

Myth: OSS community participants are top-notch developers.

Kazuki Kusama (Kyushu University),
Masanari Kondo (Kyushu University), Shinobu Saito, Yukako Iimura (NTT), Naoyasu Ubayashi, Yasutaka Kamei (Kyushu University)

Background: Growing demand for OSS devs.

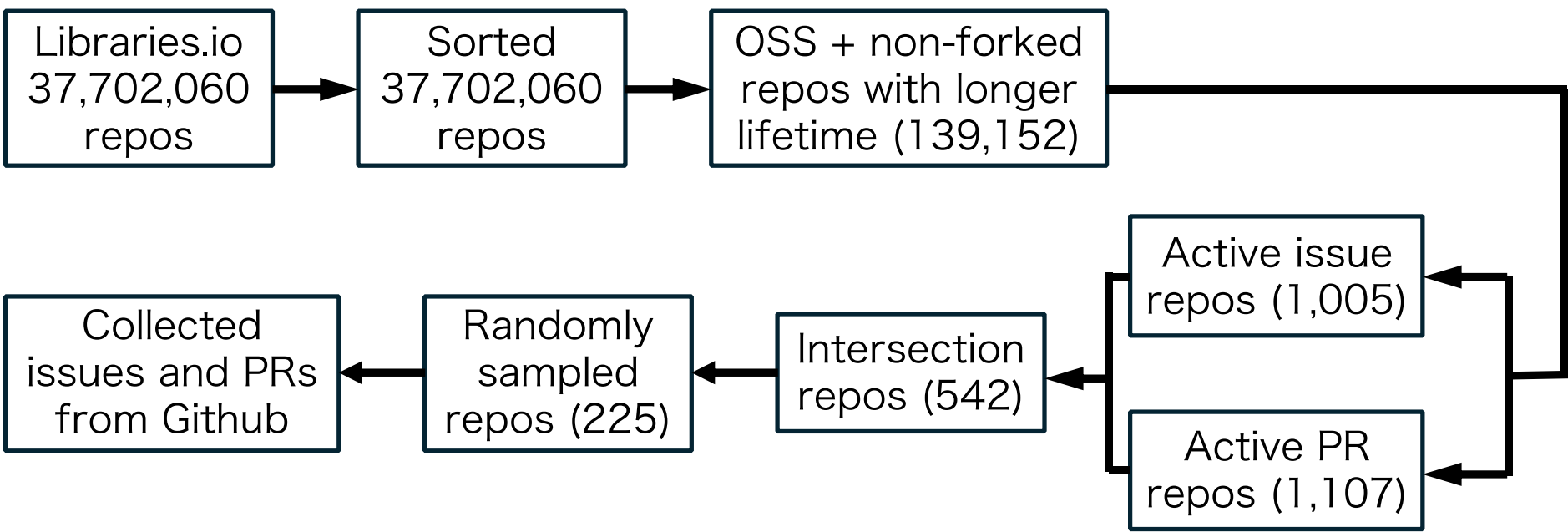


Myth “OSS community participants are top-notch developers”

[1]https://training.linuxfoundation.org/wp-content/uploads/2022/06/OpenSourceJobsReport2022_FINAL.pdf

Target data[2]

#Repos: 225 #Devs: 1,008,257



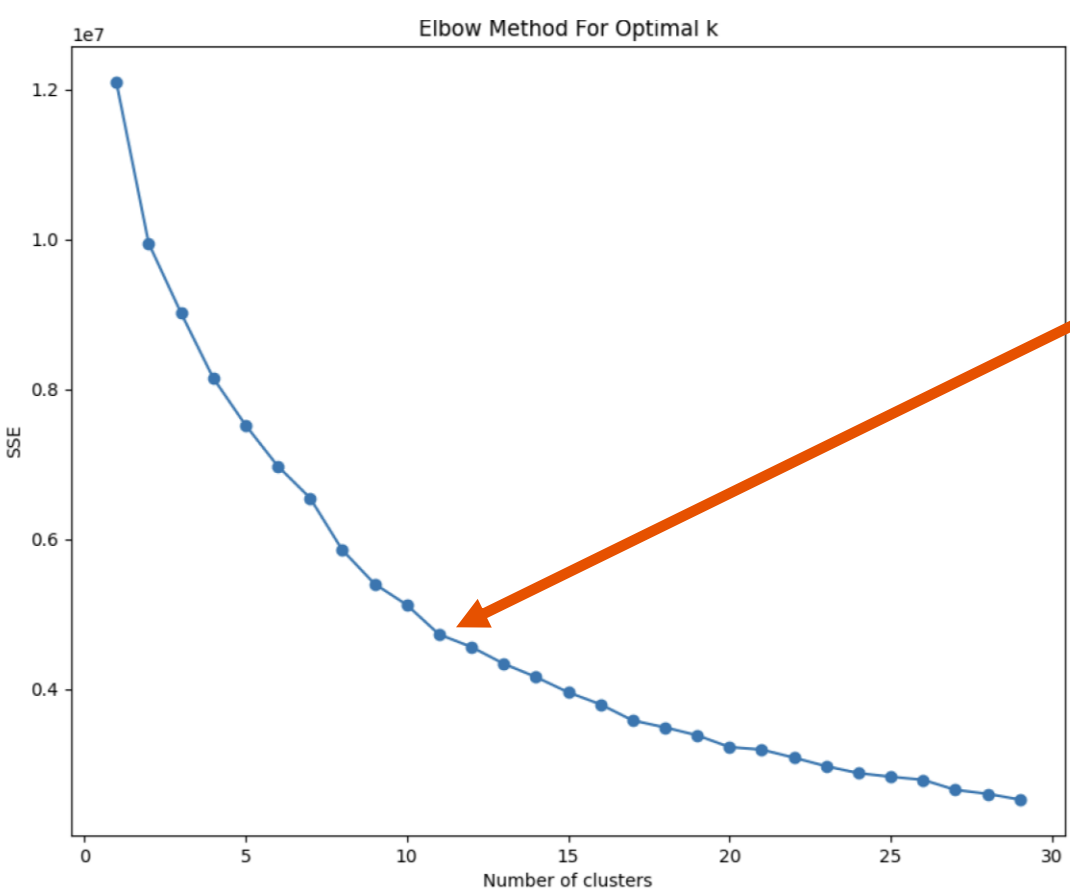
[2] M. Kondo et al., "Towards Better Online Communication for Future Software Development in Industry," 2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC), Torino, Italy, 2023, pp. 1619-1624, doi: 10.1109/COMPSAC57700.2023.00250. keywords: {Industries;COVID-19;Companies;Programming;Remote working;Data mining;Task analysis;mining software repository;online communication;open-source software;industrial development},

Methodology: Define 12 features and retrieve clusters of developers

Features

- Coding
 - #commits
 - Ave. commit int.
 - #PRs
 - Ave. PR int.
- Issue
 - Issue contri.
 - Ave. issue comment length
 - Ave. issue creat. Int.
- Code review
 - Review contri.
 - Ave. review length
 - Ave. review distance
- All
 - Aggr. contri.
 - Coding-discussion balance

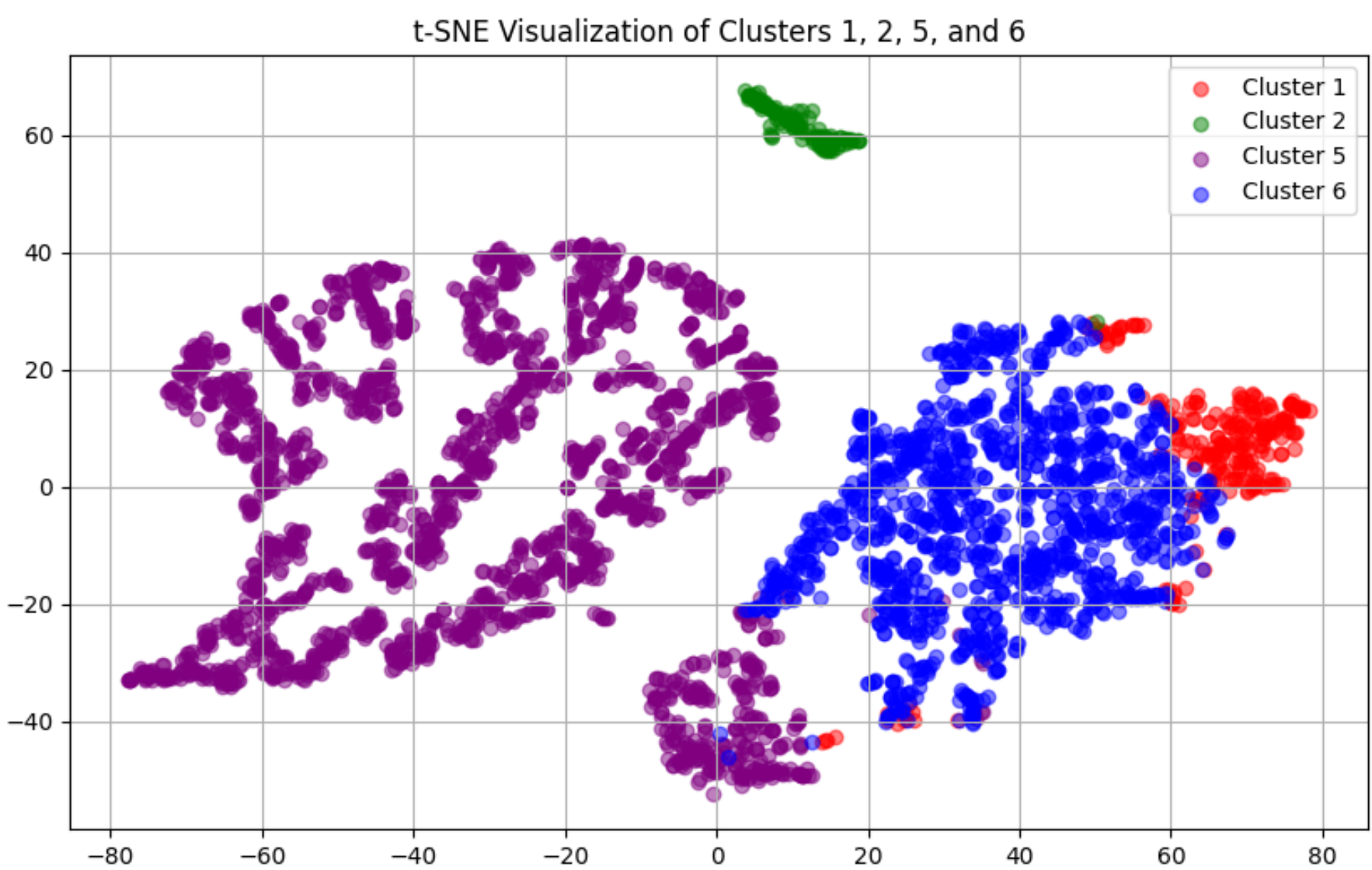
Clusters



Clusters

Labeling each cluster

1. Volunteer: All activities are less than average, accouts for 95% of all
2. User: Issue creating interval is long
3. Reviewer: Average code review activities
4. Issue contributor: Provides informative comments
5. Relaxed dev.: Large activity intervals
6. Balanced dev.: Balanced activities, more than average
7. Episodic dev.: Large commit intervals
8. Committer: Many activities are commit-related
9. Semi-committer: Similar to committer but less activities
10. Core dev.: The largest total contribution, all feature values are high
11. Semi-core dev. : Similar to core dev. but less activities



Result: Not only top-notch devs.

Finding 1: OSS development relies on the numerous anonymous volunteers

Volunteer: **95%** vs. Core dev.: **0.02%**

Finding 2: Numerous roles

- Focus on coding
- Focus on issue and review activities
- Large activity intervals

➔ **This myth is false**

Contribution: Recruiting

Use case: Recruiting activities

Current: Whether developers have experience with OSS activities

Future: The features described in this study contribute to recruiting activities