

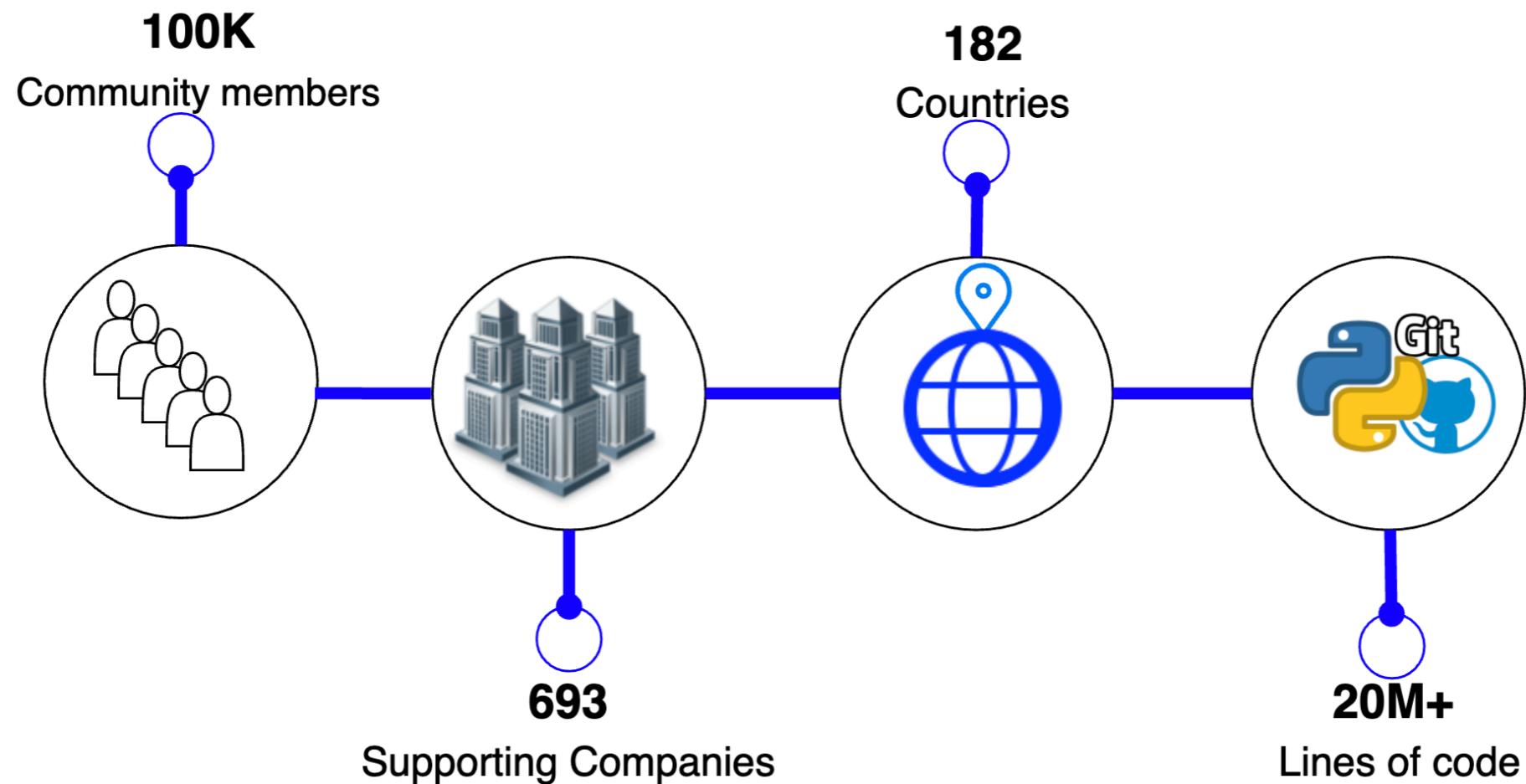
# Onboarding vs. Productivity, Diversity and Quality – Empirical Study of the OpenStack Ecosystem

Armstrong Foundjem, Ellis Eghan, Bram Adams

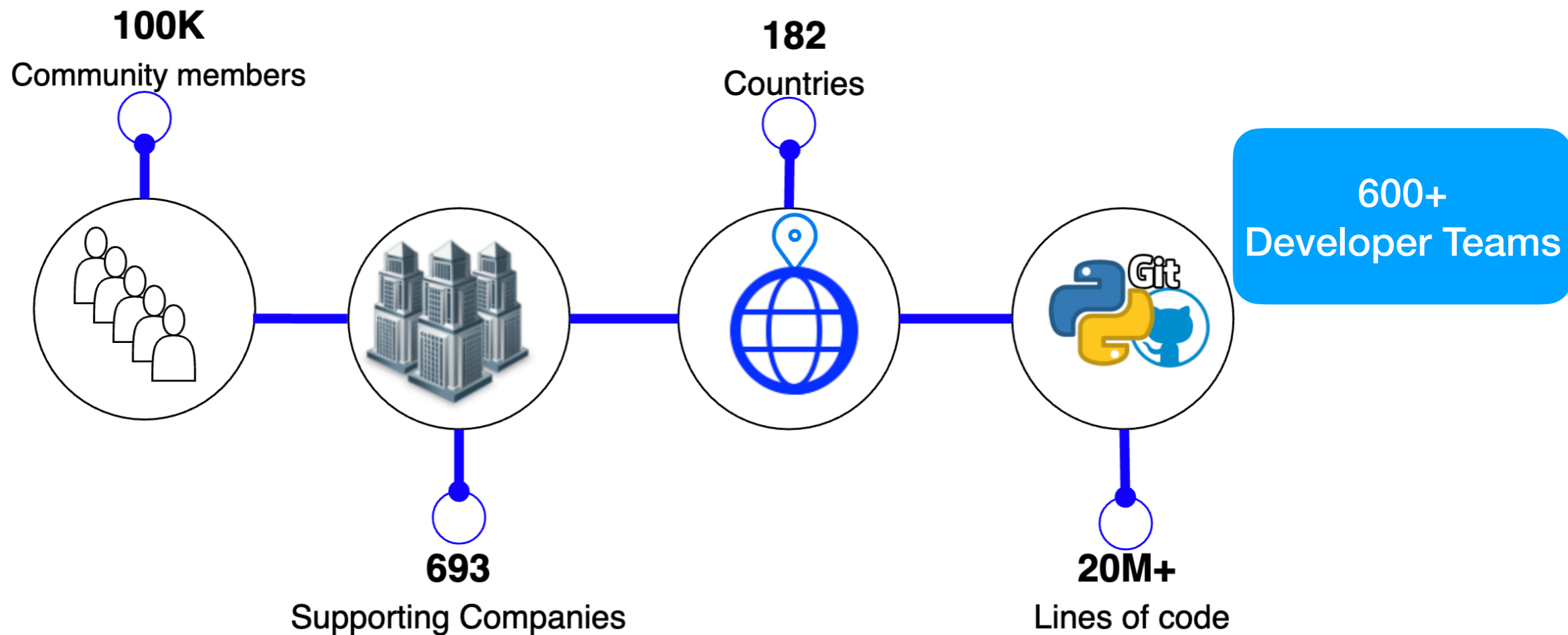
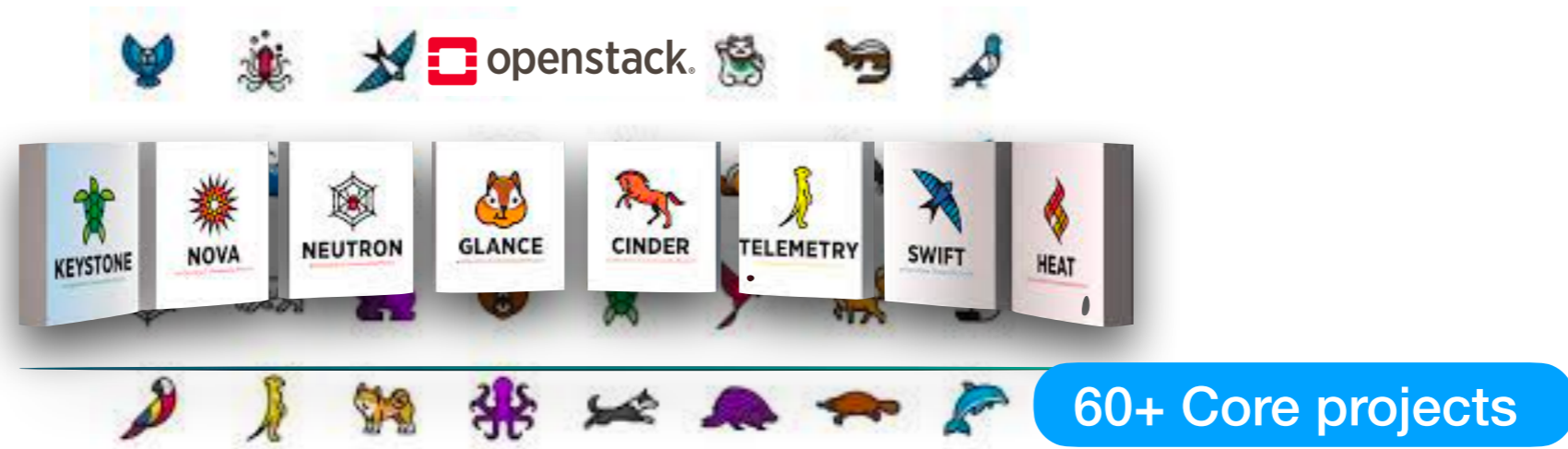
MCIS Laboratory  
a.foundjem@queensu.ca

ICSE 2021

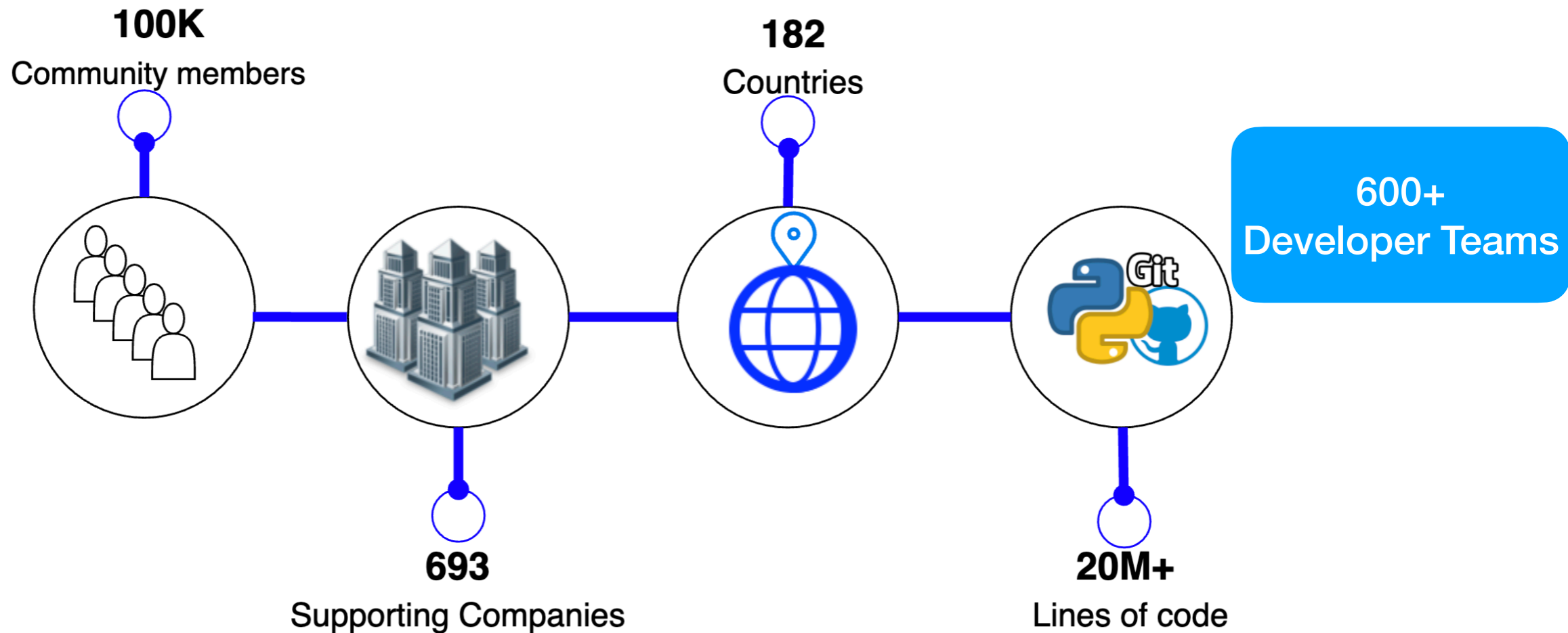
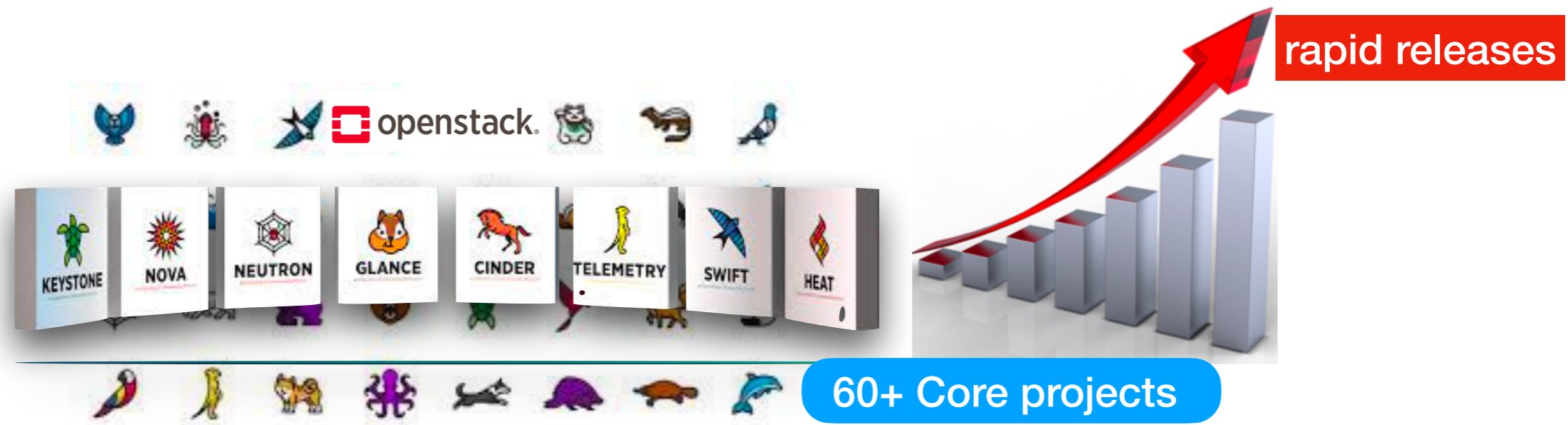
# OpenStack: An Open Source Software Ecosystem (SECO)



# OpenStack: An Open Source Software Ecosystem (SECO)



# OpenStack: An Open Source Software Ecosystem (SECO)



# How to Start Contributing to Such A SECO?

# How to Start Contributing to Such A SECO?



<https://languages.oup.com/google-dictionary-en/>

# Onboarding Program: A Socio-Technical Adaptation of New Contributors to a Software Community.



## Onboarding Pass



**on-board·ing** | 'än bôrdiNG, ä'n'bôrdiNG |

*noun*

The action or process of integrating a new employee into an organization or familiarizing a new customer or client with one's products or services: after the initial onboarding is complete, continue to offer new hires relevant training and development opportunities | *client onboarding is a critical time for any business* | [as modifier] : **a little extra effort during the onboarding period can go a long way towards improving staff retention.**



<https://languages.oup.com/google-dictionary-en/>

# Onboarding Program: A Socio-Technical Adaptation of New Contributors to a Software Community.



Thus far, all research was on onboarding in individual projects, what about SECOs?



<https://languages.oup.com/google-dictionary-en/>



# SECO-Level Onboarding Involves Two Phases

Phase 1



SECO-Level: In-person Event

Our Paper: <https://2021.msrf.org/profile/armstrongfoundjem1>

# SECO-Level Onboarding Involves Two Phases

Phase 1



SECO-Level: In-person Event

Phase 2



Project-Level: Remote One-One Mentoring

Our Paper: <https://2021.msrfconf.org/profile/armstrongfoundjem1>

# Our Goal: To Understand the Process and Impacts of Onboarding Program in Open-source SECOs



## Phase 1

Observation of 72 SECO-level onboarding event participants:

To identify the **activities**, **challenges** and **perceived benefits** of SECO onboarding on new Contributors.

Our Paper: <https://2021.msrfconf.org/profile/armstrongfoundjem1>

# Our Goal: To Understand the Process and Impacts of Onboarding Program in Open-source SECOs



## Phase 1

Observation of 72 SECO-level onboarding event participants:

To identify the **activities**, **challenges** and **perceived benefits** of SECO onboarding on new Contributors.

## Phase 2

Quantitative study of 1281 contributors:

To validate the **perceived benefits** in terms of contributors' **diversity**, **productivity**, and **quality**.

Our Paper: <https://2021.msrfconf.org/profile/armstrongfoundjem1>

# Our Goal: To Understand the Process and Impacts of Onboarding Program in Open-source SECOs



## Phase 1

Observation of 72 SECO-level onboarding event participants:

To identify the **activities**, **challenges** and **perceived benefits** of SECO onboarding on new Contributors.

## Phase 2

Quantitative study of 1281 contributors:

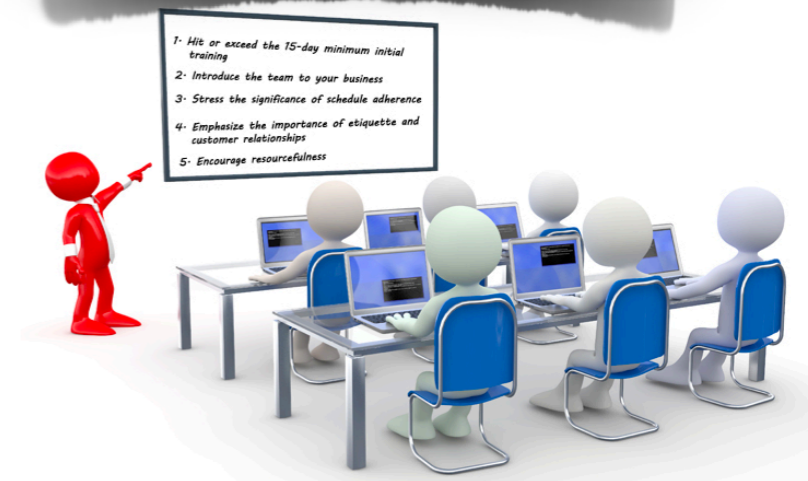
To validate the **perceived benefits** in terms of contributors' **diversity**, **productivity**, and **quality**.

Our Paper: <https://2021.msrfconf.org/profile/armstrongfoundjem1>

# Extracting Rich Set of Qualitative Data



## Theoretical knowledge



- 1- Hit or exceed the 15-day minimum initial training
- 2- Introduce the team to your business
- 3- Stress the significance of schedule adherence
- 4- Emphasize the importance of etiquette and customer relationships
- 5- Encourage resourcefulness

## Day 1





# Extracting Rich Set of Qualitative Data



Hands-on sessions

Day 2



# Identification of 43 Main Themes From the Audio-Visual Data

Categorizing codes/themes using an Affinity diagram

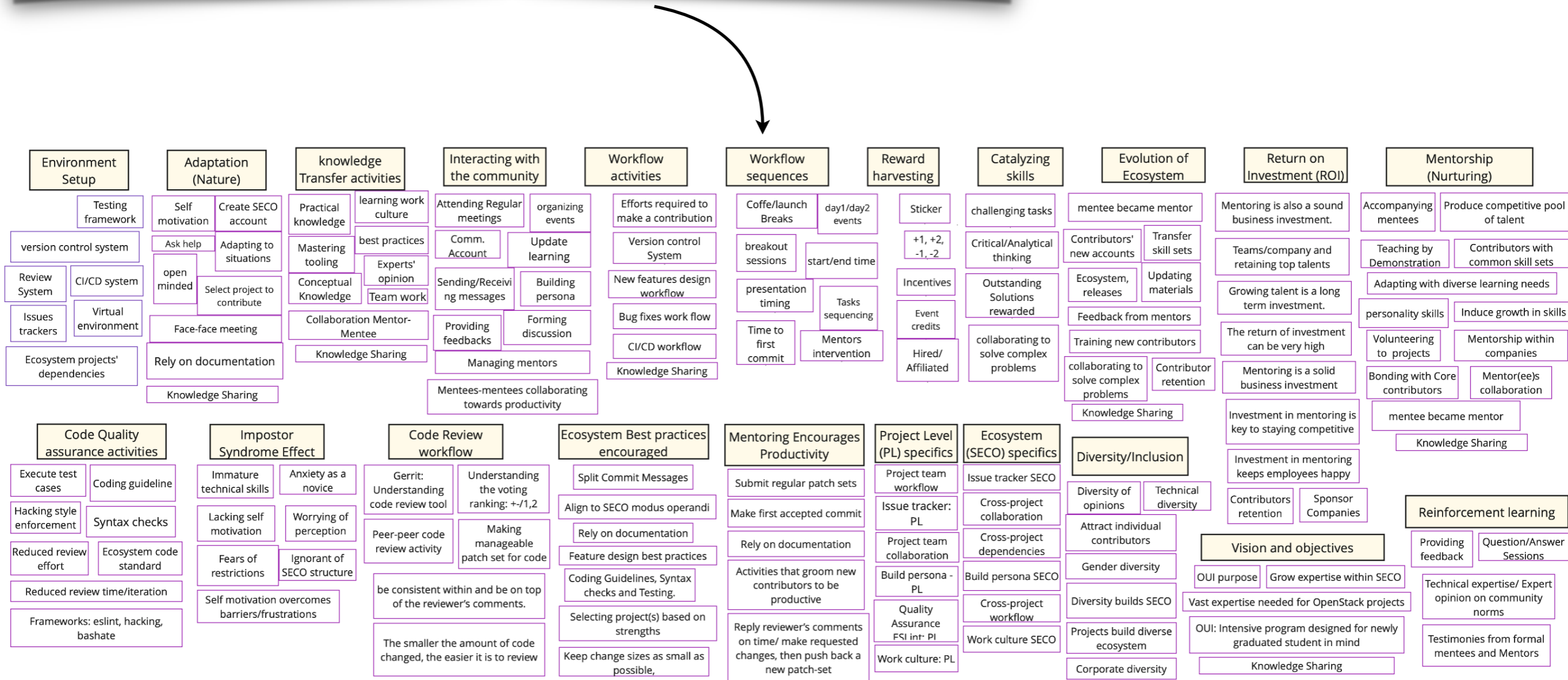


Our Paper: <https://2021.msconf.org/profile/armstrongfoundjem1>



# Identification of 43 Main Themes From the Audio-Visual Data

Categorizing codes/themes using an Affinity diagram



Separate Themes under different Categories

# Three Categories Emerged From the 43 Main Themes

**22** Teaching Content/  
Strategies

**13** Onboarding  
Challenges

**8** Onboarding  
Benefits

Our Paper: <https://2021.msrf.org/profile/armstrongfoundjem1>

# Category #1: Teaching Content Strategies

**22** Teaching Content/  
Strategies

13 Onboarding  
Challenges

8 Onboarding  
Benefits

Conceptual Knowledge (7)

Example Codes:

... SECO processes and tools concepts

# Category #1: Teaching Content Strategies

**22** Teaching Content/  
Strategies

13 Onboarding  
Challenges

8 Onboarding  
Benefits

**Example Codes:**

Conceptual Knowledge (7)

... SECO processes and tools concepts

Practical Knowledge (7)

... Code quality assurance activities

# Category #1: Teaching Content Strategies

**22** Teaching Content/  
Strategies

13 Onboarding  
Challenges

8 Onboarding  
Benefits

**Example Codes:**

Conceptual Knowledge (7)

... SECO processes and tools concepts

Practical Knowledge (7)

... Code quality assurance activities

Strategies (8)

... Adaptive teaching strategies

# Category #2: Onboarding Challenges for SECO, Companies and Contributors

22

Teaching Content/  
Strategies

13

Onboarding  
Challenges

8

Onboarding  
Benefits

... Vast Expertise needed for SECO projects

... Mentorship sustainability

... Soundness of Return on Investment

# Category #3: Onboarding Benefits for SECO, Companies and Contributors

22

Teaching Content/  
Strategies

13

Onboarding  
Challenges

8

Onboarding  
Benefits

**Mentoring Enhances Diversity and Inclusion**

**Mentoring Enhances Productivity**

**Mentoring Enhances SECO Best Practices and Quality assurance**

# We provide quantitative evidences to three Most Discussed benefits

22

Teaching Content/  
Strategies

13

Onboarding  
Challenges

8

Onboarding  
Benefits

Mentoring Enhances Diversity and Inclusion

Mentoring Enhances Productivity

Mentoring Enhances SECO Best Practices and Quality assurance



# Our Goal: To Understand the Process and Impacts of Onboarding Program in Open-source SECOs



## Phase 1

Observation of 72 SECO-level onboarding event participants:

To identify the **activities**, **challenges** and **perceived benefits** of SECO onboarding on new Contributors.

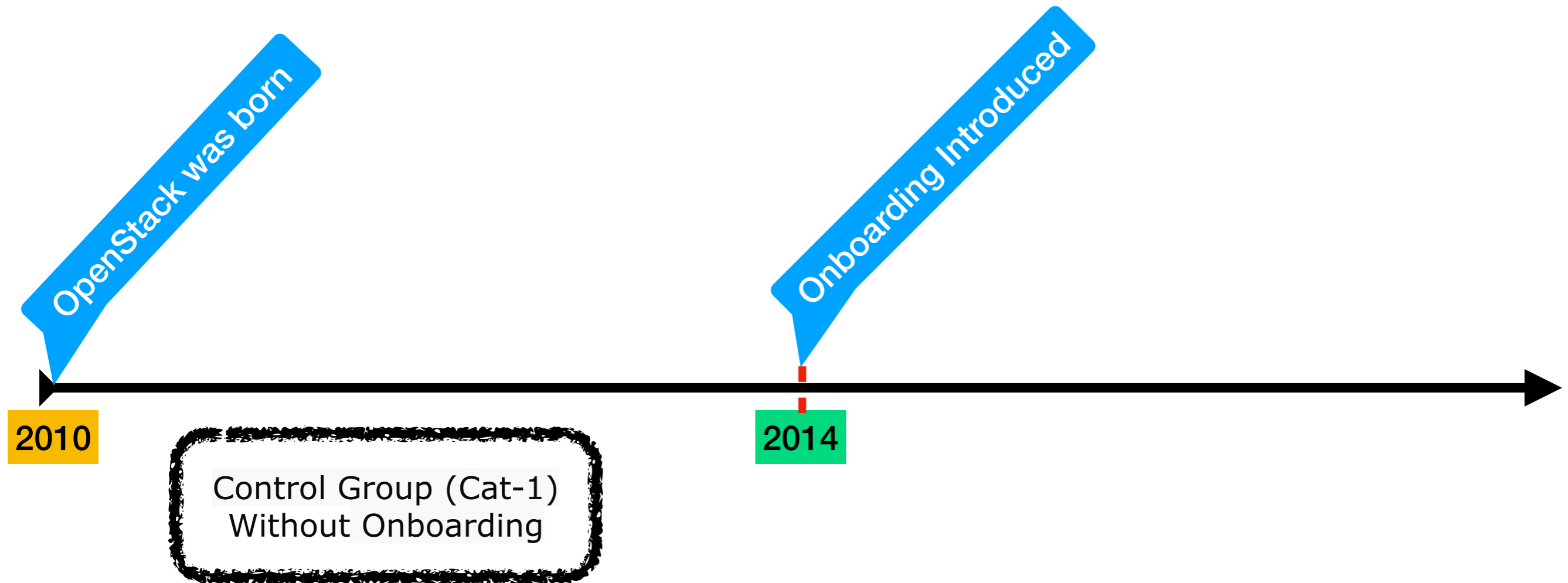
## Phase 2

Quantitative study of 1281 contributors:

To validate the **perceived benefits** in terms of contributors' **diversity**, **productivity**, and **quality**.

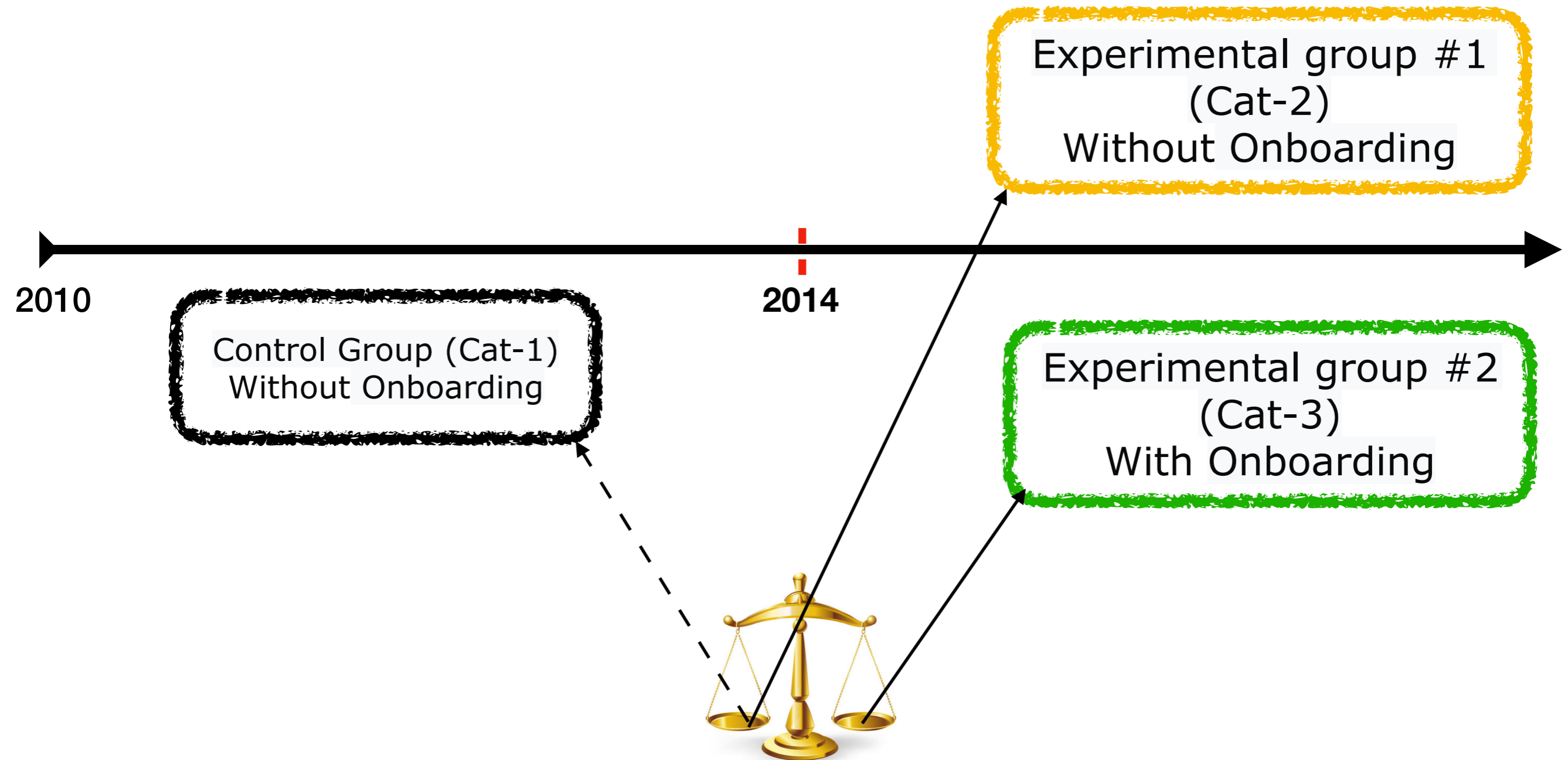
Our Paper: <https://2021.msrfconf.org/profile/armstrongfoundjem1>

# Quantitative Data Set: 3-year of Development data of 1281 Contributors

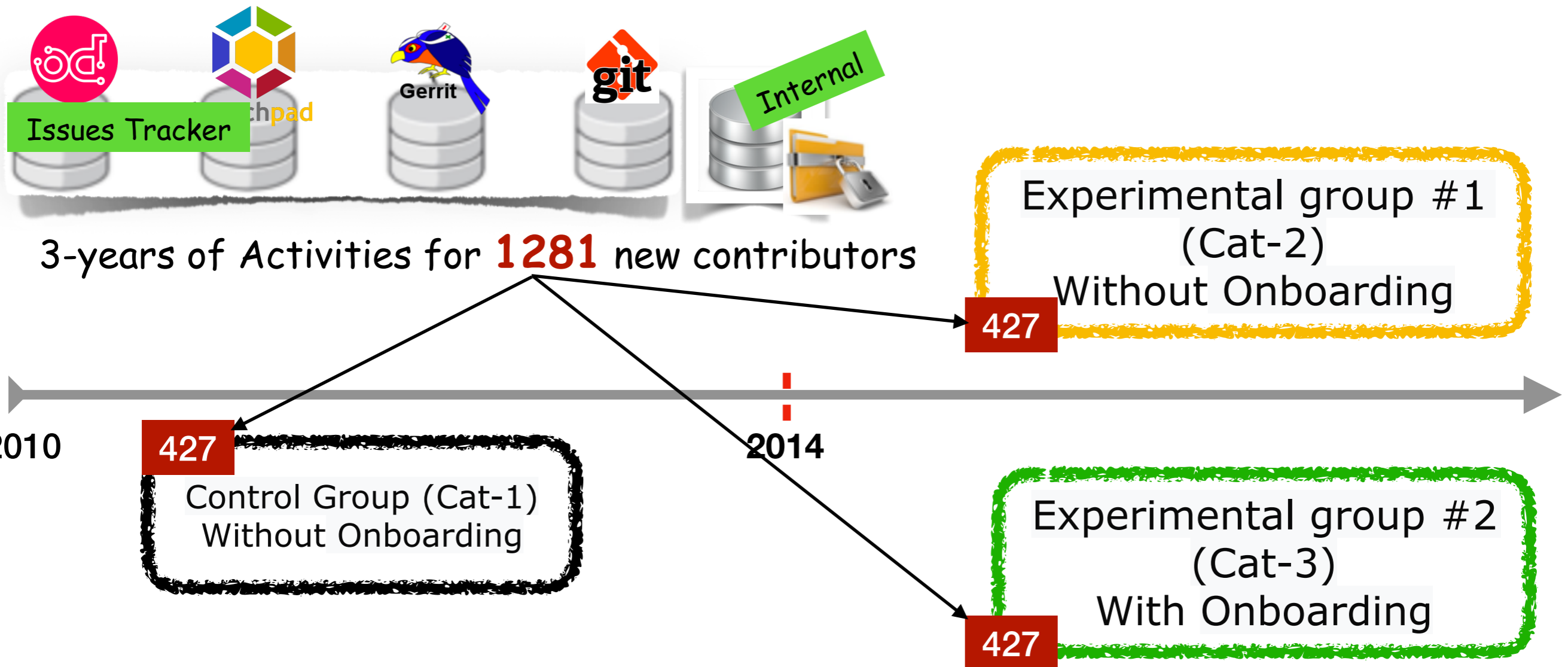


Our Paper: <https://2021.msrf.org/profile/armstrongfoundjem1>

# Quantitative Data Set: 3-year of Development data of 1281 Contributors

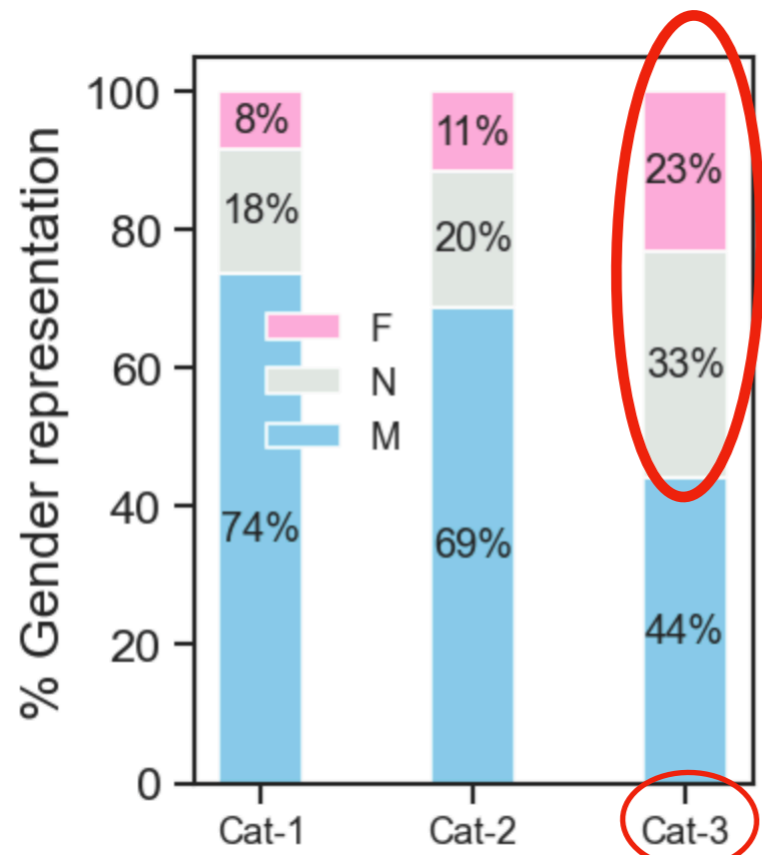


# Quantitative Data Set: 3-year of Development data of 1281 Contributors



Our Paper: <https://2021.msrfconf.org/profile/armstrongfoundjem1>

# 65% more gender diversity in Cat-3



**Cat-3 Female & non-binary 65% higher Compared to Cat-2**

Internal Datasets

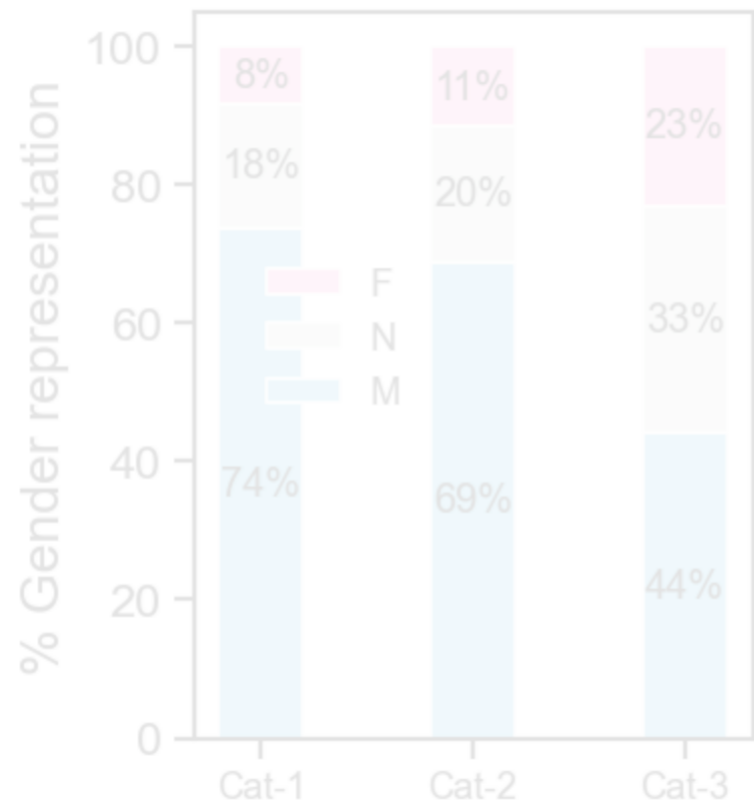


2: Median days-to-hire for Cat-3 is 33.0 vs. 212.0 days for Cat-2



3: Cat-3 (84.6%  $\geq$  3 projects), while Cat-2 (82.7% = 1 project)

# Cat-3 Contributors get Hired more Quickly by SECO Companies



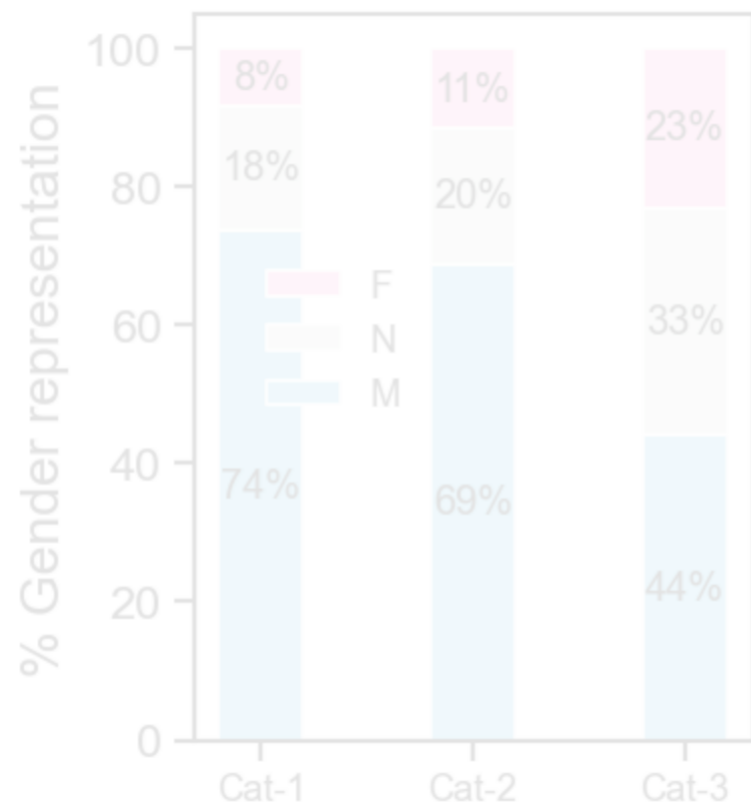
Cat-3 Female & non-binary 65% higher Compared to Cat-2

2: Median days-to-hire for Cat-3 is 33.0 vs. 212.0 days for Cat-2



3: Cat-3 (84.6%  $\geq$  3 projects), while Cat-2 (82.7% = 1 project)

# Onboarded Contributors are Active in more Projects



1: Cat-3 Female & non-binary 65% Compared to Cat-2

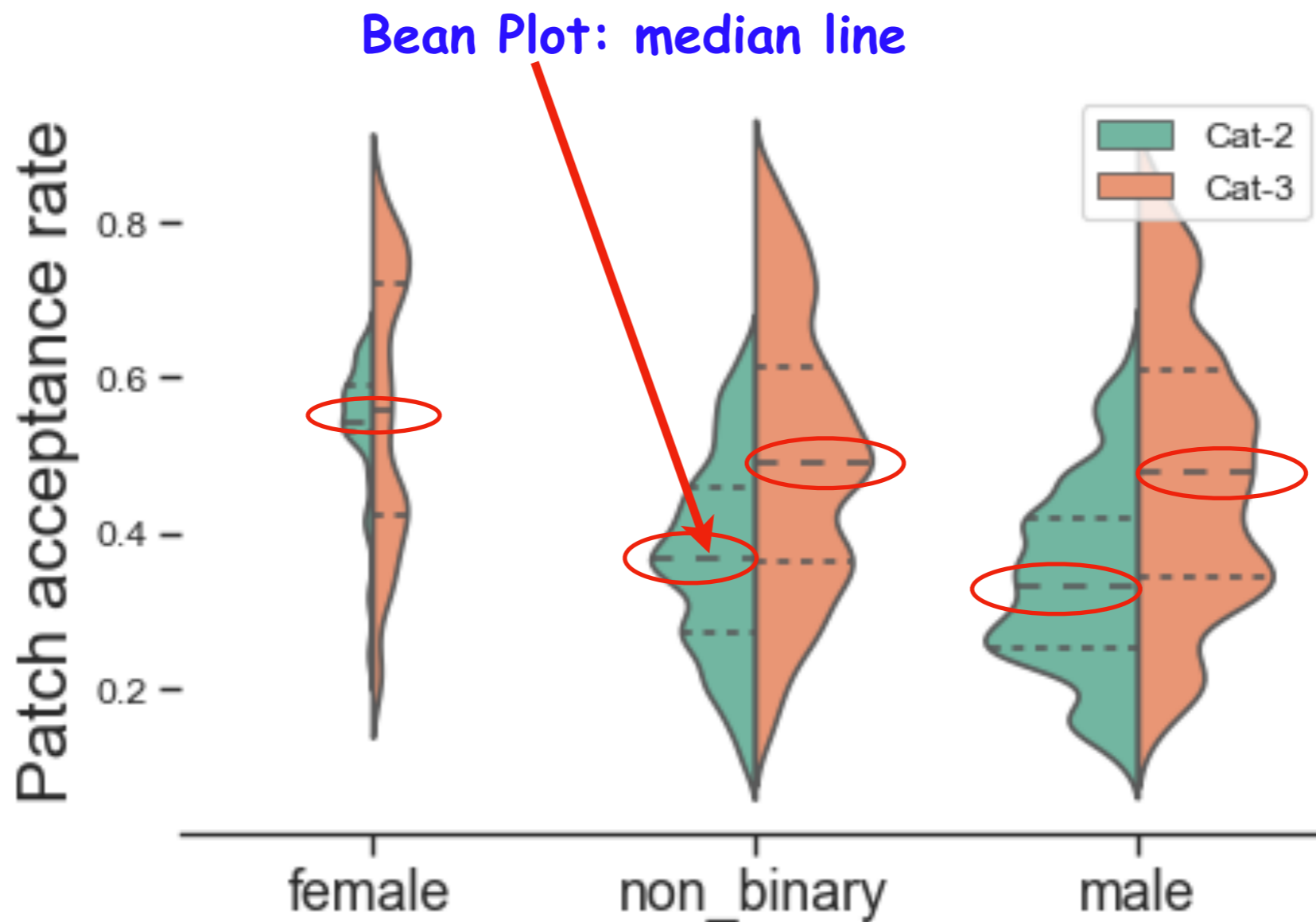
2: Median days-to-hire for Cat-3 is 33.0 vs. 212.0 days for Cat-2



3: Cat-3 (84.6%  $\geq$  3 projects), while Cat-2 (82.7% = 1 project)

# Onboarded Participants have higher Patch Acceptance Rate (PAR)

$$PAR = \frac{\#Accepted\_PRs}{\#Submitted\_PRs}$$



Cat-3 contributors saw an increase in terms of PAR.

In both Categories, females outperformed the other genders with large effect size.



# Onboarded Contributors Contribute More Quickly, and Stay Longer

## Retention Rate

Onboarding correlates with a 16% longer average retention rate.

## Time-to-First Commit (TFC)

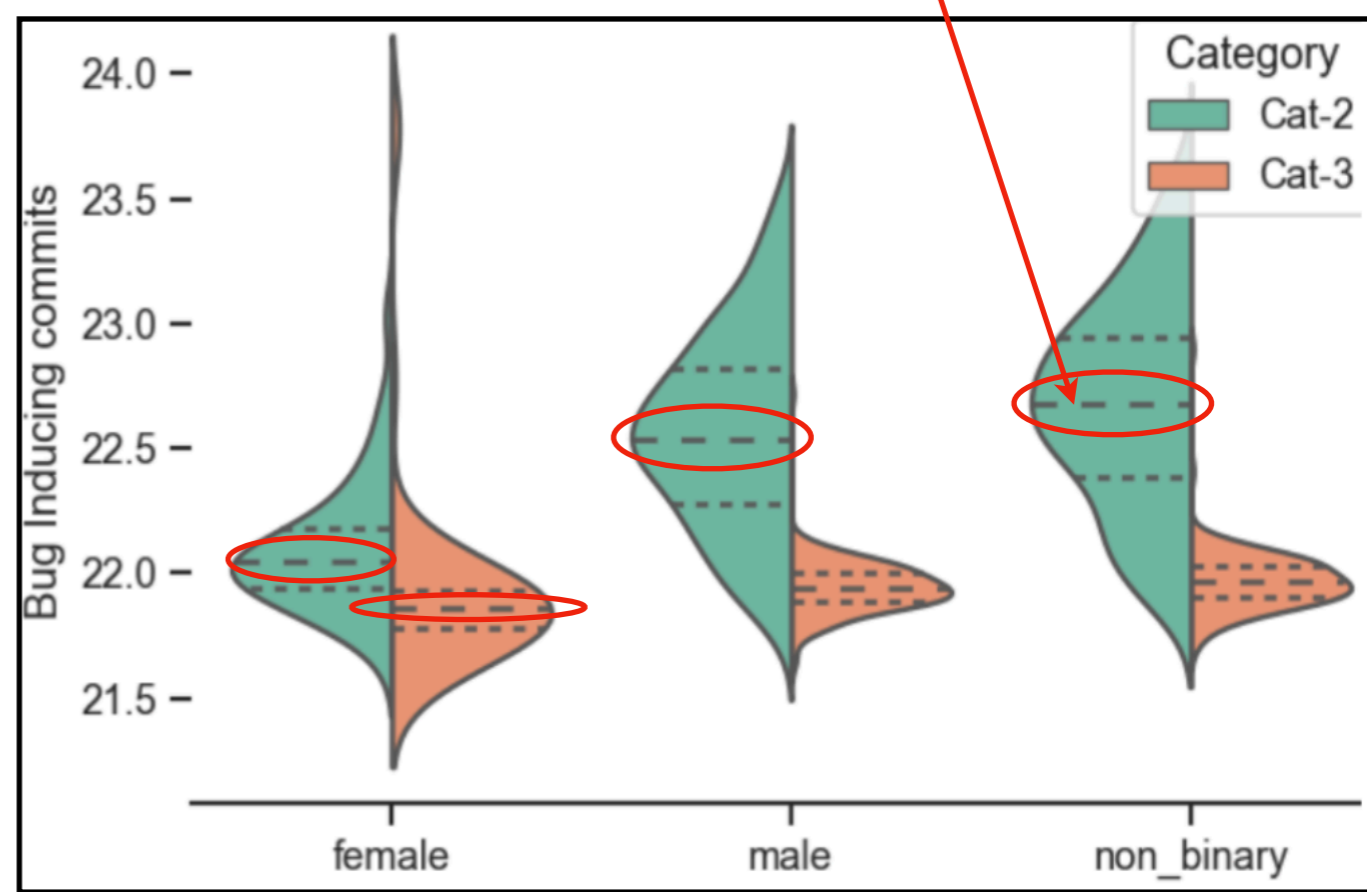
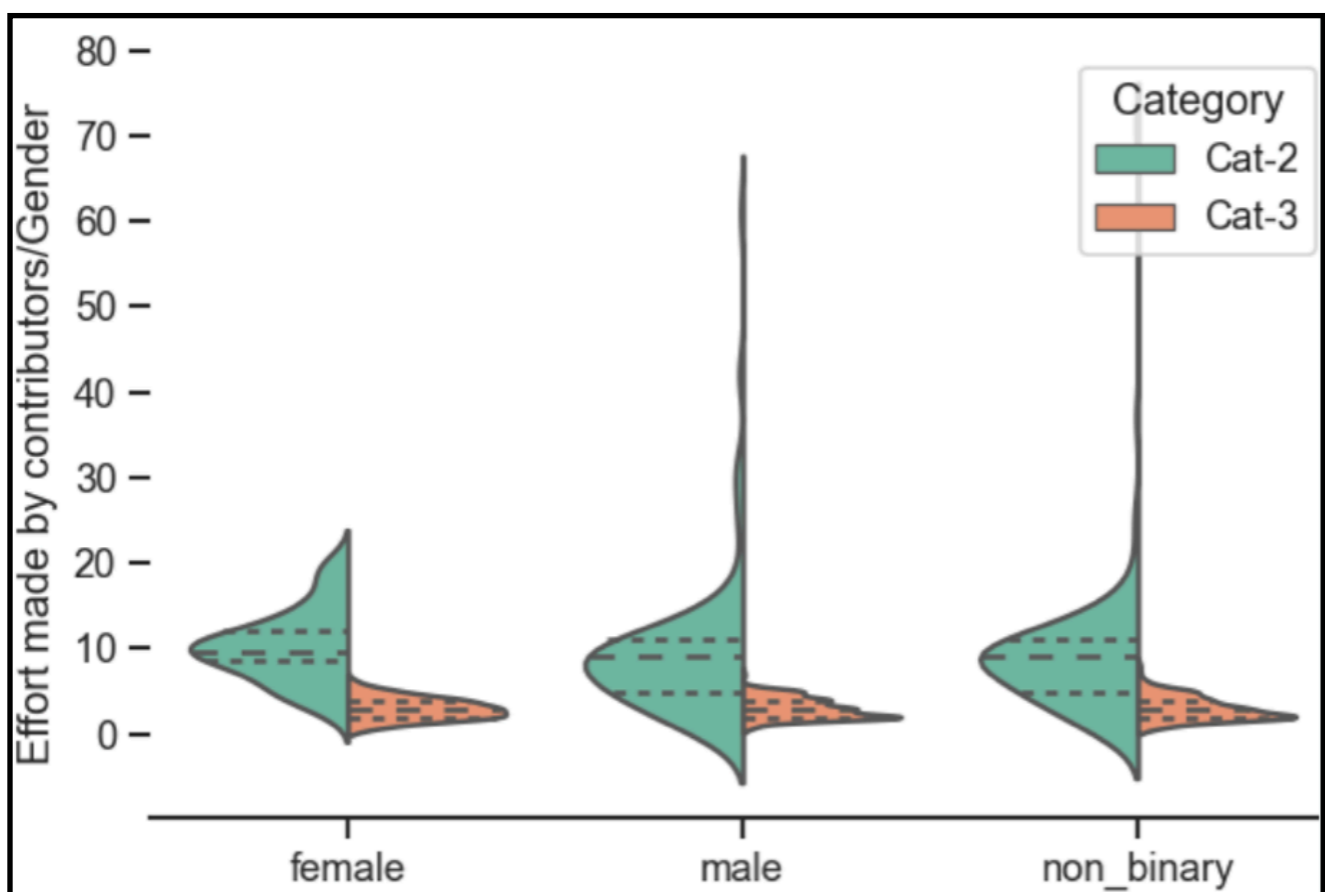


Onboarding correlates with a median 45% (Female) or 35% (male/non-binary) lower TFC.

# Onboarded Contributors Require Less Attempts to make High Quality Code Changes

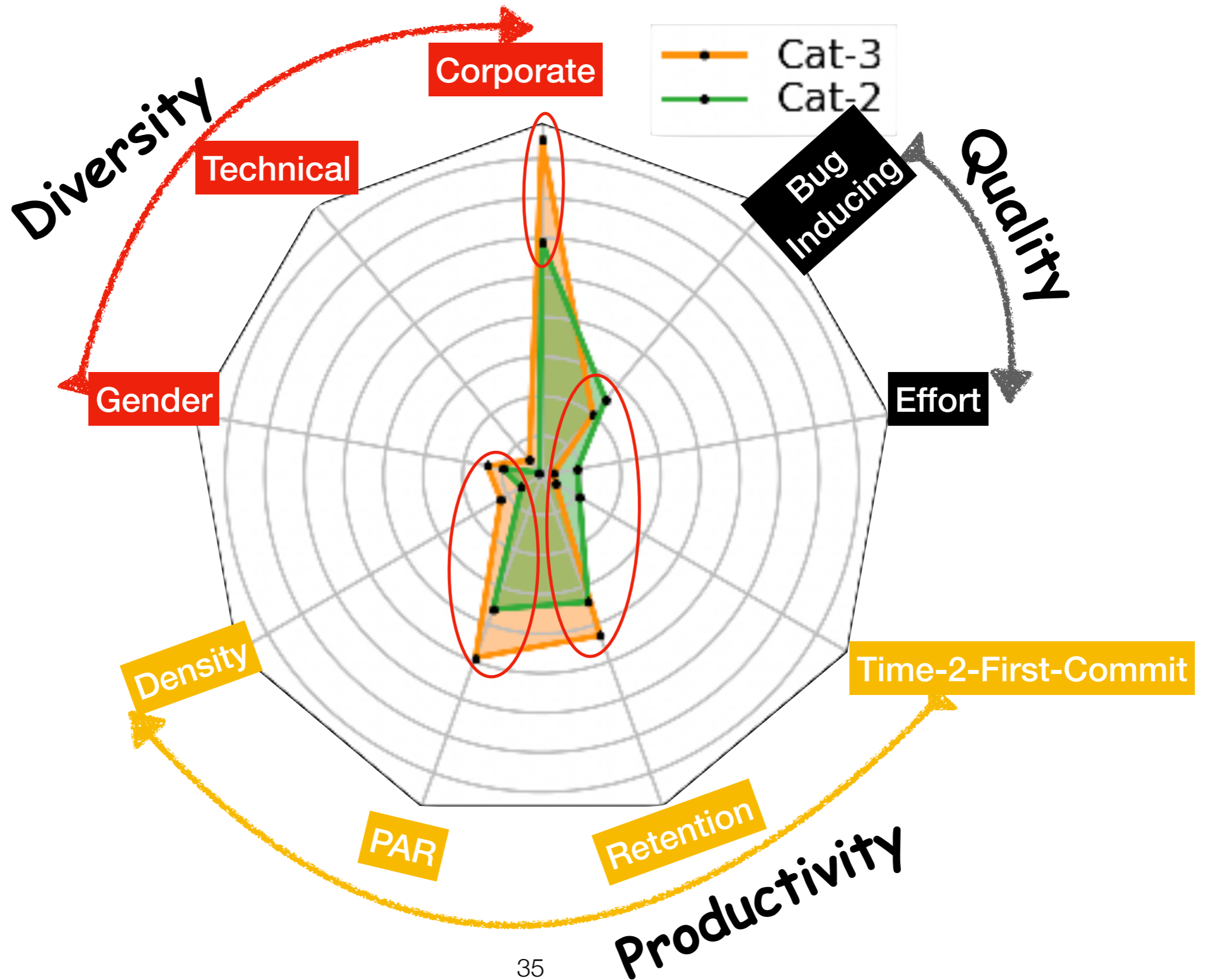
**Effort:** Onboarded contributors require less **attempts** to get their commits accepted.

Bean Plot: median line

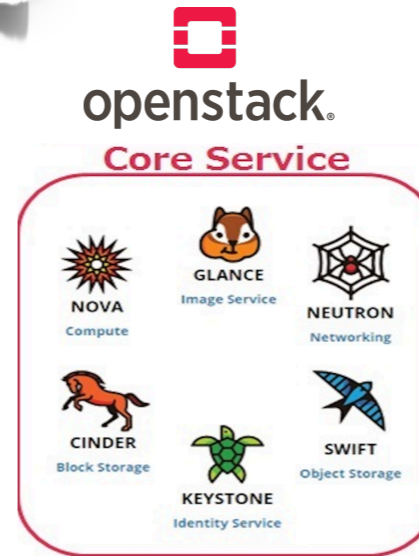
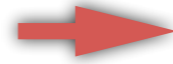
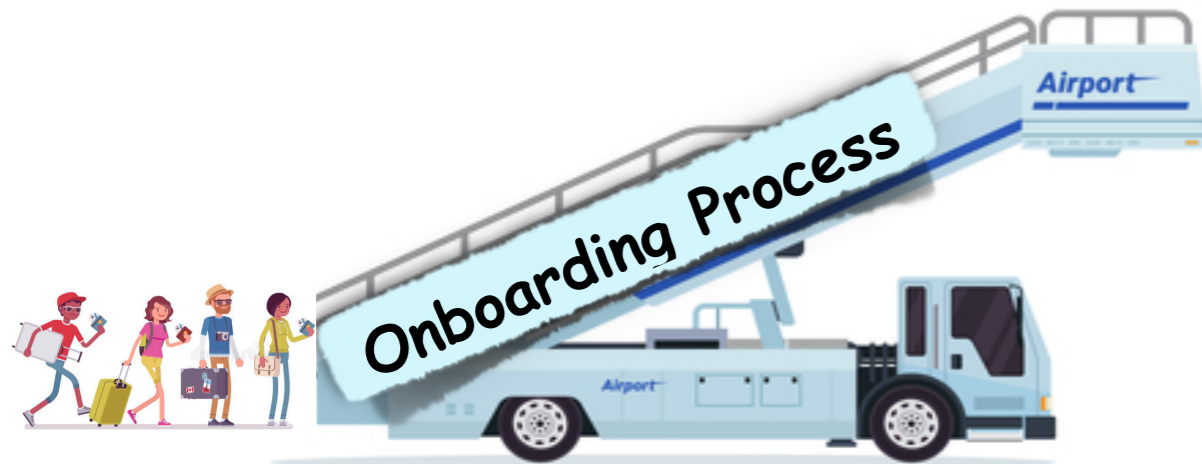


**Bug Inducing commits:** Onboarded contributors make less bug inducing commits (14%).

# Onboarding Correlates With Higher Productivity, Diversity and Quality

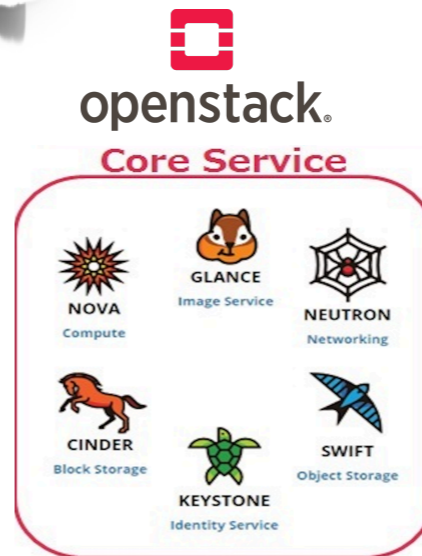
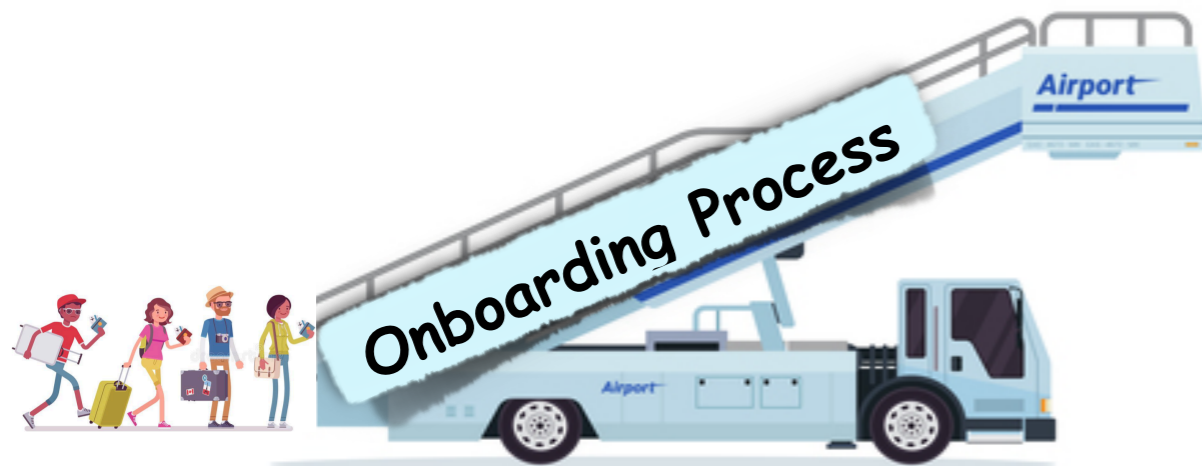


# In Sum



Our Paper: <https://2021.msrfconf.org/profile/armstrongfoundjem1>

# In Sum



Observational Study

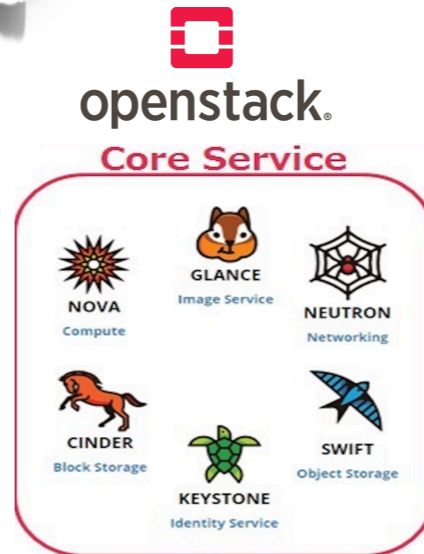
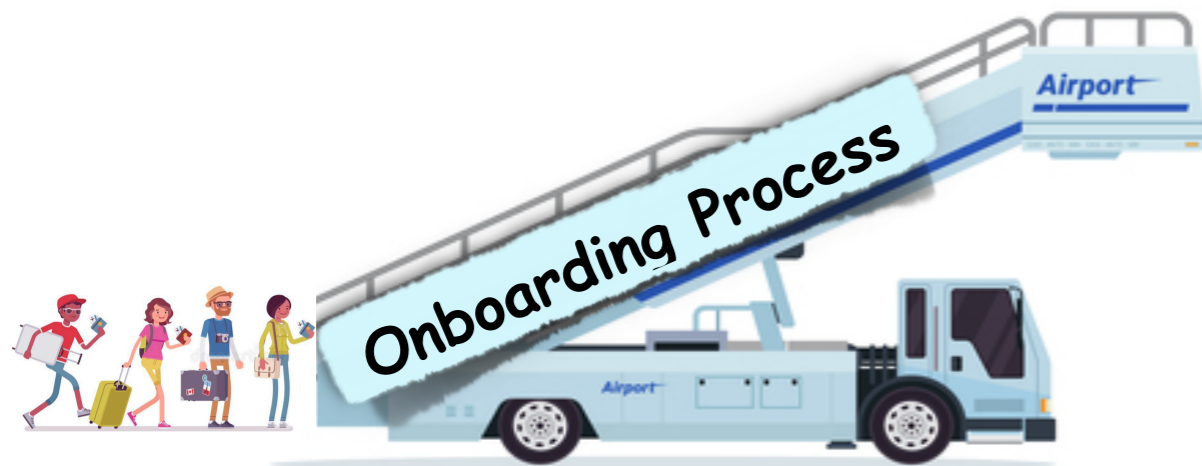
22 Teaching Content/Strategies

13 Onboarding Challenges

8 Onboarding Benefits

Our Paper: <https://2021.msrfconf.org/profile/armstrongfoundjem1>

# In Sum

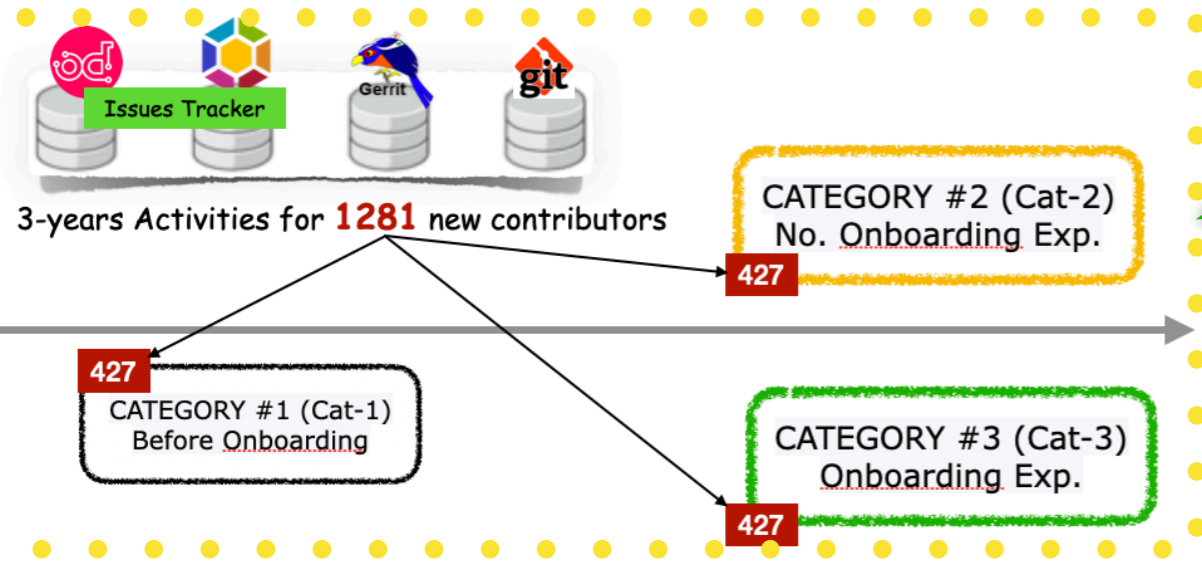


Observational Study

22 Teaching Content/Strategies

13 Onboarding Challenges

8 Onboarding Benefits



Our Paper: <https://2021.msrfconf.org/profile/armstrongfoundjem1>

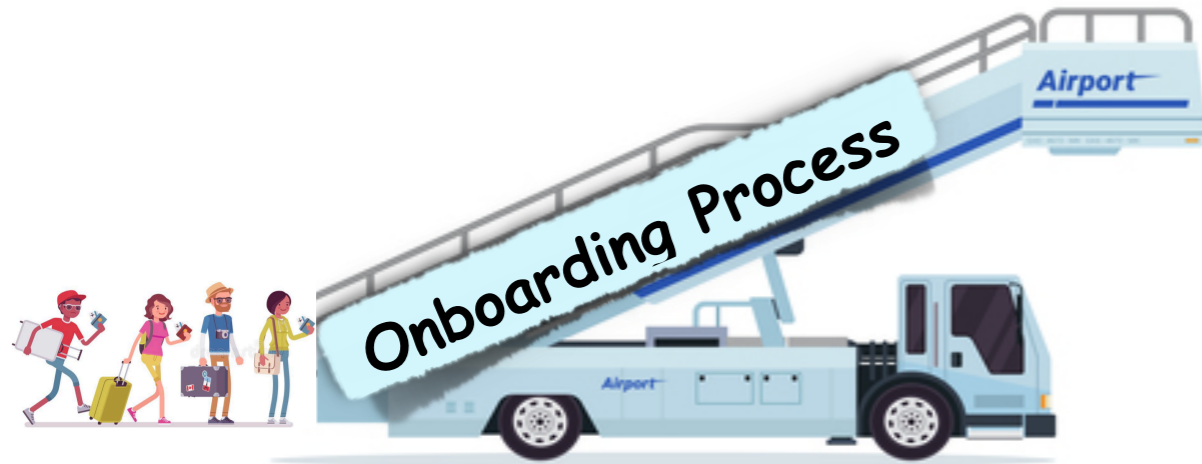
Q/A?

a.foundjem@queensu.ca

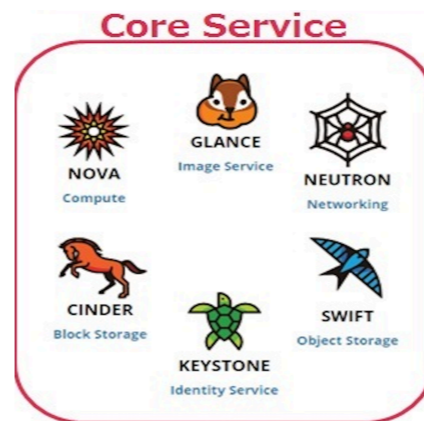
In Sum

Q/A?

a.foundjem@queensu.ca



openstack.

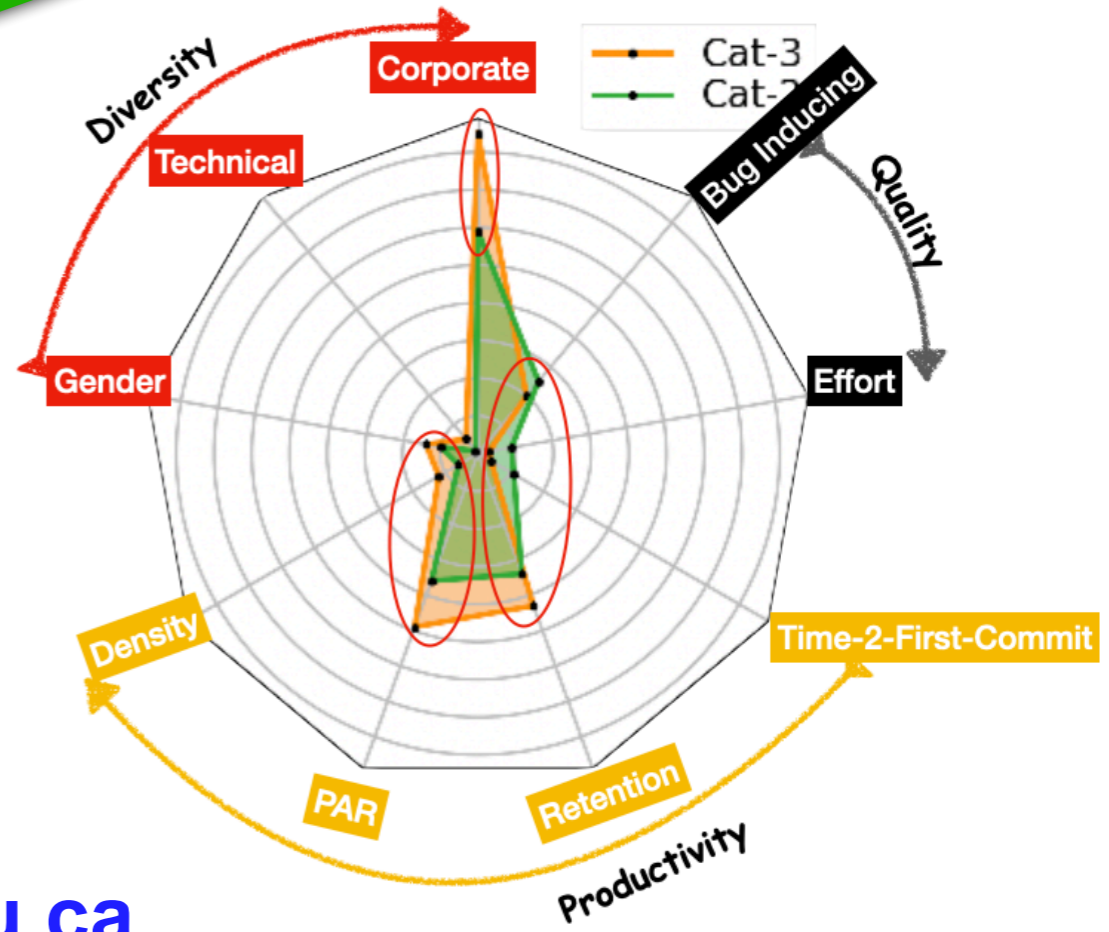
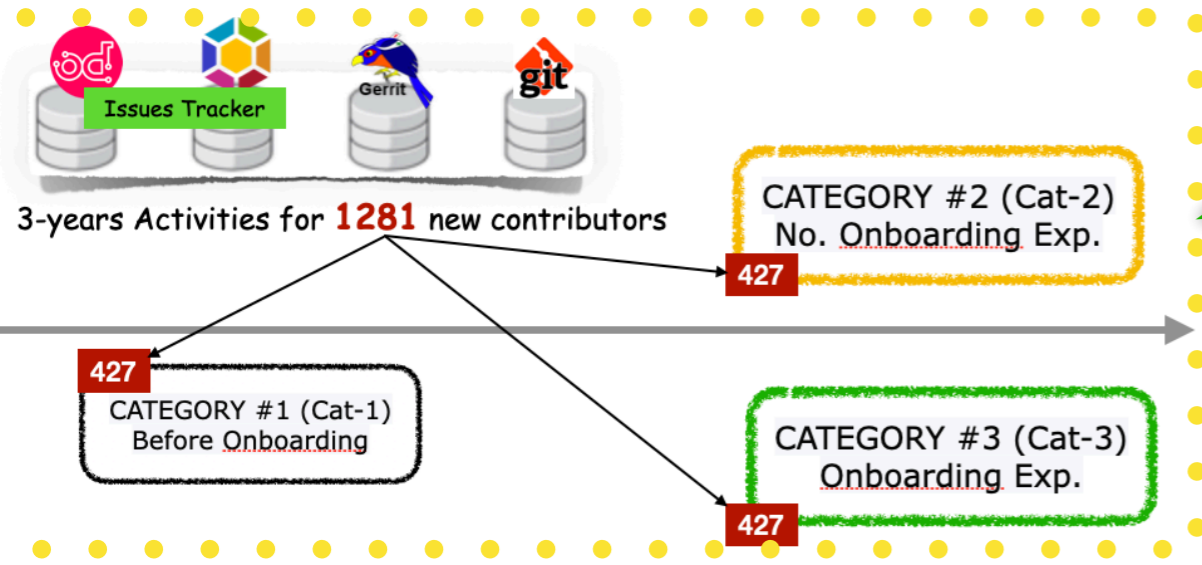


Observational Study

22 Teaching Content/Strategies

13 Onboarding Challenges

8 Onboarding Benefits



Q/A?

a.foundjem@queensu.ca