

## **Supporting the Next Generation of OSS Engineers**

### **Introduction**

The blockchain is very famous in our daily life, such as Bitcoin, even most of us don't know what blockchain exactly is. And, today, more and more open source software is written in Blockchain. Open source, as the name suggests, is a licensing agreement that allows users to freely modify their work, use the work in new ways, integrate work into larger projects, or get new jobs based on original work. It is most commonly used in the software industry. The term originates from the context of a software development environment and is used to specify a specific method of creating a computer program. Open source software (OSS) is a type of computer software whose source code is distributed under license, in which the copyright owner grants the user the right to research, modify and distribute the software to anyone and for any purpose. The idea of open source comes from the technology community. From the early days of computers, programmers and engineers have developed new technologies through collaboration. In many ways, the creation of open source licenses is a direct response to proprietary manufacturing. Community-based open source software (OSS) projects are often self-organizing and dynamic, receiving the contributions of distributed volunteers. New immigrants are critical to the survival, long-term success and continuity of these communities. However, newcomers face many obstacles when making their first contribution to the OSS project, which in many cases leads to dropouts. By eliminating barriers between innovators, open source promotes the free exchange of ideas within the community and encourages innovation, science and technological advancement.

On October 19th, 2018, Dr.Gerosa presented his research to everyone on INF501 course in Northern Arizona University. By leveraging large repositories of sociotechnical data and the human perspective, Dr.Gerosa's study systematically conceives theories and innovative support

for OSS software developers, explores the barriers that hinder the contribution of newcomers in Open source software (OSS) projects and conduct the newcomer's barriers model to overcome the contribution barriers.

### **Research Topic Overview**

Dr.Gerosa's research focuses on the systematically conceive theories and innovative support for OSS software developers, leveraging large repositories of socio-technical data and the human perspective. Open Source Software (OSS) is exciting and unusual because it is an intensively collaborative product, it is a real-world software, it has a real problem, and real people create it. For example, Github has over 67 million repositories and almost 24 million users. By placing an open source license on the original work, the individual or organization agrees to provide the entire program code to the public, allowing anyone to modify, enhance or redesign the program code, to allow the creation of derivative works, to enable the program to be used for any purpose intended by the user.

While open source software (OOS) has many advantages, new software engineers still face some obstacles in the development of open source software (OOS).In Dr.Gerosa'research, professor Gerosa explores the barriers that hinder the contribution of newcomers in Open source software (OSS) projects and conduct the newcomer's barriers model to overcome the contribution barriers. In the research, Dr.Gerosa use qualitative methods in which a researcher can understand social phenomena in their natural environment and analyze the rich data available in software repositories to discover exciting and actionable information about software systems and projects. Initially, the researchers conducted a qualitative analysis and established a model for the preliminary practitioner by analyzing the feedback from students. The researchers then did a

systematic literature review and obtained a literature review model. And by using interview analysis, the researchers established another model which is interviews' model. Finally, by using the bias analysis, the researchers merged this three models and developed a new model which helps the novice in open source software (OSS) to find barriers.

The Newcomers Barriers Model shows obstacles for novices in five different ways which are newcomer's characteristics, newcomer's orientation, communication issues, documentation problems, and technical hurdles. In Dr.Gerosa's presentation, the professor also introduces a new tool---FLOSScoach---which is a web portal created to support newcomers to OSS projects. To evaluate the portal, the professor conducted a study of 65 students, relying on qualitative data on diaries, self-efficacy questionnaires, and technical acceptance models. The results of the research indicate that the FLOSScoach plays an essential role in guiding new immigrants and reducing barriers associated with the orientation and contribution process, while it is ineffective in lowering technical obstacles. The study also found that the FLOSScoach to be very useful, easy to use, and add to the contributions of novices. The results can help project maintainers decide that more attention is needed to help newcomers to OSS projects overcome barriers to entry.

In the study, the researchers found that the portal improved the contribution process experience of new members and served as a compass in the contribution process. And newcomers who use the web portal have more confidence to decide what steps they need to take to achieve their goals. However, the researchers could not identify any significant progress in supporting newcomers to overcome technical barriers. The practical application of the study to the obstacle model led to observations leading to new versions of the model, including some rearrangements. The results of the research point to a set of features that can help OSS project novices overcoming entry barriers, including a categorized task list, a definite contribution process related to useful

documents, and a message template that supports beginners interacting for the first time on the mailing list. At the same time, the researchers in the study are currently developing a new and more collaborative version of FLOSScoach that will create and maintain entries for other projects.

## **Discussion**

Although my research topic which is pair programming based on computer programming self-efficacy is not very close to open source software (OOS), I also can distinguish some critical connections between research and Dr.Gerosa's research. Both of our research use one psychological concept which is self-efficacy. In 1997, Bandura stated that self-efficacy is used to evaluate an individual's mental state and refers to individual beliefs in different situations. Moreover, Bandura also explored that self-efficacy influences students' activity choice, including how much effort or time they will invest in solving particular tasks and situations. The difference between my research and Dr.Gerosa's is that the researchers use self-efficacy in the different stage on the experiment. In pair programming, the researchers use programming self-efficacy score at the very beginning of the experiment. More specific, based on the score of the computer programming self-efficacy, instructors placed students in level-matched for a portion of their programming exercises in the first day of the class. However, in Dr.Gerosa's research, the researchers use self-efficacy in the FLOSScoach evaluation stage. The principle of analysis methods in Dr.Gerosa's research such as analysis from multiple viewpoints, turning research results into relevant products and the necessity of open collaboration community can also be used in my study.

As a computer science student, I know that the essential functions of the Internet are built on open source software (OOS). For example, the Linux operating system manages the operation

of the web server, and the Apache web server application negotiates the data transfer between the global server infrastructure and the personal device. Numerous Internet applications are also open source. Such as large Internet companies like Facebook and Google have even opened up some proprietary innovations to the open source community. By increasing efficiency and simplifying the entire Internet, communication technologies can improve the experience for all users. Open source licenses encourage innovation through collaboration. Without it, many of the techniques that we take for granted today will never develop or be bound by patent law. The open source movement is why technology has grown at such a fantastic rate over the past few decades.

Open source will be the critical driver of innovation. For example, tools like TensorFlow have helped machine learning democratize by making machines learn what they do in the company. Open source as a means of achieving innovation and more and more companies will begin to focus on open source to drive the roadmap for the future. The future of open source is where advanced, cutting-edge technology is developed. This is a true geek, even a daily person who gathers and lets the computer do everything from thinking to manufacturing. As a computer science student, I want to see the broadest impact around AI and machine learning. The future of open source software is bright. More and more suppliers will contribute new open source software. Non-traditional software companies such as Capital One will continue to develop new open standards and provide open source software. Meanwhile, more and more open source software will be written in the AI, Blockchain. Open source organizations like the Apache Foundation will continue to drive new projects in the open source space and will create new and innovative open source communities such as [openai.com](https://openai.com).

## **Conclusion**

In this response essay, I wrote the main parts of Dr.Gerosa's research that relate to my research and my major. His experiments are based on statistical analysis viewpoints (qualitative analysis and quantitative analysis) more than software aspects of the study. I highly appreciate his research in term of the investigation of the newcomer's barriers model, because it is very beneficial for me---a computer science student. Dr.Gerosa's research worked on analyzing the obstacles for novices in five different ways which are newcomer's characteristics, newcomer's orientation, communication issues, documentation problems, and technical hurdles.

## **Reference**

- 1.Steinmacher I, Silva M A G, Gerosa MA, et al. A systematic literature review on the barriers faced by newcomers to open source software projects[J]. Information and Software Technology, 2015, 59: 67-85.
- 2.Steinmacher I, Gerosa M, Conte T U, et al. Overcoming Social Barriers When Contributing to Open Source Software Projects[J]. Computer Supported Cooperative Work (CSCW), 1-44.
3. Steinmacher I, Conte T, Gerosa MA, et al. Social barriers faced by newcomers placing their first contribution in open source software projects[C]//Proceedings of the 18th ACM conference on Computer supported cooperative work & social computing. ACM, 2015: 1379-1392.
- 4.Steinmacher I, Pinto G, Wiese I S, et al. Almost there: A study on quasi-contributors in open-source software projects[C]//2018 IEEE/ACM 40th International Conference on Software Engineering (ICSE). IEEE, 2018: 256-266.