

How to increase diversity in open source projects

Dr Maren Westermann

About me

- PhD in agricultural/ environmental science from University of Queensland
- 2 years of experience as Data Scientist
- 2 years of experience as Machine Learning Engineer
- Active contributor to open source since 2020, especially scikit-learn
- PyLadies Berlin organiser since 2021

Diversity

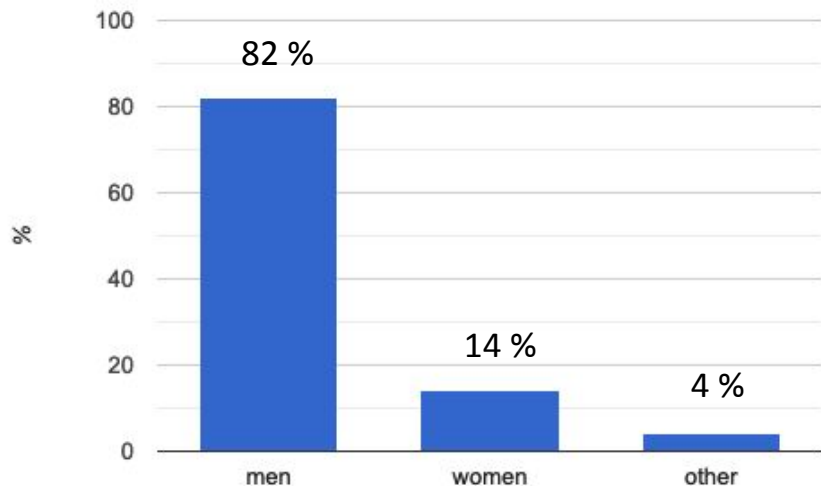
What is diversity?

“We define diversity within open source communities as a pluralism of any number of the following possibilities: gender identity and expression, race, ethnicity, sexual orientation, age, social class, caste, language, physical and neurological ability or attributes, religious beliefs, value systems, national origin, and political affiliation.”

The Linux Foundation, Diversity, Equity, and Inclusion in Open Source, 2021

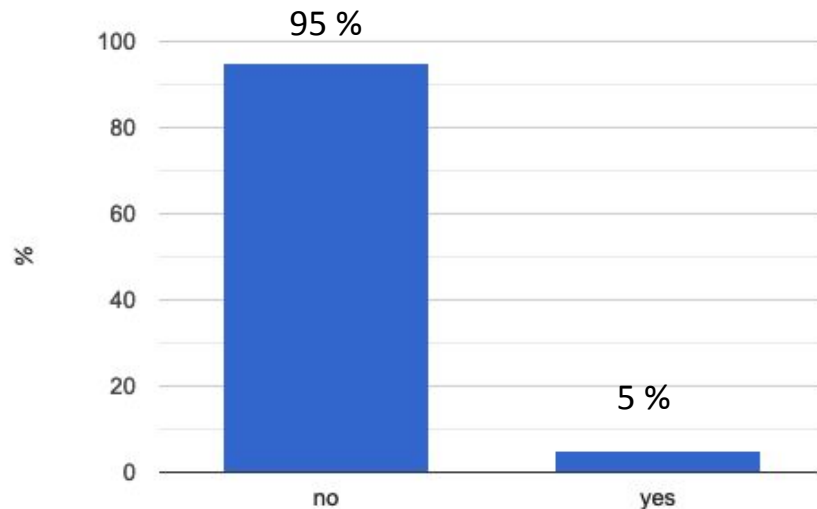
Statistics

Gender of contributors



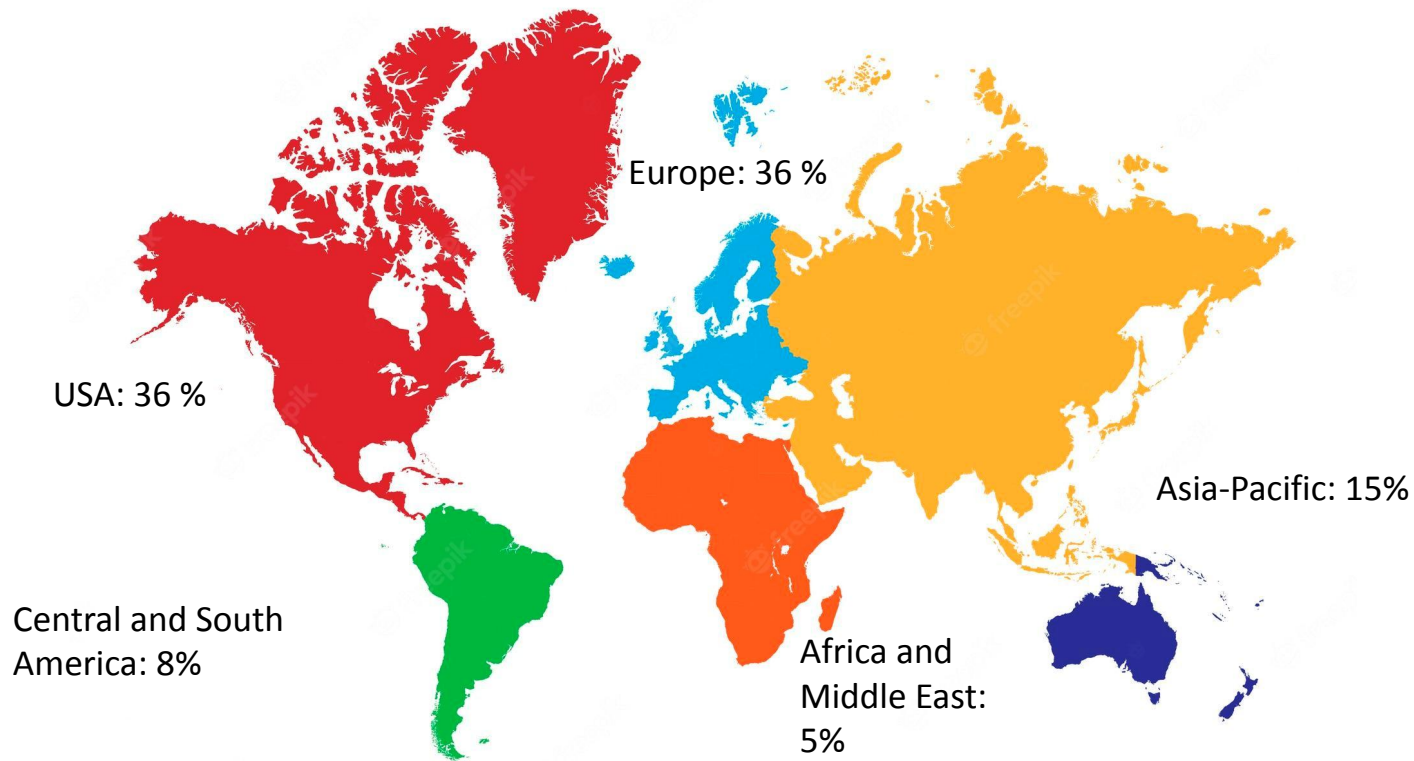
The Linux Foundation, Diversity, Equity, and Inclusion in Open Source, 2021

Female core developer?



Trinkenreich et al. 2022

Statistics



Consequences

For underrepresented groups

- Loss of opportunities
 - Skill development
 - Career advancement

For open source projects

- Lack of diversity of perspectives
 - Lower problem-solving capacity
 - Less robust technology
 - Lower productivity
 - Susceptibility to unhealthy work environments

Challenges faced

Mainly social challenges

- Lack of peer parity
- Non-inclusive communication
- Toxic culture
- Expectation issues
- Gender bias/ stereotyping
- Imposter syndrome
- Work life balance issues

Challenges faced

Other challenges

- Language access
- Educational access
 - Educational opportunities (hackathons, internships, courses)
- Geographic access
 - Time zones: ability to attend meetings and network
- Access to technologies (“digital divide”)
 - Reliable Internet access at home vs.elsewhere
- Economic and professional access
 - Compensation for open source contributions
 - Employers’ approval of open source technologies and employee time spent
 - Sponsorships, mentorships

Strategies

Use inclusive language



Guys



Master/ slave



Man hours



“So simple your mother can do it”



Folks/ people/ everyone



Primary/ secondary



Engineer/ personal hours



“User-friendly”

Strategies

Have contributing documentation in different languages (e.g. Spanish)

Ejemplo de contribución y envío de un Pull Request a [scikit-learn](#)



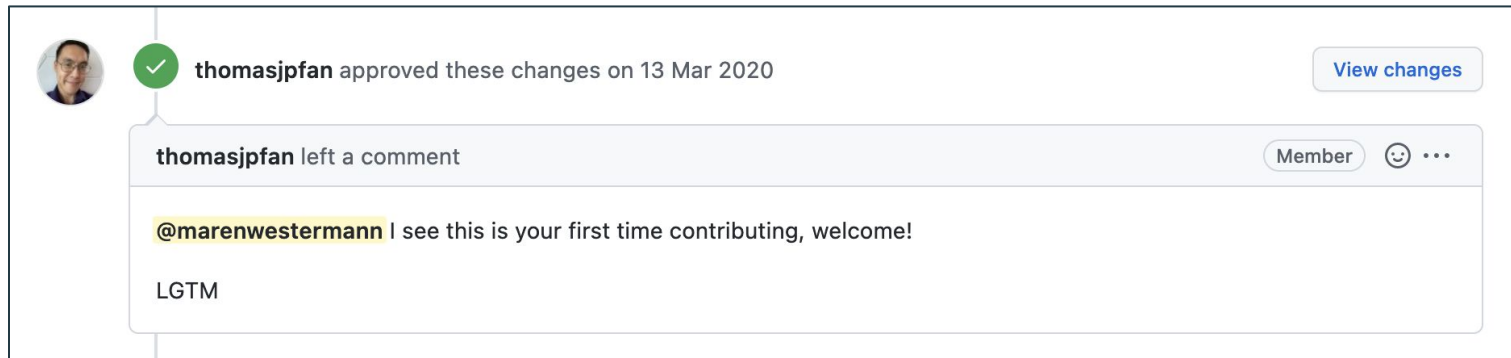
Transcripción del Video

- Ponente: [Reshama Shaikh](#)
- Video: [Scikit-learn PR Contributing Example](#) (30 minutes)
- Transcritora: [Reshama Shaikh](#)

Strategies

Encourage and be welcoming

- “Thanks a lot for your contribution, this is a valuable addition to the project.”
- “We really value this kind of contribution.”



Strategies

Promote specific groups and events

- Run events through specific groups, e.g.
 - scikit-learn: WiMLDS Paris sprint 2022
 - pandas: PyLadies London hack evenings
 - PyLadies Berlin: open source hack nights
- Social media: like and share
- Run events in schools and universities and give priority to people from underrepresented groups
- Run online events targeted at people from geographically underrepresented regions



Strategies

Promote awareness of the presence of peers

- At conferences: make sure you have speaker diversity
- Promotion on social media
- Write blog articles, e.g. about sprints, contributor interviews
- Recognize people's achievement



Strategies

De-stereotype the OSS contributor

- Don't automatically assign community building tasks to underrepresented groups; spread the workload



“women developers”



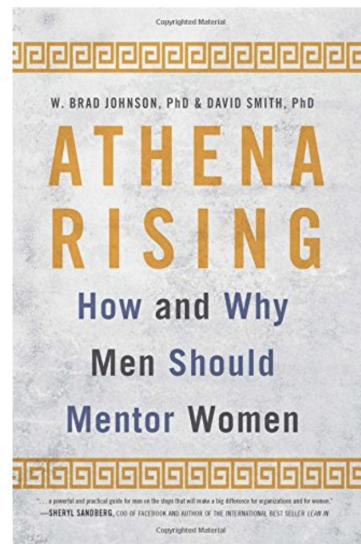
“developers”

Strategies

Mentorship

“Invaluable to improving diversity and inclusion in open source communities” - *Mentorship program leaders at the Linux Foundation*

- Lower barriers to entry and onboarding
- Networking
- Help junior contributors being able to establish themselves
- Career development
- Prepare mentors to guide women



Strategies

Governance

- Have governance board
 - Should be representative of different demographics
- Transparency and accountability around decision making
 - How are leaders determined?
 - How are contributors onboarded, managed, and supported?
- Transparent career pathways
- Promote people from underrepresented groups to leadership roles

Strategies

Create and enforce a Code of Conduct

- Can signal to marginalized people that the open source community cares
- Provides guidance for newcomers
- Provides reference for core values and norms

References

- Trinkenreich, B. et al, 2022. Women's Participation in Open Source Software: A Survey of the Literature. Manuscript submitted to ACM
- Carter, H. et al, 2021. Diversity, Equity, and Inclusion in Open Source. The Linux Foundation
- Geiger, R. S., 2017. Summary analysis of the 2017 github open source survey
- Zlotnick, F. et al, 2017. The open source survey. GitHub repository, 2017
- Finley, K., 2017. Diversity in Open Source Is Even Worse Than in Tech Overall, Wired Magazine
- Weber Mendonça, M., 2022. Beyond the basics: Contributor experience, diversity and culture in Open Source Projects. Keynote speech at PyConDE and PyData 2022
- Steinmacher, I. et al, 2021. Being a Mentor in open source projects. J Internet Serv Appl 12, 7

Thank you

Let's connect :)

- LinkedIn: de.linkedin.com/in/dr-maren-westermann-0b8575144
- Twitter: [@MarenWestermann](https://twitter.com/MarenWestermann)
- GitHub: github.com/marenwestermann