

Gender Diversity Analysis in the OpenStack Community

Contributors:

Daniel Izquierdo, Bitergia

Nicole Huesman, Intel Corporation

Allison Price, OpenStack Foundation

July 2017

Work commissioned by Intel Corporation





Foreword

I am grateful to Intel and Bitergia for sponsoring this important work and to Nicole Huesman and Daniel Izquierdo for getting this done. Thank you to the OpenStack Foundation for supporting this work and being open to recommendations. Since the technology industry started measuring and publishing numbers on diversity, the dialogue and actions have increased. We have a long way to go and publishing meaningful numbers of the open source community helps us understand the issues and apply action to moving the needle. While the technology industry has been a major source of innovation and economic growth, its ability to encourage diversity among its ranks lags. In 2016, women and underrepresented minorities accounted for 30% of the larger technology industry, while comprising 11% and nearly 10% of the OpenStack and Linux Kernel communities, respectively. We need to understand all three areas - percent in the projects, percent in leadership, and percent leaving projects and why. What is new in this report is the measurement of women in leadership positions, an important metric if we want to create change.

This research report aims at examining and informing actions that we collectively take to have a positive change to increase the diversity and inclusion across open source communities. It has also provided recommendations of actions that can be taken based on this work. We hope you will find this helpful and actionable.

Nithya Ruff, Women of OpenStack Member, Senior Director, Comcast



(cc) 2017 Bitergia, Intel. Some rights reserved.
This work licensed under Creative Commons Attribution-ShareAlike 4.0
Unported License.
To view a copy of full license, see
<http://creativecommons.org/licenses/by-sa/4.0>,
or write to Creative Commons, 559 Nathan Abbott Way, Stanford,
California 94305, USA.



Executive Summary

This research report seeks to examine and provide data about the gender diversity within the OpenStack community through April 1, 2017 from two perspectives: (1) governance/leadership, and (2) technical contributions. All research and subsequent analysis is based on publicly available data sources, including Git and Gerrit repositories, as well as the OpenStack website.

From a governance/leadership standpoint (contained within Chapter 2), this report accounts for females that hold roles on the OpenStack Board of Directors, the Technical and User Committees, the Working Groups, and the OpenStack Ambassadors, as well as Project Team Lead (PTL) positions. It also considers those who hold officer positions within the OpenStack Foundation, as well as those who have served as keynote speakers at OpenStack Summits.

Currently, women comprise nearly 17% of the Board of Directors and Working Groups, while 20% occupy roles within the User Committee, and 18% serve as OpenStack Ambassadors. A quarter of the OpenStack Foundation officers are women, while nearly 30% of women have served as keynote speakers at the previous two OpenStack Summits in Barcelona and Boston. A far lower percentage serve as PTLs at 5%, while no women sit on the Technical Committee.

In regards to technical contributions (contained within Chapters 3 & 4), this report examines source code and code review contributions, and identifies which projects have achieved the greatest gender diversity. The research shows a peak in the number of total OpenStack contributors in early 2016, with an overall decrease in contributions since then, which may signal an overall maturation of the OpenStack project itself. Within this context, female representation has remained relatively steady across both source code and code review contributions, both with respect to population, or number of female contributors, and their level of activity.

Of those projects with the greatest gender diversity, the Packaging-deb, Documentation and Infrastructure projects ranked highest, as measured by number of female contributors and their respective number of commits, followed by the Horizon, Nova and Neutron projects.

Table of Contents

Executive Summary	3
Table of Contents	4
1. Leadership & Governance	6
Board of Directors	6
Technical Committee	7
User Committee	8
Working Groups	8
OpenStack Ambassadors	9
Project Technical Leads (PTLs)	10
Other Considerations	11
Summit Keynote Representation	12
2. Technical Contributions	13
Top Projects	13
Source Code Contributions	14
Commits by Gender	15
Developers by Gender	16
Types of Contributions	17
Code Review Contributions	19
Code Review Submissions	19
Developers Submitting Changesets	21
Code Review Votes	22
Core Code Review Votes	24
3. Further Work and Recommendations	27
Appendix A: Detailed Summary of Female Developer Activity	29
Appendix B: Technical Details and Limitations	32



1. Leadership & Governance

This section focuses on the analysis of gender diversity in the OpenStack community from a leadership and governance perspective. Of total membership, women comprise a sizable percentage at 20% of the User Committee, 18% of OpenStack Ambassadors, and nearly 17% of the Board of Directors and Working Groups. A quarter of the OpenStack Foundation officers are women, while nearly 30% of women have served as keynote speakers at the previous two OpenStack Summits in Barcelona and Boston. Alternatively, female representation lags among Project Team Lead (PTL) positions at 5%, and within the Technical Committee at 0%.

The sources for the data in this section include the OpenStack website and the Project Teams Governance file. These sources provide varying data--some include company associations for individuals, whereas others do not. Subsequently, some of the charts within this report include company affiliation, while others do not.

Board of Directors

The Board of Directors "provides strategic and financial oversight of Foundation resources and staff."¹ At the time of this report, females comprise 16.66% of Board membership, or 4 of 24 members. The list of members is as follows:

Member	Co. Affiliation	Member	Co. Affiliation
Alan Clark	SUSE	Joseph Wang	inwinSTACK
Allison Randal	OSI, HPE, others	Junwei Liu	China Mobile Research Institute
Anni Lai	Huawei	Kavit Munshi	Aptira
Boris Renski	Mirantis	Kenji Kaneshige	Fujitsu
Brad Topol	IBM	Lew Tucker	Cisco Systems
Brian Stein	Rackspace	Mark Baker	Canonical Group Ltd
ChangBo Guo	EasyStack	Mark McLoughlin	Red Hat Inc.
Christopher Price	Ericsson AB	Robert Esker	NetApp

¹ <https://www.openstack.org/foundation/board-of-directors/>

Egle Sigler	Rackspace	Russell Bryant	Red Hat Inc.
Gnanavelkandan Kathirvel	AT&T	Shane Wang	Intel
Imad Sousou	Intel	Steven Dake	Cisco Systems
Jessica Field	Aptira	Tim Bell	CERN

Table: OpenStack Foundation Board of Directors

Technical Committee

The Technical Committee is a “fully-elected committee that represents the contributors to the project” whose primary purpose is “to provide technical leadership for OpenStack as a whole.” At the time of this report, this committee is comprised of 13 people, all males².

Current roster is as follows:

Member	Co. Affiliation	Member	Co. Affiliation
Davanum Srinivas	Futurewei Tech., Mirantis	Matthew Treinish	IBM
Dean Troyer	Intel	Mike Perez	OpenStack Foundation
Doug Hellmann	Red Hat Inc.	Monty Taylor	Red Hat Inc.
Emilien Macchi	Red Hat Inc.	Sean Dague	IBM
Flavio Percoco	Red Hat Inc.	Steve Martinelli	IBM
Jeremy Stanley	OpenStack Foundation	Thierry Carrez	OpenStack Foundation
John Garbutt	-	-	-

Table: OpenStack Foundation Technical Committee

² <https://www.openstack.org/foundation/tech-committee/>

User Committee

The User Committee is "led by a core group of five individuals, who provide oversight and guidance to a number of working groups that target specific areas for improvement."³ At the time of this report, the percentage of female committee participants is 20%. The list of members is as follows:

Member	Co. Affiliation
Edgar Magana	Workday
Jonathan Proulx	MIT CSAIL
Melvin Hillsman	Rackspace
Shamail Tahir	athenaHealth
Shilla Saebi	Comcast

Table: OpenStack Foundation User Committee

Working Groups

There are a number of working groups that the User Committee assists.⁴ These working groups and their co-chairs are listed in the chart below. At the time of this report, of the total 18 co-chairs, 3, or 16.6%, are women.

Working Group	Co-Chairs
App Dev Enablement Working Group	Patricia Montenegro, Christopher Aedo
Enterprise Working Group	Yih Leong Sun
Fault-Genes Working Group	Nemat Bidokhti, Rochelle (Rocky) Grober
LCOO	Jamey McCabe, Sundar Krishnamurthy

³ <https://www.openstack.org/foundation/user-committee/>

⁴ https://wiki.openstack.org/wiki/Governance/Foundation/UserCommittee#Working_Groups

Large Deployment Team	Matt Van Winkle
Massively Distributed Clouds	Adrien Lebre
Ops Tags Team	Tom Fifield, Edgar Magana, Jon Proulx, Shilla Saebi
Operators Telecom/NFV	Curtis Collicutt
Product Working Group	Yih Leong Sun, Shamail Tahir
Scientific Working Group	Stig Telfer, Blair Blethwaite

Table: OpenStack Foundation Working Groups

OpenStack Ambassadors

OpenStack ambassadors⁵ are globally distributed and help “tie user groups together, and work with each one to mentor it to be the best it can be”. Of the total 11 ambassadors, 2, or 18% are women.

Current list of ambassadors:

Ambassador	Ambassador
Akihiro Hasegawa	Kavit Munshi
Akira Yoshiyama	Lisa-Marie Namphy
Christian Berendt	Lu Ye
Erwan Gallen	Marcelo Dieder
Ilya Alekseyev	Márton Kiss
Jaesuk Ahn	-

⁵ <https://groups.openstack.org/ambassador-program>

Project Technical Leads (PTLs)

Project Team Leads (PTLs) are elected every six months to govern each OpenStack project, as specified in the Governance YAML file by the Foundation.⁶ At the time of this report, of the total 60 PTLs, 3, or 5% are women. The following is the list of PTLs:

PTL Name	Project	PTL Name	Project
Dave McCowan	Barbican	Michal Jastrzebski	Kolla
Jan Klare	Chef OpenStack	Antoni Segura Puimedon	Kuryr
Sean McGinnis	Cinder	Adrian Otto	Magnum
Christophe Sauthier	Cloudkitty	Ben Swartzlander	Manila
Eric Kao	Congress	Renat Akhmerov	Mistral
Tim Simmons	Designate	Roland Hochmuth	Monasca
Alexandra Settle	Documentation	Felipe Monteiro	Murano
Omer Anson	Dragonflow	Kevin Benton	Neutron
Alexandre Levine	Ec2-api	Matt Riedemann	Nova
Saad Zaher	Freezer	Michael Johnson	Octavia
Vladimir Kuklin	Fuel	James Page	OpenStack Charms
Brian Rosmaita	Glance	Andy McCrae	OpenStackAnsible
Rico Lin	Heat	Dean Troyer	OpenStackClient
Rob Cresswell	Horizon	ChangBo Guo	Oslo
Ian Y. Choi	I18n	Thomas Goirand	Packaging-deb
Jeremy Stanley	Infrastructure	Igor Yozhikov	Packaging-rpm

⁶ <http://git.openstack.org/cgit/openstack/governance/tree/reference/projects.yaml>

Dmitry Tantsur	Ironic	Alex Schultz	Puppet OpenStack
Yuval Brik	Karbor	Andrea Frittoli	Quality Assurance
Lance Bragstad	Keystone	Andrey Kurilin	Rally
Catherine Diep	RefStack	John Dickinson	Swift
Thierry Carrez	Release Management	Gongysh Gongysh	Tacker
Tony Breeds	Requirements	Julien Danjou	Telemetry
Telles Mota Vidal Nóbrega	Sahara	Chaoyi Huang	Tricircle
Steve McLellan	Searchlight	Emilien Macchi	Tripleo
Robert Clark	Security	Amrith Kumar	Trove
Qiming Teng	Senlin	Ifat Afek	Vitrage
Monty Taylor	Shade	Alexander Chadin	Watcher
Devdatta Kulkarni	Solum	Claudiu Betu	Winstackers
Tony Breeds	Stable Branch Maintenance	Fei Long Wang	Zaqar
Eran Rom	Storlets	Hongbin Lu	Zun

Table: OpenStack Foundation PTLs


Other Considerations

OpenStack Foundation Officers

The OpenStack Foundation’s goal is “to serve developers, users, and the entire ecosystem.” At the time of this report, a quarter of OpenStack Foundation officers are women⁷.

Current List of Officers:

⁷ <https://www.openstack.org/foundation/staff/>



Officer	Job Title
Jonathan Bryce	Executive Director
Mark Collier	Chief Operating Officer
Lauren Sell	VP, Marketing & Community Services
Thierry Carrez	VP, Engineering

Table: OpenStack Foundation Officers

Summit Keynote Representation

During the previous two OpenStack Summits, in Barcelona and Boston, women comprised 27%, or 12 of a total 45, keynote speakers. Also notable is that during these bi-annual events, females account for approximately 10-12% of attendees.

2. Technical Contributions

This report examines technical activity as measured by source code contributions and code review contributions. It also identifies which projects rank highest in gender diversity, as measured by population, or number of female contributors, and their respective activity.

The number of total OpenStack contributors peaked at the beginning of 2016 and has seen an overall decrease since then. Where the data presented in this report shows slight decreases in the number of female contributors within the OpenStack community, this trend is line with the overall decrease of total OpenStack contributors. The same is true when examining the level of activity. These trends may indicate the overall maturation of the OpenStack project itself.

Across both source code and code review contributions, female representation has remained relatively steady, as measured by population, or the number of female contributors, and their respective number of commits. While the level of source code activity among females over the last year slightly exceeds this activity in aggregate over the last four years, the level of code review activity among women over the last year falls slightly below its respective four-year aggregate.

Sources used in this section include Gerrit repositories, where code reviews take place, and Git repositories, where the pieces of code that have been accepted in Gerrit are merged. The report provides an analysis of all Git and Gerrit repositories available within the OpenStack Foundation

governance file,⁸ which contains pointers to all of the projects, repositories and PTL names used in this report.

Top Projects

Of the approximately 60 projects within the OpenStack Foundation, the following tables summarize the projects with the greatest gender diversity, as measured by population, or number of female contributors, and their respective activity, or number of commits submitted by females.

Overall, the Packaging-deb, Documentation and Infrastructure projects ranked highest, followed by the Horizon, Nova and Neutron projects. From a population perspective, the Quality Assurance, Cinder, Keystone, and Oslo projects were notable, while from an activity level standpoint, the Fuel, Ironic, Murano, and Puppet projects ranked high.

When examining female population and activity level relative to their male counterparts within projects, the Documentation, Murano, Horizon, and Ironic projects ranked highest overall. Also notable were the Neutron and Nova projects with respect to female population, and the Fuel and Keystone projects with regard to female activity level.

project	authors	commits	ratio_authors	ratio_commits
Packaging-deb	623.0	13175.0	14.361457	8.536736
Documentation	145.0	3080.0	20.393812	37.211550
Infrastructure	145.0	2303.0	11.214230	4.849442
nova	121.0	891.0	16.285330	6.913944
Quality Assurance	101.0	650.0	13.760218	7.747318
neutron	99.0	858.0	15.840000	7.279824
horizon	96.0	949.0	22.748815	25.333689
cinder	77.0	532.0	13.898917	9.646419
oslo	59.0	416.0	11.706349	4.804250
Telemetry	57.0	281.0	21.590909	5.624500
glance	57.0	234.0	18.811881	10.077519

Table: List of top 10 projects with the greatest gender diversity, as measured by population, or number of female contributors. Source: Git repositories.

⁸ <http://git.openstack.org/cgit/openstack/governance/tree/reference/projects.yaml>

project	authors	commits	ratio_authors	ratio_commits
Packaging-deb	623.0	13175.0	14.361457	8.536736
fuel	48.0	4518.0	14.634146	17.884570
Documentation	145.0	3080.0	20.393812	37.211550
Infrastructure	145.0	2303.0	11.214230	4.849442
ironic	43.0	1172.0	14.478114	24.209874
murano	31.0	1002.0	21.379310	27.073764
horizon	96.0	949.0	22.748815	25.333689
Puppet OpenStack	33.0	894.0	9.455587	11.149913
nova	121.0	891.0	16.285330	6.913944
neutron	99.0	858.0	15.840000	7.279824

Table: List of top 10 projects with the greatest gender diversity, as measured by the highest number of commits contributed by women. Source: Git repositories.

Source Code Contributions

This report provides an analysis of source code contributions along the following dimensions:

- Number of commits by gender over time
- Number of developers by gender authoring those commits over time
- Types of contributions (code and others) by gender within those commits over time

Commits by Gender

This section reflects activity by gender, and aggregates this data over the last four years. During this time period, females contributed 6.7% of total commits. Over the last year, this percentage slightly exceeded the aggregate at 7.15%.

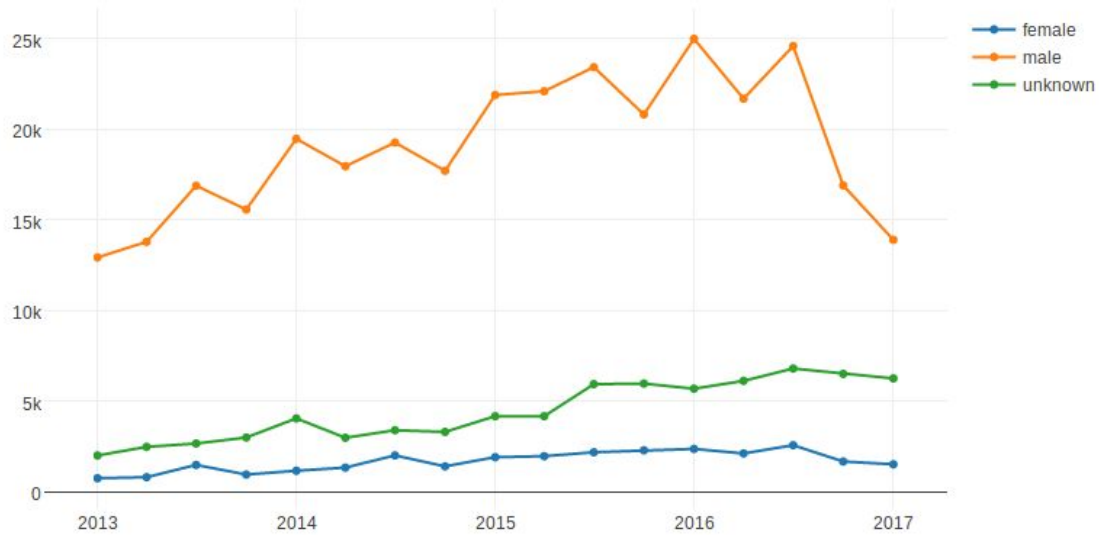


Chart: Number of commits by gender. Source: Git repositories.

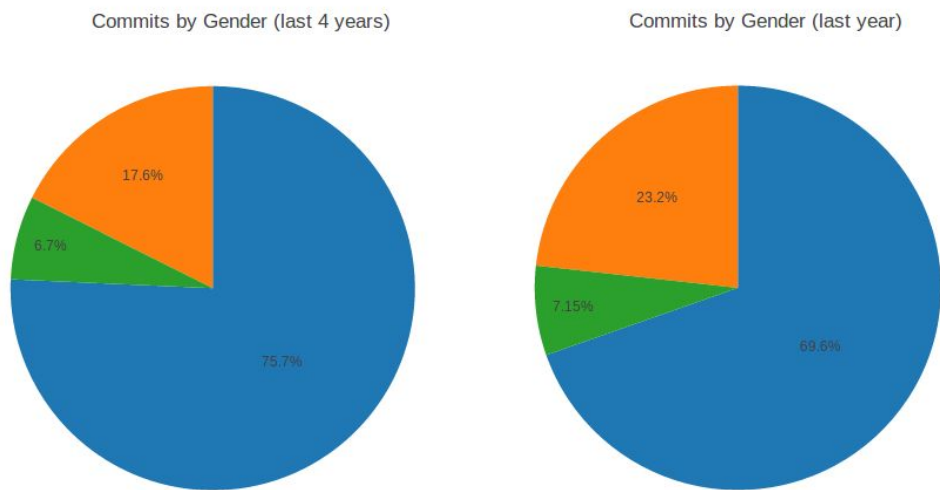


Chart: Number of commits by gender over the last 4 years (left) and over the last year (right). Blue denotes male developers, green denotes female developers, and orange denotes developers unaffiliated with a particular gender. Source: Git repositories.

Developers by Gender

This section provides an account of individuals who have authored commits, by gender, and aggregates this data over the last four years. During this time period, women comprised 9.8% of the total population across OpenStack project teams. Over the last year, this percentage slightly exceeded the aggregate at 10.4%.

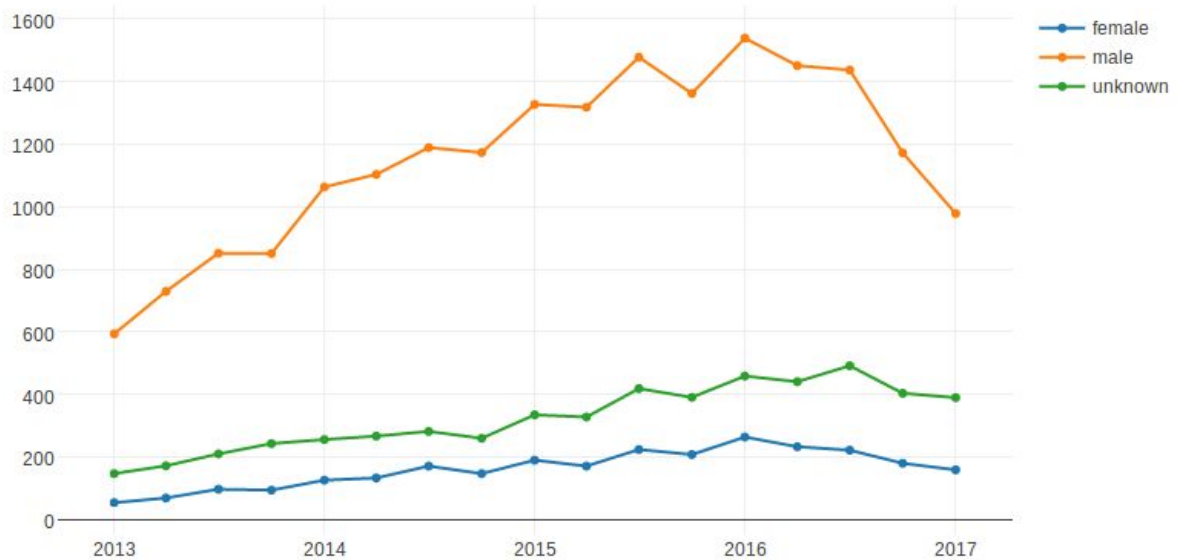
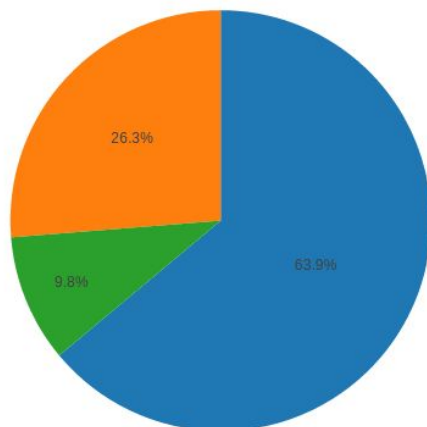


Chart: Number of individuals who have authored commits by gender. Source: Git repositories.

Authors by Gender (last 4 years)



Authors by Gender (last year)

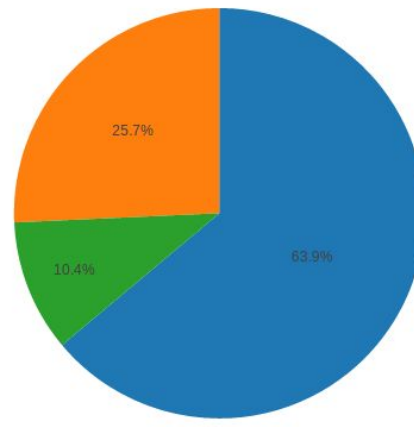


Chart: Number of individuals who have authored commits, by gender, over the last 4 years (left) and over the last year (right). Blue denotes male developers, green denotes female developers, and orange denotes developers unaffiliated with a particular gender. Source: Git repositories.

Types of Contributions

This section illustrates the types of contributions, by gender, as measured by the types of files--code and non-code--displayed in aggregate over the last four years. During this time, data suggests a decrease in overall activity among women, which again, is in line with the general trend of the community.

When comparing the types of contributions as measured by percentage of code and non-code activity, women have been active across both types of activities over the last year, as measured by 7.57% of code files and 7.71% of non-code files. Alternatively, males show a higher prevalence of code activity than non-code activity during this same time period, at 75% and 70.6% respectively.

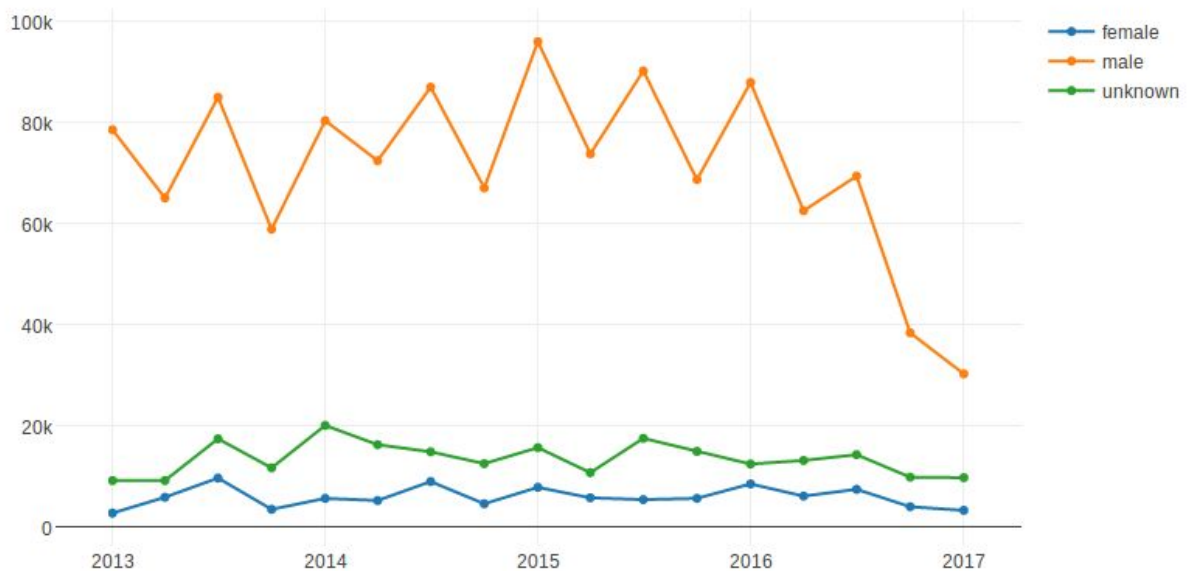


Chart: Number of 'touched' code files by gender. Source: Git repositories.

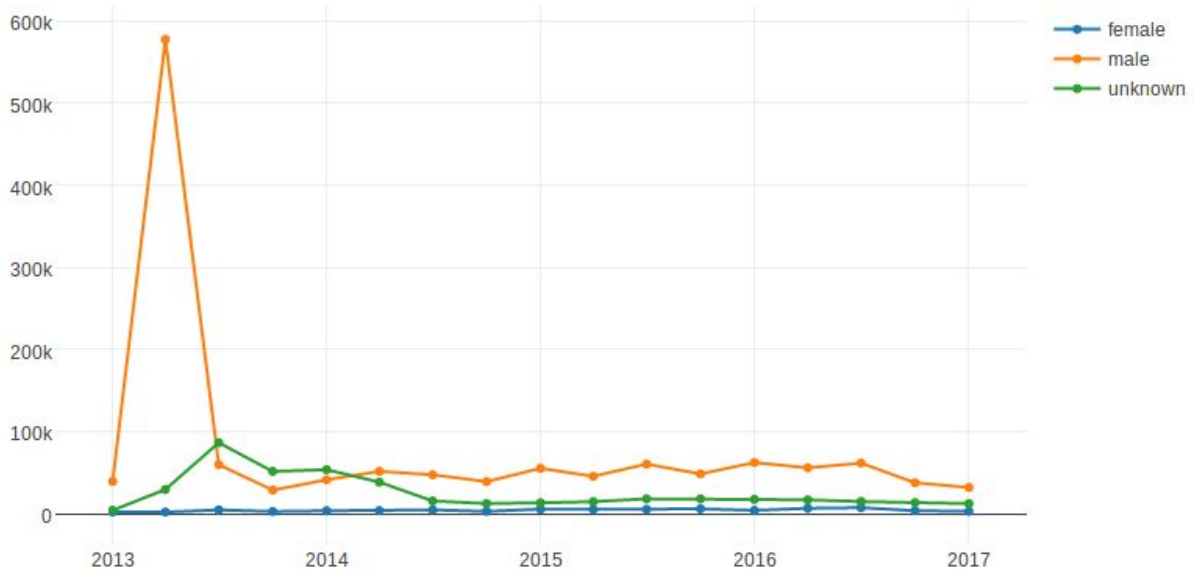


Chart: Number of 'touched' non-code files by gender. Source: Git repositories.

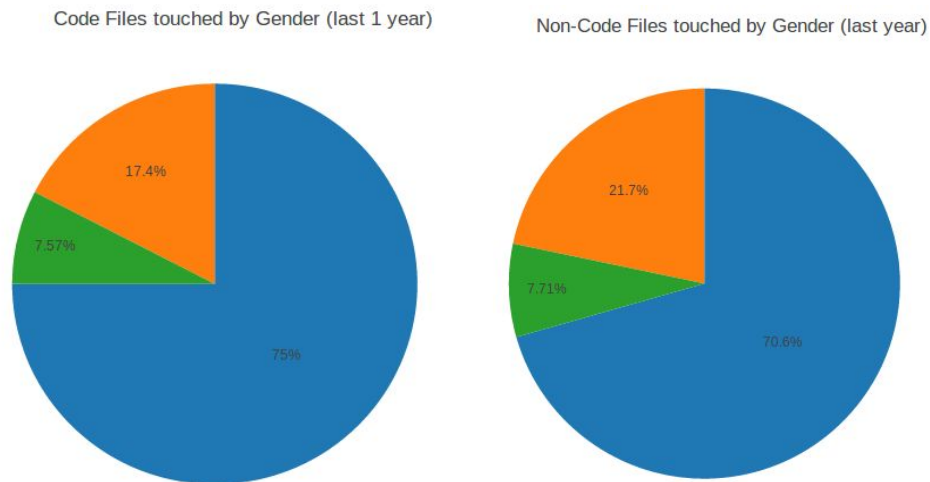


Chart: Percentage of 'touched' code files (left) and 'touched' non-code files (right) for the last year of activity by gender. Source: Git repositories.

Code Review Contributions

This report provides an analysis of code review contributions along the following dimensions:

- Number of code review submissions by gender over time
- Number of developers who have submitted changesets, by gender over time
- Number of code review votes undertaken by developers, by gender over time
- Number of developers who have voted in a code review process, by gender over time
- Number of core code review votes undertaken by developers, by gender over time
- Number of core code review votes undertaken by developers, by gender over time
- Number of developers who have participated in a core code review process, by gender over time

Code Review Submissions

This section reflects the number of changeset submissions by gender, and aggregates this data over the last four years. These numbers do not imply that these changesets have been accepted, only that they have been submitted for review. Within the OpenStack community, 83% of changeset submissions are merged into code, while 17% are abandoned. Of these total submissions, those submitted by women in aggregate over the last four years has reached 8.75%. This percentage over the last year falls slightly below the aggregate at 7.9%. This trend is also reflected among male contributors, with a four-year aggregate of 72.8% compared to 69.8% over the last year.

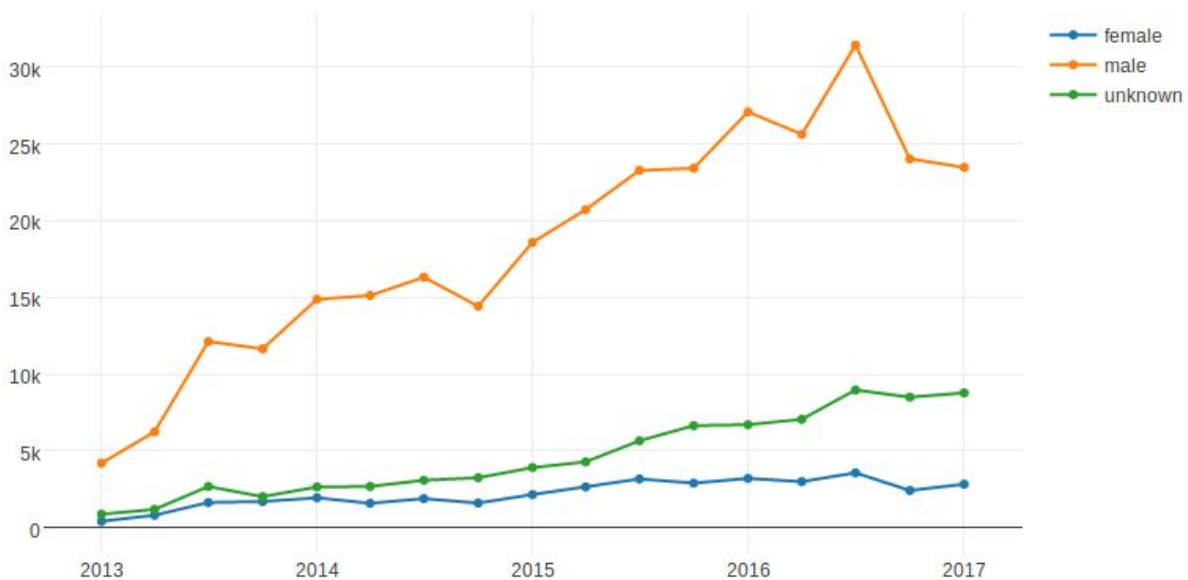
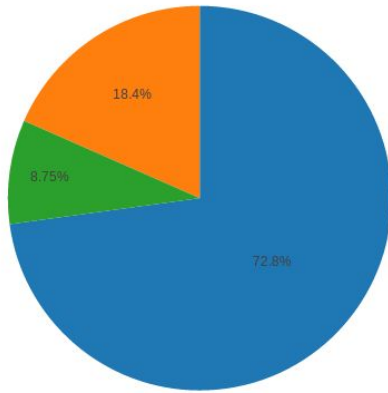


Chart: Number of changeset submissions by gender. Source: Gerrit repositories.

Changeset Submissions by Gender (last 4 years)



Changeset Submissions by Gender (last year)

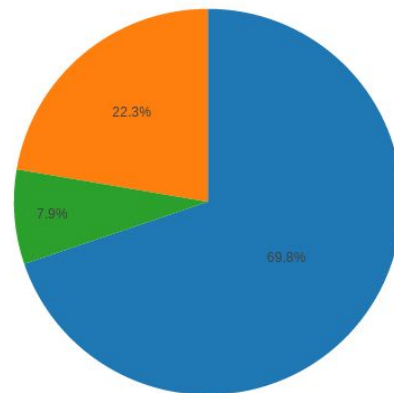


Chart: Number of changeset submissions by gender over the last 4 years (left) and over the last year (right). Blue denotes male developers, green denotes female developers, and orange denotes developers unaffiliated with a particular gender. Source: Gerrit repositories.

Developers Submitting Changesets

This section reveals the number of developers by gender who have submitted changesets in aggregate over the last four years. Of the total population who have submitted code for review over the last four years, women have represented 12.4% in aggregate. This percentage over the last year has fallen slightly under the aggregate at 11.9%.

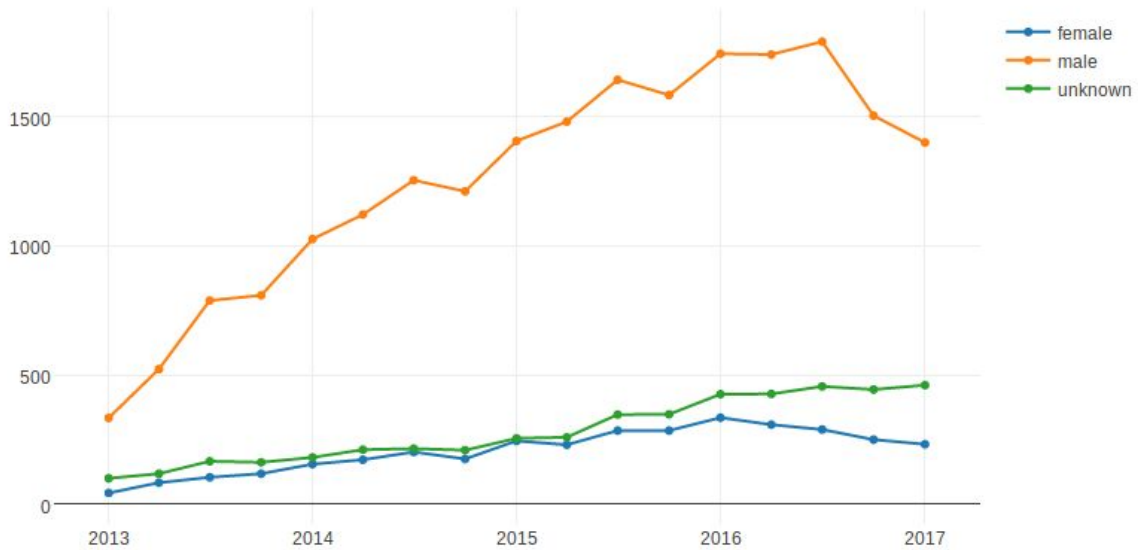
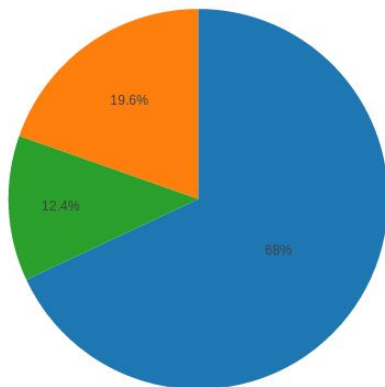


Chart: Number of developers by gender who have submitted changesets. Source: Gerrit repositories.

Changeset Submitters by Gender (last 4 years)



Changeset Submitters by Gender (last year)

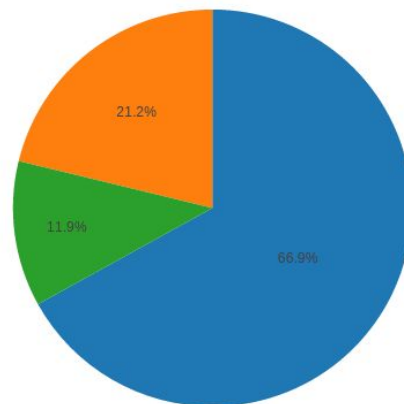


Chart: Number of developers by gender who have submitted changesets over the last 4 years (left) and over the last year (right). Blue denotes male developers, green denotes female developers, and orange denotes developers unaffiliated with a particular gender. Source: Gerrit repositories.

Code Review Votes

This section details the number of code review votes by gender, as well as the number of developers by gender who have voted in a code review process, and aggregates this data over the last four years. A code review is conducted by a developer when a -1, +1, -2 or +2 is provided as a response to a piece of code that has been submitted.

Of the total population who have submitted code review votes, the percentage of women has remained steady at 11.5% when comparing the four-year aggregate with that of the last year. When examining the male population using this same lens, the trend remains consistent, with approximately 1% variance, at 68.9% and 67.5%, respectively. In terms of the activity, the trend is pretty similar as being stable for women at 8.5% of the total votes when comparing the four-year aggregate with the last year activity.

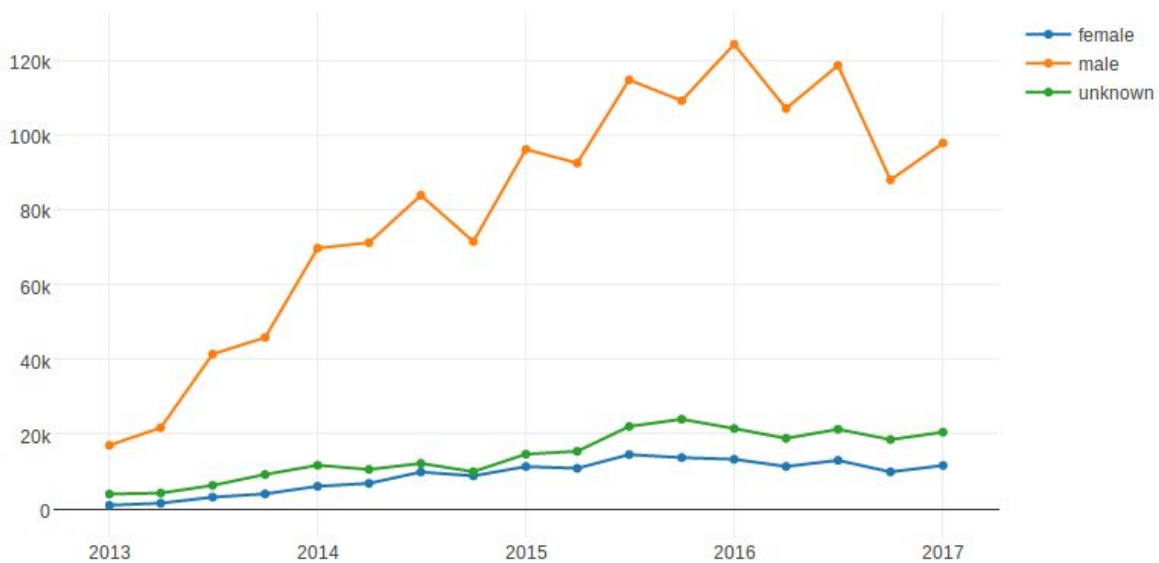


Chart: Number of code review votes by gender (votes as -2, +2, -1 and -1). Source: Gerrit repositories.

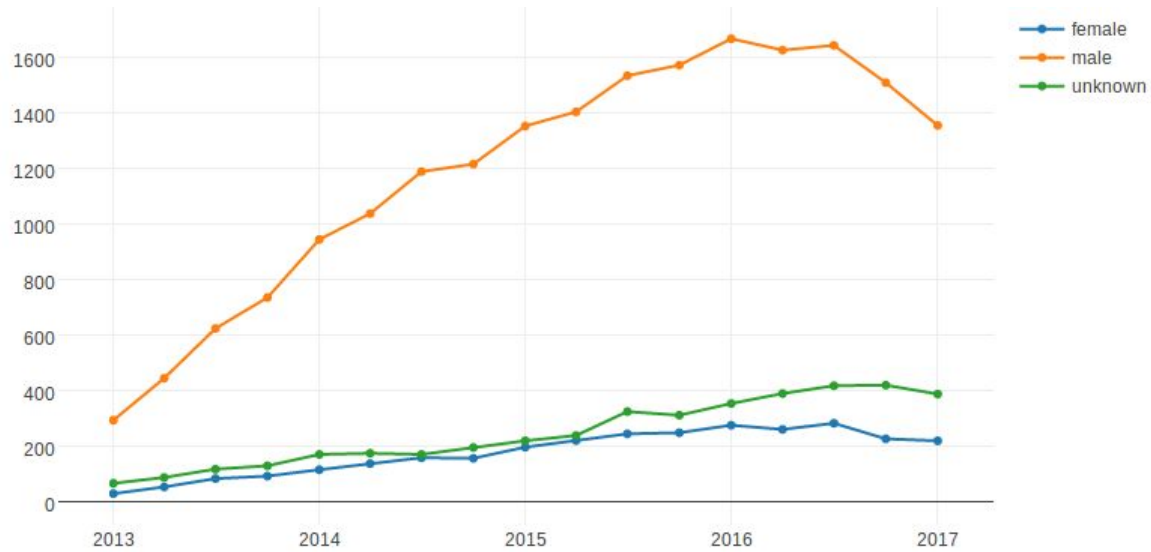


Chart: Number of developers by gender who have voted in a code review process (votes as -2, +2, -1 and -1). Source Gerrit repositories.

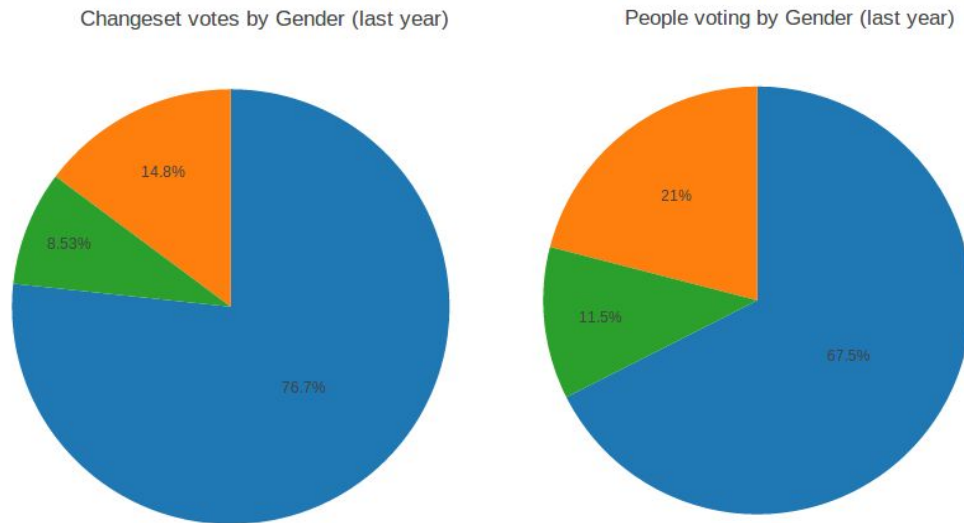


Chart: Number of code review votes by gender (left) and number of people voting in a code review process as -2, -1, +1, +2 (right) for the last year. Last 4 year charts shows a pretty similar trend. Blue denotes male developers, green denotes female developers, and orange denotes developers unaffiliated with a particular gender. Source: Gerrit repositories.

Core Code Review Votes

This section illustrates the number of core code review votes by gender, as well as the number of developers by gender who have participated in a core code review process, and aggregates this data over the last four years. A core code review is delineated as any vote in the Gerrit system, being +2 or -2. Only core reviewers are allowed to vote in core code reviews, and to allow the piece of code to be merged into the master branch or abandoned.

Of the total population who have submitted core code review votes, again, the percentage of women has remained steady at 11-12% (11.9% and 11.1%, specifically) when comparing the four-year aggregate with that of the last year. This trend mirrors that of the male population when making the same comparison, at 73.6% and 74%, respectively.

This trend holds steady when examining the level of activity, or number of core review votes submitted by women. This activity level has remained steady at 7.5% (7.48% and 7.84%, specifically) when comparing the four-year aggregate with that of the last year. Again, this trend is on par with the activity level of the male population using the same lens, at 82% and 81.6%, respectively.

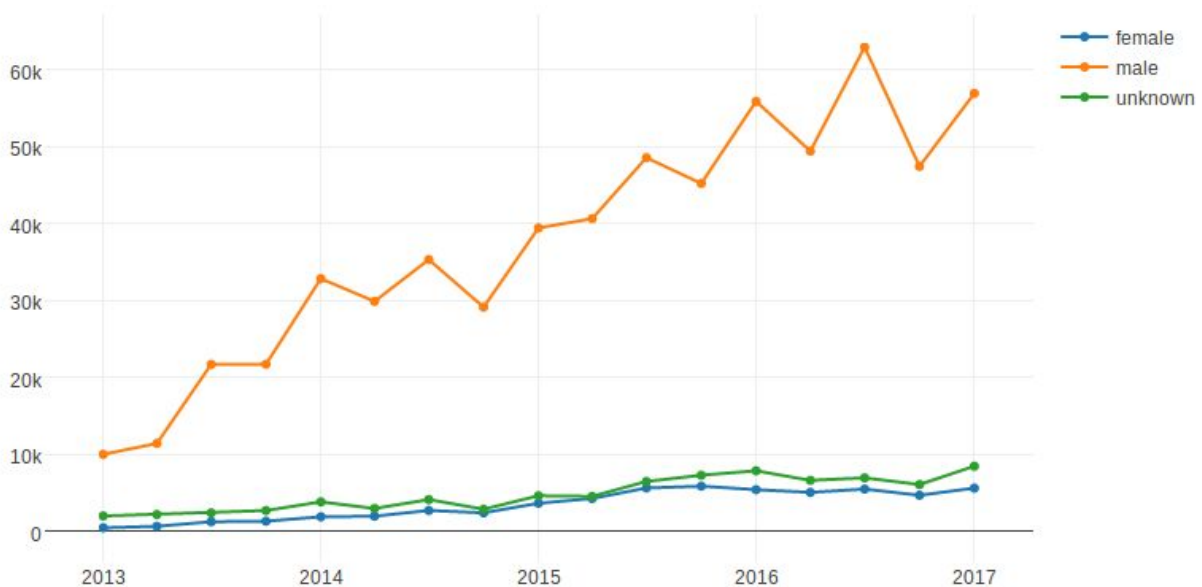
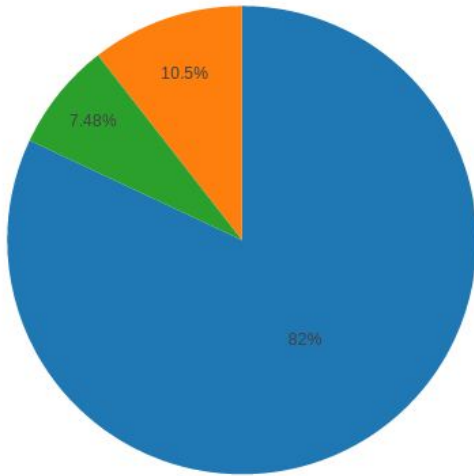


Chart: Number of core code review votes (votes as +2 or -2). Source: Gerrit repositories.



Core Reviews by Gender (last 4 years)



Core Reviews by Gender (last year)

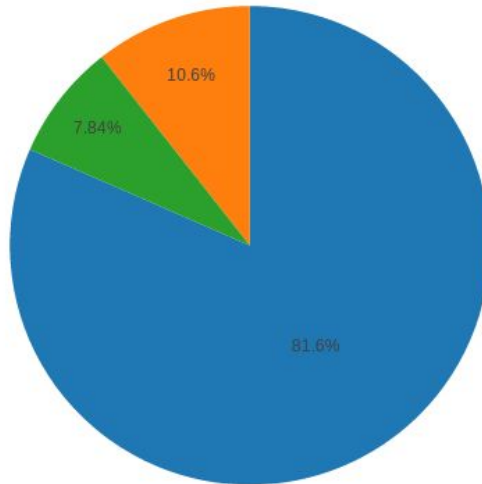


Chart: Core review activity (as -2 or +2) by gender. Last four years of activity (left) and last year of activity (right). Blue denotes male developers, green denotes female developers, and orange denotes developers unaffiliated with a particular gender. Source: Gerrit repositories.

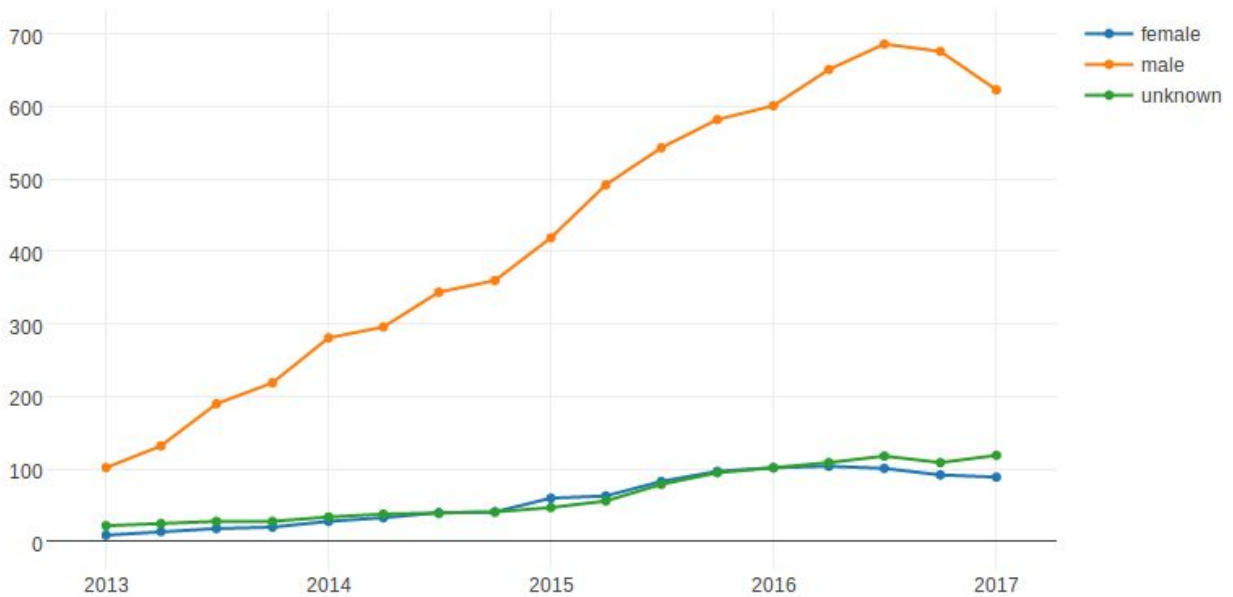


Chart: Number of core reviewers (those allowed to vote +2 or -2). Source: Gerrit repositories.

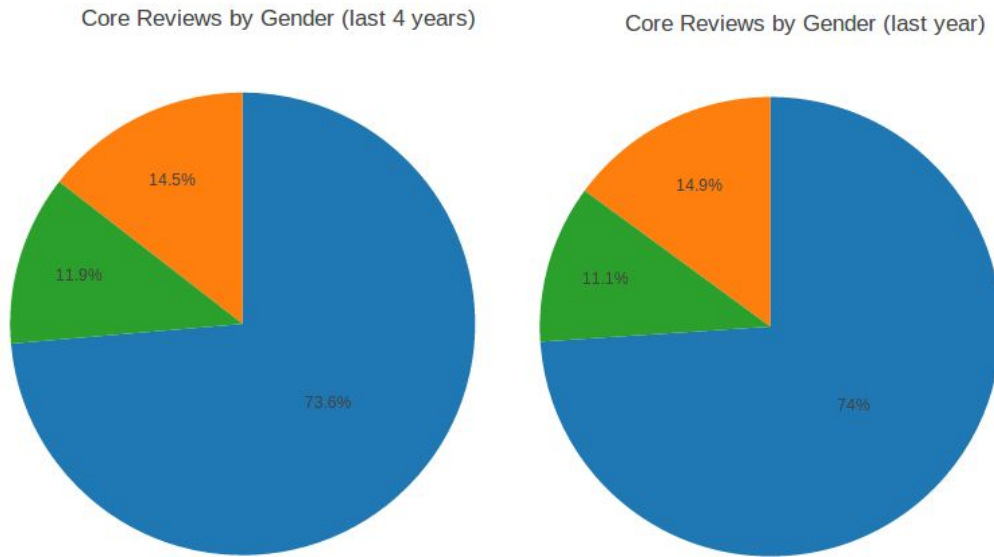


Chart: Core reviewer activity (as -2 or +2) by gender. Last four years of activity (left) and last year of activity (right). Blue denotes male developers, green denotes female developers, and orange denotes developers unaffiliated with a particular gender. Source: Gerrit repositories.

3. Further Work and Recommendations


This research report has examined gender diversity within the OpenStack community from two perspectives: (1) governance/leadership, and (2) technical contributions.

Companies have placed a steady importance on diversity and inclusion to further innovation, beginning with visible measurement and reporting of data in 2012-2013, which has spurred a high degree of focus, accountability and discussion on increasing the numbers of women and underrepresented minorities within the technology industry. A thoughtful study of diversity and inclusion across open source communities like OpenStack, Linux and Apache, in which much of today's innovation and collaboration is taking place, allows us to extend this awareness and prompt collective action to step up and make a difference. This report provides a starting point.

While diversity focuses on bringing women and underrepresented minorities into the community, inclusion emphasizes the importance of environments that welcome and support a diverse community. The data in this report suggests some recommendations as further work, divided into focuses and diversity, followed by inclusion.

Recommendations to increase **gender diversity**:

- Continue to track women's participation in governance/leadership and technical activities within the OpenStack community, while extending this tracking to include other forms of contribution such as marketing, community building, and participation in question-and-answer forums. All contributions--both technical and non-technical--must be recognized, and women often contribute more heavily in non-technical areas.
- Collaborate with the OpenStack project teams with the highest diversity (as measured by the highest number of female developers and highest number of activity from female developers, and other parameters) to collect, document and publicize best practices.
- Study the impact of specific, diversity-related policies and initiatives undertaken by the OpenStack Foundation to determine their impact on the pipeline and entrance of women and underrepresented minorities into the community.
- Work with the community to understand their needs more deeply, and build follow-on action plans based on this data to address these needs.
- Assist PTLs who express a desire for greater diversity within their project team with recruitment activities to achieve this stated objective.



Recommendations to foster **more inclusive teams and communities**:

- Track both the tenure and attrition of women in the OpenStack community, and study the impact of specific policies and initiatives undertaken by the OpenStack Foundation, such as the Outreachy mentor program, to determine their impact on these factors. It is important to create an inviting and supportive environment for these individuals to land, once they have decided to join the community.
- Collaborate with the OpenStack project teams with the most inclusive environments (as measured by the highest tenures and retention rates among women and underrepresented minorities, and other parameters) to collect, document and publicize best practices.
- Invest in increasing the number of women and underrepresented minorities who participate in technical leadership, such as the Technical and User Committees, as well as PTL positions. Consider a mentorship program that aims to provide mentorship and shadowing opportunities to women with PTL potential. Consider extending invitations to these females and underrepresented minorities to attend and observe technical meetings, so that they gain first-hand experience and knowledge about how technical leadership teams work.
- Continue to support the Women of OpenStack program and associated onboarding activities, such as the Upstream University, Long-Term Mentoring and Speed Mentoring programs, and event speakerships, with diversity as a foundation. These programs have all been well accepted and are succeeding in increasing diversity and inclusion in the OpenStack community.
- Develop ways to ensure that the community is well aware of how important diversity and inclusion are to the leadership of the OpenStack Foundation and to the success of the OpenStack Project.
- Continue to enforce the OpenStack Foundation's code of conduct to reinforce the importance of diversity and inclusion within OpenStack project teams.
- Lastly, it has been shown that inclusive communities have good documentation, onboarding processes and mentors. Ensure all projects have these elements as a baseline.

Appendix A: Detailed Summary of Female Developer Activity

This appendix tracks the activity by female developers within each of the OpenStack project teams.

Each of the columns within the table below are computed as described below:

- **Ratio_Authors:** The number in this column denotes the ratio of 100 female developers for every 100 male developers for a given project.
- **Ratio_Commits:** The number in this column denotes the ratio of 100 commits submitted by female developers for every 100 commits submitted by male developers for a given project.
- **Authors:** The number in this column denotes the total number of individuals identified as females who have contributed to a given project, as tracked through the Git repositories.
- **Commits:** The number in this column denotes the total number of changes to the source code for a given project. A commit is usually submitted through a code review process.

project	authors	commits	ratio_authors	ratio_commits
Packaging-deb	623.0	13175.0	14.361457	8.536736
Documentation	145.0	3080.0	20.393812	37.211550
Infrastructure	145.0	2303.0	11.214230	4.849442
nova	121.0	891.0	16.285330	6.913944
Quality Assurance	101.0	650.0	13.760218	7.747318
neutron	99.0	858.0	15.840000	7.279824
horizon	96.0	949.0	22.748815	25.333689
cinder	77.0	532.0	13.898917	9.646419
oslo	59.0	416.0	11.706349	4.804250
Telemetry	57.0	281.0	21.590909	5.624500
glance	57.0	234.0	18.811881	10.077519
keystone	56.0	621.0	15.217391	10.482782
heat	54.0	217.0	15.652174	3.174371
fuel	48.0	4518.0	14.634146	17.884570
octavia	44.0	312.0	11.956522	8.208366
ironic	43.0	1172.0	14.478114	24.209874
requirements	41.0	109.0	11.581921	5.025357
tripleo	40.0	663.0	11.204482	4.850746
OpenStackClient	37.0	385.0	18.407960	19.201995
swift	35.0	209.0	14.583333	7.706490
tacker	33.0	180.0	22.758621	22.988506
Puppet OpenStack	33.0	894.0	9.455587	11.149913
kolla	32.0	154.0	16.410256	3.985507
trove	32.0	214.0	19.277108	11.536388
rally	31.0	190.0	15.196078	9.099617
murano	31.0	1002.0	21.379310	27.073764
sahara	28.0	379.0	15.819209	8.212351
barbican	27.0	142.0	19.424460	8.722359
Security	26.0	117.0	22.033898	7.758621
magnum	26.0	315.0	17.931034	22.292994
manila	24.0	282.0	16.216216	15.537190
OpenStackAnsible	23.0	377.0	13.372093	3.270866
Chef OpenStack	20.0	64.0	19.417476	2.550817
Release Management	19.0	49.0	11.377246	1.724745
zaqar	19.0	184.0	17.757009	18.852459
monasca	18.0	167.0	13.533835	4.693648
senlin	16.0	171.0	32.653061	5.765341
dragonflow	13.0	226.0	30.952381	25.799087
mistral	13.0	138.0	13.000000	8.151211
watcher	13.0	46.0	23.636364	9.368635
vitrage	11.0	167.0	45.833333	28.842832
zun	10.0	97.0	38.461538	47.783251

designate	10.0	59.0	7.751938	5.263158
congress	10.0	239.0	12.987013	24.044266
kuryr	9.0	111.0	19.565217	33.035714
OpenStack Charms	9.0	639.0	8.181818	3.438072
notavailable	9.0	36.0	8.108108	1.066983
tricircle	8.0	41.0	72.727273	128.125000
karbor	8.0	12.0	38.095238	7.272727
RefStack	7.0	54.0	18.421053	12.676056
searchlight	7.0	41.0	14.285714	8.932462
Community App Catalog	6.0	17.0	10.714286	4.632153
solum	6.0	45.0	10.169492	5.415162
freezer	6.0	15.0	12.000000	1.726122
winstackers	5.0	33.0	15.151515	7.951807
cloudkitty	4.0	6.0	10.256410	1.973684
storlets	3.0	47.0	16.666667	16.607774
Packaging-rpm	2.0	22.0	6.896552	2.906209
ec2-api	2.0	5.0	13.333333	1.798561
I18n	1.0	18.0	7.142857	20.000000

Table: List of female developer activity within OpenStack project teams.

Appendix B: Technical Details and Limitations

This appendix outlines the sources and methodologies for the research within this report. It also identifies potential limitations of this research, and proposes ways to counter these limitations.

- The data for this report from Git repositories is retrieved from the governance file using the command below, and is parsed and stored in an ElasticSearch instance:
`git log --raw --numstat --pretty=fuller --decorate=full --parents --reverse --topo-order -M -C -c --remotes=origin --all`
- The data for this report from Gerrit repositories is retrieved from the governance file using an SSH interface, and is parsed and stored in an ElasticSearch database.
- The data for this report from both Git and Gerrit repositories is retrieved using Perceval⁹, a 100% open source software tool under the GrimoireLab¹⁰ umbrella.
- The code used to enrich and visualize the raw information in this report is provided by Perceval, which is available in the GitHub repository of one of the authors of this study¹¹. Charts and tables can be visualized in the same repository¹².
- In this report, gender identification is based on the individual's first name, which relies on the genderize.io API. However, this report also required manual manipulation of the datasets (e.g. manual updates and improvements), which proved to be extremely time-consuming. One of the goals of this project is to capture an increasingly more curated dataset, which will require assistance and support from the community. The need for protection of privacy of the data brings complexities, such as the need for handling of the data by a third party and the restriction of access to the dataset.
- In this report, technical contributions are defined as commits, code reviews and code review vote actions. This is a starting point for this ongoing research, based on input from the community. Other sources for measurement of female activity, engagement and contributions can be added in the future, such as mailing lists, wiki editions, and more.
- The analysis in the Governance/Leadership section of this report relies on data in the OpenStack Foundation governance file¹³ and the wiki sites that contain information about the Board of Directors, the Technical and User Committees and others. Therefore, the


⁹ <https://github.com/grimoirelab/perceval>

¹⁰ <http://grimoirelab.github.io/>

¹¹ <https://github.com/dicortazar/ipython-notebooks/tree/master/projects/openstack-diversity>

¹² <https://htmlpreview.github.io/?https://github.com/dicortazar/ipython-notebooks/blob/master/projects/openstack-diversity/OpenStack%20Diversity%20Metrics.html>

¹³ <http://git.openstack.org/cgit/openstack/governance/tree/reference/projects.yaml>



analysis of OpenStack governance/leadership is only as accurate and updated as these sources are.