

Leonardo da Silva Sousa

Curriculum Vitae

Professional Address Carnegie Mellon Silicon Valley
NASA Research Park, Building 23, Room 105 B
Moffett Field, CA 94035

Electronic Address leo.sousa@sv.cmu.edu
ssousaleo@gmail.com

Web Page <https://ssousaleo.github.io/>

Biography

I am an Assistant Teaching Professor in the Department of Electrical and Computer Engineering (ECE) at Carnegie Mellon University's (CMU) Silicon Valley location. I received the Ph.D. in Computer Science from the Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). I also earned an MSc in Computer Science from Universidade Federal de Goiás and a BS in Computer Science from the Universidade Federal de Mato Grosso. My main research interests include: Code Smells, Design Problems, Refactoring, Software Architecture, Software Testing, Empirical Software Engineering and Machine Learning. I work daily in empirical software engineering, focusing on applying both quantitative and, especially qualitative analysis based on grounded theory procedures. During my Ph.D., my work was published in multiple prestigious papers, journals and international conferences with strong reputations and visibility in the Software Engineering field including the top two international academic conferences - FSE (ACM SIGSOFT Symposium on the Foundations of Software Engineering) and ICSE (International Conference on Software Engineering). In fact, my work has been published in the last three ICSEs (2016, 2017, and 2018), and I am proud to report that last year, I was the main author of a paper which received the "**ACM SIGSOFT Distinguished Paper Award**." I have also been co-advising master and doctoral students in our research group for the past 3 years. Additionally, I have acquired real-world software development experience on innovative projects intended to help society through my work on the **VazaZika platform** -- a system to assist with the surveillance and detection of dengue-carrying mosquitos and outbreaks. Our team created a new platform that is currently in use in high-risk areas, allowing users to contribute to etymological surveillance by reporting mosquito breeding sites.

Academic Background

- 2014 - 2018 PhD in Computer Science (Major in Software Engineering).
Pontifical Catholic University of Rio de Janeiro, PUC-Rio, Rio De Janeiro, Brazil
Advisor: Alessandro Garcia
Scholarship: Conselho Nacional de Desenvolvimento Científico e Tecnológico (National Council for Scientific and Technological Development)
Exchange Program at University of Southern California (USC), Los Angeles – USA
PhD. Thesis: Understanding How Developers Identify Design Problems in Practice
- 2012 - 2014 MSc. in Computer Science
Federal University of Goiás, UFG, Goiania, Goiás, Brazil
Advisor: Cássio Leonardo Rodrigues
Co-advisor: Márcio Eduardo Delamaro
Scholarship: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Coordination of Superior Level Staff Improvement)
Master Thesis: A Service-Based Infrastructure for Evolution of Mutation Testing
- 2008 - 2011 BSc. in Computing Science
Federal University of Mato Grosso, UFMT, Barra do Garças, Mato Grosso, Brasil
Advisor: Maxweel Silva Carmo
Undergraduate Thesis: WIKISENSE - A System for Management of Points of Interest for the Android Platform
-

Main Programming Languages

Java, C, C++ and Python

Topics of Interest

Software Design and Architecture, Code Smells, Refactoring, Software Testing, Machine Learning (Deep Learning)

Distinguished Paper Award

1. **ACM SIGSOFT Distinguished Paper Award:** Leonardo Sousa, Anderson Oliveira, Willian Oizumi, Simone Barbosa, Alessandro Garcia, Jaejoon Lee, Marcos Kalinowski, Rafael de Mello, Balduino Neto, Roberto Oliveira, Carlos Lucena, Rodrigo Paes. 2018. *Identifying Design Problems in the Source Code: A Grounded Theory*. In Proceeding of the 40th International Conference on Software Engineering, May 27-June 3, 2018, Gothenburg, Sweden. ACM, New York, NY, USA, 921-931. DOI: <https://doi.org/10.1145/3180155.3180239>
2. **SBCARS Best Paper Award:** Willian Oizumi, Leonardo Sousa, Alessandro Garcia, Roberto Oliveira, Anderson Oliveira, O. I. Anne Benedicte Agbachi, and Carlos Lucena. 2017. Revealing design problems in stinky code: a mixed-method study. In Proceedings of the 11th Brazilian Symposium on Software Components, Architectures, and Reuse (SBCARS '17). ACM, New York, NY, USA, Article 5, 10 pages. DOI: <https://doi.org/10.1145/3132498.3132514>

Bibliographical Production

Articles published in Journal

1. **Leonardo Sousa**, Rafael de Mello, Diego Cedrim, Alessandro Garcia, Paolo Missier, Anderson Uchôa, Anderson Oliveira, Alexander Romanovsky. *VazaDengue: An Information System for Preventing and Combating Mosquito-Borne Diseases with Social Networks*. In Information Systems, 2018, ISSN 0306-4379, <https://doi.org/10.1016/j.is.2018.02.003>
2. Willian Oizumi, **Leonardo Sousa**, Anderson Oliveira, Alessandro Garcia, Anne Benedicte Agbachi, Roberto Oliveira, Carlos Lucena. *On the Identification of Design Problems in Stinky Code: Experiences & Tool Support*. In: Journal of the Brazilian Computer Society, 2018, url="<https://doi.org/10.1186/s13173-018-0078-y>

Book Chapters Published

1. Paolo Missier, Callum McClean, Jonathan Carlton, Diego Cedrim, **Leonardo Silva**, Alessandro Garcia, Alexandre Plastino, Alexander Romanovsky. *Recruiting from the Network: Discovering Twitter Users Who Can Help Combat Zika Epidemics*. In: Cabot J., De Virgilio R., Torlone R. (eds) Web Engineering. ICWE 2017. Lecture Notes in Computer Science, vol 10360. Springer, 2017
2. Paolo Missier, Alexander Romanovsky, Tudor Miu, Atinder Pal, Michael Daniilakis, Alessandro Garcia, Diego Cedrim, **Leonardo Sousa**. *Tracking Dengue Epidemics Using Twitter Content Classification & Topic Modelling*. In: Proceedings of the International Conference on Web Engineering (ICWE 2016). Springer, 2017
3. Eduardo Fernandes, Gustavo Vale, **Leonardo Sousa**, Eduardo Figueiredo, Alessandro Garcia, Jaejoon Lee. *No Code Anomaly is an Island*. In: Botterweck G., Werner C. (eds) Mastering Scale and Complexity in Software Reuse. ICSR 2017. Lecture Notes in Computer Science, vol 10221. Springer, 2017

Articles Published in Annals of Events

1. Willian Nalepa Oizumi, **Leonardo da Silva Sousa**, Anderson Oliveira, Luiz Carvalho, Alessandro Garcia, Thelma Elita Colanzi, Roberto Felicio Oliveira. *On the Density and Diversity of Degradation Symptoms in Refactored Classes: A Multi-case Study*. In: Proceedings of 30th IEEE International Symposium on Software Reliability Engineering (ISSRE 2019). Pgs. 346-357.
2. Andre Eposhi, Willian Nalepa Oizumi, Alessandro Garcia, **Leonardo da Silva Sousa**, Roberto Felicio Oliveira, Anderson Oliveira. *Removal of design problems through refactorings: are we looking at the right symptoms?* In: Proceedings of the International Conference on Program Comprehension (ICPC 2019). Pgs. 148-153.

3. Anderson Oliveira, **Leonardo da Silva Sousa**, Willian Nalepa Oizumi, Alessandro Garcia. *On the Prioritization of Design-Relevant Smelly Elements: A Mixed-Method, Multi-Project Study*. In: Proceedings of the XIII Brazilian Symposium on Software Components, Architectures, and Reuse (SBCARS 2019). Pgs. 83-92
- 4.
5. **Leonardo Sousa**, Anderson Oliveira, Willian Oizumi, Simone Barbosa, Alessandro Garcia, Jaejoon Lee, Marcos Kalinowski, Rafael de Mello, Baldoino Neto, Roberto Oliveira, Carlos Lucena, Rodrigo Paes. 2018. *Identifying Design Problems in the Source Code: A Grounded Theory*. In Proceeding of the 40th International Conference on Software Engineering, May 27-June 3, 2018, Gothenburg, Sweden. ACM, New York, NY, USA, 921-931. DOI: <https://doi.org/10.1145/3180155.3180239>
6. **Leonardo Sousa**, Auri Vincenzi, Marcio Eduardo Delamaro, Igor R. Vieira, Vinicius R. Mendenca, Cassio L. Rodrigues. *Reducing the Cost of Mutation Testing Using the Semantic Size of Mutant*. In: 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 2018, pp. 2675-2680. DOI: <https://doi.org/10.1109/SMC.2018.00457>
7. **Leonardo Sousa**, Roberto Oliveira, Alessandro Garcia, Jaejoon Lee, Tayana Conte, Willian Oizumi, Rafael de Mello, Adriana Lopes, Natasha Valentim, Edson Oliveira, and Carlos Lucena. 2017. *How Do Software Developers Identify Design Problems? A Qualitative Analysis*. In: Proceedings of the 31st Brazilian Symposium on Software Engineering (SBES'17). ACM, New York, NY, USA, 54-63. DOI: <https://doi.org/10.1145/3131151.3131168>
8. Willian Oizumi, **Leonardo Sousa**, Alessandro Garcia, Roberto Oliveira, Anderson Oliveira, O. I. Anne Benedicte Agbachi, and Carlos Lucena. 2017. *Revealing Design Problems in Stinky Code: a Mixed-Method Study*. In: Proceedings of the 11th Brazilian Symposium on Software Components, Architectures, and Reuse (SBCARS '17). ACM, New York, NY, USA, Article 5, 10 pages. DOI: <https://doi.org/10.1145/3132498.3132514>
9. Diego Cedrim, Alessandro Garcia, Melina Mongiovi, Rohit Gheyi, **Leonardo Sousa**, Rafael de Mello, Baldoino Fonseca, Márcio Ribeiro, and Alexander Chávez. 2017. *Understanding the Impact of Refactoring on Smells: A Longitudinal Study of 23 Software Projects*. In: Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering (ESEC/FSE 2017). ACM, New York, NY, USA, 465-475. DOI: <https://doi.org/10.1145/3106237.3106259>
10. Roberto Oliveira, **Leonardo Sousa**, Rafael de Mello, Natasha Valentim, Adriana Lopes, Tayana Conte, Alessandro Garcia, Edson Oliveira, and Carlos Lucena. 2017. *Collaborative Identification of Code Smells: A Multi-Case Study*. In: Proceedings of the 39th International Conference on Software Engineering: Software Engineering in Practice Track (ICSE-SEIP '17). IEEE Press, Piscataway, NJ, USA, 33-42. DOI: <https://doi.org/10.1109/ICSE-SEIP.2017.7>
11. Rafael de Mello, Roberto Oliveira, **Leonardo Sousa**, and Alessandro Garcia. 2017. *Towards Effective Teams for the Identification of Code Smells*. In: Proceedings of the 10th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE '17). IEEE Press, Piscataway, NJ, USA, 62-65. DOI: <https://doi.org/10.1109/CHASE.2017.11>
12. Willian Oizumi, Alessandro Garcia, **Leonardo da Silva Sousa**, Bruno Cafeo, and Yixue Zhao. 2016. *Code Anomalies Flock Together: Exploring Code Anomaly Agglomerations for Locating Design Problems*. In: Proceedings of the 38th International Conference on Software Engineering (ICSE '16). ACM, New York, NY, USA, 440-451. DOI: <https://doi.org/10.1145/2884781.2884868>
13. Diego Cedrim, **Leonardo Sousa**, Alessandro Garcia, and Rohit Gheyi. 2016. *Does Refactoring Improve Software Structural Quality? A Longitudinal Study of 25 Projects*. In: Proceedings of the 30th Brazilian Symposium on Software Engineering (SBES '16), Eduardo Santanda de Almeida (Ed.). ACM, New York, NY, USA, 73-82. DOI: <http://dx.doi.org/10.1145/2973839.2973848>
14. **Leonardo Sousa**. *Analyzing and Supporting Behavior Preservation in Software Refactorings*. In: *Congresso Brasileiro de Software: Teoria e Prática (CBSOFT), 2014, Maceió*. ICSE 2017 PhD and Young Researchers Warm Up Symposium. , 2014. v.1.

Expanded Summary Published in Proceedings of Conferences

1. **Leonardo da Silva Sousa**. 2016. *Spotting Design Problems with Smell Agglomerations*. In: Proceedings of the 38th International Conference on Software Engineering Companion (ICSE '16). ACM, New York, NY, USA, 863-866. DOI: <https://doi.org/10.1145/2889160.2889273>
2. **Leonardo da Silva Sousa**. 2014. *Analyzing and Supporting Behavior Preservation in Software Refactorings, 2014, Maceió*. In: ICSE 2017 PhD and Young Researchers Warm Up Symposium. v. 1

Workshop Paper

1. Willian Oizumi, Alessandro Garcia, **Leonardo Sousa**, Dannylo Albuquerque, Diego Cedrim Gomes. *Towards the Synthesis of Architecturally-Relevant Code Anomalies* In: 11th Workshop on Software Modularity, 2014, Maceió, AL. 11th Workshop on Software Modularity. 2014.
2. Igor Vieira, **Leonardo da Silva Sousa**, Vinícius Mendonca, Cássio Rodrigues, Auri Vincenzi. *Dívida Técnica: Um Estudo de Caso com Produtos de Código Aberto*. In: I Simpósio Brasileiro de Qualidade de Software, Salvador - Ba. SIMPÓSIO BRASILEIRO DE QUALIDADE DE SOFTWARE - SBQS'2013, 2013, 2013. v.XII
3. Auri Vincenzi, Cassio Rodrigues, Igor Vieira, **Leonardo Sousa**, Vinícius Mendonca, Jacson Barbosa, Mário Diaz. *Evolução de Software Livre Baseada em Métricas de Qualidade: um Estudo de Caso* In: IX Simpósio Brasileiro de Sistemas de Informação (SBSI), 2013, João Pessoa - PB. IX Simpósio Brasileiro de Sistemas de Informação - SBSI'2013, 2013

Research Projects

- 2015 – Current
- Academic Cooperation in Recommendation Systems for Collaborative Software Maintenance (CARECo)**
- Description:* The CARECo project aims to investigate how recommendation systems can effectively support software maintenance tasks and to overcome limitations of the state-of-practice tools. Given the cooperative nature of this project, the team members comprise researchers from prestigious Brazilian universities (PUC-Rio, UFAL, UFCG, and UFAM): experts in each relevant area covered by the project. The project team is made up of researchers from postgraduate programs at these four institutions, located in three different Brazilian regions: Southeast, Northeast, and North. The general coordinator of the project is Prof. Carlos José Pereira de Lucena, who has extensive experience in project coordination. The expected technical contributions of the CARECo project are: (i) a recommendation system to assist collaborative maintenance of software systems, (ii) methods that support the use of recommendation systems, (iii) new collaboration mechanisms integrated with development environments (IDEs), (iv) application of machine learning for development of recommendation systems that support collaborative software maintenance, and (v) design and evaluation of recommendation systems to teach software maintenance. The expected academic contributions of this project are: (i) strengthening the cooperation between researchers of the different universities in the project; (ii) promoting the regional research exchange in different areas of Computer Science, such as Software Engineering, Human-Computer Interaction, Artificial Intelligence, Databases, and Collaborative Systems; (iii) increasing the productivity of all universities involved in the project; (iv) constructing cooperative research networks to leverage teaching, research and development in these areas of Computer Science.
- Status:* in Progress.
Nature: Research and Development.
- 2016 – Current
- Leveraging Gamification and Social Networks for Improving Prevention and Control of Zika**
- Description:* This project seeks to develop a platform for promoting online, virtual communities to prevent and combat the spread of Zika. A little background on the problem: For years now, the Brazilian population has not responded well to prevention programs that aim to combat arboviral diseases, such as Zika and Dengue. Although the Brazilian Health System requires that health agents report each Zika case, it takes several days to process and publish this information. And even when this information is published, it tends not to easily reach poorer regions, which are more likely to be affected by the diseases. Lately, however, increasing concern with such diseases, and the ever-increasing ubiquity of online and social media have led an overwhelming number of citizens to increasingly share strategic information online, including the discovery of mosquito breeding sites in public locations. The term “social sensors” refers to this online population, motivated to contribute relevant information on social media channels. Recently, smartphone and social media usage has made its way even into poorer Brazilian communities, and because strategic information in these locations is harder to gather for health agents in traditional ways, VazaZika offers a unique solution to the problem. Health agencies in the Northeast, for instance, could explore this collective knowledge generated by citizens to improve prevention and combat actions in their area. VazaZika will use geolocation and gamification technologies to entice citizens to confirm or deny Aedes breeding sites, finding social sensors, updating users in real time about danger sites and the corresponding actions being taken by health agents.

Status: in Progress.
Nature: Research and Development.

2016 – 2017

ReSoM: Restructuring Software Systems Implemented with Multiple Programming Languages

Description: Maintenance and evolution are frequent tasks during the software development process. When modifications occur to a system, architectural problems are either introduced or amplified, resulting in a decrease in software quality. An architectural problem is the result of bad architectural decisions that have a negative effect on non-functional requirements. Given this scenario, we noticed the need to investigate how to better support the developers in solving architectural problems that manifest in the source code. Identifying and solving architectural problems in system implemented with multiple programming languages is critical since developers must deal with both the characteristics of each language, and how they interact to each other. To that end, this project aims to investigate and develop a strategy to support developers in restructuring software systems developed with multiple programming languages. The strategy is based on the automatic identification of indicators of architectural problems that can help developers to identify and solve these problems. Code smells are the indicators that we have been investigating. A code smell is a micro-structure that appears in code implementation that can be a partial indication of an architectural problem. In this project, we have been investigating how to improve the current detectors of code smells that can indicate architectural problems. In addition, we proposed techniques that support the synthesis and prioritization of groups of code smells that are potentially responsible for architectural problems. Moreover, we also proposed and implemented heuristics to support the removal of code smells in software systems.

Status: Concluded.
Nature: Research.

2014 – 2016

A Software Infrastructure for Promoting Efficient Entomological Monitoring of Dengue Fever

Description: Dengue is an endemic problem in many areas where public health services assistance is inefficient, and sometimes not even present. The Brazilian public health system cannot meet the demands of these areas due to the scarcity of resources available and the number of at-risk areas that require monitoring. To make matters worse, it is very difficult to identify and control dengue outbreaks in their initial stages, due to extended processing times. In this context, to control and prevent the spread of Dengue, the health system needs to rely on reports provided by the citizens. Unfortunately, Brazilian citizens have not historically actively engaged in reporting Dengue cases and mosquito breeding sites. To assist in the surveillance and detection of dengue mosquitos and outbreaks, we proposed an integrated system to entice the population to act as an etymological surveillance agent. The goal is to allow citizens to easily report observations that can help the public health system. To accomplish this goal, we developed a software system, named VazaDengue, to collect and transmit geo-referenced data. VazaDengue is composed of (i) a mobile application that allow citizens to report mosquito-related information, (ii) a web portal for centralizing that data and (iii) an algorithm to mine existing social media platforms to extract data and add to the monitoring of dengue outbreaks. The computational solution aims to gather information on the reporting of mosquitos that transmit dengue, allowing us to organize that data and plot it on maps. The objective is to make it available to the government and the population.

Status: Concluded.
Nature: Development.

Participation in Events

1. 40th International Conference on Software Engineering (**ICSE 2018**)
2. IEEE/ACM Automated Software Engineering (**ASE 2018**)
3. VIII Brazilian Software Congress (**CBSoft 2017**)
4. 39th International Conference on Software Engineering (**ICSE 2017**)
5. 38th International Conference on Software Engineering (**ICSE 2016**)
6. V Brazilian Software Congress (**CBSoft 2015**)

Organization of Events

1. **Student Volunteer** on the 40th International Conference on Software Engineering (ICSE 2018)
2. **Student Volunteer** on the 39th International Conference on Software Engineering (ICSE 2017)
3. **Student Volunteer** on the 38th International Conference on Software Engineering (ICSE 2016)

Languages – Interagency Language Roundtable scale (ILR scale)

Portuguese	ILR Level 5 – Native or bilingual proficiency
English	ILR Level 4 – Full professional proficiency
Spanish	ILR Level 2 – Limited working proficiency

ICSE 2018 Paper Presentation

Link: <https://youtu.be/oqJvGdiuEtQ>