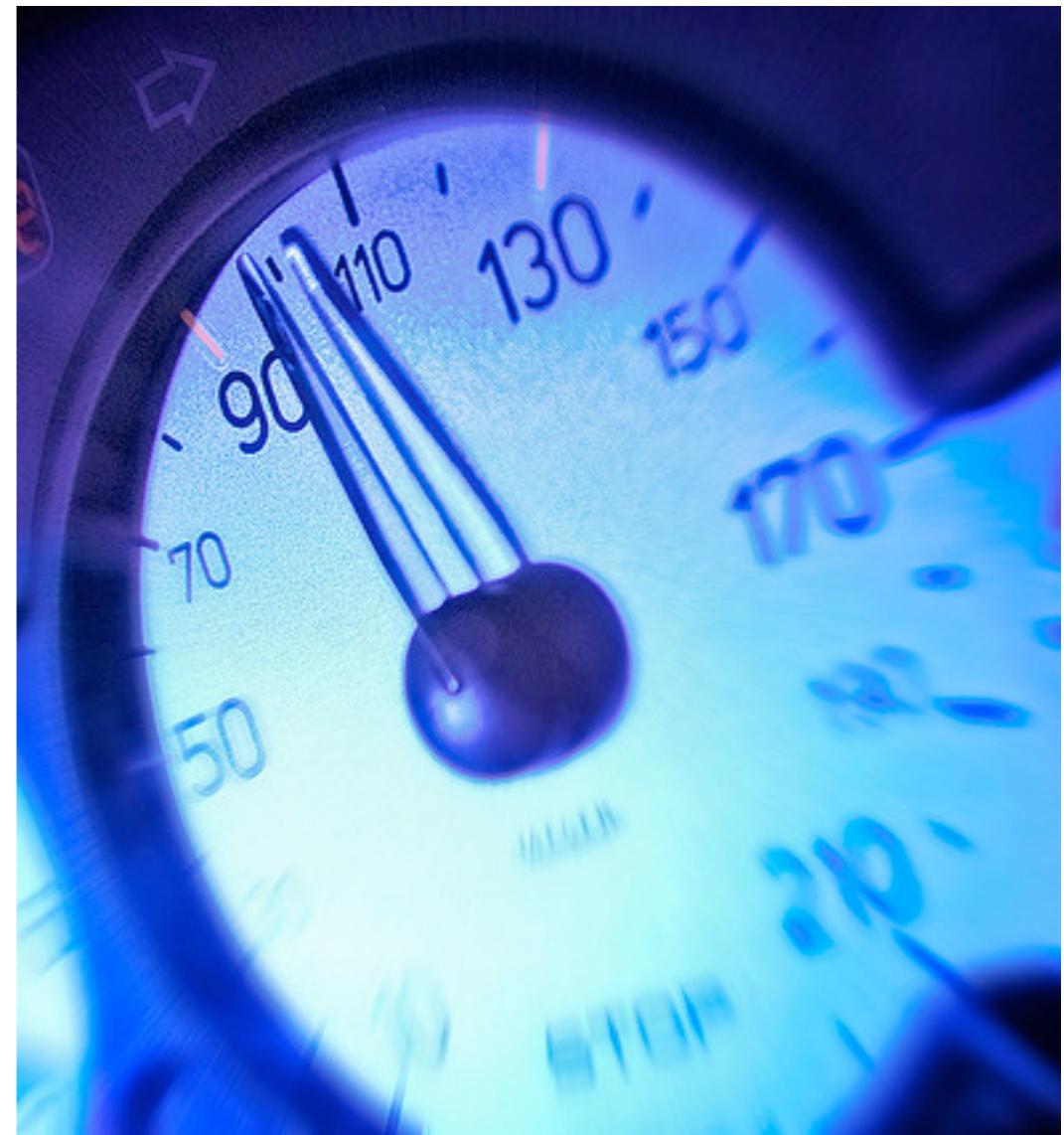
Is participation in SO related to productivity of developers?

Beneficial:

- good technical solutions
- fast answers (median 11 mins)

[Parnin et al. “*Crowd documentation: Exploring the coverage and the dynamics of API discussions on Stack Overflow*,” Georgia Institute of Technology, Tech. Rep., 2012]

[Mamykina et al. “*Design lessons from the fastest Q&A site in the west*,” in CHI. ACM, 2011, pp. 2857–2866]



Is participation in SO related to productivity of developers?

Detrimental:

- **competes for time**

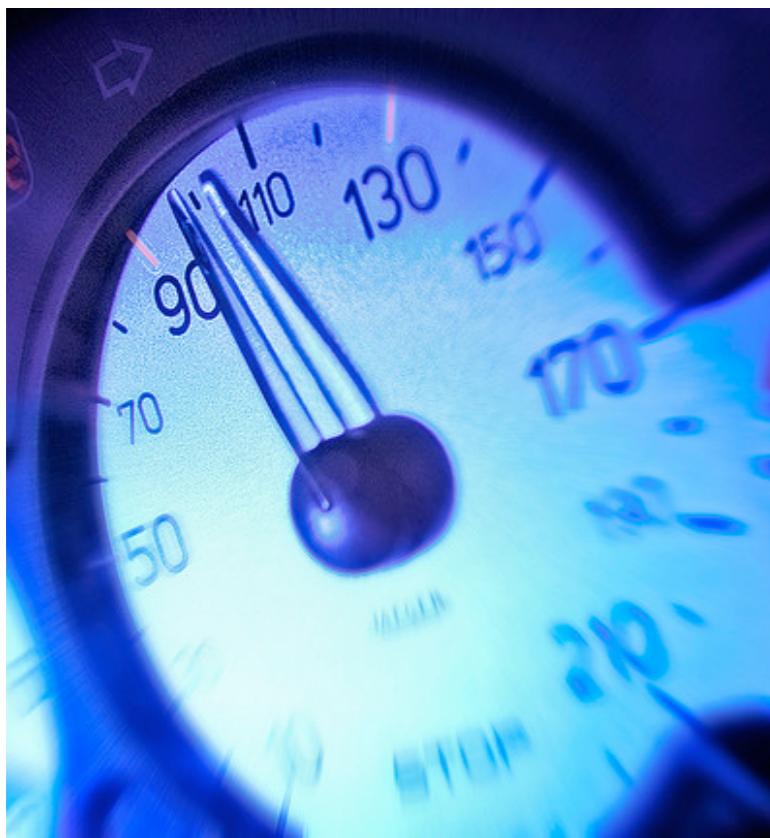
[Storey et al. “*The impact of social media on software engineering practices and tools*,” FoSER. ACM, 2010, pp. 359–364]
- **gamified, thus addictive**

[Deterding, “*Gamification: designing for motivation*,” *Interactions*, vol. 19, no. 4, pp. 14–17, 2012]
- **context switches are expensive**

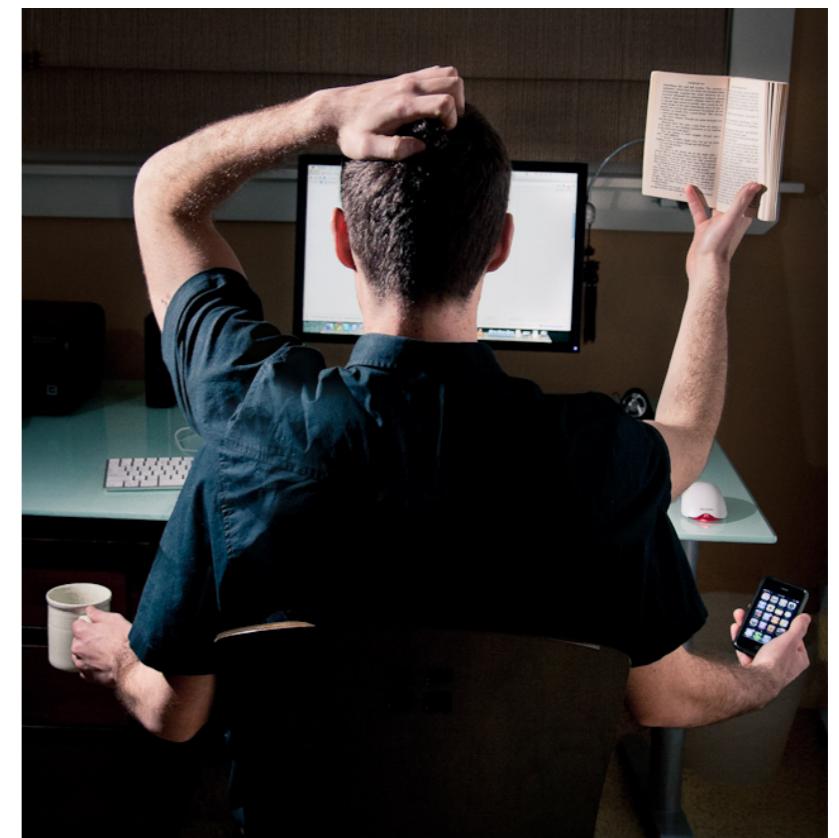
[Bacchelli et al. “*Harnessing Stack Overflow for the IDE*,” in RSSE. IEEE, 2012, pp. 26–30]



Is participation in SO related to productivity of developers?



Asset
or
burden?



Dataset

Largest code host
in the world



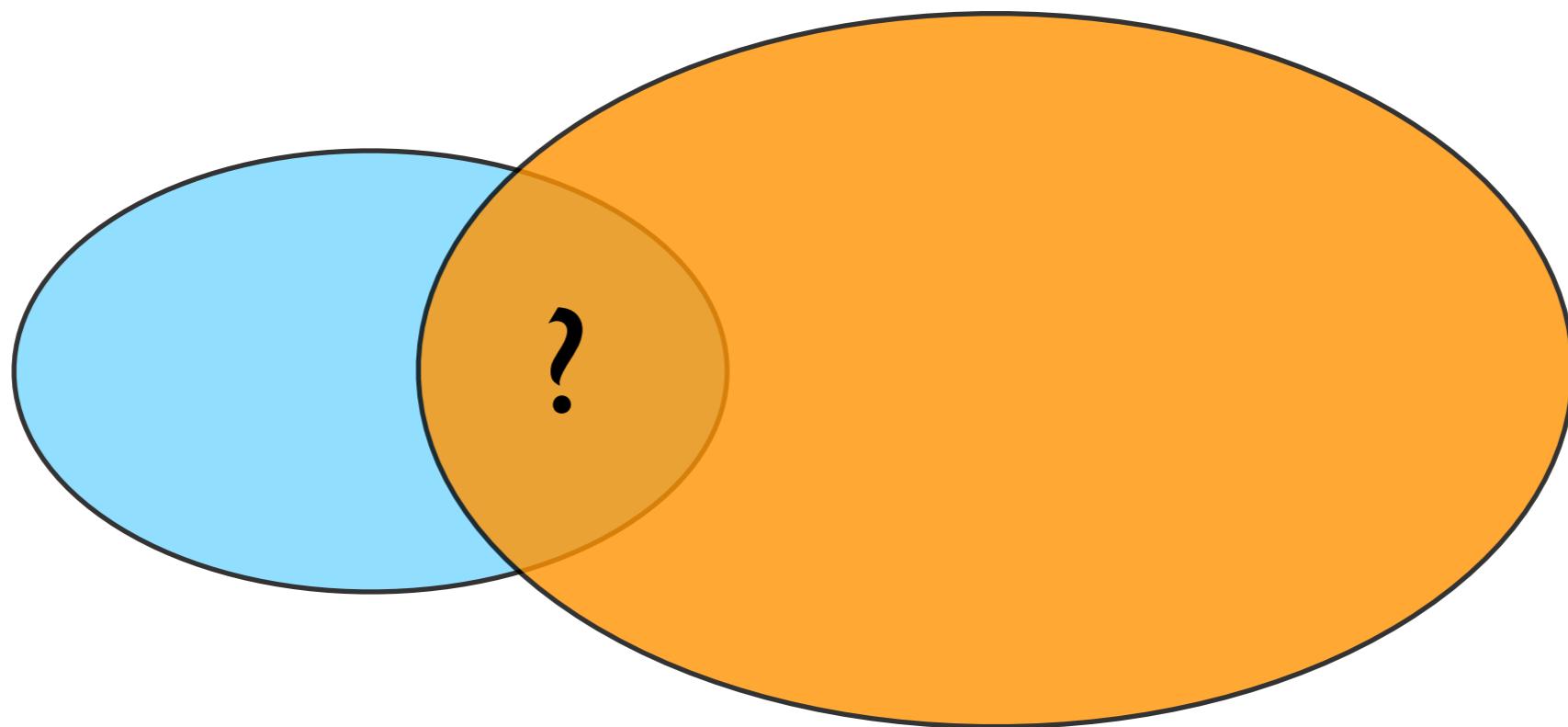
~400k users
July 2011 - April 2012

Largest programming
Q&A site in the world



~1.3M users
July 2008 - August 2012

Dataset



July 2011 - April 2012

July 2008 - August 2012

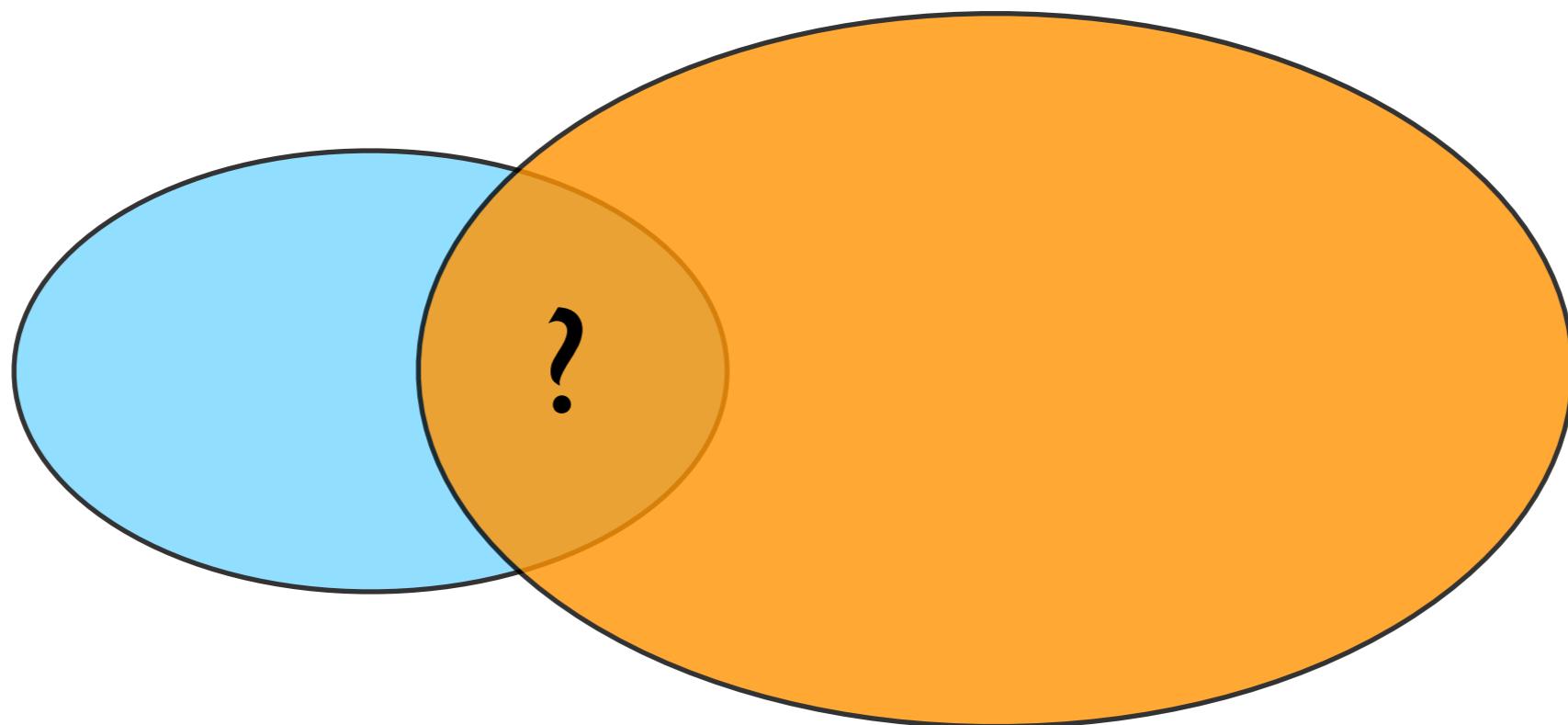
Dataset



Email address (plain text)



Email address (MD5 hash)



July 2011 - April 2012

July 2008 - August 2012

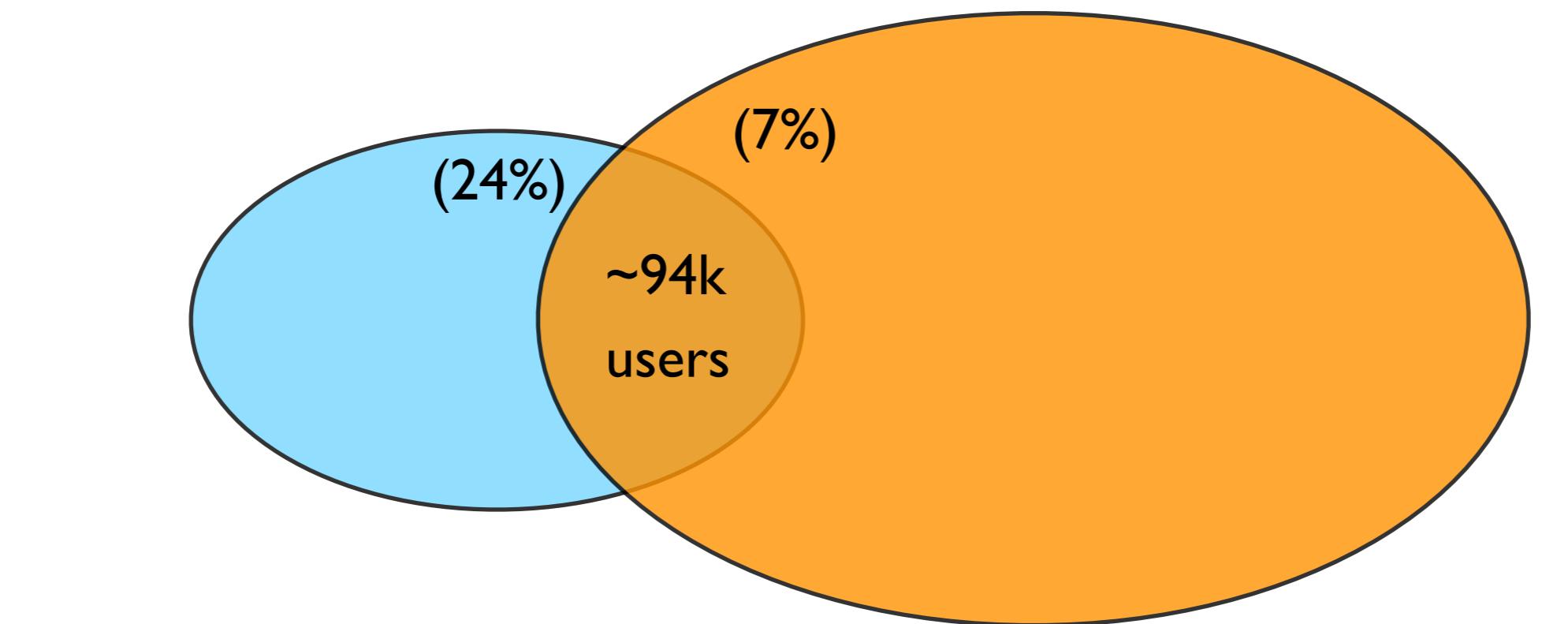
Dataset



Email address (plain text)



Email address (MD5 hash)



July 2011 - April 2012

July 2008 - August 2012

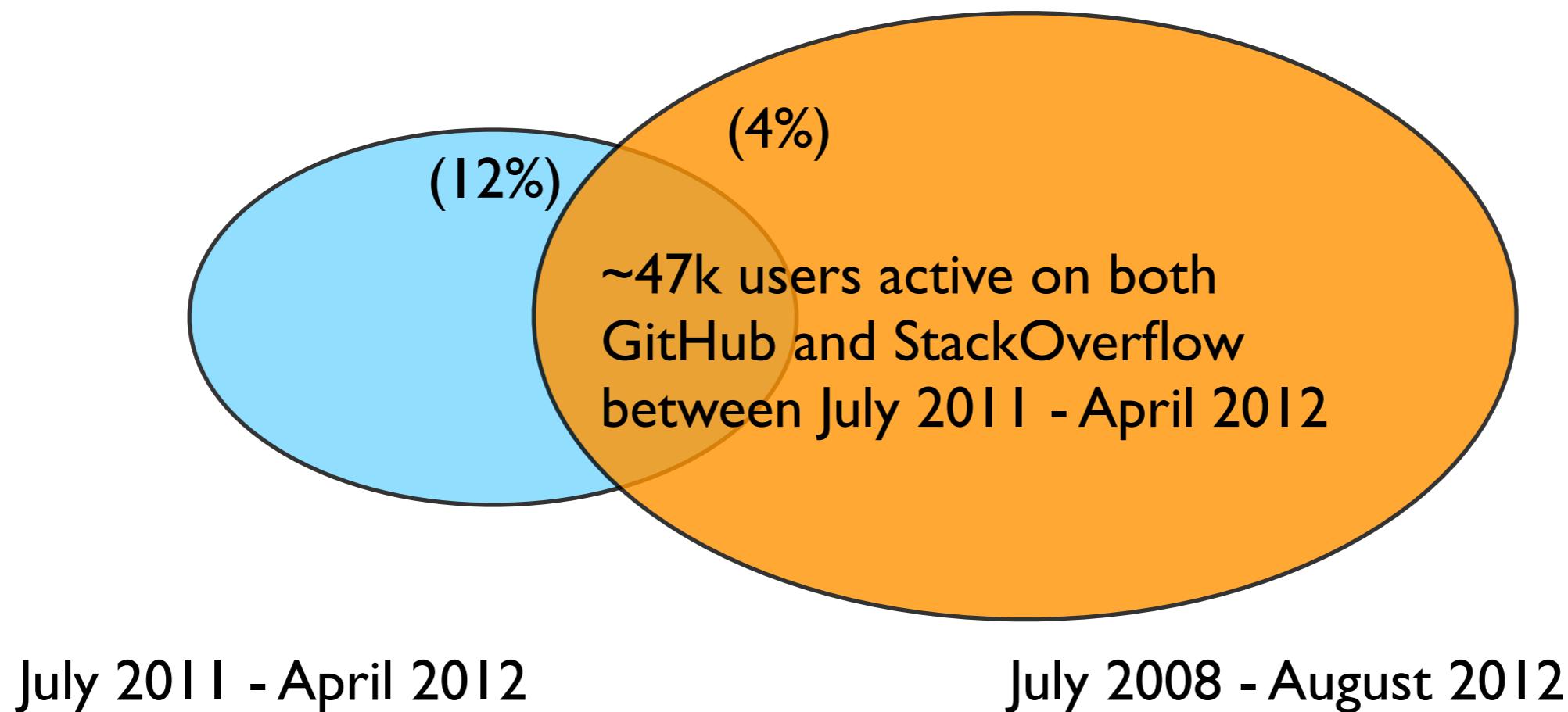
Dataset



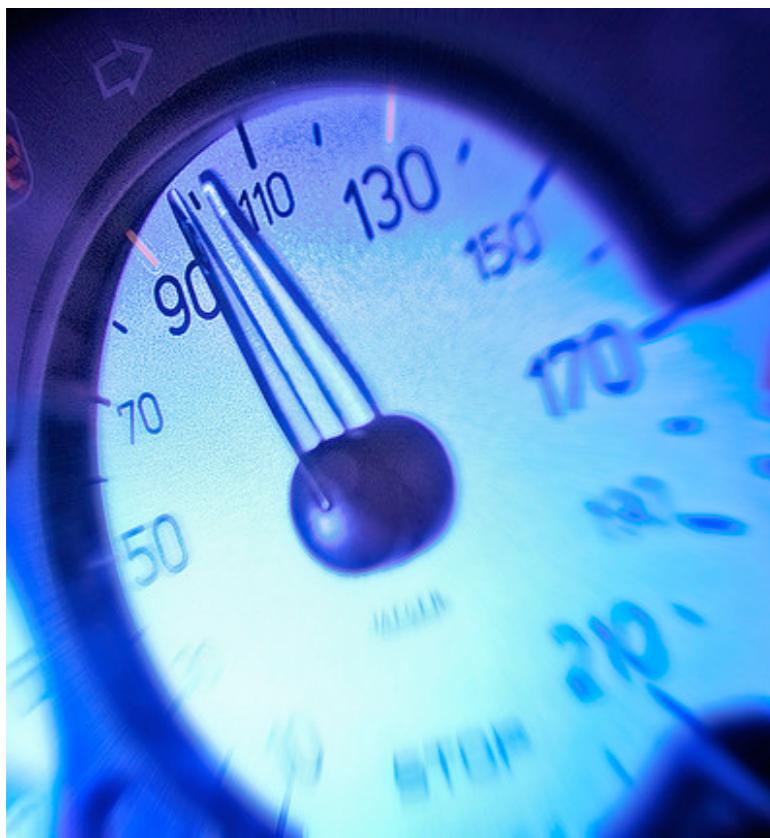
Email address (plain text)



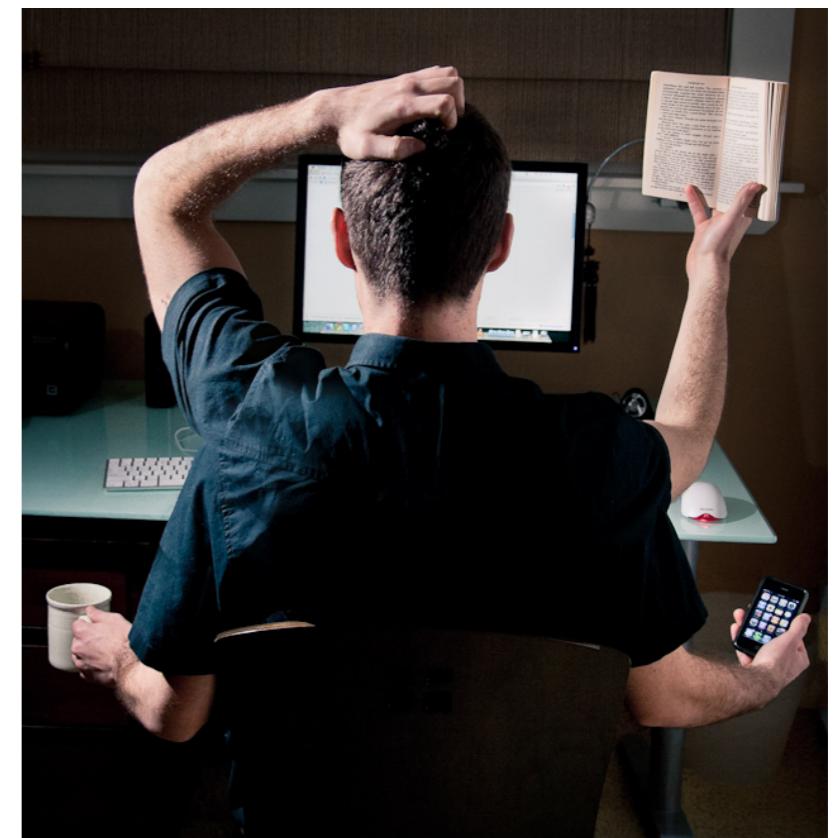
Email address (MD5 hash)



Is participation in SO related to productivity of developers?



Asset
or
burden?



Macro: overall activity levels



To what extent can activity (expertise) on one platform be used as a **proxy** for activity (expertise) on the other?

- social signals (e.g., open source projects, professional social media) ~ career advancement

[Capiluppi et al. “Assessing technical candidates on the social web,” IEEE Software, vol. 30, no. 1, pp. 45–51, 2013]

Intermediate: working rhythms



Is attention **focused** (bursts of commits) or
divided between the two platforms?

- working rhythms of developers ~ software quality

[Eyolfson et al. “Correlations between bugginess and time-based commit characteristics,”
Empirical Software Engineering, pp. 1–31, 2013]

Micro: coordination between commits and Q&A



Do StackOverflow activities **accelerate** or
slow down GitHub commits?

[Storey et al. “*The impact of social media on software engineering practices and tools*,” FoSER. ACM, 2010, pp. 359–364]



Macro

Overall activity



	#Commits	#Questions	#Answers
 Dave	100	5	50
 Stuart	10	75	15
 Kevin	25	10	75



Macro

Overall activity



#Commits

#Questions

#Answers

	Dave	100	5	50
	Kevin	25	10	75
	Stuart	10	75	15



Macro

Overall activity

Fix, sort



#Commits



Dave

100



Kevin

25



Stuart

10

Quartiles/Deciles, compare



stack**over**flow

#Questions

5

10

75

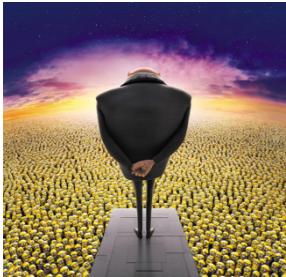
#Answers

50

75

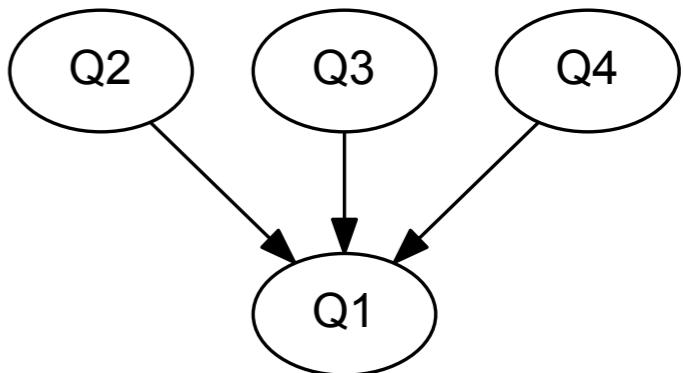
15

Not restricted to monotonic relations!

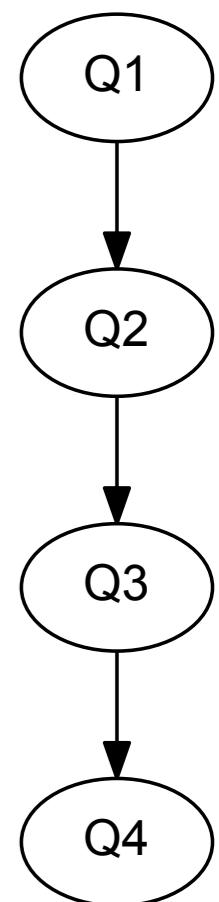


Findings

- Active GitHub committers are experienced developers:
 - few StackOverflow questions
 - many StackOverflow answers



Quartiles (#Commits)
Compare #Questions

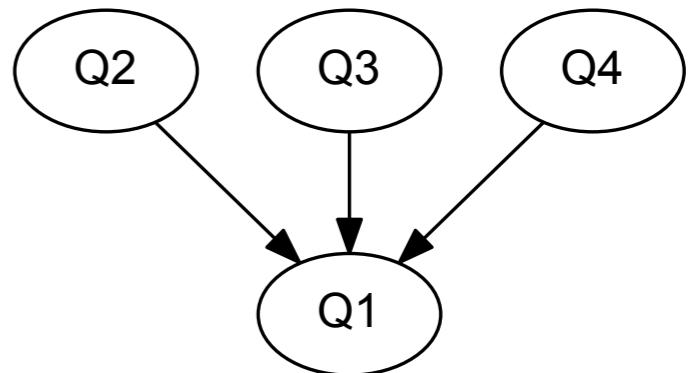


Quartiles (#Commits)
Compare #Answers

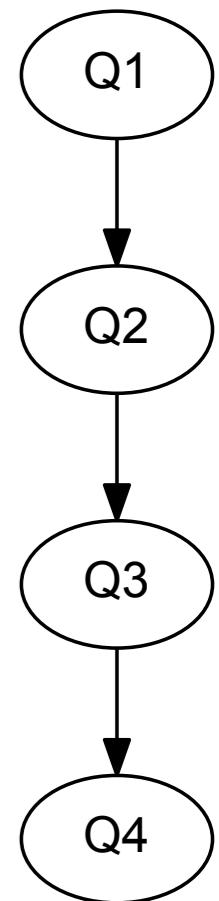


Findings

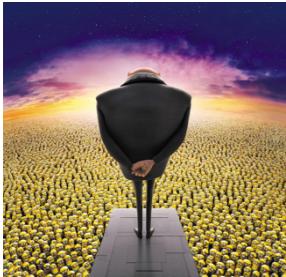
- Active GitHub committers are experienced developers:
 - few StackOverflow questions
 - many StackOverflow answers
- Top StackOverflow users are **superstars** rather than **slackers!**



Quartiles (#Commits)
Compare #Questions



Quartiles (#Commits)
Compare #Answers



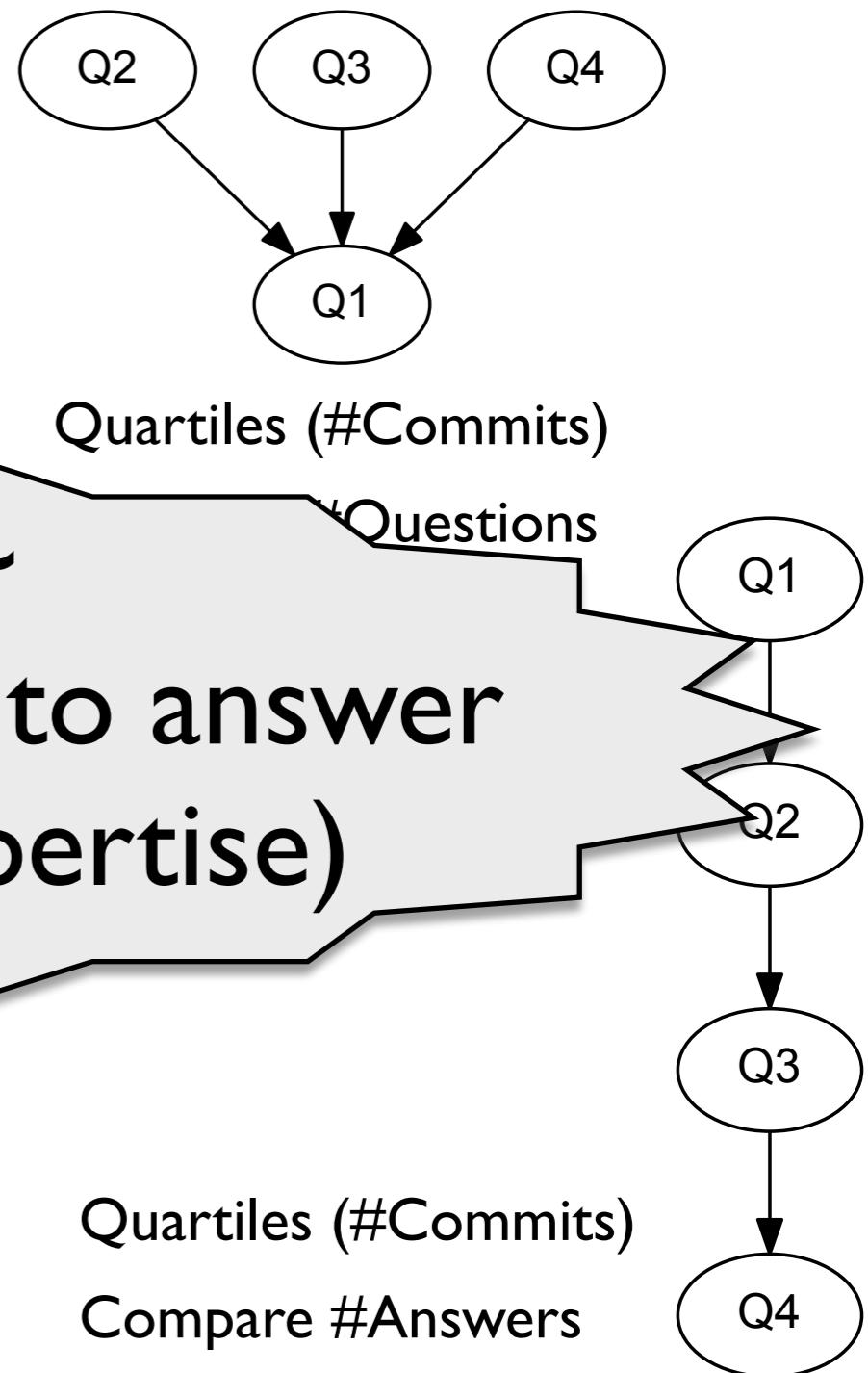
Findings

- Active GitHub committers are experienced developers.

GitHub activity ~

StackOverflow willingness to answer technical questions (expertise)

Top StackOverflow users are superstars rather than slackers!





Intermediate Working rhythms

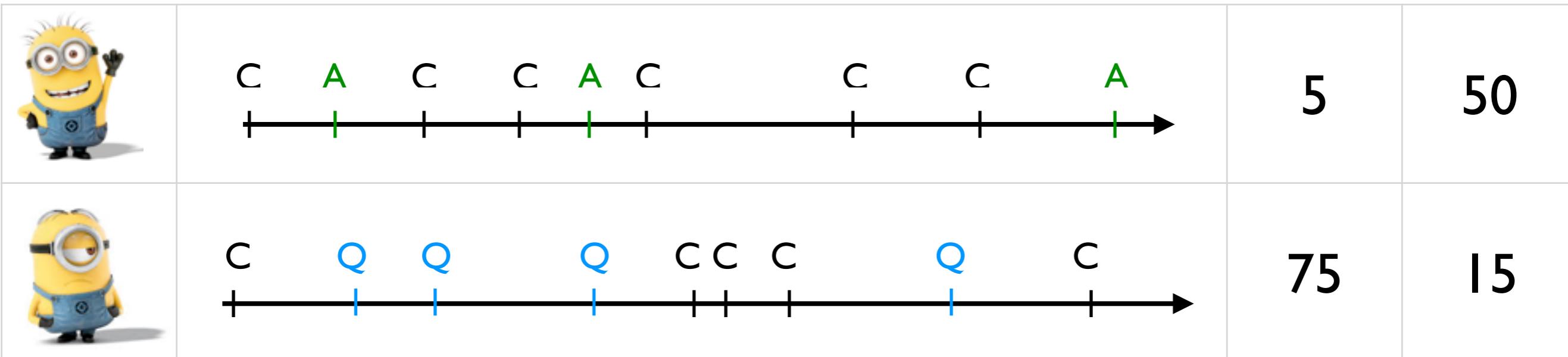
Committing rhythm: series of inter-commit time intervals

[Xuan et al. “Measuring the effect of social communications on individual working rhythms: A case study of open source software,” in Social Informatics. ASE/IEEE, 2012]

Gini (Committing rhythms): focused vs. distributed attention



Gini (Committing rhythms)





Findings

- Asking questions on StackOverflow influences how developers distribute their GitHub commits:
 - ▶ heavy askers: bursts of intense commit activity followed by longer periods of inactivity (focused attention)



Findings

- Asking questions on StackOverflow

influence their

Learning from

StackOverflow (by asking) and

committing to GitHub

activity, and periods of

inactivity (focused attention)

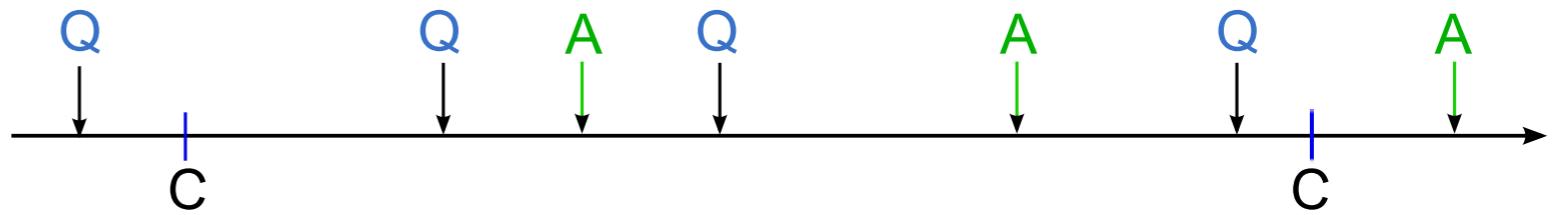


Micro

Who benefits from participating in SO



Dave



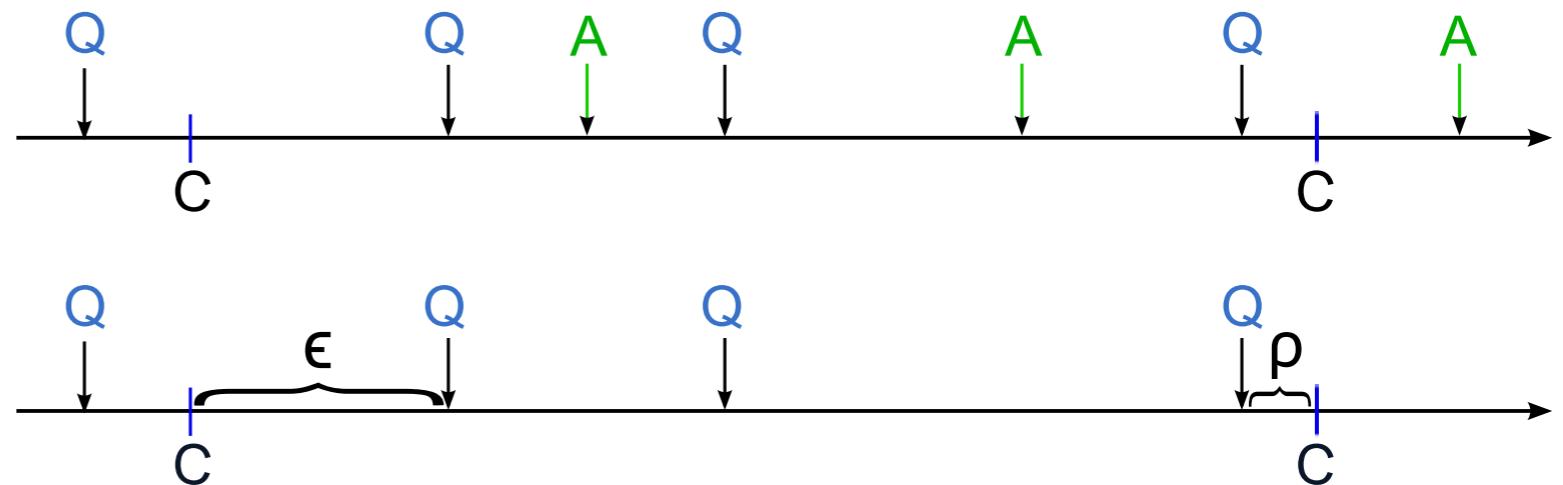


Micro

Who benefits from participating in SO



Dave



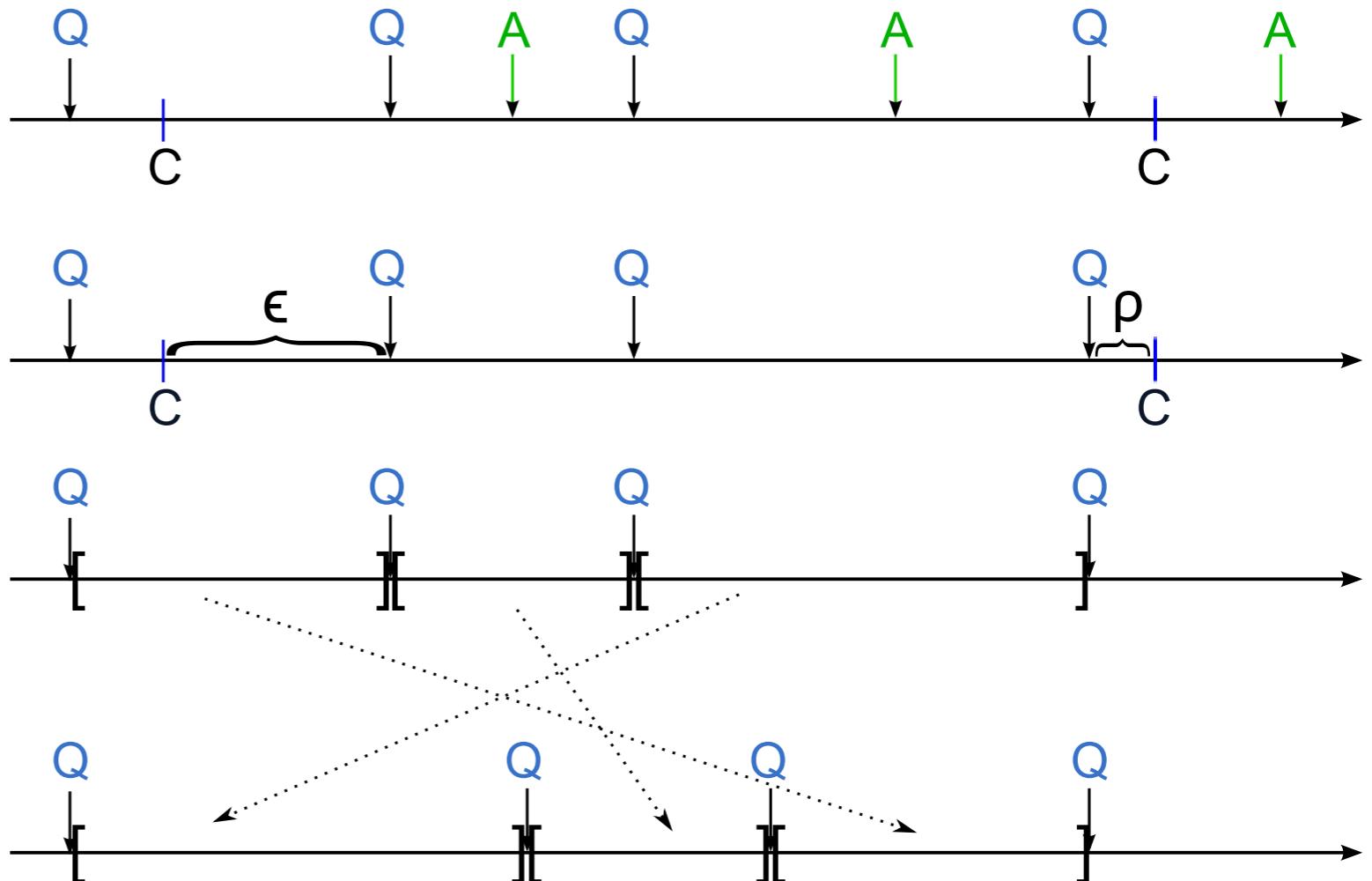


Micro

Who benefits from participating in SO



Dave





Micro

Who benefits from participating in SO

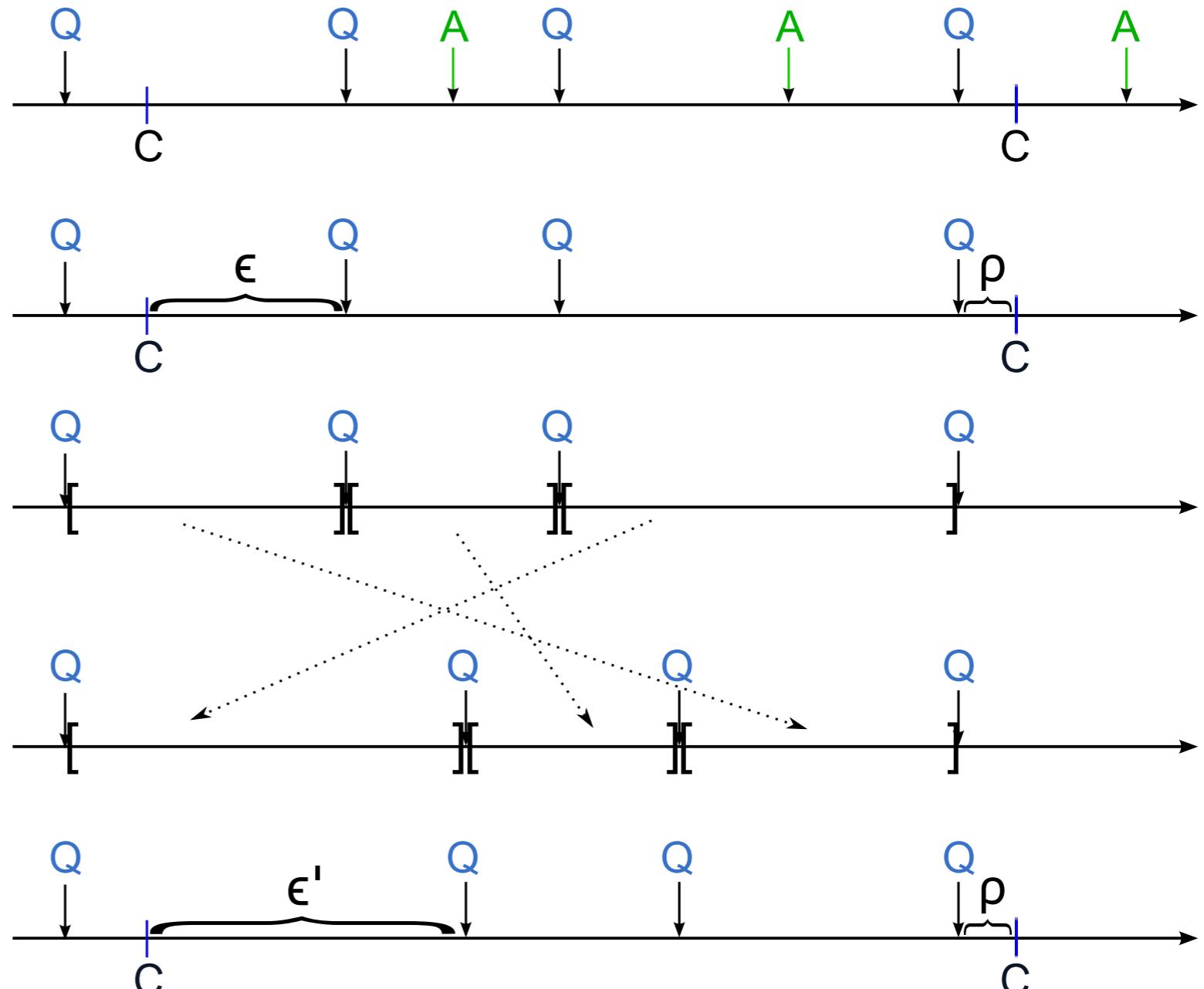


Dave

Compare actual and shuffled series:

actual < shuffled:
acceleration

actual > shuffled:
impediment





Findings

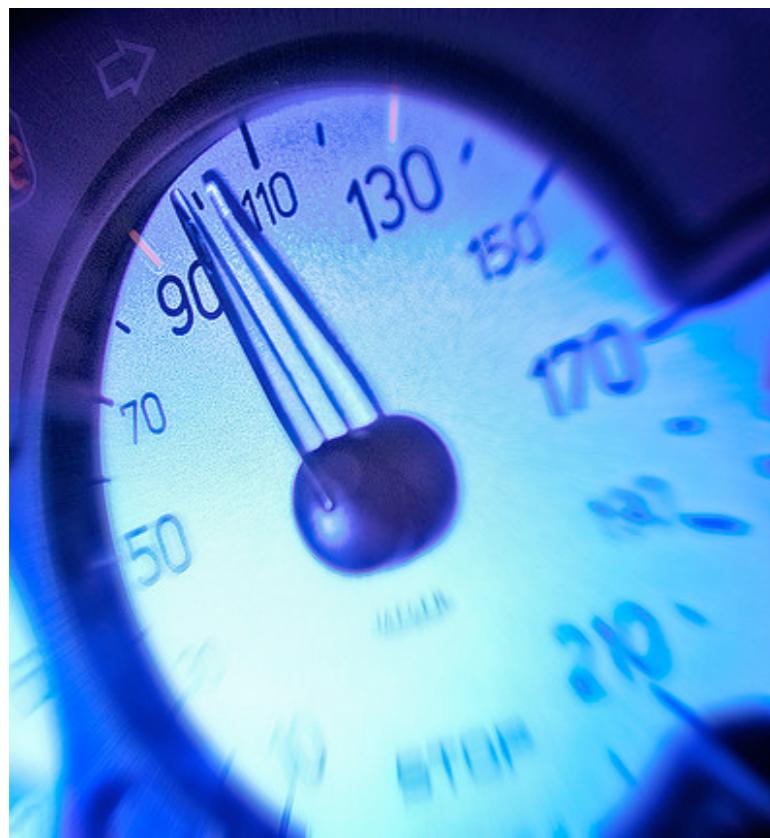


Findings

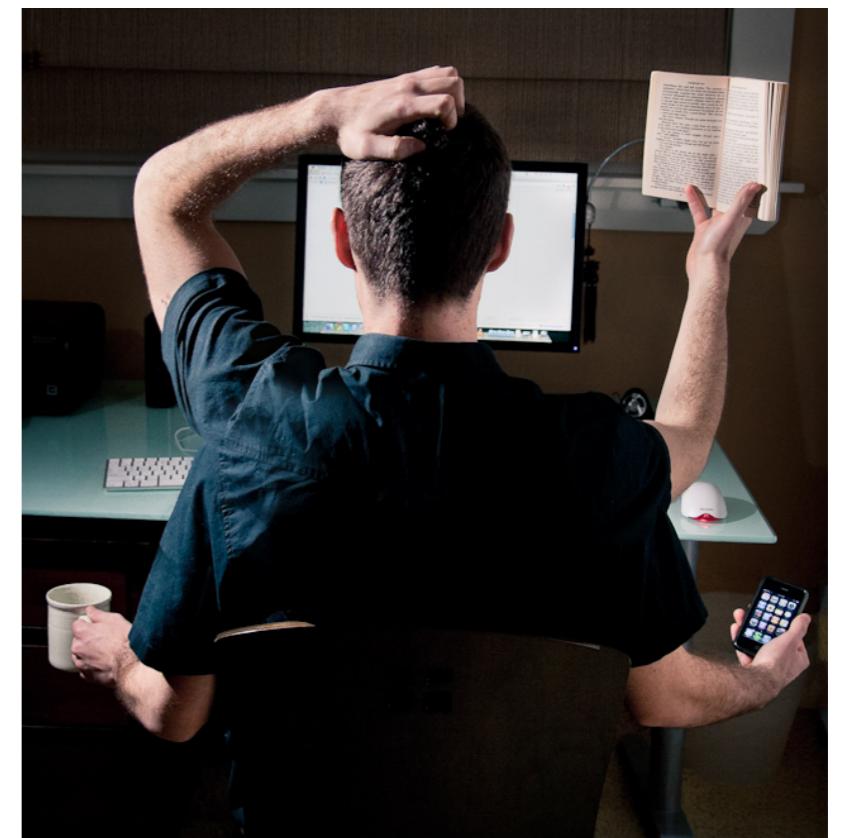
For active committers, asking and answering questions on StackOverflow catalyses committing on GitHub.

For no group is participating in StackOverflow detrimental!

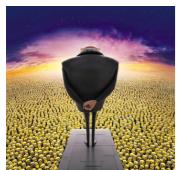
Summary: Is participation in SO related to productivity of GitHub dev's?



Asset
or
burden?



Summary



Active  committers are also active answerers (knowledge providers) on  stackoverflow



Different working rhythms for novices (focused attention) and experts



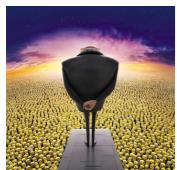
Participating in  stackoverflow reinforces commit activities on  github

Experts are experts everywhere!

Going to  stackoverflow is “costlier” for novices

 stackoverflow asset or burden

Summary



Active  committers are also active answerers (knowledge providers) on  stackoverflow



Different working rhythms for novices (focused attention) and experts



Participating in  stackoverflow reinforces commit activities on  github

Experts are experts everywhere!

Going to  stackoverflow is “costlier” for novices

 stackoverflow asset or burden