RESEARCH STATEMENT

Igor Steinmacher
School of Informatics, Computing, and Cyber Systems
Northern Arizona University (NAU)
igor.steinmacher@nau.edu

My research is at the intersection of Software Engineering (SE) and Computer Supported Cooperative Work (CSCW), specifically Open Source Software, Human Aspects of Software Engineering, Empirical Software Engineering, Data Analysis, Social Network Analysis, and Mining Software Repositories techniques. I have experience in both quantitative and qualitative research, applying the methods and approaches most suitable for the problem and research.

The Open Source Software (OSS) model has become an important driving force in today's software development, resulting in many prominent projects that are extensively used throughout the development stack, from kernels to sophisticated end-user applications. Even startups and commercial projects are increasingly participating in OSS. The 10th Annual Future of Open Source Survey[1] showed that 65% of the surveyed companies leverage OSS to speed up application development and 55% leverage OSS for production infrastructure.

It is therefore no surprise that the OSS movement attracts a large, globally distributed community of volunteers. Developers want to participate in OSS as their contributions help them learn, gain visibility, benefit society, and even get jobs. With hundreds of thousands of active OSS projects, comprising billions of lines of code and involving millions of developers, the survival, long term success, and continuity of these projects requires a continuous influx of newcomers. Although there are many who want to volunteer, newcomers face many difficulties entering OSS, and are left to learn on their own. This affects not only large and well-known projects like application servers, browsers, and frameworks, but also scientific software projects, deserving a specific workshop to discuss the topic.[2] These projects are usually built and maintained by scientists from interdisciplinary areas (e.g., bio-informatics, astronomy, and geology) who depend on software systems to conduct data analysis, simulation, and visualization.

Given the importance of OSS, and the challenge faced by newcomers joining a project, my research interests relate to understanding the social aspects of OSS communities, with a focus on new developer onboarding. Over the past 6 years, I have been analyzing the behavior of open source software developers using quantitative data mining and statistical techniques (e.g. [1–4]). as well as systematic literature review [6, 7], qualitative methods relying on interviews (e.g. [8, 9]), surveys (e.g. [1, 10, 11]), and diary studies (e.g. [12]). This research benefits OSS communities by providing tools and guidelines to better onboard newcomers and by demonstrating how to create an active and sustainable community. My results can potentially be applied to support scientific software projects for engaging new developers.

During my Ph.D., I used a mixed method approach and different data sources to create a model of barriers that newcomers face while onboarding to OSS projects. Based on this model, I developed and assessed a portal (http://www.flosscoach.com) that reduces the identified barriers by providing awareness information for newcomers [12]. The main results were published at the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW 2015) [8], and ICSE 2016 (International Conference on Software Engineering) [12], the top conferences for CSCW and SE fields, respectively.

My Ph.D. studies triggered several different research topics, which I am working on with different collaborators: Rafael Prikladnicki (PUCRS, Brazil), Sabrina Marczak (PUCRS, Brazil), Tayana Conte (UFAM, Brazil), Christoph Treude (University of Adelaide, Australia), Anita Sarma (OSU, USA), Thelma Colanzi (UEM, Brazil), and Gustavo Pinto (UFPA, Brazil). In addition, I am advising 2 M.Sc students, and co-advising 2 M.Sc. students and 3 Ph.D. students.

One of the topics relates to understanding **the role of social coding environments, such as GitHub and BitBucket, in attracting new contributors** to a project. These environments changed the way developers contribute to OSS projects, by providing social and collaborative features with a high degree of social transparency. These new features, together with a single/cross-project contribution model, fostered newcomers to participate in the process and boosted the number of casual contributors, developers who are only casually or briefly interested in a project and make a single contribution with no further commitment. Our goal is to better understand the casual contributor phenomenon, as well as the benefits and problems behind it, including how communities can make the best use of them. Preliminary results based on data qualitative and quantitative analysis of data gathered from GitHub and on answers to a complementary survey conducted with developers were published at ERA track of ICSME 2016 [2] and SANER 2016 [10].

---

[1] https://www.blackducksoftware.com/2016-future-of-open-source

[2] http://wssspe.researchcomputing.org.uk/

Another topic, which I am now conducting with my Ph.D. student Jefferson Silva (co-supervised by Dr. Marco Gerosa)**,** is **the long-term effects of offering money to new developers to start contributing to an open source project in programs such as Google Summer of Code**. One goal of GSoC program is to help OSS projects to identify and bring in new developers. However, little is known about the long-term effects of this kind of program. In this research, we are also taking advantage of mixed-method approaches, including data analysis approaches to analyze data extracted from software repositories and interviews and surveys with mentors and participants of GSoC. This research is being conducted in cooperation with Daniel German (University of Victoria, Canada) and a paper with preliminary results was just accepted at ICSME 2017 [5].

Recently, I began analyzing **how to take advantage of gamification approaches to improve the engagement and motivation of newcomers t**o OSS projects. This work is being conducted in cooperation with Sabrina Marczak (PUCRS) and our co-supervised M.Sc. student Carolina Toscani. We are designing game elements and implementing them in our newcomers' portal (FLOSScoach) to evaluate which elements are useful in this context [13]. I conducted a preliminary study in collaboration with an undergraduate student on gamifying GitLab and evaluating it as a way to engage students to OSS [14]. The results showed that some elements play a role in better orienting and motivating the students.

I have been teaching Open Source Software related courses for six years to undergraduates in Brazil. Because of my successful history teaching OSS to undergraduate courses and the impact of my research on OSS newcomers, I was invited to be a member of the Education Committee of the Brazilian Computing Society (SBC) to defend the idea of using OSS to teach computing in general, and, particularly, Software Engineering. The goal is to make OSS part of the SBC guidelines for undergraduate programs, by presenting how to use OSS as a tool to teach different Computer Sciences disciplines.

Aligned with this idea of using OSS in education, I started researching the benefits of using OSS projects to train the future workforce of software engineers [15]. This research is being conducted in collaboration with Dr. Anita Sarma (Oregon State University, USA) with the goal to create a support structure for familiarizing students with the technical and social aspects of an ecosystem of projects, thereby generating a workforce of contributors who can transfer knowledge across projects. We propose that worked examples and near-peer-mentoring can scaffold skill acquisition for newcomers, mainly in the beginning of their journey, when the knowledge gap between them and core members is too wide. Past, closed issues, and associated changes are archived in the software repository and can serve as worked examples.

Future research

A clear future direction for new research topics and projects is improving FLOSScoach and enabling its use in real settings. For example, I am involved studying the role of gamification to foster the engagement of new developers (in collaboration with Dr. Sabrina Marczak), and on developing tools to mining specific data sources to automatically feed FLOSScoach (in collaboration with Dr. Christoph Treude). In addition, a potential new research direction would be exploring how to engage new developers to scientific projects hosted or lead by NAU researchers (such as QIIME2[3]).

A potential opportunity involves Open Source Scientific Software projects, gamification, and Software Engineering Education and Training, encompassing different areas of knowledge and bringing benefits for multiple stakeholders. We aim to design and implement a gamified platform—continuing a preliminary work [13, 14]—that faculty can use to provide students with an assignment with real-world impact. By better understanding the specifics of scientific software projects, and taking advantage of my background on newcomers to OSS projects [7-12], we will propose strategies and tools to meet these projects' requirements, ultimately creating a more attractive and welcoming environment for new contributors. The gamified environment will foster new contributions and motivate existing members to keep contributing.

The broader impact of this work may be significant in several regards. (1) It allows for advancing knowledge while promoting teaching and learning by integrating the research findings and tools into graduate courses and supporting learning among newcomers to OSS. This will train the future workforce of software developers. (2) It provides infrastructure for researching Open Source Scientific Software, producing an integrated, open sourced gamified environment for use by researchers and communities. (3) We will disseminate our results to researchers and industry through publications, presentations, and blog posts. (4) The scientific community will benefit from our results, since our focus is on sustainability of Open Source Scientific Software. (5) The research benefits society by empowering thousands of newcomers to become successful contributors and yield tangible benefits, since participation in OSS is now also used for recruitment. (6) Moreover, our work has societal implications: OSS has become an important economic driving force and newcomers are critical for their growth and continuity. Our work will help in the continuity of OSS projects.

Besides further exploring my current research lines on the social aspects of software development, and on the educational use of OSS, I plan to use and apply my current background in OSS communities to different contexts and communities. I am interested in exploring new domains, and investigating how my research interconnects with different domains. For example, it would be interesting to study how to foster the onboarding the new volunteers to citizen science projects, by analyzing the role of gamifying these environments or by studying the potential barriers to new members.

[3] https://github.com/qiime2

TEACHING STATEMENT

Igor Steinmacher
School of Informatics, Computing, and Cyber Systems
Northern Arizona University (NAU)
Igor.steinmacher@nau.edu

I have taught university courses for 12 years and enjoy teaching and mentoring. Even while working in the software industry, I continued teaching undergraduate courses and mentoring students afterhours. My teaching philosophy is based in hands-on experience, collaboration, and constant challenge. Students are always motivated to learn by doing, and their work should extend the classroom. By applying this philosophy, I usually flip the classroom and put the students in charge of learning. When I am presenting, I put all my efforts into making myself clear by using the whiteboard and always requesting feedback from students. My classroom is a space of dialogue and practice, which goes well beyond the traditional "memorize and understand" approach to train better professionals.

During my teaching career, I've noticed that the steady flow of new techniques, tools, and processes for developing software systems, along with the interdisciplinary nature of software development, makes the software engineering discipline incredibly challenging to teach and learn. Addressing this, I make use of the following strategies in my teaching:

More than technical skills. In addition to teaching technical skills, which are important for the Computer Science career, it is essential for students to acquire a set of soft skills that include: critical thinking, the ability to work in teams, proactivity, and communication. I usually give group assignments that require practicing negotiation, oral and written communication, discussing ideas backed in evidence, and creativity for solving real-world problems.

Real-world experience. To keep students motivated while training them for their career, I bring to the classroom real, non-trivial problems that the software industry faces by taking advantage of Open Source Software (OSS). The motivation for using OSS is many-fold: (1) there is a plethora of choices, with different domains, sizes, and complexities; (2) well-known OSS projects often exhibit the breadth of scope necessary for real problems; (3) well-known OSS projects are maintained by an active global community of software developers; (4) the students work on an assignment with real-world impact. Exposing students to OSS gives them the opportunity to face and overcome social and technical barriers that they confront when joining software projects.

Engaging Undergraduate Students in Research. I believe that bringing research into the classroom is a beneficial way to challenge students and introduce them to the academic career. Encouraging students to read and criticize scientific papers published in top-venues is an excellent exercise to train critical thinking. I usually start collaborating with undergraduate students by bringing research into the classroom. This kind of approach has already brought fruitful results, which include 12 papers published and presented by my undergraduate students in Brazilian workshops and conferences [16-27], one paper published at the IEEE International Conference Software Maintenance and Evolution (ERA track) [2], and one paper published at the International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE) [14]. Five of the students I collaborated with (or advised), chose an academic career and have already received their M.Sc. degree.

Openness and trust. I believe in close and personal interaction with students I teach and advise. I always try to be honest and create a trusting relationship with my students. Whenever possible, I admit I do not know the answer for some questions and give students the opportunity to discuss and show that I can be wrong. This kind of approach allows the student to feel more comfortable and confident in participating, asking questions, and engaging with the course.

I have 12 years of teaching university courses in different institutions, and am quite flexible regarding subjects. I can teach courses such as Software Engineering, Agile Methods, Project Management, Research Method, Programming, Data Structures, and Human-Computer Interaction.

REFERENCES

(underlined authors identify undergraduate students I advised)

[1] **STEINMACHER, Igor**; WIESE, Igor; CHAVES, Ana Paula; GEROSA, Marco. Why do newcomers abandon open source software projects?. In: *6th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE 2013)*, 2013. pp. 25-32.

[2] DIAS, Luiz F.; **STEINMACHER, Igor;** PINTO, Gustavo H.; COSTA, Daniel A.; GEROSA, Marco Aurelio. How Does the Shift to GitHub Impact Project Collaboration? In*: 32nd IEEE International Conference on Software Maintenance and Evolution (ICSME)*, 2016. pp. 473-477.

[3] OLIVA, Gustavo A.; **STEINMACHER, Igor;** WIESE, Igor; GEROSA, Marco. What can commit metadata tell us about design degradation?. In: *Proceedings of the 2013 International Workshop on Principles of Software Evolution (IWPSE 2013*). ACM. pp. 18-27.

[4] WIESE, Igor Scaliante; RÉ, Reginaldo; **STEINMACHER, Igor;** KURODA, Rodrigo T.; OLIVA, Gustavo Ansaldi; TREUDE, Christoph; GEROSA, Marco. Using contextual information to predict co-change. The Journal of Systems and Software, v. Online, p. 1, 2016.

[5] SILVA, Jefferson O.; WIESE, Igor S.; GERMAN, Daniel; **STEINMACHER, Igor;** GEROSA, Marco. 2017. How Long and How Much: What to Expect from Summer of Code Participants? 33rd IEEE International Conference on Software Maintenance and Evolution (ICSME 2017) (2017).

[6] **STEINMACHER, Igor;** CHAVES, Ana Paula; GEROSA, Marco. Awareness Support in Distributed Software Development: A Systematic Review and Mapping of the Literature. Computer Supported Cooperative Work, v. 22, p. 113-158, 2013.

[7] **STEINMACHER, Igor;** GRACIOTTO SILVA, Marco; GEROSA, Marco; REDMILES, David F. A systematic literature review on the barriers faced by newcomers to open source software projects. *Information and Software Technology*, v. 59, p. 67-85, 2015.

[8] **STEINMACHER, Igor**; CONTE, Tayana U.; REDMILES, David F.; GEROSA, Marco. Social Barriers Faced by Newcomers Placing Their First Contribution in Open Source Software Projects. In: *The 18th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2015)*. ACM, 2015. pp. 1-13.

[9] **STEINMACHER, Igor**; CHAVES, Ana Paula; CONTE, Tayana U.; GEROSA, Marco. Preliminary Empirical Identification of Barriers Faced by Newcomers to Open Source Software Projects. In: *2014 Brazilian Symposium on Software Engineering (SBES)*, 2014. pp. 51-60.

[10] PINTO, Gustavo H. L.; **STEINMACHER, Igor**; GEROSA, Marco. More Common Than You Think: An In-Depth Study of Casual Contributors. In: *23rd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER'2016)*, 2016. pp. 112-123.

[11] **STEINMACHER, Igor**; WIESE, Igor; CONTE, Tayana U.; GEROSA, Marco. Increasing the self-efficacy of newcomers to Open Source Software projects. In: *Proceedings of the 29th Brazilian Symposium on Software Engineering,* 2015, pp. 160–169.

[12] **STEINMACHER, Igor**; CONTE, Tayana U.; TREUDE, Christoph; GEROSA, Marco. Overcoming Open Source Project Entry Barriers with a Portal for Newcomers*. In: International Conference on Software Engineering, 2016 (ICSE 2016)*, 2016, pp. 273–284.

[13] TOSCANI, Carolina; **STEINMACHER, Igor**; MARCZAK, Sabrina. Gamificando o Portal para Novatos em Software Livre FLOSSCoach: Insigths de uma Avaliação Preliminar. In: *XIV Simpósio Brasileiro de Sistemas Colaborativos (SBSC 2017)*.

[14] DINIZ, Guilherme; GRACIOTTO SILVA, Marco A.; GEROSA, Marco; **STEINMACHER, Igor**. Using Gamification to Orient and Motivate Students to Contribute to OSS projects. In*: 10th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE 2017)*. 2017. pp. 1-7.

[15] SARMA, Anita; GEROSA, Marco; **STEINMACHER, Igor**; Leano, Rafael . Training the Future Workforce through Task Curation in an OSS Ecosystem. In: Foundations on Software Engineering (FSE'16) - Visions and Reflections track, 2016, Seattle, WA. 2016. p. 934-937.

[16] ALVES, Luiz Phillipe Serrano; **STEINMACHER, Igor**; CHAVES, Ana Paula. Um aplicativo baseado em inteligência coletiva para compartilhamento de rotas em redes sociais. In*: Anais do VIII Simposio Brasileiro de Sistemas Colaborativos (SBSC 2011)*, 2011. pp. 41-44.

[17] ANTONIO, Marlon F.; COGO, Filipe R.; **STEINMACHER, Igor**; CHAVES, Ana Paula. Filtragem Colaborativa de Rotas de Ônibus Usando Dados sobre a Acessibilidade das Vias Urbanas. In: *Anais do XI Simpósio Brasileiro de Sistemas Colaborativos (SBSC2014)*, 2014.

[18] DA LUZ, João Alessandro; CAIRES, Vinicius; **STEINMACHER, Igor**; WIESE, Igor Scaliante. Uso do GeoGebra no Ensino de Matemática: Avaliação de Usabilidade e de Aprendizado. In: *Anais do II Encontro Nacional de Informática e Educação*, 2011.

[19] DIAS, Luiz F.; **STEINMACHER, Igor**; WIESE, IGOR; PINTO, Gustavo H. L.; COSTA, Daniel A.; GEROSA, Marco. Migrar para o Ambientes de Codificação Social é Sinônimo de Mais Contribuições? Uma Análise Preliminar do Github. In: Anais do Workshop on Software Visualization, Evolution and Maintenance (VEM 2016), 2016.

[20] FONTOURA, Vidal Daniel; WIESE, Igor Scaliante; **STEINMACHER, Igor**; RÉ, Reginaldo; SILVA, José Teodoro; Gerosa, Marco. Uma abordagem para classificação de construção de software utilizando redes de comunicação e cooperação. In: *WMSWM 2013 - X Workshop de Manutenção de Software Moderna*, 2013, pp. 1-8.

[21] FONTOURA, Vidal Daniel; **STEINMACHER, Igor**; WIESE, Igor Scaliante; GEROSA, Marco. Predição de falhas em projetos de software baseado em métricas de redes sociais de comunicação e cooperação. In: *Anais do VIII Simposio Brasileiro de Sistemas Colaborativos (SBSC 2011)*, 2011. pp. 37-40.

[22] LUCIO, Diego Rafael; CHAVES, Ana Paula; **STEINMACHER, Igor**. Um aplicativo para dispositivos móveis voltado para usuários de transporte público. In: *Anais do IV Encontro Paranaense de Computação*, 2011.

[23] MACIEL, Ana. C.; STEINMACHER, Igor.; GRACIOTTO SILVA, Marco A. Padrões de Socialização de Novatos em Projetos de Software Livre. In: XV Workshop Internacional de Software Livre, 2014, Porto Alegre, RS. p. 1-10.

[24] NISHI, Marcelo K.; CHAVES, Ana Paula; **STEINMACHER, Igor**. A Collective Intelligence Based System for Visualizing Problems in Public Roads. In: *2012 Brazilian Symposium on Collaborative Systems (SBSC)*, 2012, pp. 94-98.

[25] PRUDENCIO, Mariana M. V.; COGO, Filipe R.; CHAVES, Ana Paula; **STEINMACHER, Igor**. Participação de Mulheres em Projetos de Software Livre: Caso da Comunidade Brasileira da Mozilla Foundation. In: *XVI Workshop de Software Livre*, 2015. pp. 1-8.

[26] RAMOS, Felipe V.; CHAVES, Ana Paula; GEROSA, Marco; **STEINMACHER, Igor**. Um estudo exploratório sobre contribuições casuais em projetos de software livre: caso do projeto LibreOffice. In: Anais do IX Workshop em Desenvolvimento Distribuído de Software, Ecossistemas de Software e Sistemas-de-Sistemas, 2015.

[27] SILVA, José Teodoro; WIESE, Igor Scaliante; STEINMACHER, Igor; GEROSA, Marco. MinerAll: Uma ferramenta para extração e mineração de dados de repositórios de software livre. In: *XII Workshop de Software Livre.* Pelotas-RS: Editora Universitária UFPel, 2011.