Advancing Sustainable Communities in Scientific OSS: A Replication Study with Astropy

Jiayi Sun[1] **Dr. Aarya Patil**[2] **Youhai Li**[1] **Dr. Jin Guo**[3] **Dr. Shurui Zhou**[1]

[1]University of Toronto [2]Max Planck Institute for Astronomy [3]McGill University

Background: Scientific Software

Definition: software used in scientific disciplines, including "source **code files**, **algorithms**, **scripts**, computational **workflows**, and **executables** created during the research process or for a research purpose" [1, 2].

[1] Arvanitou et al., "Software engineering practices for scientific software development: A systematic mapping study," JSS 2021.
[2] Morane, et al. "Defining Research Software: a controversial discussion." 2021. https://zenodo.org/records/5504016

Background: Scientific Software

Definition: software used in scientific disciplines, including "source **code files**, **algorithms**, **scripts**, computational **workflows**, and **executables** created during the research process or for a research purpose" [1, 2].

Astronomy

[1] Arvanitou et al., "Software engineering practices for scientific software development: A systematic mapping study," JSS 2021.
[2] Morane, et al. "Defining Research Software: a controversial discussion." 2021. https://zenodo.org/records/5504016

Background: Scientific Software

Definition: software used in scientific disciplines, including "source **code files**, **algorithms**, **scripts**, computational **workflows**, and **executables** created during the research process or for a research purpose" [1, 2].

Astronomy

Drug discovery

[1] Arvanitou et al., "Software engineering practices for scientific software development: A systematic mapping study," JSS 2021.
[2] Morane, et al. "Defining Research Software: a controversial discussion." 2021. https://zenodo.org/records/5504016

Background: Scientific Software

Definition: software used in scientific disciplines, including "source **code files**, **algorithms**, **scripts**, computational **workflows**, and **executables** created during the research process or for a research purpose" [1, 2].


Astronomy


Drug discovery


Seismology


Genomics


Meteorology

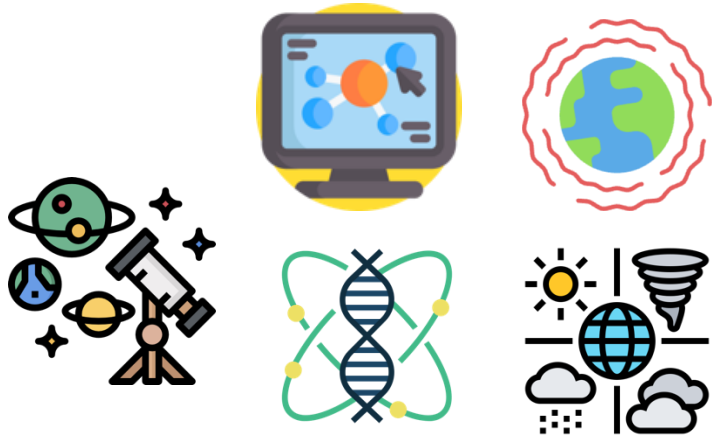
[1] Arvanitou et al., "Software engineering practices for scientific software development: A systematic mapping study," JSS 2021.
[2] Morane, et al. "Defining Research Software: a controversial discussion." 2021. https://zenodo.org/records/5504016

Scientific Software Adopted Open Source Model

+

Scientific Software Adopted Open Source Model

Sci-OSS: Scientific software developed openly and collaboratively, with source code freely available for use, modification, and contribution



+

Scientific Software Adopted Open Source Model

Sci-OSS: Scientific software developed openly and collaboratively, with source code freely available for use, modification, and contribution

Benefits

- Open collaboration
- Promote code sharing and reusability

Event Horizon Telescope (EHT) Project

M87 - the first image of a black hole [1]

[1] Case Study: First Image of a Black Hole. 2019. https://numpy.org/case-studies/blackhole-image/

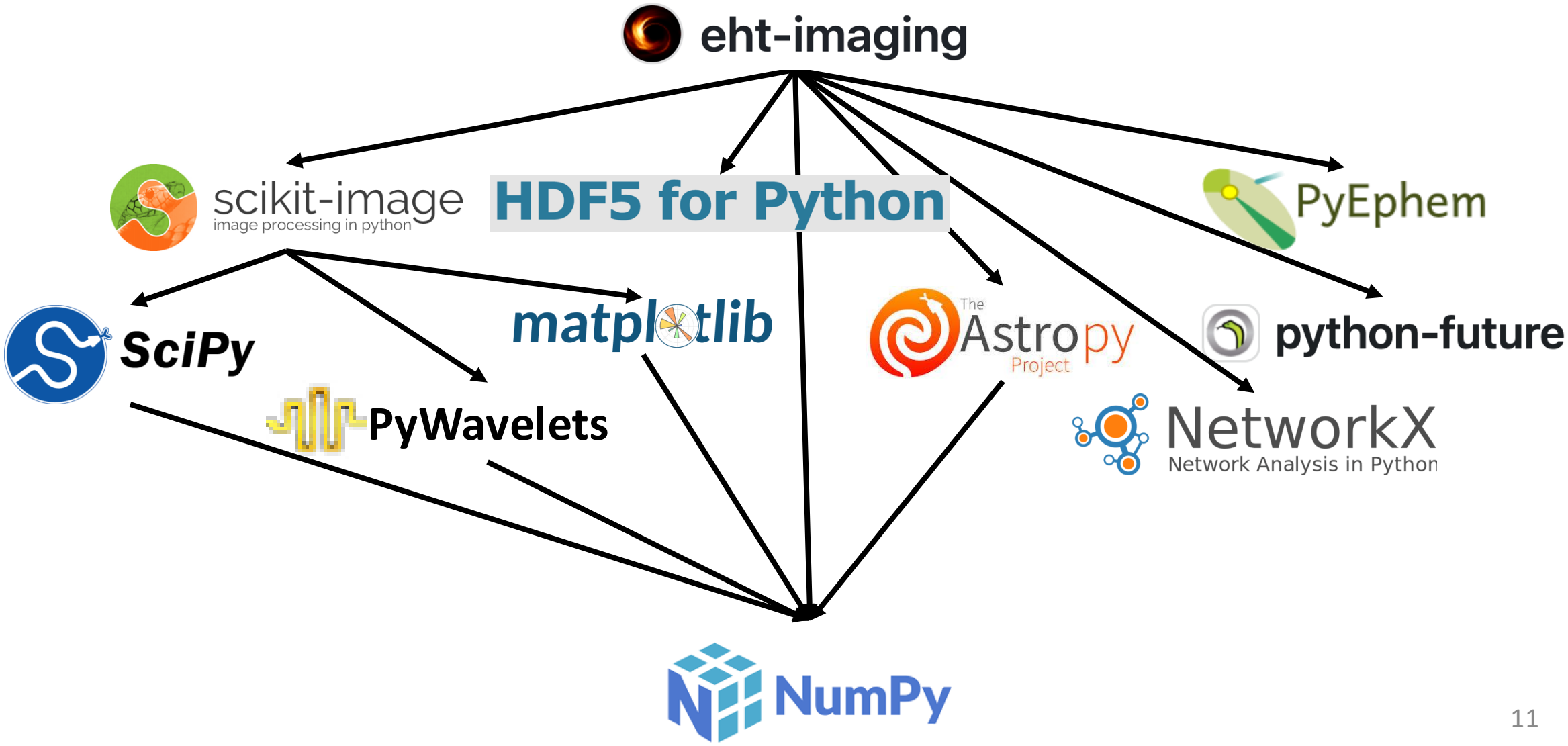
Event Horizon Telescope (EHT) Project

M87 - the first image of a black hole [1]

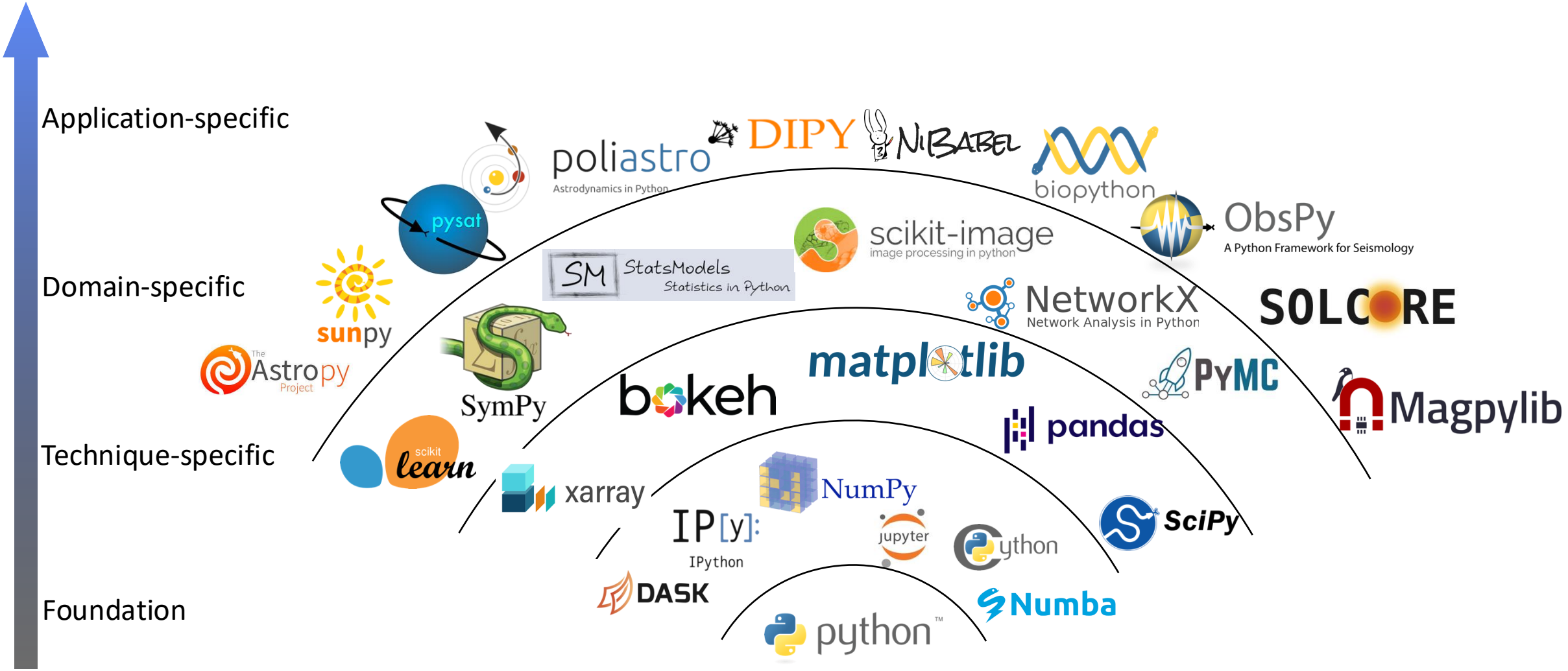
eht-imaging

[1] Case Study: First Image of a Black Hole. 2019. https://numpy.org/case-studies/blackhole-image/

Sci-OSS in EHT Project



Python Scientific Ecosystem Landscape

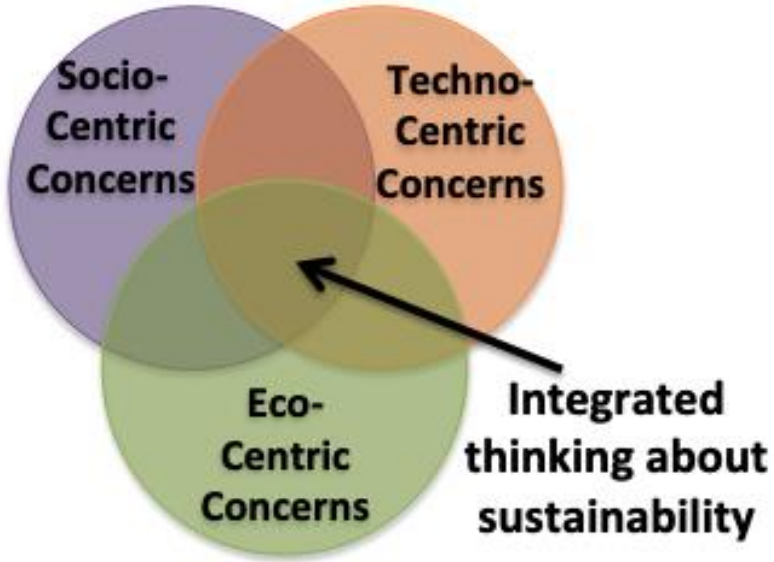


[1] Adopted from “The Unexpected Effectiveness of Python in Science”, Jake VanderPlas. 2017
[2] Adopted from Harris, Charles R., et al. "Array programming with NumPy." Nature. 2020

Sustainability of Scientific OSS is important !

Sustainability of Scientific OSS

Karlskrona Manifesto on software sustainability [1]



Software sustainability in scientific research context

"The ability to maintain the software in a state where scientists can **understand**, **replicate**, and **extend** previously reported results that depend on that software" [2]

[1] Becker, Christoph, et al. "Sustainability design and software: The karlskrona manifesto." ICSE, 2015.
[2] Trainer, Erik H., et al. "Community code engagements: summer of code & hackathons for community building in scientific software." Proceedings of the 2014 ACM International Conference on Supporting Group Work. 2014.

Sustainability of Scientific OSS

Software-aspect

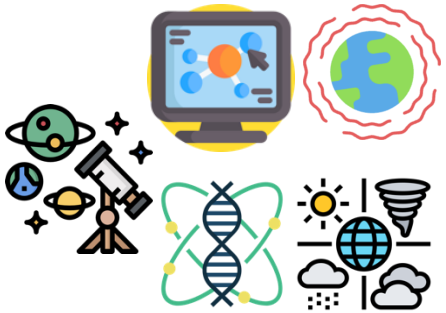
Sustainability of Scientific OSS

- **Software-aspect**
  - Reusability
  - Accessibility
  - Reliability
  - ...

Sustainability of Scientific OSS

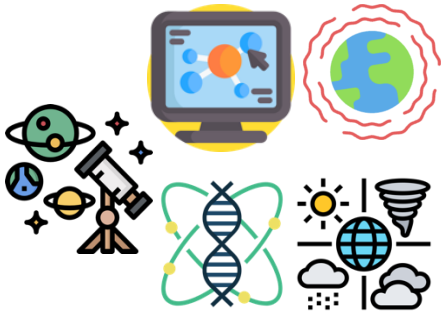
- **Software-aspect**
  - Reusability
  - Accessibility
  - Reliability
  - ...
- **Community-aspect**
  - Contributor retention
  - Funding resource
  - Culture
  - ...

Prior Work on Sustaining Scientific Software



Prior Work on Sustaining Scientific Software

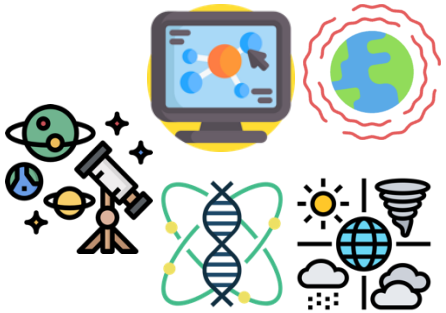
- Development requires domain knowledge [1].



[1] Kelly, “Scientific software development viewed as knowledge acquisition: Towards understanding the development of risk-averse scientific software”, JSS 2015.

Prior Work on Sustaining Scientific Software

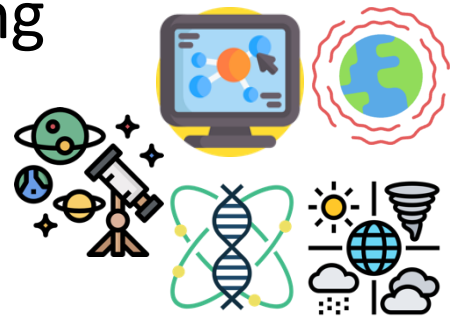
- Development requires domain knowledge [1].
- Scientists may lack SE background and best practices are often not prioritized [2, 3].



[1] Kelly, “Scientific software development viewed as knowledge acquisition: Towards understanding the development of risk-averse scientific software”, JSS 2015.
[2] Bozho: “The low quality of scientific code,” 2014. https://techblog.bozho.net/the-astonishingly-low-quality-of-scientific-code
[3] Merali, “Computational science:... error.” Nature 2010

Prior Work on Sustaining Scientific Software

- Development requires domain knowledge [1].
- Scientists may lack SE background and best practices are often not prioritized [2, 3].
- Technical challenges: uncertainty in requirement and testing difficulties (lack of test oracles) [4].



[1] Kelly, “Scientific software development viewed as knowledge acquisition: Towards understanding the development of risk-averse scientific software”, JSS 2015.
[2] Bozho: “The low quality of scientific code,” 2014. https://techblog.bozho.net/the-astonishingly-low-quality-of-scientific-code
[3] Merali, “Computational science:... error.” Nature 2010
[4] Nguyen-Hoan et al., “A survey of scientific software development.” ESEM 2010

Prior Work on Sustaining General OSS Communities

- Contributor retention [1,2]
  - Maintainers burnout and turnover
  - Difficulties in attracting newcomers
- Funding, sponsorship, corporates' participations [3,4]
- Culture (toxicity) [5]
- …

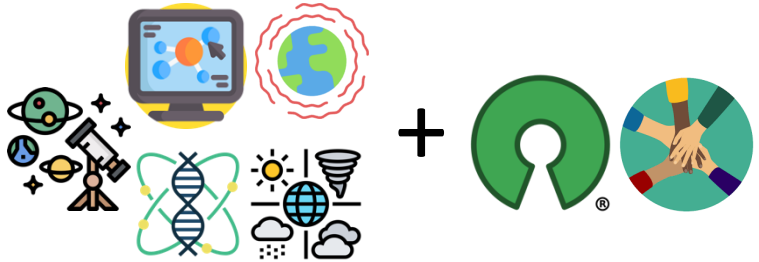
[1] Raman et al., "Stress and burnout in open source: Toward finding, understanding, and mitigating unhealthy interactions." ICSE-NIER 2020.
[2] Steinmacher et al., "A systematic literature review on the barriers faced by newcomers to open source software projects." IST 2015
[3] Shimada et al., "Github sponsors: exploring a new way to contribute to open source." ICSE 2022.
[4] Zhang, et al., "Companies' participation in oss development–an empirical study of openstack." TSE 2019
[5] Miller et al., "" Did you miss my comment or what?" understanding toxicity in open source discussions." ICSE 2022.

Study Goal: Is it the same for scientific OSS?

General OSS

VS

Scientific OSS



Prior Work on Sustaining General OSS Communities

- Contributor retention [1,2]
  - Maintainers burnout and turnover
  - Difficulties in attracting newcomers

Our Focus

- Funding, sponsorship, corporates' participations [3,4]
- Culture (toxicity) [5]
- …

[1] Raman et al., "Stress and burnout in open source: Toward finding, understanding, and mitigating unhealthy interactions." ICSE-NIER 2020.
[2] Steinmacher et al., "A systematic literature review on the barriers faced by newcomers to open source software projects." IST 2015
[3] Shimada et al., "Github sponsors: exploring a new way to contribute to open source." ICSE 2022.
[4] Zhang, et al., "Companies' participation in oss development–an empirical study of openstack." TSE 2019
[5] Miller et al., "" Did you miss my comment or what?" understanding toxicity in open source discussions." ICSE 2022.

Comparing Prior Work on General OSS with Sci-OSS

The Shifting Sands of Motivation: Revisiting What Drives Contributors in Open Source

Marco Gerosa,[1] Igor Wiese,[2] Bianca Trinkenreich,[1] Georg Link,[3] Gregorio Robles,[4] Christoph Treude,[5]
Igor Steinmacher,[1,2] Anita Sarma[6]
[1]*Northern Arizona University, USA*, [2]*Universidade Tecnológica Federal do Paraná, Brazil*, [3]*Bitergia, USA*,
[4]*Universidad Rey Juan Carlos, Spain*, [5]*University of Adelaide, Australia*, [6]*Oregon State University, USA*

marco.gerosa@nau.edu, igor@utfpr.edu.br, bt473@nau.edu, georglink@bitergia.com, grex@gsyc.urjc.es
christoph.treude@adelaide.edu.au, igor.steinmacher@nau.edu, anita.sarma@oregonstate.edu

***Abstract*—Open Source Software (OSS) has changed drastically over the last decade, with OSS projects now producing a large ecosystem of popular products, involving industry participation, and providing professional career opportunities. But our field's understanding of what motivates people to contribute to OSS is still fundamentally grounded in studies from the early 2000s. With the changed landscape of OSS, it is very likely that motivations to join OSS have also evolved. Through a survey of 242 OSS contributors, we investigate shifts in motivation from three perspectives: (1) the impact of the new OSS landscape, (2) the impact of individuals' personal growth as they become part of OSS communities, and (3) the impact of differences in individuals' demographics. Our results show that some motivations related to social aspects and reputation increased in frequency and that some intrinsic and internalized motivations, such as learning and intellectual stimulation, are still highly relevant. We also found that contributing to OSS often transforms extrinsic motivations to intrinsic, and that while experienced contributors often shift toward altruism, novices often shift toward career, fun, kinship, and learning. OSS projects can leverage our results to revisit current strategies to attract and retain contributors, and researchers and tool builders can better support the design of new studies and tools to engage and support OSS development.**

***Index Terms*—open source, motivation, incentive**

RQ1a: What motivates contributors to OSS today?
RQ1b: How has motivation to contribute shifted as OSS has matured?

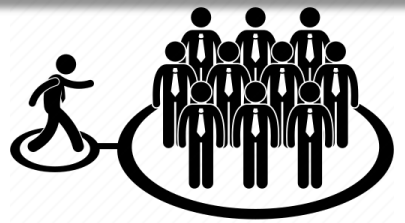
Besides understanding what motivates individuals now, so we can better support them, we also aim to identify the ways in which people's motivations have shifted in response to the changing landscape, so that OSS communities can rethink their strategies to attract and retain contributors.

Shifts in motivation occur not only because of changes to the OSS landscape, but might also reflect the journey an individual makes and their growth since first joining [4]. Currently, we lack an understanding of the differences in motivation for the early joiners compared to those who are well-entrenched in OSS. To support both the attraction of new members and the retention of existing contributors, we need to understand how the motivation changes after the members join OSS. This leads us to our next research question:

RQ2: How does motivation to contribute to OSS shift as OSS contributors gain tenure?

What motivates people and shifts their motivation as they gain experience in OSS may also depend on their individual

[Gerosa et al., 2021]



Why do People Give Up FLOSSing? A Study of Contributor Disengagement in Open Source

Courtney Miller[1⋆], David Widder[2], Christian Kästner[2], and Bogdan Vasilescu[2]
[1] New College of Florida, USA
[2] Carnegie Mellon University, USA

Abstract. Established contributors are the backbone of many free/libre open source software (FLOSS) projects. Previous research has shown that it is critically important for projects to retain contributors and it has also revealed the motivations behind why contributors choose to participate in FLOSS in the first place. However, there has been limited research done on the reasons why established contributors disengage, and factors (on an individual and project level) that predict their disengagement. In this paper, we conduct a mixed-methods empirical study, combining surveys and survival modeling, to identify the reasons and predictive factors behind established contributor disengagement. We find that different groups of established contributors tend to disengage for different reasons, however, overall contributors most commonly cite some kind of transition (e.g., switching jobs or leaving academia). We also find that factors such as the popularity of the projects a contributor works on, whether they have experienced a transition, when they work, and how much they work are all factors that can be used to predict their disengagement from open source.

[Miller et al., 2019]

Comparing Prior Work on Contribution Motivation

The Shifting Sands of Motivation: Revisiting What Drives Contributors in Open Source

Marco Gerosa,[1] Igor Wiese,[2] Bianca Trinkenreich,[1] Georg Link,[3] Gregorio Robles,[4] Christoph Treude,[5] Igor Steinmacher,[1,2] Anita Sarma[6]
[1]Northern Arizona University, USA, [2]Universidade Tecnológica Federal do Paraná, Brazil, [3]Bitergia, USA, [4]Universidad Rey Juan Carlos, Spain, [5]University of Adelaide, Australia, [6]Oregon State University, USA

marco.gerosa@nau.edu, igor@utfpr.edu.br, bt473@nau.edu, georglink@bitergia.com, grex@gsyc.urjc.es
christoph.treude@adelaide.edu.au, igor.steinmacher@nau.edu, anita.sarma@oregonstate.edu

Abstract—**Open Source Software (OSS) has changed drastically over the last decade, with OSS projects now producing a large ecosystem of popular products, involving industry participation, and providing professional career opportunities. But our field's understanding of what motivates people to contribute to OSS is still fundamentally grounded in studies from the early 2000s. With the changed landscape of OSS, it is very likely that motivations to join OSS have also evolved. Through a survey of 242 OSS contributors, we investigate shifts in motivation from three perspectives: (1) the impact of the new OSS landscape, (2) the impact of individuals' personal growth as they become part of OSS communities, and (3) the impact of differences in individuals' demographics. Our results show that some motivations related to social aspects and reputation increased in frequency and that some intrinsic and internalized motivations, such as learning and intellectual stimulation, are still highly relevant. We also found that contributing to OSS often transforms extrinsic motivations to intrinsic, and that while experienced contributors often shift toward altruism, novices often shift toward career, fun, kinship, and learning. OSS projects can leverage our results to revisit current strategies to attract and retain contributors, and researchers and tool builders can better support the design of new studies and tools to engage and support OSS development.**

Index Terms—**open source, motivation, incentive**

RQ1a: What motivates contributors to OSS today?
RQ1b: How has motivation to contribute shifted as OSS has matured?

Besides understanding what motivates individuals now, so we can better support them, we also aim to identify the ways in which people's motivations have shifted in response to the changing landscape, so that OSS communities can rethink their strategies to attract and retain contributors.

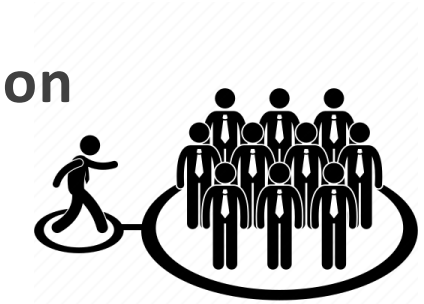
Shifts in motivation occur not only because of changes to the OSS landscape, but might also reflect the journey an individual makes and their growth since first joining [4]. Currently, we lack an understanding of the differences in motivation for the early joiners compared to those who are well-entrenched in OSS. To support both the attraction of new members and the retention of existing contributors, we need to understand how the motivation changes after the members join OSS. This leads us to our next research question:

RQ2: How does motivation to contribute to OSS shift as OSS contributors gain tenure?

What motivates people and shifts their motivation as they gain experience in OSS may also depend on their individual

[Gerosa et al., 2021]

Motivations for contribution



- Ideology
- Altruism
- Fun
- Kinship
- Reputation
- Reciprocity
- Learning
- Own-use
- Career
- Pay
- Google Summer of Code
- Coursework

Comparing Prior Work on Contributor Disengagement

Why do People Give Up FLOSSing? A Study of Contributor Disengagement in Open Source

Courtney Miller[1]*, David Widder[2], Christian Kästner[2], and Bogdan Vasilescu[2]
[1] New College of Florida, USA
[2] Carnegie Mellon University, USA

Abstract. Established contributors are the backbone of many free/libre open source software (FLOSS) projects. Previous research has shown that it is critically important for projects to retain contributors and it has also revealed the motivations behind why contributors choose to participate in FLOSS in the first place. However, there has been limited research done on the reasons why established contributors disengage, and factors (on an individual and project level) that predict their disengagement. In this paper, we conduct a mixed-methods empirical study, combining surveys and survival modeling, to identify the reasons and predictive factors behind established contributor disengagement. We find that different groups of established contributors tend to disengage for different reasons, however, overall contributors most commonly cite some kind of transition (e.g., switching jobs or leaving academia). We also find that factors such as the popularity of the projects a contributor works on, whether they have experienced a transition, when they work, and how much they work are all factors that can be used to predict their disengagement from open source.

Disengagement reasons

- Occupational reasons
  - Changed role/project, got new job…
- Social reasons
  - Lost interest in OSS, lack of peer support…
- Technical reasons
  - Issues with GitHub, feature complete project…

[Miller et al., 2019]

Existing Strategies for Sustaining OSS Communities

Sponsors

GitHub Sponsors

Programming events

Google Summer of Code

Good First Issues

Add social media "share" button for blogpost

enhancement good first issue

Study Goal: Is it the same for Scientific OSS?

General OSS

Contribution **motivation**
[Gerosa et al.,2021]

Reasons for **disengagement**
[Miller et al.,2019]

Strategies for improving community sustainability

Study Goal: Is it the same for Scientific OSS?

General OSS

Contribution **motivation**
[Gerosa et al.,2021]

Reasons for **disengagement**
[Miller et al.,2019]

Strategies for improving community sustainability

Scientific OSS

RQ1: Motivations for contributing to scientific OSS?

RQ2: Reasons for **disengaging** from scientific OSS?

RQ3: Suggestions for sustaining scientific OSS communities?

Study Goal: Is it the same for Scientific OSS?

General OSS

Contribution **motivation**
[Gerosa et al.,2021]

Reasons for **disengagement**
[Miller et al.,2019]

Strategies for improving community sustainability

Conceptual Replication

Scientific OSS

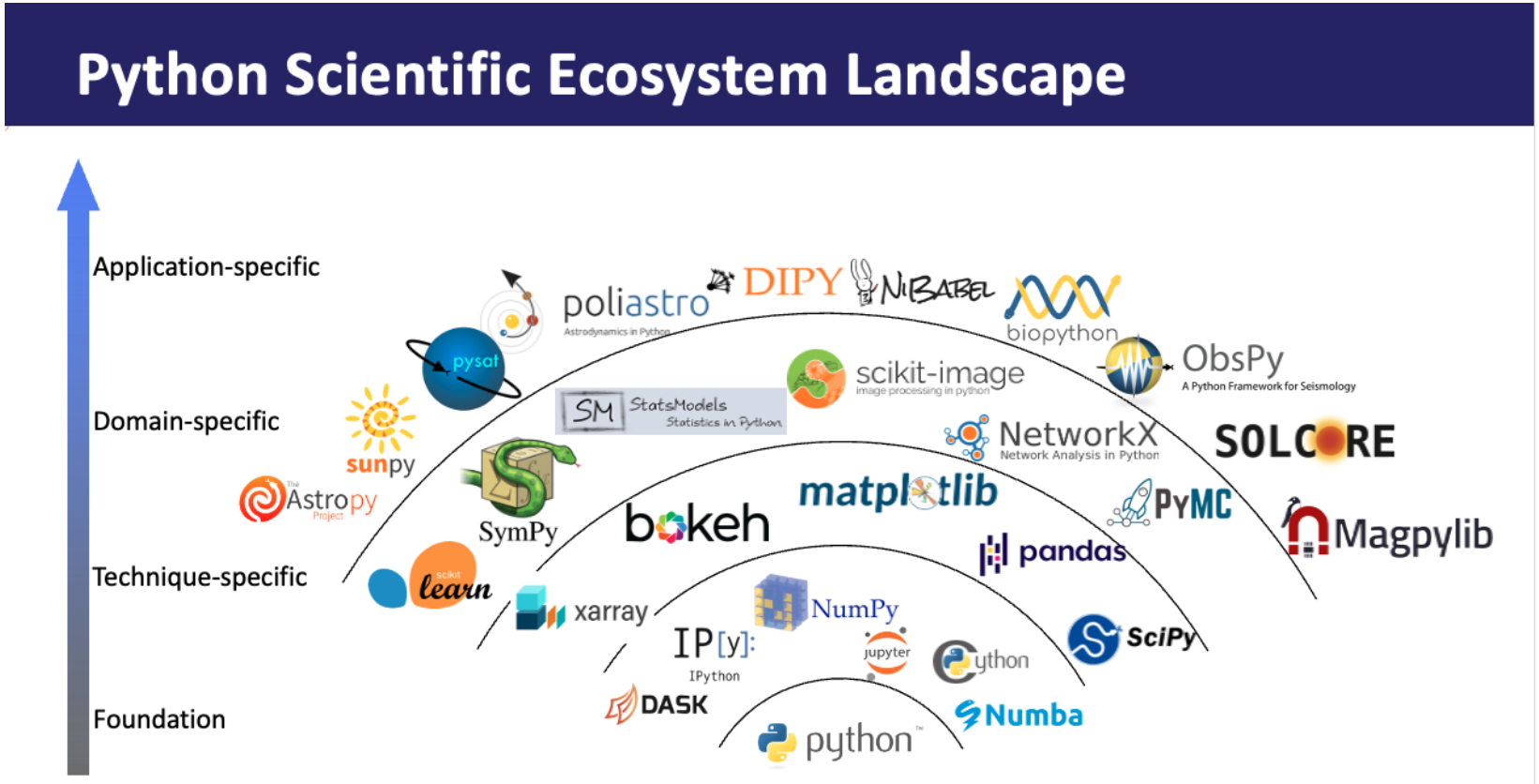
RQ1: Motivations for contributing to scientific OSS?

RQ2: Reasons for **disengaging** from scientific OSS?

RQ3: Suggestions for sustaining scientific OSS communities?

Study Subject

Wide range of scientific disciplines and projects…
Where do we start?



Study Subject – The Astropy Project

- Python software ecosystem for astronomy

Study Subject – The Astropy Project

- Python software ecosystem for astronomy
- One core package: **astro**py
  - Age >= 10 years of age
  - > 400 contributors

Study Subject – The Astropy Project

- Python software ecosystem for astronomy
- One core package: **astro**py
  - Age >= 10 years of age
  - > 400 contributors

- 50 other interoperable packages

Study Subject – The Astropy Project

- Python software ecosystem for astronomy
- One core package: **astropy**
  - Age >= 10 years of age
  - > 400 contributors

- 50 other interoperable packages

☑ **Maturity** ☑ **Community-oriented**

☑ **Popularity** ☑ **Development artifacts**

Study Goal: Is it the same for Scientific OSS?

General OSS

Contribution **motivation**
[Gerosa et al.,2021]

Reasons for **disengagement**
[Miller et al.,2019]

Strategies for improving community sustainability

Conceptual Replication

Scientific OSS

RQ1: Motivations for contributing to scientific OSS?

RQ2: Reasons for **disengaging** from scientific OSS?

RQ3: Suggestions for sustaining scientific OSS communities?

The Astropy Project

Method: Survey Disengaged Contributors

Method: Survey Disengaged Contributors

51 repositories

1116 code
commit contributors

Method: Survey Disengaged Contributors

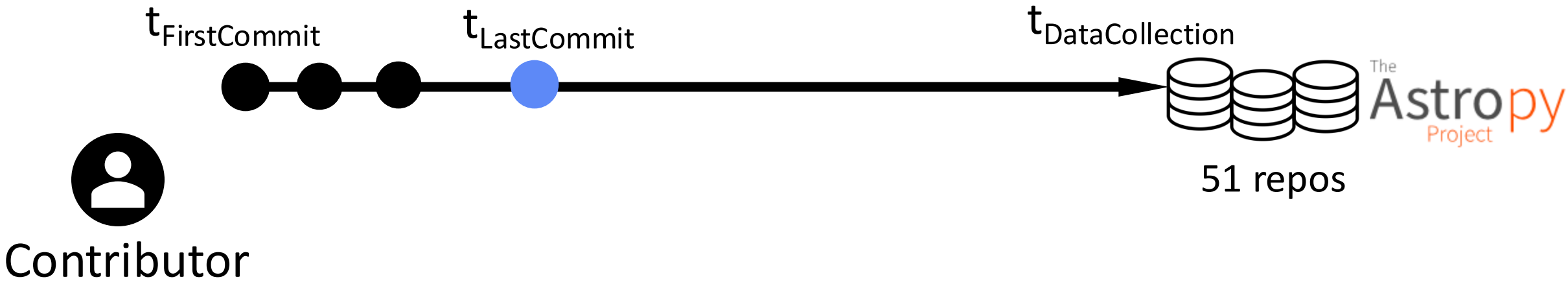
Definition of a disengaged contributor:

Contributor

Method: Survey Disengaged Contributors

Definition of a disengaged contributor:

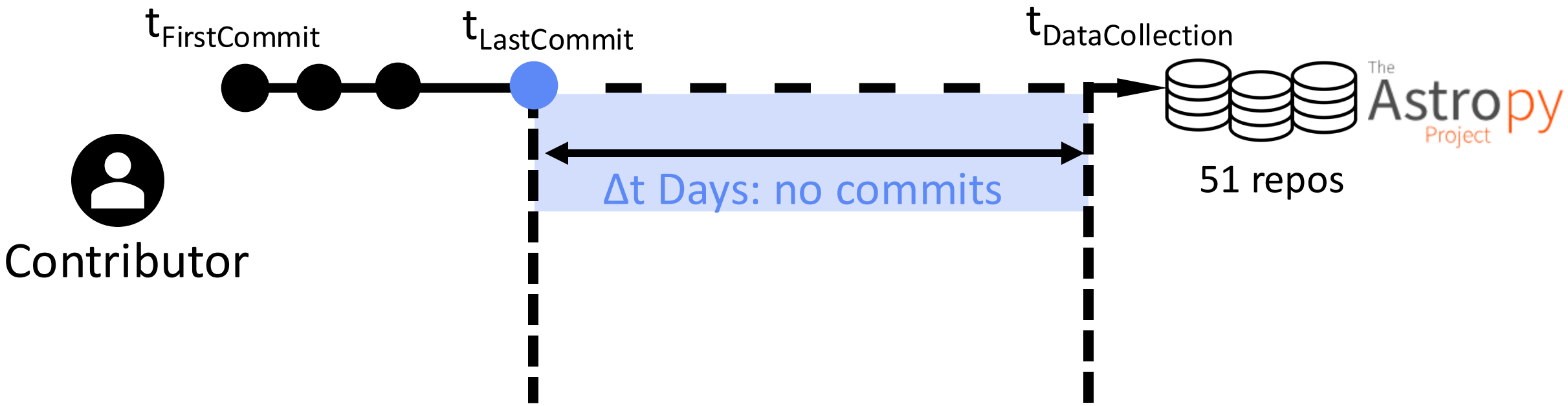
1. Aggregate the commit history in 51 repos, and identify $t_{LastCommit}$



Method: Survey Disengaged Contributors

Definition of a disengaged contributor:

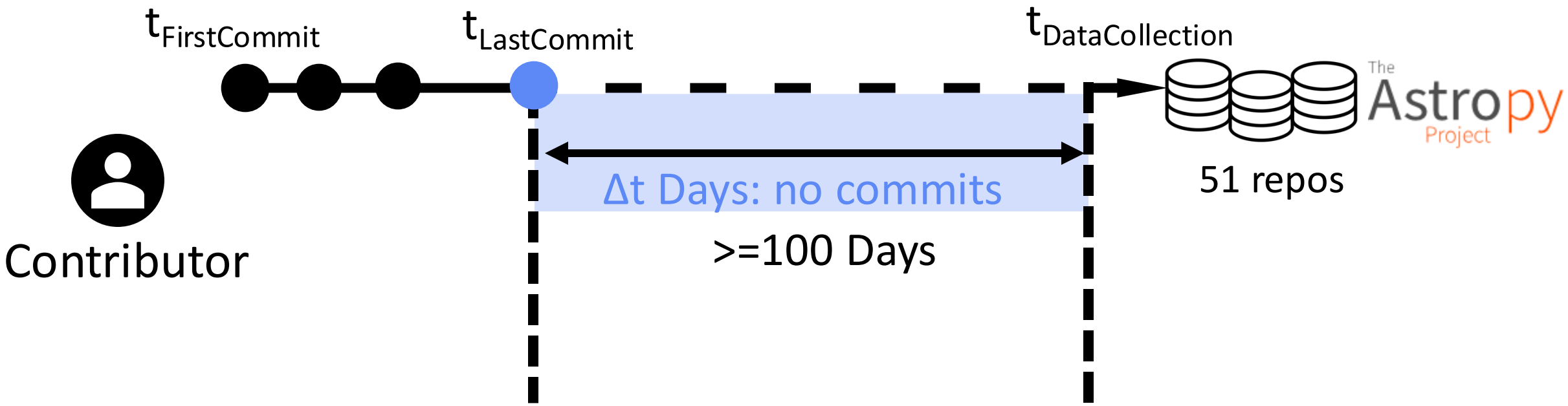
1. Aggregate the commit history in 51 repos, and identify $t_{LastCommit}$
2. Calculate $\Delta t = t_{DataCollection} - t_{LastCommit}$



Method: Survey Disengaged Contributors

Definition of a disengaged contributor:

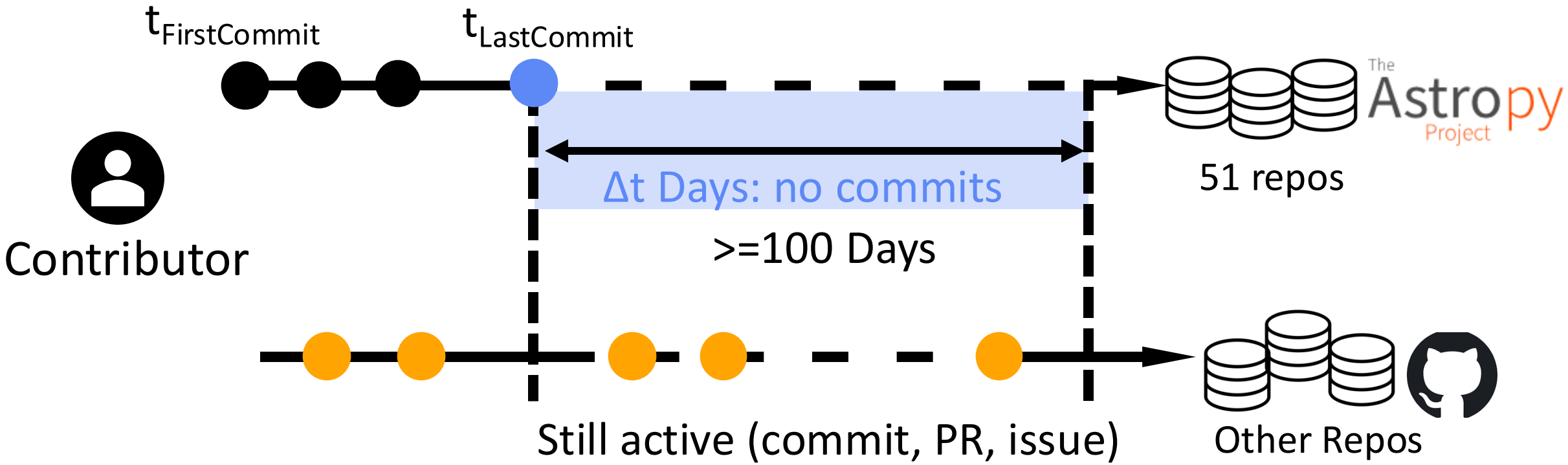
1. Aggregate the commit history in 51 repos, and identify $t_{LastCommit}$
2. Calculate $\Delta t = t_{DataCollection} - t_{LastCommit}$
3. If the contributor is **inactive** in The Astropy Project repos for >=100 days



Method: Survey Disengaged Contributors

Definition of a disengaged contributor:

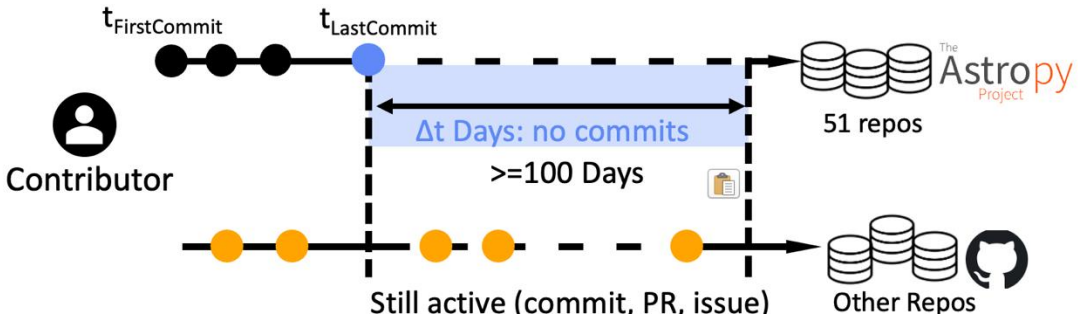
1. Aggregate the commit history in 51 repos, and identify $t_{LastCommit}$
2. Calculate $\Delta t = t_{DataCollection} - t_{LastCommit}$
3. If the contributor is **inactive** in The Astropy Project repos for >=100 days, but still active in other GitHub repos.



Method: Survey Disengaged Contributors

51 repositories

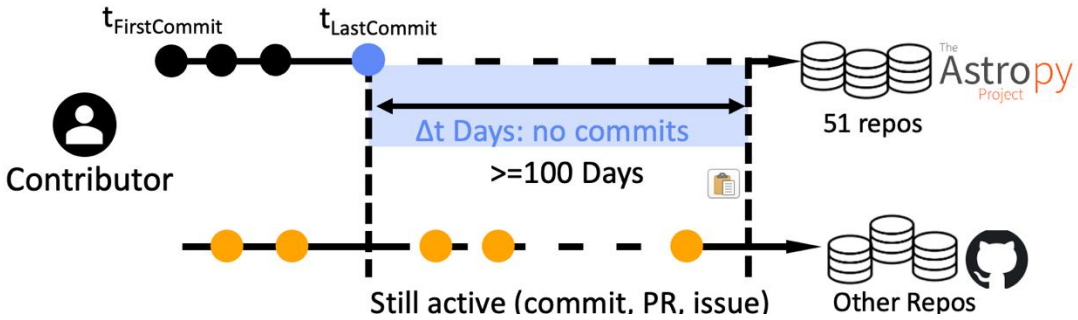
1116 code
commit contributors



Method: Survey Disengaged Contributors

51 repositories

1116 code commit contributors

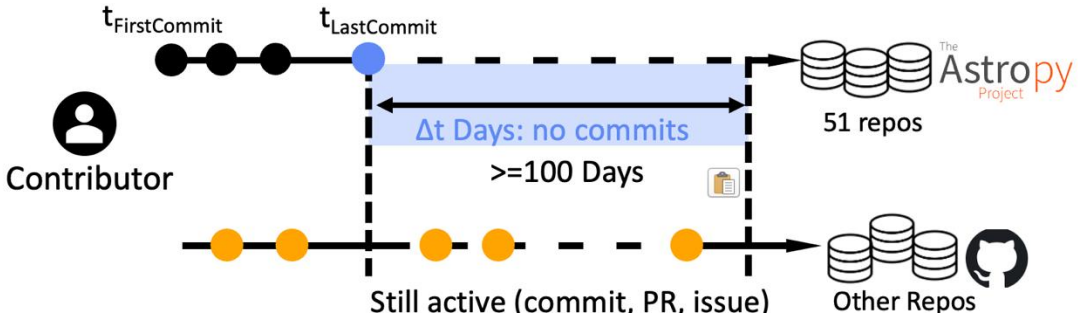


292/469 disengaged contributors with public email contact

Method: Survey Disengaged Contributors

51 repositories

1116 code commit contributors



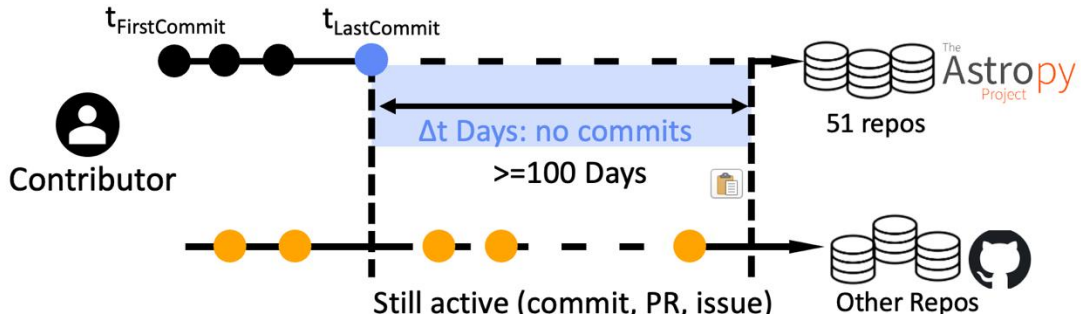
292/469 disengaged contributors with public email contact

3 Open-ended survey questions
RQ1: Motivation for contributing
RQ2: Reasons for disengaging
RQ3: Suggestions for improvement

Method: Survey Disengaged Contributors

51 repositories

1116 code commit contributors



292/469 disengaged contributors with public email contact

3 Open-ended survey questions

RQ1: Motivation for contributing

RQ2: Reasons for disengaging

RQ3: Suggestions for improvement

Qualitatively analyzed 80 Reponses from 292 recipients

Study Goal: Is it the same for Scientific OSS?

General OSS

Contribution **motivation**
[Gerosa et al.,2021]

Reasons for **disengagement**
[Miller et al.,2019]

Strategies for improving community sustainability

Conceptual Replication

Scientific OSS

RQ1: Motivations for contributing to scientific OSS?

RQ2: Reasons for **disengaging** from scientific OSS?

RQ3: Suggestions for sustaining scientific OSS communities?

Study Goal: Is it the same for Scientific OSS?

General OSS

Contribution **motivation**
[Gerosa et al.,2021]

Reasons for **disengagement**
[Miller et al.,2019]

Strategies for improving community sustainability

Conceptual Replication

Scientific OSS

RQ1: Motivations for contributing to scientific OSS?

RQ2: Reasons for **disengaging** from scientific OSS?

RQ3: Suggestions for sustaining scientific OSS communities?

RQ1: Motivations for Contributing to Sci-OSS

RQ1: Motivations for Contributing to Sci-OSS

Own use

Pay

Altruism

Learning

Invitation

GSoC

"...those functions were essential for my research, I developed & committed onto Astropy."

RQ1: Motivations for Contributing to Sci-OSS

Own use

Pay

Altruism

Learning

Invitation

GSoC

“I was employed as a programmer in a lab.”

RQ1: Motivations for Contributing to Sci-OSS

Own use

Pay

Altruism

Learning

Invitation

GSoC

"Sharing my work so others can benefit thereby."

RQ1: Motivations for Contributing to Sci-OSS

Own use

Pay

Altruism

Learning

Invitation

GSoC

“Learn how to structure my programming better.”

RQ1: Motivations for Contributing to Sci-OSS

Own use

Pay

Altruism

Learning

Invitation

GSoC

"It was part of a Hacktoberfest event and I knew someone involved in the project"

RQ1: Motivations for Contributing to Sci-OSS

Own use

Pay

Altruism

Learning

Invitation

GSoC

Google Summer of Code

"I was looking for a summer intern, so I started contributing to Astropy for GSoC."

RQ1: Motivations for Contributing to Sci-OSS

General OSS

Primary reasons for contribution are ***learning*** *and* ***altruism,*** while contributing because of ***own use*** has dropped in ranking compared to prior studies.

1st	**Learning**
2nd	**Altruism**
…	
4th	**Own use**

[Gerosa et al., 2021]

RQ1: Motivations for Contributing to Sci-OSS

Own use

Pay

Altruism

Learning

Invitation

GSoC

General OSS

Primary reasons for contribution are ***learning*** *and* ***altruism,*** while contributing because of ***own use*** has dropped in ranking compared to prior studies.

1st **Learning**

2nd **Altruism**

…

4th **Own use**

[Gerosa et al., 2021]

Different to general OSS, contribution to sci OSS is mainly driven by ***own use*** for contributors' research work.

Study Goal: Is it the same for Scientific OSS?

General OSS

Contribution **motivation**
[Gerosa et al.,2021]

Reasons for **disengagement**
[Miller et al.,2019]

Strategies for improving community sustainability

Conceptual Replication

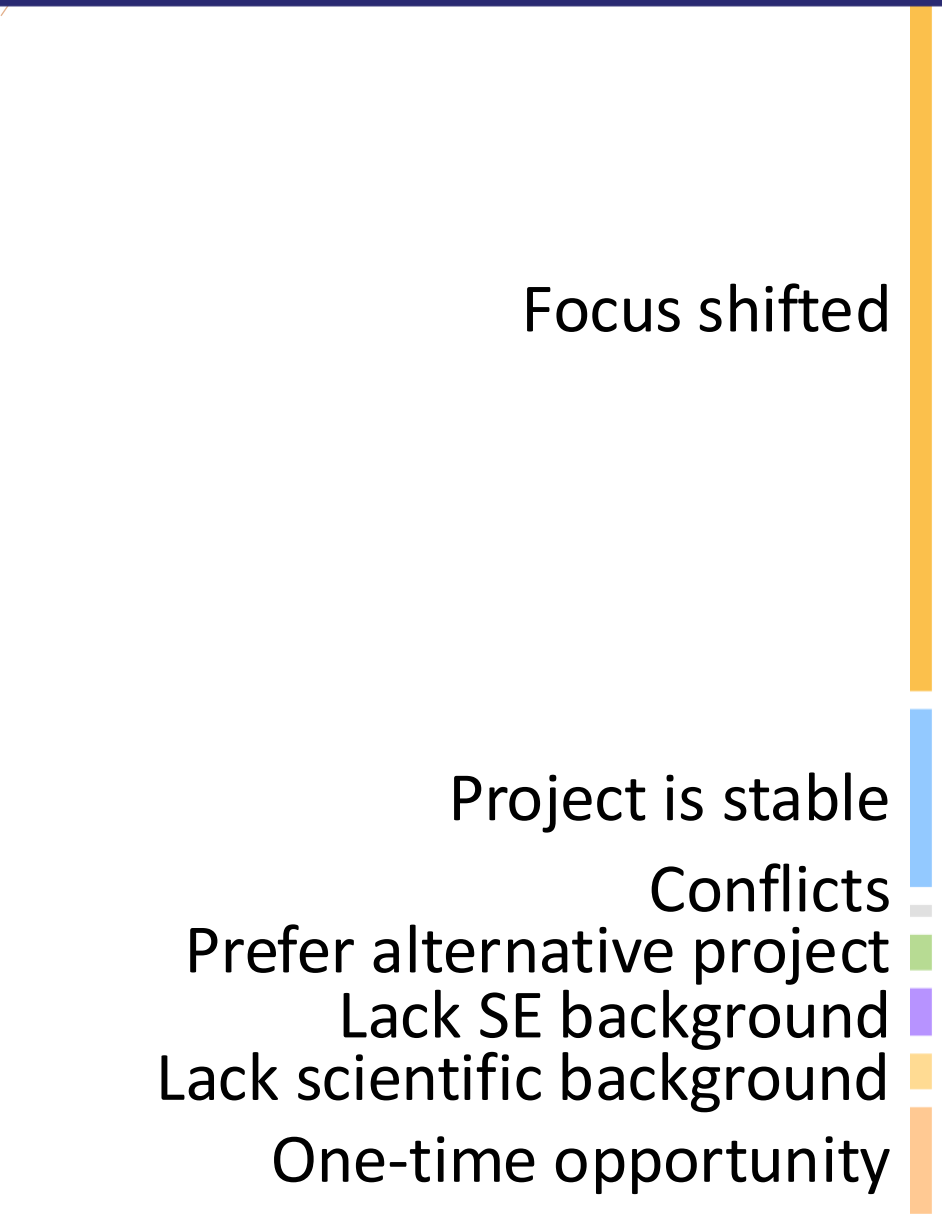
Scientific OSS

RQ1: Motivations for contributing to scientific OSS?

RQ2: Reasons for **disengaging** from scientific OSS?

RQ3: Suggestions for sustaining scientific OSS communities?

RQ2: Reasons for Disengaging from Sci-OSS



RQ2: Reasons for Disengaging from Sci-OSS

Focus shifted

“After graduating, I took a postdoc working on a different open-source astronomy project“

Project is stable

Conflicts

Prefer alternative project

Lack SE background

Lack scientific background

One-time opportunity

RQ2: Reasons for Disengaging from Sci-OSS

Focus shifted

Project is stable

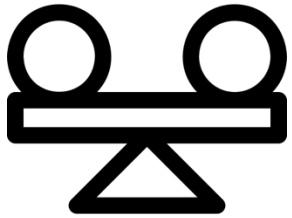
Conflicts

Prefer alternative project

Lack SE background

Lack scientific background

One-time opportunity



“I believe that most of the features were mature enough and the package (is maintained) as stable as possible at that point.“

RQ2: Reasons for Disengaging from Sci-OSS

Focus shifted

Project is stable

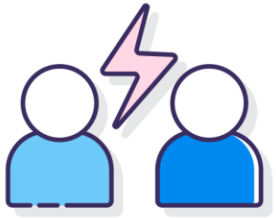
Conflicts

Prefer alternative project

Lack SE background

Lack scientific background

One-time opportunity



“I stopped because my vision of the package did not aligned with the other core team members.“

RQ2: Reasons for Disengaging from Sci-OSS

Focus shifted

Project is stable

Conflicts

Prefer alternative project

Lack SE background

Lack scientific background

One-time opportunity

"I've made more contributions to another package with similar functionality…has a lot of flexibility and is maintained more regularly."

RQ2: Reasons for Disengaging from Sci-OSS

Focus shifted

Project is stable

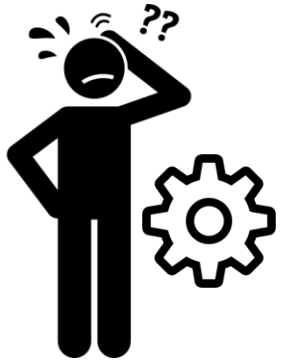
Conflicts

Prefer alternative project

Lack SE background

Lack scientific background

One-time opportunity



“my software development skills are very poor…GitHub very non-intuitive, so I just stayed in my comfort zone as far as coding goes.“

RQ2: Reasons for Disengaging from Sci-OSS

Focus shifted

Project is stable

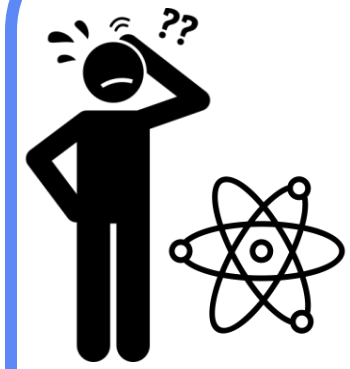
Conflicts

Prefer alternative project

Lack SE background

Lack scientific background

One-time opportunity



"I was comfortable in churning out code, but the logic behind my contributions was limited to resources that were suitable for (astronomy) Ph.D. folks, my lack of depth in understanding became a bottleneck."

RQ2: Reasons for Disengaging from Sci-OSS

Focus shifted

Project is stable

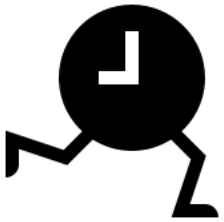
Conflicts

Prefer alternative project

Lack SE background

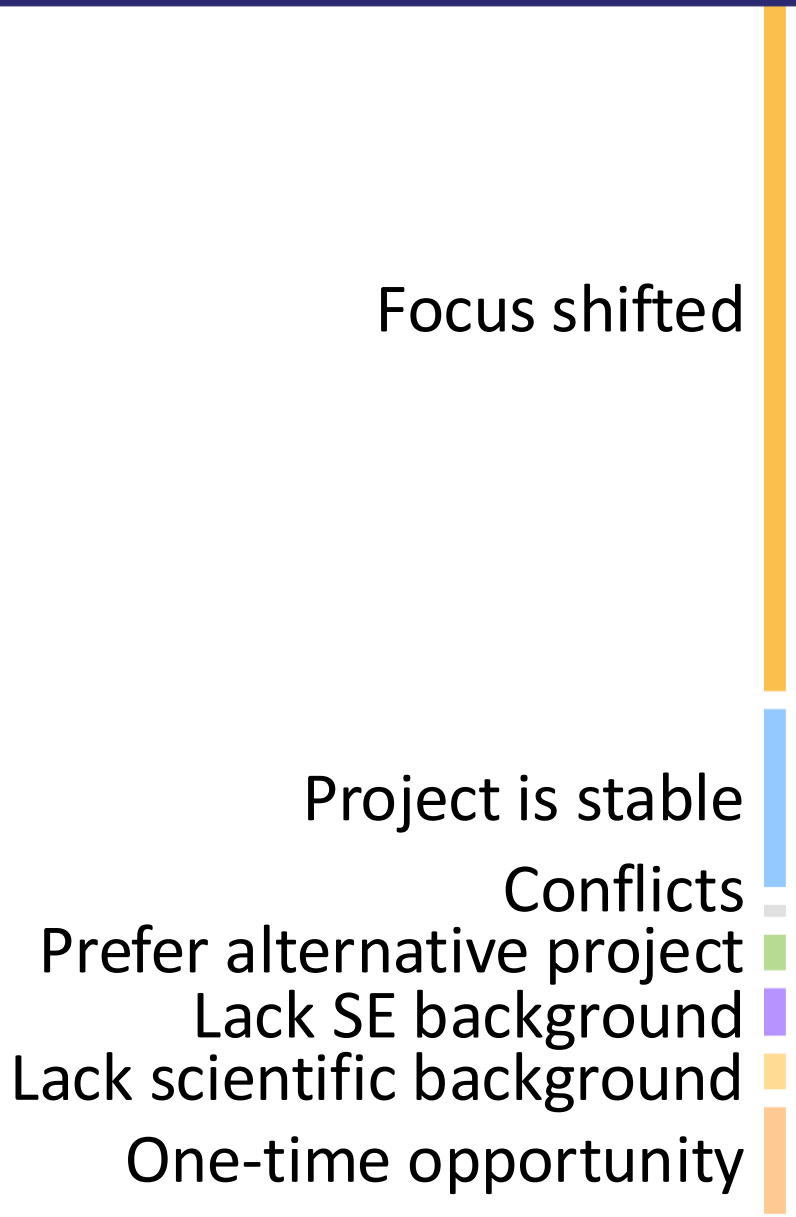
Lack scientific background

One-time opportunity



"I stopped contributing because I was not accepted in Google Summer of Code. So there was no motivation left."

RQ2: Reasons for Disengaging from Sci-OSS



General OSS

Occupational reasons such as employment transitions were the most common reasons for disengagement.

[Miller et al., 2019]

RQ2: Reasons for Disengaging from Sci-OSS

Focus shifted

Project is stable

Conflicts

Prefer alternative project

Lack SE background

Lack scientific background

One-time opportunity

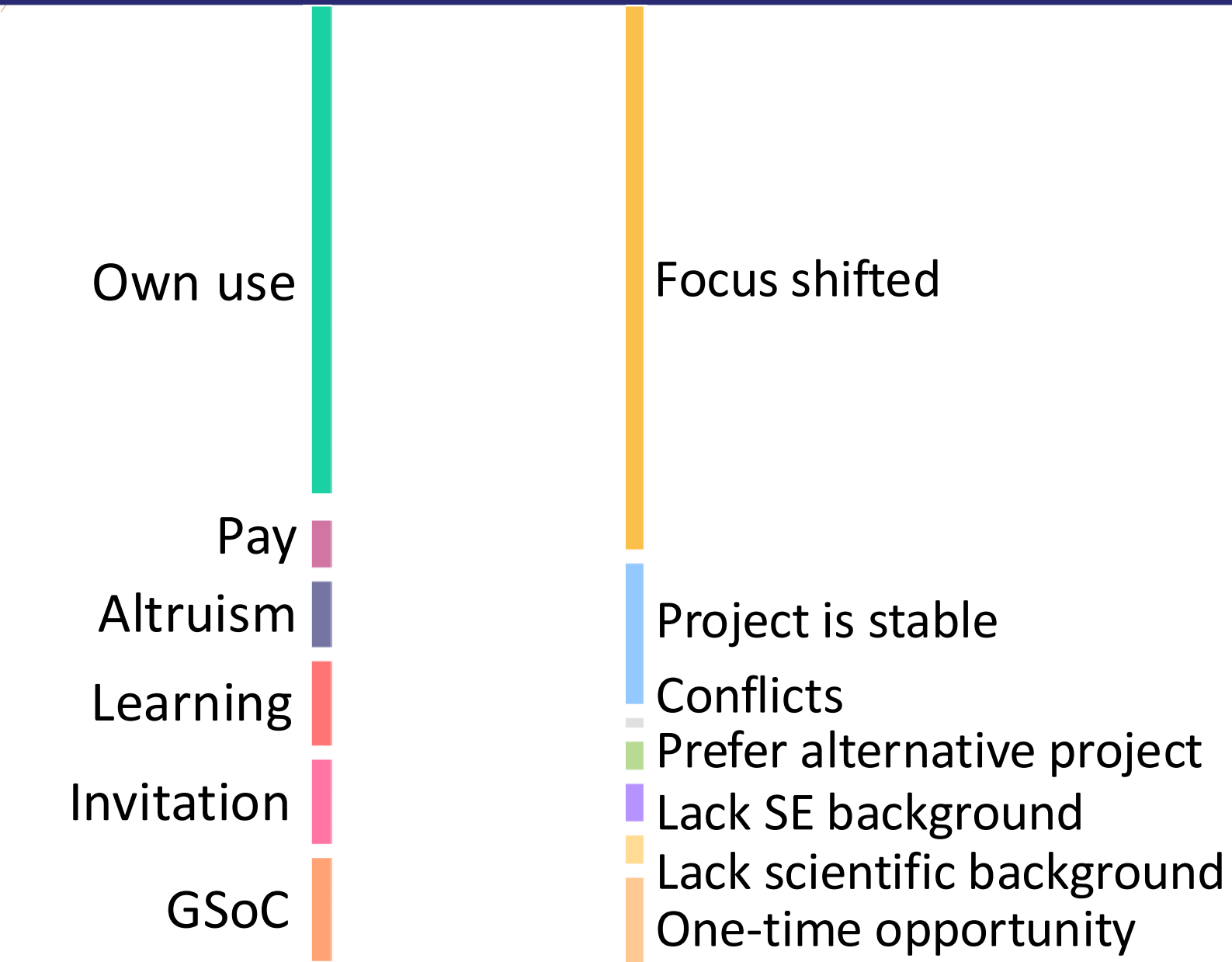
General OSS

Occupational reasons such as employment transitions were the most common reasons for disengagement.

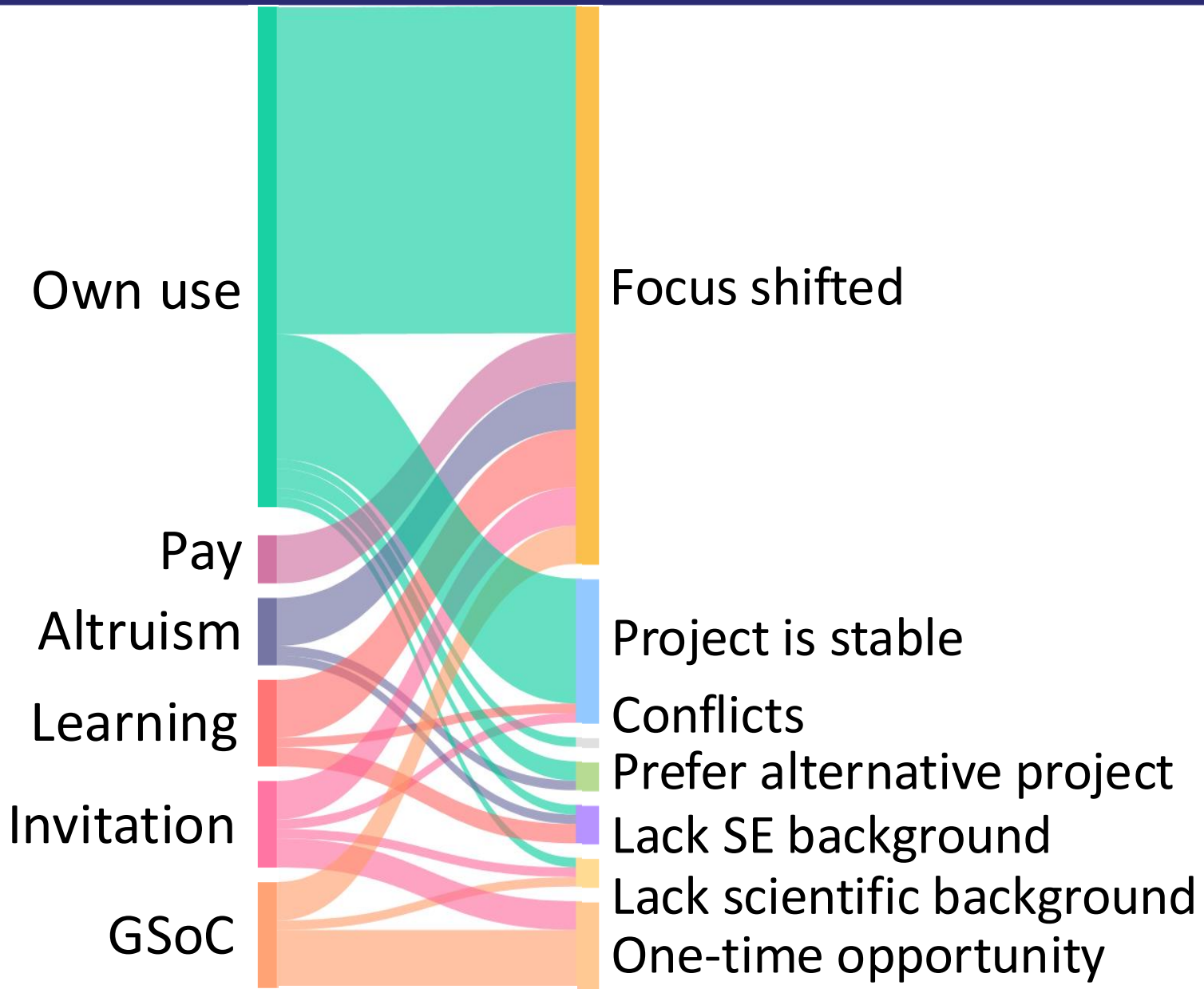
[Miller et al., 2019]

Similar to general OSS, scientific OSS has occupational reasons as a main cause of disengagement but also faces additional domain-specific and technical barriers.

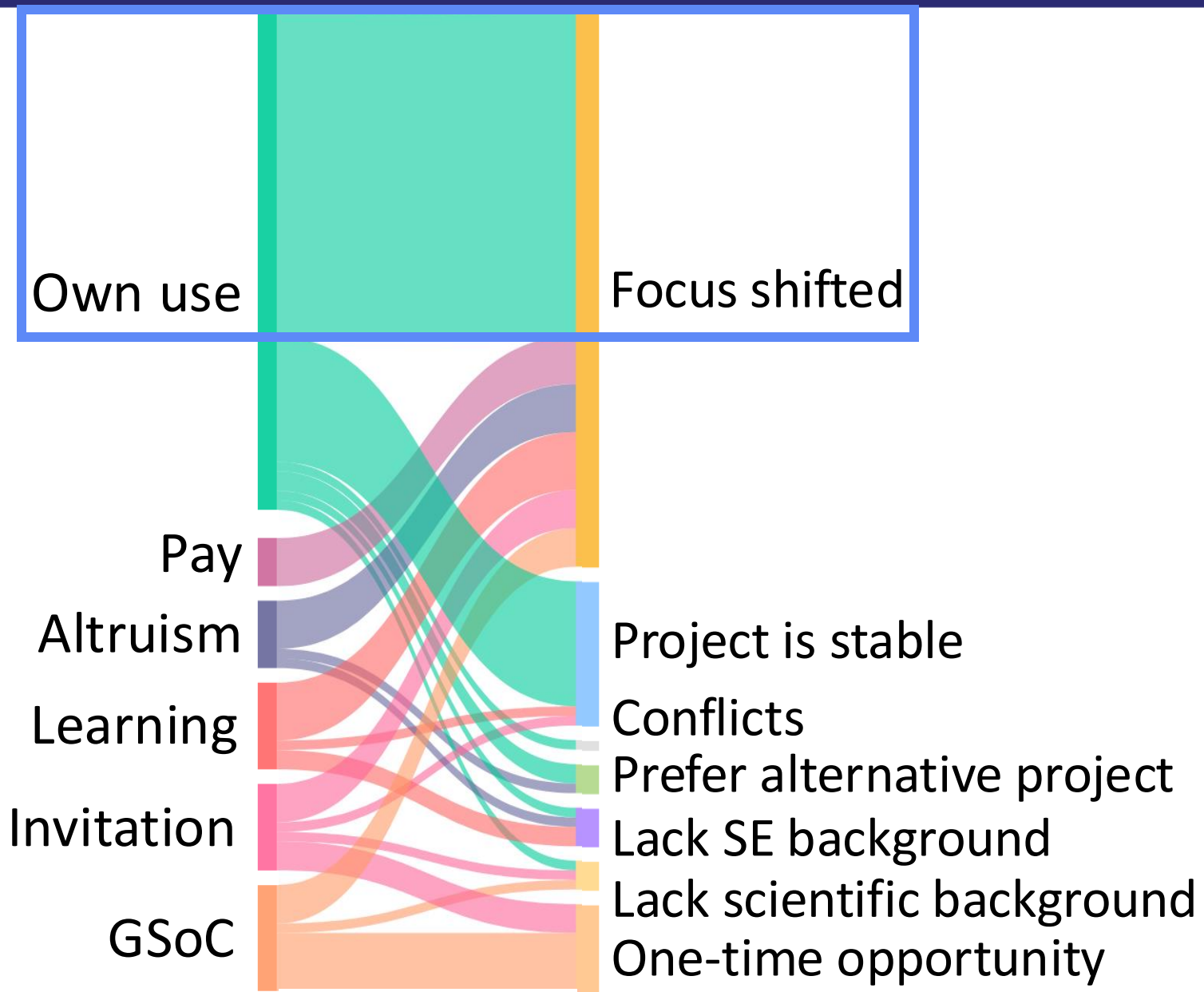
Mapping: from Contributing to Disengaging



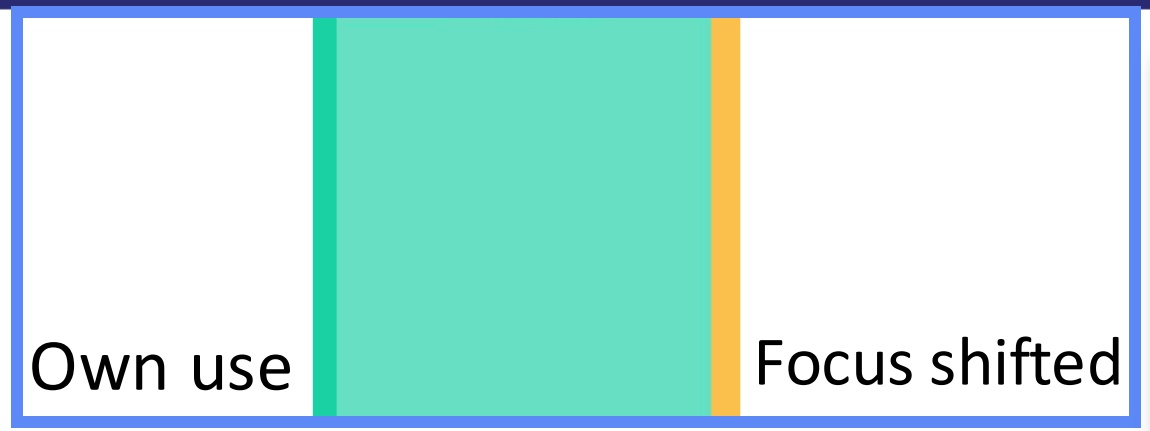
Mapping: from Contributing to Disengaging



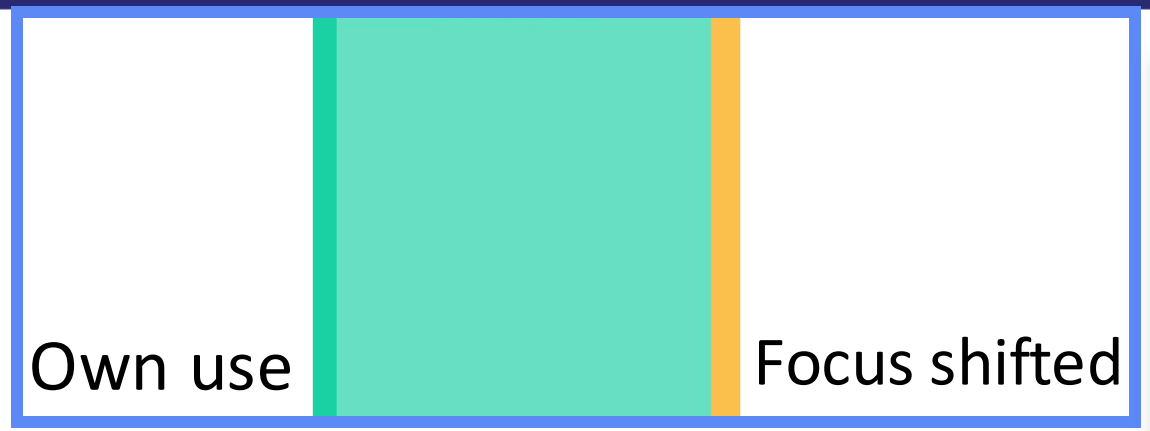
Mapping: from Contributing to Disengaging

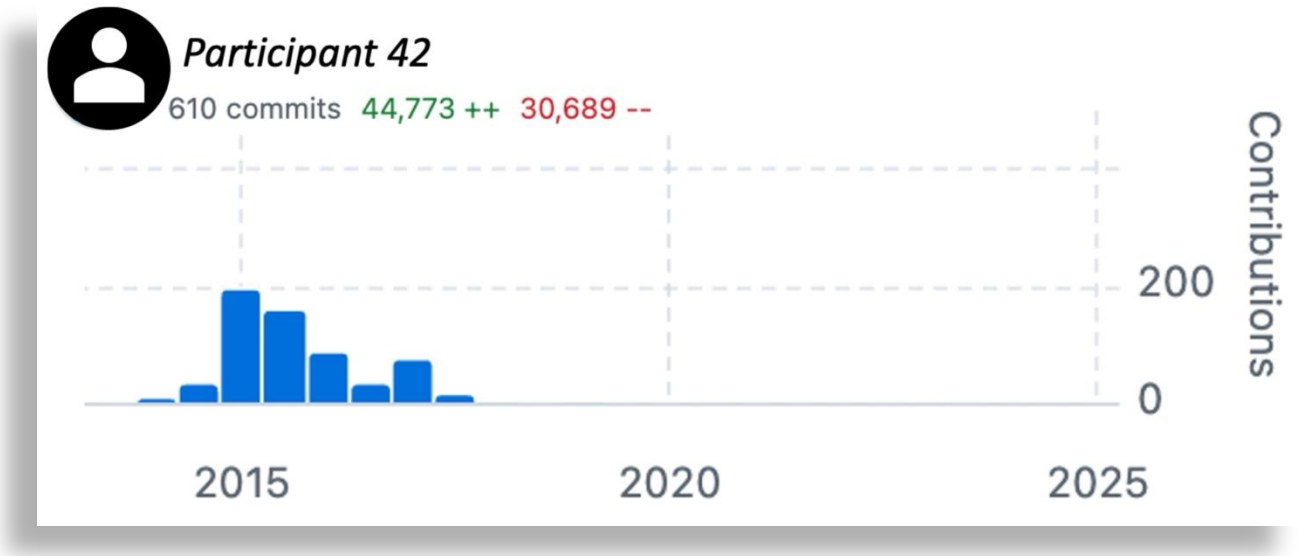


Mapping: from Contributing to Disengaging



Mapping: from Contributing to Disengaging



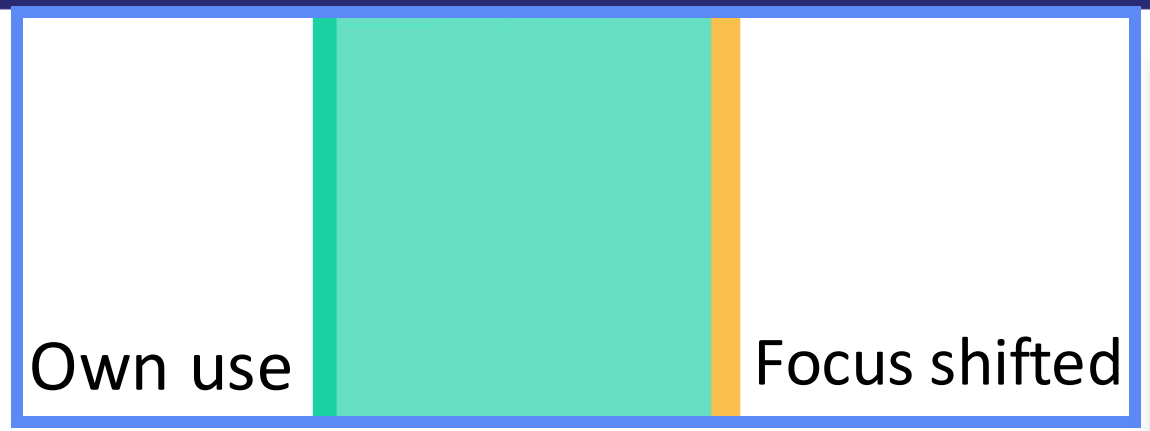


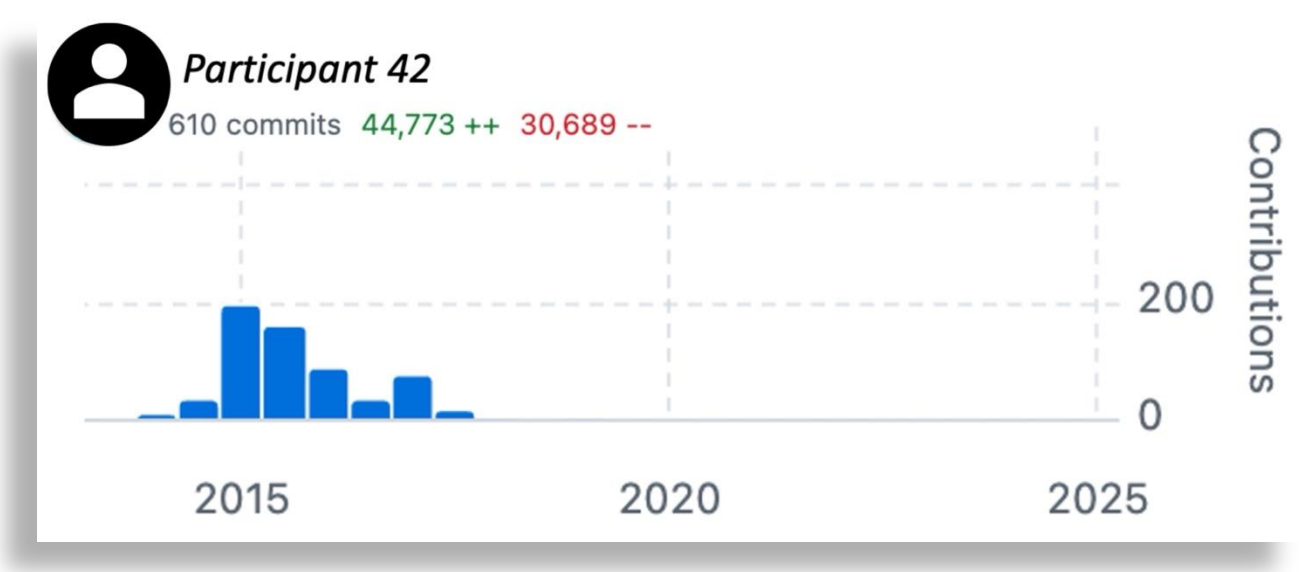
Why did you contribute?

Own use

"... it was essential for my research as a grad student." [2014]

Mapping: from Contributing to Disengaging





Why did you contribute?

Own use

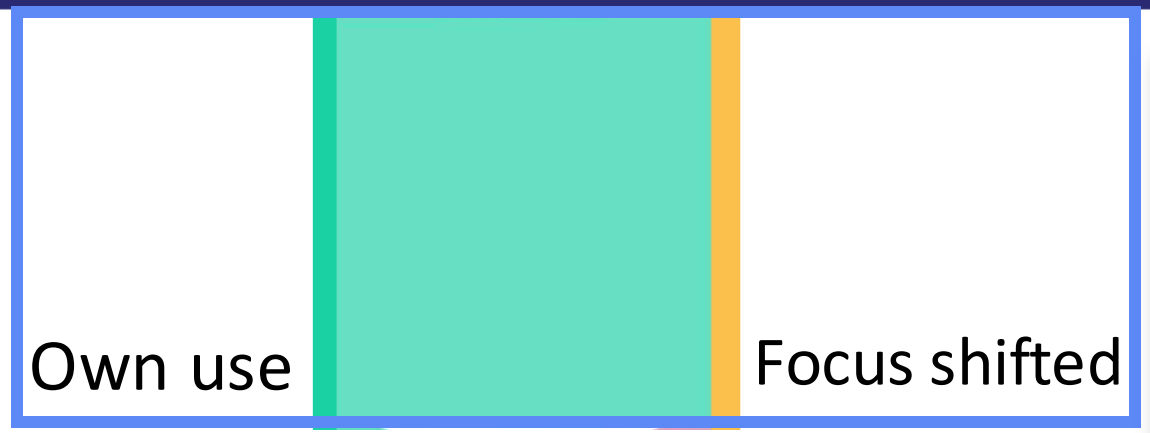
"... it was essential for my research as a grad student." [2014]

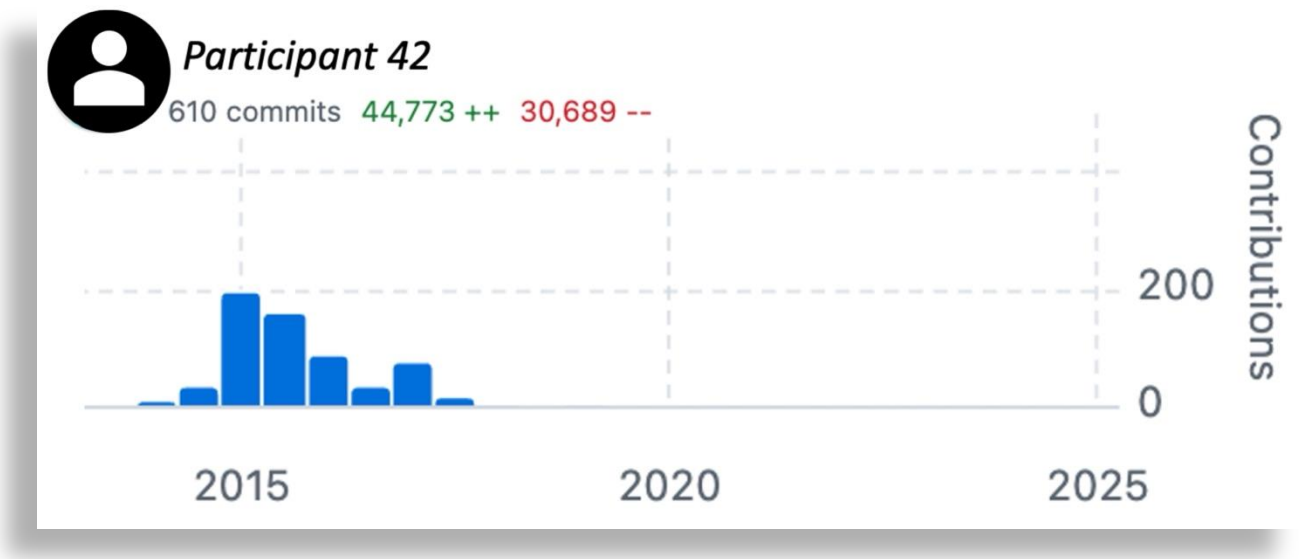
Why did you STOP contribute?

Focus shifted

"My research focus shifted away toward areas with more funding." [2019]

Mapping: from Contributing to Disengaging





Why did you contribute?

Own use

"... it was essential for my research as a grad student." [2014]

Why did you STOP contribute?

Focus shifted

"My research focus shifted away toward areas with more funding." [2019]

Mapping: from Contributing to Disengaging

"There wasn't much the community could have done — my career trajectory simply changed"

Retaining contributors in scientific OSS community can be difficult: mostly motivated by ***own use*** to contribute, inevitable disengagement and turnover will happen due to ***focus shift*** (graduation/research change).

Mapping: from Contributing to Disengaging

How to sustain the Sci-OSS communities? (RQ3)

Retaining contributors in scientific OSS community can be difficult: Mostly motivated by own use to contribute, inevitable disengagement and turnover happen due to focus shift (graduation/research change).

Study Goal: Is it the same for Scientific OSS?

General OSS

Contribution **motivation**
[Gerosa et al.,2021]

Reasons for **disengagement**
[Miller et al.,2019]

Strategies for improving community sustainability

Conceptual Replication

Scientific OSS

RQ1: Motivations for contributing to scientific OSS?

RQ2: Reasons for **disengaging** from scientific OSS?

RQ3: Suggestions for sustaining scientific OSS communities?

RQ3: Suggestions for Sustaining Sci-OSS Communities

Onboarding newcomer & retaining contributors

Undervalued engineering work

Building inclusive &engaging community

Fragmented communities

RQ3: Suggestions for Sustaining Sci-OSS Communities

Onboarding newcomer & retaining contributors

Undervalued engineering work

Building inclusive &engaging community

Fragmented communities

RQ3: Suggestions for Sustaining Sci-OSS Communities

Onboarding newcomer & retaining contributors

Undervalued engineering work

Building inclusive &engaging community

Fragmented communities

Missing connection

Code **Scientific theories**

- Need to have both sufficient knowledge about the code base and scientific domain knowledge to contribute.
- Time consuming for Sci OSS maintainers to make such documentations.

RQ3: Suggestions for Sustaining Sci-OSS Communities

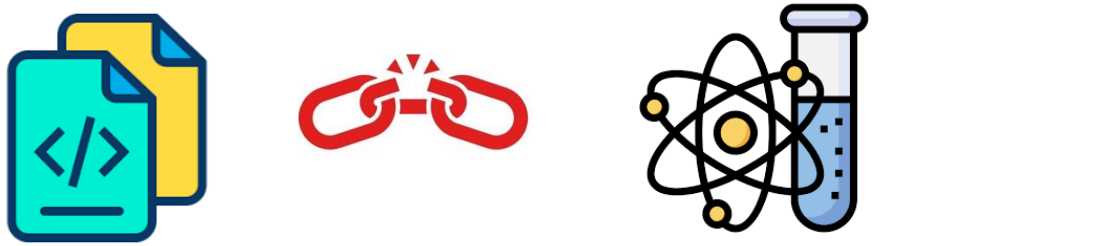
Onboarding newcomer & retaining contributors

Undervalued engineering work

Building inclusive &engaging community

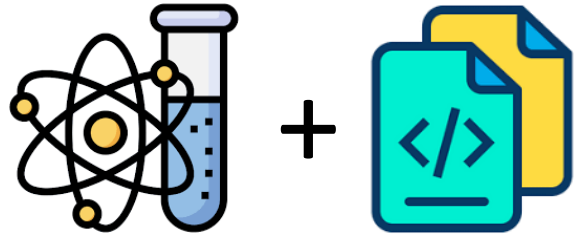
Fragmented communities

Missing connection



Code **Scientific theories**

Better documentation



SE researchers could look into building automated tools to help generate documentations connecting Sci + Code knowledge

RQ3: Suggestions for Sustaining Sci-OSS Communities

Onboarding newcomer & retaining contributors

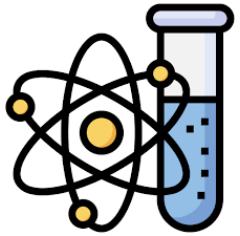
Undervalued engineering work

Building inclusive &engaging community

Fragmented communities

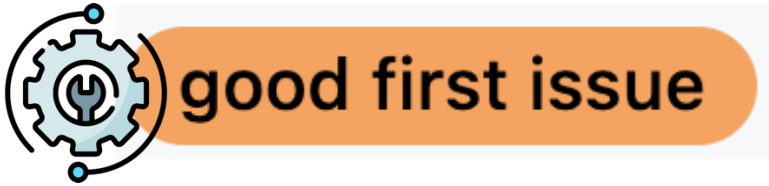
Missing connection

Code



Scientific theories

Tooling support



SE researchers could look into building automated tools to break down the domain specific tasks to help narrow down the scope and create more concrete guidance on the task.

RQ3: Suggestions for Sustaining Sci-OSS Communities

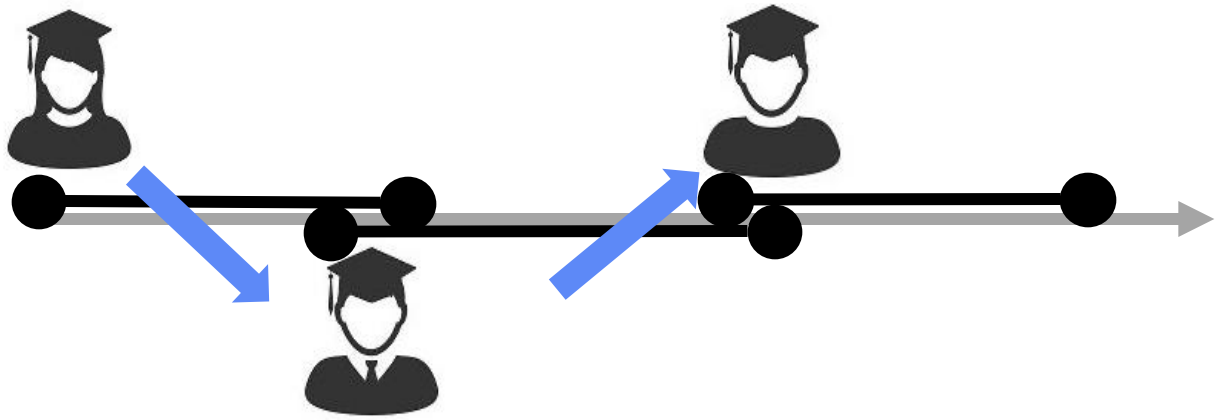
Onboarding newcomer & retaining contributors

Undervalued engineering work

Building inclusive &engaging community

Fragmented communities

Smooth transition between contributor turnover



As the inevitable turnover happen, respondents also suggest to pair graduate students in the field to scientific OSS project to support smooth turnover of contributors.

RQ3: Suggestions for Sustaining Sci-OSS Communities

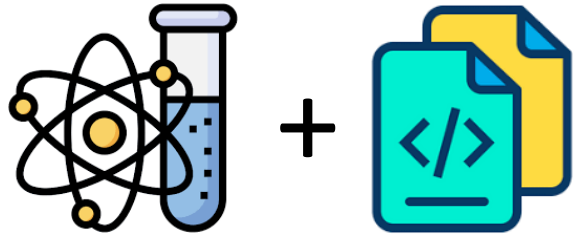
Onboarding newcomer & retaining contributors

Undervalued engineering work

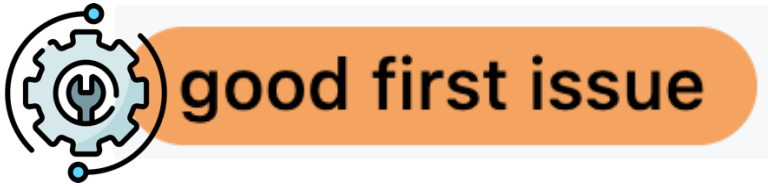
Building inclusive &engaging community

Fragmented communities

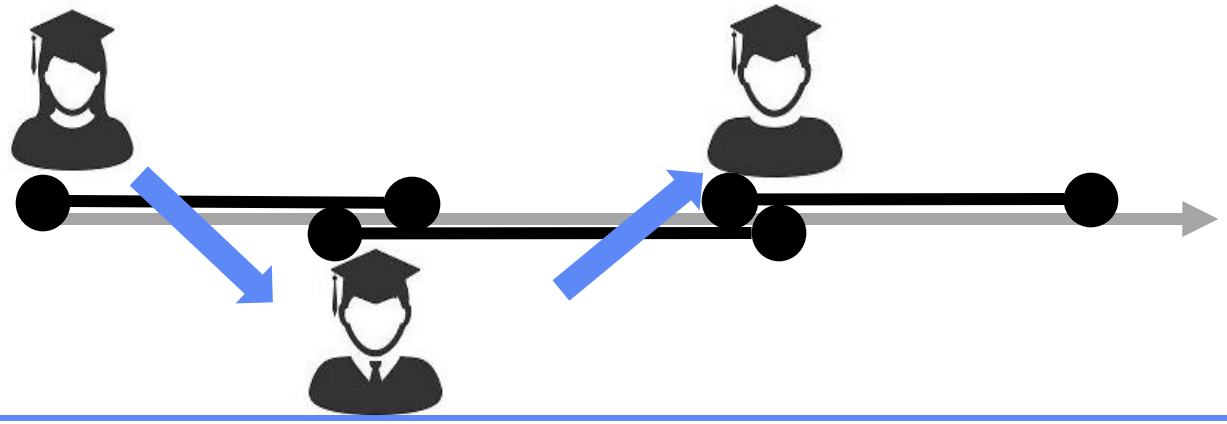
Better documentation



Tooling support



Smooth transition between contributor turnover



RQ3: Suggestions for Sustaining Sci-OSS Communities

Onboarding newcomer & retaining contributors

Undervalued engineering work

Building inclusive &engaging community

Fragmented communities

RQ3: Suggestions for Sustaining Sci-OSS Communities

Onboarding newcomer & retaining contributors

Undervalued engineering work

Building inclusive &engaging community

Fragmented communities

Publish papers on Sci-OSS projects

RQ3: Suggestions for Sustaining Sci-OSS Communities

Onboarding newcomer & retaining contributors

Undervalued engineering work

Building inclusive &engaging community

Fragmented communities

Publish papers on Sci-OSS projects

Astropy: A community Python package for astronomy
TP Robitaille, EJ Tollerud, P Greenfield… - Astronomy & …, 2013 - aanda.org
We present the first public version (v0.2) of the open-source and community-developed Python package, **Astropy**. This package provides core astronomy-related functionality to the …
☆ Save 99 Cite Cited by 13571 Related articles All 23 versions Web of Science: 9280

[PDF] aanda.org
Get full text

The **astropy** project: Building an open-science project and status of the v2. 0 core package
…, A Contributors, **Astropy** Collaboration… - The Astronomical …, 2018 - iopscience.iop.org
… of the **Astropy** community and the **astropy** core package and … We start by describing the way the **Astropy** Project functions … by the **Astropy** Project itself: a core package called **astropy** (…
☆ Save 99 Cite Cited by 10110 Related articles All 36 versions Web of Science: 1437 »

[PDF] iop.org
Full View

RQ3: Suggestions for Sustaining Sci-OSS Communities

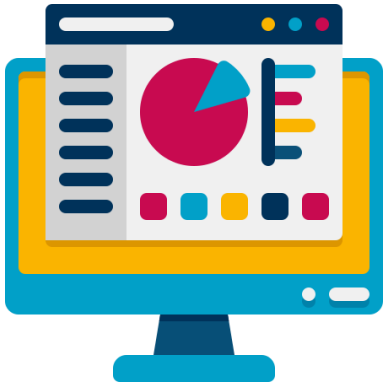
Onboarding newcomer & retaining contributors

Undervalued engineering work

Building inclusive &engaging community

Fragmented communities

Publish papers on Sci-OSS projects



Quantify impact of code contribution on science

RQ3: Suggestions for Sustaining Sci-OSS Communities

Onboarding newcomer & retaining contributors	Better support for contribution workflow
	Train scientists on SE best practices
	Offer more fellowship/internship opportunities
	Lower entry barrier in science
	Smooth transition of contributor turnover
Undervalued engineering work	Provide financial incentives
	Gain recognition through the academic reward system
	Acknowledge the impact of contribution
	Publish papers for the scientific OSS
Building inclusive &engaging community	Foster a welcoming community culture
	Organize gatherings to keep community engaged
	Accept contribution at all levels inclusively
Fragmented communities	Call for collaboration to reduce duplicated efforts

Future Work: Broader Scope + Beyond Sustainability

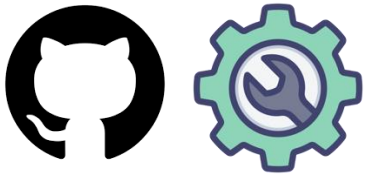
Future Work: Broader Scope + Beyond Sustainability

- Generalizing to broader scope

  - Hear from active contributors.
  - Compare across scientific OSS domains.

Future Work: Broader Scope + Beyond Sustainability

- Generalizing to broader scope
  - Hear from active contributors.
  - Compare across scientific OSS domains.
- Automated tools to help contribution workflow.



Future Work: Broader Scope + Beyond Sustainability

- Generalizing to broader scope
  - Hear from active contributors.
  - Compare across scientific OSS domains.
- Automated tools to help contribution workflow.

Accepted to CSCW2025

Collaboration Challenges and Opportunities in Developing Scientific Open-Source Software Ecosystems: A Case Study on Astropy

JIAYI SUN, University of Toronto
AARYA PATIL, Max Planck Institute for Astronomy
YOUHAI LI, Carnegie Mellon University
JIN L.C. GUO, McGill University
SHURUI ZHOU, University of Toronto

Scientific open-source software (OSS) has greatly benefited research communities through its transparent and collaborative nature. Given its critical role in scientific research, ensuring the efficiency of collaboration within development teams has become vital. Earlier

Advancing Sustainable Communities in Scientific OSS: A Replication Study with Astropy

Are Contributor Retention Challenges the Same in General vs Scientific OSS?

General OSS

Contribution **motivation** [Gerosa et al.,2021]

Reasons for **disengagement** [Miller et al.,2019]

Strategies for improving community sustainability

Conceptual Replication

Scientific OSS

RQ1: Motivations for contributing to scientific OSS?

RQ2: Reasons for **disengaging** from scientific OSS?

RQ3: Suggestions for sustaining scientific OSS communities?

51 repositories

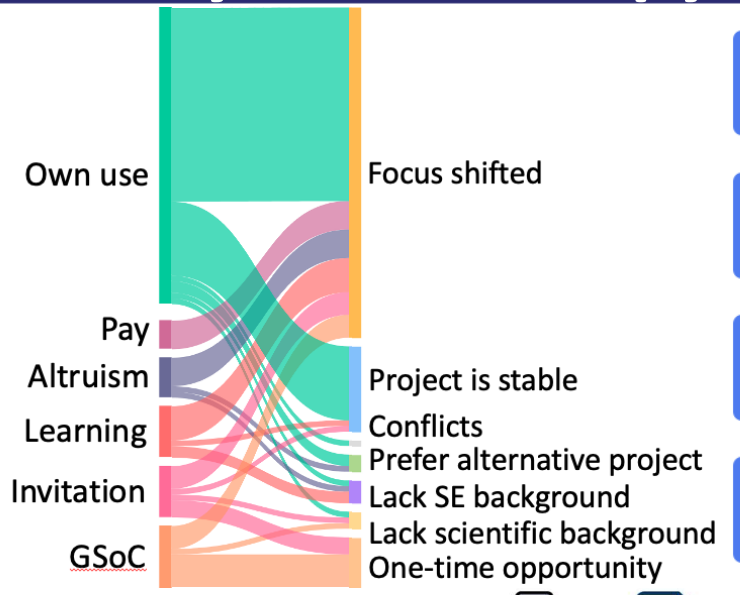
1116 code commit contributors

The Astropy Project

292/469 disengaged contributors with public email contact

3 Open-ended survey questions
RQ1: Motivation for contributing
RQ2: Reasons for disengaging
RQ3: Suggestions for improvement

Qualitatively analyzed 80 Reponses from 292 recipients



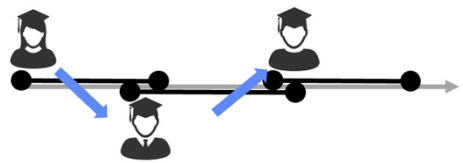
- Onboarding newcomer & retaining contributors
- Undervalued engineering work
- Building inclusive &engaging community
- Fragmented communities

Better documentation

Tooling support

good first issue

Smooth transition between contributor turnover



UNIVERSITY OF TORONTO

FORCOLAB

Contact me & read our papers!

Collaboration Challenges and Opportunities in Developing Scientific Open-Source Software Ecosystems: A Case Study on Astropy

JIAYI SUN, University of Toronto
AARYA PATIL, Max Planck Institute for Astronomy
YOUHAI LI, Carnegie Mellon University
JIN L.C. GUO, McGill University
SHURUI ZHOU, University of Toronto

CSCW2025

Scientific open-source software (OSS) has greatly benefited research communities through its transparent and collaborative nature. Given its critical role in scientific research, ensuring the efficiency of collaboration within development teams has become vital. Earlier