

A Eashaan Rao



Department: Computer Science and Engineering

Ph. D Advisor: Dr. Sridhar Chimalakonda

1. Academic Background

- **Under Graduation** (2014-2018)
Department: Computer Science and Engineering
77% (First class with Distinction)
College: International Institute of Information Technology
University: Savitribai Phule Pune University
- **Post-Graduation** (2019 - onwards)
Department: Computer Science and Engineering
MS (Research) - Software Engineering
Indian Institute of Technology Tirupati
Converted to the Ph.D. programme after 1.5 years
- **Ph.D.** (2021 - onwards)
Department: Computer Science and Engineering
Indian Institute of Technology Tirupati
Date of Conversion: 1st Jan 2021

Note: Applied for MS + Ph.D Dual Degree

The first time I got a taste of Software Engineering as a field was during my BE final year project, titled "Software Bug Prediction using Machine Learning Techniques." Following which I joined IIT Tirupati as an MS (Research) student working in the domain of Software Engineering along with my supervisor Dr. Sridhar Chimalakonda. During my Master's, while working with the Software Engineering group called [RISHA](#) (Research in Intelligent Software and Human Analytics) Lab, I was fortunate to work on a variety of problems. For instance, in one of the works, the task is to visualize the changes in the architecture of a software project across its

multiple versions (over a decade) using call graphs and collaboration graphs. In another work, we did an exploratory analysis on the usage of “Lambda Expressions,” a programming construct in open-source Python repositories. The exposure I got to a variety of problems during this time made me very interested in Software Engineering research, and I got converted into the Ph.D. program under my supervisor. Software Engineering made me curious to venture into new or less-known areas, which acts as one of the factors for pursuing my Ph.D.

2. Research Interest/Broad Area

- **Broad Research Area:** Software Engineering
- **Research Interests:** Cross-Project Learning, Rapid Releases, Mining Software Repositories, Bug Localization

Software is pervasive today with significant impact in every aspect of our lives and society. On the other hand, the intangible and social-technical nature of “software” makes it hard to understand the intricate quality of “software”. However, with the ever-increasing complexity of software and the availability of open-source projects and data today, there are tremendous opportunities to investigate and improve software quality from multiple perspectives. This is much more critical in the Indian context, which has a massive presence in the software industry. It is here, I am fascinated to do cutting-edge research and develop systems in the area of software engineering towards improving the quality of software development and maintenance activities.

We are currently working on problems, such as understanding the nature of open-source repositories, which are releasing new versions at a smaller time scale, and exploring the similarity between different software artifacts related to *Cross-Project Learning* (CPL). The core premise of CPL is to figure out novel ways to leverage, learn and reuse knowledge from existing software artifacts to develop and maintain new software in an automated or semi-automated way.

My Ph.D. primarily involves understanding and figuring out the feasibility of CPL in the context of various software engineering tasks such as *code comprehension*, *code summarization*, *bug localization* and so on. Essentially, the goal is to improve the quality of software development and maintenance with reduced effort. To this end, we are exploring ways for identifying the relationship between different software artifacts such as source codes, issues, pull requests, documentation, and so on. A few interesting and open-ended questions in the proposed project are:

- *For what tasks of software engineering, is cross-project learning feasible?*
- *What is the notion of similarity of projects in the context of CPL?*
- *What kinds of novel techniques are required to extract relevant knowledge from software projects for application to target projects?*

3. Advisor Information

Supervisor Name: Dr. Sridhar Chimalakonda

Position: Assistant Professor

Department: Computer Science & Engineering

Areas of Interest:

- **Software Engineering:** Empirical Software Engineering, Modernizing legacy code, Novel Source Code Representations, Semantic Code Search, Code Smells, Summarization, Software Documentation, Architecture, Quality, APIs, Cross-Project Learning, Bug Localization, Knowledge Graphs, AI for SE, SE for AI, Energy-Aware Software Engineering and Tools
- **Computing for Society** - Educational Technologies, Human Computer Interaction

Dr. Sridhar Chimalakonda is a passionate researcher who works primarily in the area of Software Engineering. His vision is to push the boundaries of Software Engineering research from India and venture to create the best research lab at IIT Tirupati, India. He leads the “Research in Intelligent Software & Analytics” ([RISHA Lab](#)), which primarily works towards expanding the boundaries of Software Engineering, with a focus on cutting-edge research and building tools. He also works towards leveraging the power of computing for society (Educational Technologies and Human-Computer Interaction).

He obtained his M.S and Ph.D. from the International Institute of Information Technology (IIIT), Hyderabad. Before joining IIT Tirupati, he taught in IIIT Sri City as a Visiting Faculty. Prior to IIIT-Hyderabad, he had brief stints at a research startup and a management consulting firm.

4. Thesis objectives

Cross-project learning is a relatively new area in Software Engineering where there are few studies and work in the literature beyond defect localization. However, we plan to venture out in this to see its feasibility in the context of software projects.

Following are the proposed objectives of my thesis.

Objective 1: A systematic literature review on Cross-Project Learning.

An initial literature survey is being conducted in this direction, which will be one of the first in the literature and shall be used to identify the existing techniques, strengths, and drawbacks of cross-project learning (CPL) approaches.

Objective 2: Empirical Studies

A series of empirical studies is going to be conducted to understand the various aspects of cross-project learning settings through empirical evidence.

Objective 3: Taxonomy

We would attempt to propose a taxonomy of similar projects that can act as a guide to select projects for cross-project learning purposes.

Objective 4: Cross-Project Learning Approaches

During our CPL exploration, we will propose certain CPL approaches and techniques in specific software engineering problems such as bug localization, method naming, and many more.

Objective 5: Cross-Project Learning Tools

Based on the results obtained from CPL empirical studies and techniques, we will be developing appropriate tools to demonstrate the proposed approaches.

5. Teaching Assistantships

CS2105: Programming Methodology Lab (Odd Semester) - 2019, 2020 & 2021 (continuing)

CS3294: Software Engineering Lab (Even Semester) - 2020 & 2021

6. Future Plans

Along with my thesis problem, I would like to venture into other Software Engineering research problems and other areas related to software. I want to spread awareness among the students related to Software Engineering research and opportunities available in academia and industry.

7. List of Patents/Publications

1. Rao, A. E., & Chimalakonda, S. (2020). [An Exploratory Study Towards Understanding Lambda Expressions in Python.](#) In *Proceedings of the Evaluation and Assessment in Software Engineering* (pp. 318-323).
2. Rao, A. E., Vagavolu, D., & Chimalakonda, S. (2021, August). [AC²: towards understanding architectural changes in Python projects.](#) In *Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering* (pp. 1555-1559).