

# Understanding Newcomer Activities Prior to Onboarding Open Source Software (OSS) Projects on GitHub.

Name: Ifraz Rehman

Laboratory's name: Software Engineering

Supervisor's name: Kenichi Matsumoto

Abstract ([should be within 1st page](#))

Understanding newcomer activities prior onboarding Open Source Software (OSS) Projects on GitHub holds significance for researchers and practitioners seeking insights into their preparatory practices prior to onboarding. Firstly, I systematically map related work on onboarding of developers to 102 studies. I presents key topics offering insights into current trends and gaps in developer onboarding, and motivate this thesis. To fill the gap, I conduct an empirical study that is broken into three parts: (i) identifying Newcomers through survey who are potential candidate to OSS projects, I find 171 Newcomer OSS-candidates (i.e., 85%) with no prior experience contributing OSS, and have (i.e., 82%) intention to later onboard to OSS projects. (ii) validate their pre-onboarding activities and characterize them through mixed method approach, finding shows Newcomer OSS-candidates like to target software-based repositories (i.e., 66%), their first contributions are mainly associated with development (commits) and maintenance (PRs), and are less likely to practice social coding. (iii) Finally, analyzing the proportion of them who are onboard to OSS projects in GitHub, I find that Newcomer OSS-candidates eventually end up onboarding (i.e., 30% quantitative, 70% follow-up survey) an OSS project. Furthermore, they cite *finding a way to start* as the most challenging barrier to contribute. Suggestions for newcomers, should not be afraid to individually contribute to their own code, contribute adding new content or making documentation to upstream software repositories, or fork OSS projects before attempting to onboard. For OSS projects, it might start with tasks to update the documentation, formatting or cleaning up code. For Researchers, the majority of targeted repositories are software-based, this insight helps to understand the role of software based experimental, documentation, and web based application libraries and frameworks repositories in platforms like GitHub, that should cater for developers.