OpenStack Diversity Report
Lessons Learned

Nicole Huesman @ Intel [@uoduckswtd]
Daniel Izquierdo @ Bitergia [@dizquierdo]
General Overview
Technical Details
Wrap-Up
General Overview
• Increase awareness
• Lead positive change
• Employ quantitative + qualitative approach
• Help projects & teams (PTLs) [WIP]
• Gauge effectiveness of existing policies & programs [WIP]
• Identify best practices [WIP]
Current Data  (as of Apr 2017)

• Leadership & Governance:
  – Studied OpenStack Foundation, Board of Directors, Technical & User Committees, Ambassadors, Working Groups

• Technical Contributions:
  – Studied population & activity across commits & code reviews
Leadership & Governance

- Board of Directors and Working Groups: 17%
- User Committee: 20%
- OpenStack Ambassadors: 18%
- Keynote Speakers: 30%
- PTLs: 5%
- Technical Committee: 0%
Technical Contributions

**GIT DEVELOPERS BY GENDER**
- Male Developers: 63.9%
- Female Developers: 25.7%
- Unaffiliated Developers: 10.4%

**GIT ACTIVITY BY GENDER**
- Male Activity: 69.6%
- Female Activity: 23.2%
- Unaffiliated Activity: 7.15%

**GERRIT SUBMISSIONS BY GENDER**
- Male Submissions: 70.6%
- Female Submissions: 21.7%
- Unaffiliated Submissions: 7.71%
“Top” Projects

*those with greatest gender diversity, as measured by population & activity level*

**TOP 5 PROJECTS**

- Documentation
- Infrastructure
- Nova
- Quality Assurance
- Neutron
Current Data

• Other communities
  – Data for Linux Kernel & OpenStack communities is similar
  – Hadoop ecosystem: 8% of the population
  – Python programming language (5% [WIP])

• Technology industry claims up to 20%
Community Feedback

- Removal of certain repos (packaging-deb)
- Accuracy of data / incomplete data on website
- Open the project to the community (privacy)
- Guidance on expansion of metrics
- Receptiveness—request for assistance
So What?

- Is this data useful in your community?
- How can this data be useful to the Linux Foundation?
- Can we conduct similar analysis within the Linux & adjacent communities?

Insights can be applied across communities!
Technical Details
Grimoire Lab Architecture (draft)

Configuration file

Data Collect
- Arthur (*)
- Perseval

Data Enrichment
- Raw Data
- Identities
- Sorting Hat
- Merlin (*)

Data Consumption
- Enriched Data

Project Data
- Git
- GitHub
- Bugzilla
- ...

(*) Under development

ElasticSearch
REST API
These charts display the trend of female and male developers in the OpenStack project teams. The red lines follow a least square distance algorithm to analyze the trend of the number of commits and population of developers.
Wrap-Up
Lessons Learned & Open Questions

- Community is receptive to the data
- Bright lights & areas for improvement
- Leadership v tech contributions
- Application across communities
- Next step is understanding the “why”
A Few Recommendations

• Expand research parameters
• Develop mentoring & shadowing programs
• Collaborate with “top” projects
• Study the impact of existing programs & policies
How Can You Contribute?

• Would this analysis be useful in your work?
• What is missing from this analysis?
• How can we effectively apply these findings to increase diversity & inclusion across open source communities?

Join us at this afternoon’s break-out discussion!
OpenStack Diversity Report
Lessons Learned

Nicole Huesman @ Intel [@uoduckswtd]
Daniel Izquierdo @ Bitergia [@dizquierdo]