Understanding and Supporting the Choice of an Appropriate Task to Start With In Open Source Software Communities

Igor Steinmacher  Tayana Conte  Marco A. Gerosa
“Newcomers are explorers who must orient themselves within an unfamiliar landscape...”

[Degenais et al. 2010]

... and need support on their first steps

Context

- **Open collaboration** communities leverage and depend on new contributors
  - Innovation
  - Sustainability

- Ideally, open collaboration communities present a low barrier to entry
  - However...
Newcomers are susceptible to several barriers when they are onboarding:

- Learning curve
- Coordination
- Communication
- Cooperation
Context

- **Open Source Software** communities are open collaboration communities
  - Newcomers are usually *developers* willing to place a contribution
  - Delivering a contribution is a complex, multi-step process

- Multiple steps $\rightarrow$ multiple barriers $\rightarrow$ increase chance to giving up
Scope

- Community-based OSS projects
- First contribution to a project
- Contribution $\rightarrow$ source code
- Newcomer $\rightarrow$ developer trying to place the first contribution
In a previous research, we identified 57 barriers faced by newcomers when onboarding to OSS projects from a literature review, a survey and interviews with OSS practitioners.

“Finding a task to start with” was recurrently reported and highlighted by some interviewees.
Finding a task to start with

- It is hard to find an appropriate task in OSS communities, like:
  - Scratch their own itch
  - Find an existing task to accomplish
    - skills
    - complexity
    - high amount of tasks
    - outdated information
Finding a task to start with

- Issues
  - 314 Open
  - 580 Closed

  - google showing gephi.org as "This site may be hacked."
    - #947 opened 22 hours ago by jschoch
  - Impossible to install it on my MacBook Air
    - #946 opened 6 days ago by plemonner
  - 0.8.1 - 0.8.2 - Neo4j Plugin
    - #945 opened 10 days ago by dav009
  - Application will not open
    - #944 opened 13 days ago by smnuarmy
  - Problematic gephi installation on ubuntu 12.04
    - #943 opened 17 days ago by varun7rs
Finding a task to start with
Back to the goal

Better understanding the barrier “Finding a task to start with” when a newcomer is trying to place their first code contribution to Open Source Software projects
Method

- Systematic Literature Review
  - 20 primary studies
- Feedback from students after an assignment
  - 9 feedbacks
- Survey with practitioners
  - 24 answers
- Interviews with practitioners
  - 36 interviews

Reported finding a task as a barrier:
- 4 studies
- 6 students
- 3 answers
- 9 interviewees
Method - In-depth understanding

- Semi-structured interview with OSS practitioners
  - Recruitments → Email to mailing lists and forums of projects
  - 36 interviews
    - 11 experienced members
    - 16 newcomers that succeeded
    - 6 dropout newcomers
    - 3 newcomers that were still trying to place their first contributions
  - 14 projects
  - Text-based chat tools
Method - In-depth understanding

- Procedures of Grounded Theory to analyze the answers
  - Open coding + Axial coding
- 3 iterations
  - Iteration 1 → 36 interviews
    - What are the barriers that hinder newcomers’ onboarding to OSS projects?
  - Iteration 2 → 6 interviews
    - Clarification of doubts and gaps found in Iteration 1
  - Iteration 3 → 5 in-depth interviews
    - Better understanding the barrier “finding a task to start with”
    - Included a re-analysis of data
Results

- Newcomers don't know how to find the right place to change
- Newcomers choose a huge/wrong task
- Newcomer need to change the issue they are working on
- Issues present wrongly set difficulty
- Outdated issues

Documentation Issues

- Lack of documentation on 'how to contribute'
- Tutorials are not complete

Difficulty to choose/find their tasks

- Newcomers don't know how to reduce the scope
- Newcomers' lack of confidence to choose the right task
- Lack of direction on the beginning

Problems with the information provided in the issues

- Lack of information about required skills
- Lack of information about difficulty level
Results

- Lack of direction at the beginning
  - **Newcomers** expect that someone indicate them what are the tasks they should start with
    - “We feel more secure when someone is guiding us. I think this is related to the experience with large projects and Open Source...”
    - “If I could have got some hand holding that might have helped...”
Results

- **Lack of confidence**
  - **Newcomers reported** that they do not feel confident enough to choose the task they should choose and to handle it
    - “I really didn’t know which one to pick, which one I was ‘authorized’ to pick... whether they are important.”
    - “A college colleague gave up... He did not find a task that he felt confidence to try.”
    - “it is frustrating, there is a bunch of issues, but I don’t have the proper knowledge to judge what exactly I can touch...”
Results

- **Newcomers do not know how to reduce the scope**
  - Mainly reported by experienced members
    - "sometimes they want to contribute but don’t know how to reduce the scope when starting”
    - “the task chosen [by the newcomer] makes some sense, but is huge, and the newbie thinks she’ll be able to implement it in a few days.”
Results

- Problems with the information provided in the issues
  - Recurrently reported by newcomers and experienced members
    - “the issues that I can contribute were not clearly defined so it was my job as a newcomer to find out how to contribute there were no easy or junior bugs as in some projects.”
    - “the issues should indicate the area of knowledge, like C++, build, shell script”
Results

• **Documentation Issues**
  - “*there is no good guide for starter... I could not find how to choose a task*”
  - “*A proper up-to-date guide with tutorials would have really helped. As in when I began, I was totally new to open source, but the tutorials skipped over certain instructions and I had no idea what I was doing wrong.*”
Existing solutions

- We looked for solutions:
  - Existing literature (systematic literature review)
  - Qualitative analysis in data from practitioners
  - State-of-the-practice
    - Manual analysis of 6 pages from 6 projects:
      - Mozilla
      - MediaWiki
      - Ubuntu
      - GNOME
      - Apache Open Office
      - Libre Office
## Existing solutions (state-of-the-practice)

<table>
<thead>
<tr>
<th>Strategy adopted</th>
<th>MediaWiki</th>
<th>Mozilla</th>
<th>Ubuntu</th>
<th>Gnome</th>
<th>Open Office</th>
<th>LibreOffice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific page (tutorial) for newcomers</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>List of easy tasks by difficulty</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>List of easy tasks by required skills</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>List of easy tasks by topic/subproject</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>List of mentored bugs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mentorship per subproject</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Form to send an email with your goals and skills requesting a task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Strategy</td>
<td>Literature</td>
<td>Identified in Qualitative Analysis</td>
<td>State-of-the-practice</td>
<td>Can solve (hypotheses)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------------</td>
<td>-------------------------------------</td>
<td>-----------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific page (tutorial) for newcomers</td>
<td></td>
<td>Suggestion</td>
<td></td>
<td>Problems with information provided in the issues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>List of easy tasks by difficulty</td>
<td>•</td>
<td>Existing</td>
<td></td>
<td>Problems with information provided in the issues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>List of easy tasks by required skills</td>
<td></td>
<td></td>
<td>•</td>
<td>Problems with information provided in the issues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>List of easy tasks by topic/subproject</td>
<td></td>
<td></td>
<td></td>
<td>Problems with information provided in the issues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>List of mentored bugs</td>
<td>•</td>
<td></td>
<td>•</td>
<td>Problems with information provided in the issues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mentorship per subproject</td>
<td>•</td>
<td></td>
<td></td>
<td>Lack of confidence to choose a task</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Form to send an email with your goals and</td>
<td>•</td>
<td></td>
<td></td>
<td>Newcomers don’t know how to reduce the scope</td>
<td></td>
<td></td>
</tr>
<tr>
<td>skills requesting a task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examples in the easy bugs</td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommend artifacts that can be related to</td>
<td></td>
<td></td>
<td></td>
<td>Problems with information provided in the issues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>the issues</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommend similar bugs and artifacts</td>
<td></td>
<td></td>
<td></td>
<td>Newcomers don’t know where to place the changes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conclusions

• Finding a task to start is complex and depends on many variables

• Lack of confidence to choose a task appeared as a relevant concept in our analysis.
  ▫ Newcomers need to be properly guided
  ▫ Support their decision about which task is more suitable
Conclusions

- We identified some already existent and suggested strategies
  - Some are already used
  - Applying appropriate filters
    - newcomers can find a subset of tasks that can fit their goals, interests and skills
  - Pointing them a possible mentor

- Strategies found in the literature were not identified in any project or interview

- We conceived a portal to support newcomers on their first steps:
  - www.flosscoach.com
THANKS!
The author are deeply sorry they could not get funds to attend the conference and present the paper.

Questions, concerns and suggestions are very welcome. Please direct them to:

Igor Steinmacher
igorfs@utfpr.edu.br