

# Preetha Chatterjee

## Curriculum Vita

✉ [preethac@udel.edu](mailto:preethac@udel.edu)  
📄 [sites.udel.edu/preethac/](https://sites.udel.edu/preethac/)  
LinkedIn: [www.linkedin.com/in/preetha-chatterjee/](https://www.linkedin.com/in/preetha-chatterjee/)

### Research Interests

My research interest is in **software engineering**, with an emphasis on improving software engineers' tools through *data mining*, *natural language processing* and *machine learning*. I am currently interested in **mining software repositories** and analyzing the information in developer communications. I am also interested in **empirical software engineering**.

### Education

- Expected - Spring 2021 **PhD., Computer Science**, *Department of Computer & Information Sciences*, University of Delaware, USA.  
Thesis Title: Mining Information from Developer Chats Towards Building Software Maintenance Tools,  
Advisor: Prof. Lori Pollock
- 2016 **MS., Computer Science**, *Department of Computer & Information Sciences*, University of Delaware, USA.
- 2008 **B.Tech., Information Technology**, *Department of Information Technology*, West Bengal University of Technology, India.

### Research Experience [Google Scholar Profile]

- Sept 2014–Present **Graduate Research Assistant**, *Department of Computer and Information Sciences*, University of Delaware.
- Developed automatic techniques to understand context of conversations and extract opinions from software developer chat communications using deep learning architectures. [ICSE '2021]
  - Designed and implemented approaches towards automatically analyzing quality of information in software developer chat conversations using supervised machine learning techniques and natural language analysis. [TOSEM '2021]
  - Customized a supervised machine-learning based algorithm that automatically extracts/disentangles developer conversations from downloaded chat transcripts. Additionally, created and published a publicly available dataset of software-related Q&A chat conversations, for further reuse by the research community. [MSR '2020]
  - Conducted an empirical study to learn how software engineers direct their efforts and what kinds of information they focus on within a post selected from the results returned in response to a search query on Q&A forums, especially on Stack Overflow. [JSS '2019]
  - Conducted an exploratory study to investigate the potential usefulness and challenges of mining software developer chat conversations, especially on Slack, for supporting software maintenance and evolution tools. [MSR '2019]
  - Developed a heuristics-based approach to identify and locate code examples and the natural language text describing the code segments in research articles. [MSR '2017]
  - Conducted an exploratory study to find different types of information embedded in a variety of software-related documents including blog posts, API documentation, code reviews, public chats, etc., and gain insight into the potential value and difficulty of mining the natural language text associated with the code snippets in these documents. [SANER '2017]
  - PhD Prelims Project: Explored the generality of an approach to automatically determine the high level actions of loop constructs (in Java), based on their structure, data flow and linguistic characteristics. [Technical Report]

### Publications

#### Refereed Conferences.

- ICSE '2021 **Preetha Chatterjee, Kostadin Damevski, and Lori Pollock, "Automatic Extraction of Opinion-based Q&A from Online Developer Chats"** In Proceedings of the 43rd International Conference on Software Engineering (ICSE), Technical Track, May 2021, Acceptance Rate: 138/615 (22.4%).
- MSR '2020 **Preetha Chatterjee, Kostadin Damevski, Nicholas A. Kraft, and Lori Pollock, "Software-related Slack Chats with Disentangled Conversations,"** In Proceedings of the 17th International Conference on Mining Software Repositories (MSR), Data Showcase Track, May 2020, Acceptance Rate: 14/36 (38.8%).
- ICSE '2020 **Preetha Chatterjee, "Extracting Archival-Quality Information from Software-Related Chats,"** In Proceedings of the 42nd International Conference on Software Engineering (ICSE), Doctoral Symposium Track, May 2020, Invited-talk Acceptance rate: 9/26 (34.6%).
- MSR '2019 **Preetha Chatterjee, Kostadin Damevski, Lori Pollock, Vinay Augustine, and Nicholas A. Kraft, "Exploratory Study of Slack Q&A Chats as a Mining Source for Software Engineering Tools,"** In Proceedings of the 16th International Conference on Mining Software Repositories (MSR), Research Track, May 2019, Acceptance rate: 32/126 (25.4%).
- MSR '2017 **Preetha Chatterjee, Benjamin Gause, Hunter Hedinger, and Lori Pollock, "Extracting Code Segments and Their Descriptions from Research Articles,"** In Proceedings of the 14th International Conference on Mining Software Repositories (MSR), Research Track, May 2017, Acceptance rate: 37/121 (30.6%).
- SANER '2017 **Preetha Chatterjee, Manziba Akanda Nishi, Kostadin Damevski, Vinay Augustine, Lori Pollock, and Nicholas A. Kraft, "What Information about Code Snippets Is Available in Different Software-Related Documents? An Exploratory Study,"** In Proceedings of the 24th IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER), Early Research Achievements Track, Feb 2017, Acceptance rate: 14/43 (32.5%).

## Journal Publications.

- TOSEM '2021 **Preetha Chatterjee, Kostadin Damevski, Nicholas A. Kraft, and Lori Pollock, "Automatically Identifying the Quality of Developer Chats for Post Hoc Use"**, Transactions on Software Engineering and Methodology (TOSEM), Feb 2021, Impact Factor: 2.071.
- JSS '2020 **Preetha Chatterjee, Minji Kong, and Lori Pollock, "Finding Help with Programming Errors: An Exploratory Study of Novice Software Engineers' Focus in Stack Overflow Posts"**, Journal of Systems and Software (JSS), Research Paper, Volume 159, Jan 2020, Impact factor: 2.559.

## Other.

- 2015 **Preetha Chatterjee, "Exploring the Generality of a Java-based Loop Action Model for the Quorum Programming Language"**, PhD Prelims Manuscript, University of Delaware.

---

## Research Presentations

- 2021 **Automatic Extraction of Opinion-based Q&A from Online Developer Chats**, *43rd International Conference on Software Engineering (ICSE)*, Paper Presentation Talk, via Zoom.
- 2021 **Finding Help with Programming Errors: An Exploratory Study of Novice Software Engineers' Focus in Stack Overflow Posts**, *Journal of Systems and Software (JSS)*, Featured Paper Video.
- 2020 **Extracting Archival-Quality Information from Software-Related Chats**, *42nd International Conference on Software Engineering (ICSE)*, Paper Presentation Talk, via Zoom.
- 2020 **Software-related Slack Chats with Disentangled Conversations**, *17th International Conference on Mining Software Repositories (MSR)*, Paper Q&A Session, via Zoom.  
Link to recorded talk on Youtube: <https://www.youtube.com/watch?v=iyL3NPa9utg>
- 2019 **Exploratory Study of Slack Q&A Chats as a Mining Source for Software Engineering Tools**, *16th International Conference on Mining Software Repositories (MSR)*, Paper Presentation Talk, Montreal, Canada.
- 2019 **Extracting Archival-Quality Information from Software-Related Chats**, *University of Delaware*, PhD Proposal Presentation Talk, Newark, USA.
- 2018 **Finding My Solution in the Haystack of a Post: An Exploratory Study of Stack Overflow**, *CIS Department 55 Years Anniversary Showcase at University of Delaware*, Poster Presentation, Newark, USA.
- 2017 **Mining Code Examples with Descriptive Text from Software Artifacts**, *Special Interest Group in Software Systems at University of Delaware*, Invited Talk, Newark, USA.
- 2017 **Extracting Code Segments and Their Descriptions from Research Articles**, *14th International Conference on Mining Software Repositories (MSR)*, Paper Presentation Talk, Buenos Aires, Argentina.
- 2017 **What Information about Code Snippets Is Available in Different Software-Related Documents? An Exploratory Study**, *CRA-W Grad Cohort Workshop*, Poster Presentation, Washington D.C., USA.
- 2015 **Exploring the Generality of a Java-based Loop Action Model for the Quorum Programming Language**, *University of Delaware*, PhD Prelims Presentation Talk, Newark, USA.

---

## Teaching Experience

- 2019 **Instructor**, *Department of Computer and Information Sciences*, University of Delaware, Summer 2019.  
Taught (as the sole instructor) introductory Java programming course in computer science for **undergraduates**. Developed course syllabus, created and presented lectures, facilitated in-class discussions, built classroom community, created labs/assignments and exams, and supervised teaching assistant for the following course:  
Introduction to Computer Science II (Course Code: CISC 181)
- 2017, 2018 **Substitute Instructor**, *Department of Computer and Information Sciences*, University of Delaware.  
Facilitated classes in absence of the instructor for the following **graduate-level** courses:
- Communication Skills for CS Researchers (Course Code: CISC 667)
  - Advanced Architecture/Software Systems: Text Analysis for Software Engineering (Course Code: CISC 879)
  - Intro to Computer Science Research (Course Code: CISC 367)
- 2014 - 2016 **Graduate Teaching Assistant**, *Department of Computer and Information Sciences*, University of Delaware.  
Led weekly labs, guided and provided assistance to students during laboratory/office hours. Explained concepts, graded assignments and projects, and gave lectures as per requirement in the following courses:
- Java Programming (Course code: CISC181)
  - Python Programming (Course Code: CISC106)
  - Intro to Computer Science with Web Applications (Course Code: CISC 103)

---

## Undergraduate Student Mentoring

Under supervision and collaboration with my advisor Prof. Lori Pollock, served as graduate research mentor for seven undergraduate students in performing data analysis, and developing software research & maintenance tools. Collaborated in designing research projects, worked closely with students to guide them through the research process, designing and implementing an approach, evaluating the approach, and presenting their research in paper and oral presentation formats.

- 2019-2020 Brian Phillips, Humphrey Owusu, Kevin Mason, *Performed data analysis and case study on software developer communications.*

- 2018 Minji Kong, *Co-authored Research Paper "Finding Help with Programming Errors: An Exploratory Study of Novice Software Engineers' Focus in Stack Overflow Posts"*, Journal of Systems and Software (JSS).
- 2017 Qilin Ma, *Developed an in-house Python-based research tool for mining developer discussions on Stack Overflow*.
- 2016 Benjamin Gause and Hunter Hedinger, *Co-authored Full Research Paper "Extracting Code Segments and Their Descriptions from Research Articles"*, 14th International Conference on Mining Software Repositories (MSR).

## Service to Profession

**Conference Program Committee**, of the *Software Engineering Education and Training (SEET) track for the 44th International Conference on Software Engineering (ICSE)*, Review submitted papers and participate in the discussion of papers for inclusion in the conference proceedings, serve as lead reviewer for a subset of papers and write a meta-review, and publicize the conference and SEET track, 2022.

**Conference Program Committee**, of the *Tool Demo track for the 37th International Conference on Software Maintenance and Evolution (ICSME)*, Review submitted papers and participate in the discussion of papers for inclusion in the conference proceedings, 2021.

**Conference Program Committee**, of the *Mining Challenge track for the 18th International Conference on Mining Software Repositories (MSR)*, Review submitted papers and participate in the discussion of papers for inclusion in the conference proceedings, 2021.

**Conference Program Committee**, of the *Early Research Achievements (ERA) track for the 28th IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*, Review submitted papers and participate in the discussion of papers for inclusion in the technical program, and promote the conference and the ERA track, 2021.

**Conference Organizing Committee**, *Social Media Chair*, for the *17th International Conference on Mining Software Repositories (MSR)*, Operated and managed the conference's social media accounts and promoted the conference, 2020.

**Research Grant Contributor**, *Participated in discussion and writing of NSF research grant*, with PIs Lori Pollock and Kostadin Damevski, "Automatically Enhancing Quality of Social Communication Channels to Support Software Developers and Improve Tool Reliability", 2017.

**Sub-reviewer**, for the *Journal of Software: Practice and Experience*, (1 paper), 2018.

**Sub-reviewer**, for the *Foundations of Software Engineering (FSE, 2017)*, (2 papers), 2017.

**Founder and Chair**, *University of Delaware ACM-W Student Chapter*, Increase recruitment and retention of women in computing fields at the university level and organize student activities and projects that aim to improve the working and learning environments for women in computing, 2016 - 2017.

**Vice-Chair**, *CISters, Women in Computer Science, University of Delaware*, Provide a supportive community for all women in technology-driven fields at University of Delaware by initiating relevant programs, such as mentoring and early/pre-major outreach and informative seminars about career and research opportunities, 2015 - 2016.

**GHC Travel Graduate Mentor**, *Department of Computer and Information Sciences, University of Delaware*, Selected as a mentor to the undergraduate students to attend Grace Hopper Celebration of Women in Computing, 2015.

**Technical Administrator**, *Indian Graduate Student Association (IGSA), University of Delaware*, Providing overall technical support for social activities, maintaining website, social media accounts, and providing technical help to other committee members, 2015 - 2016.

## Industry Experience

- 2013–2014 **Quality Analyst**, *Integral India Software Development*, Bangalore, India, Domain: Foreign Exchange Trading.
- Performed Web-based Testing, Functional and Regression Testing of in-house Foreign Exchange Trading tool, and ensured adherence to quality norms throughout the testing process.
  - Performed debugging of multiple logs of application workflows for failed scenarios.
  - Contributed to building the in-house automation test framework using Groovy.
- 2012–2013 **Software Engineer**, *Theorem India Private Limited*, Bangalore, India, Domain: Computational Advertising.
- Single point of contact for the testing team of BRT (Business Response Team), which handled bug/enhancement requests from production. Analyzed software requirement specifications, prepared test plans, test scenarios (using TestLink), created and executed test cases, logged and tracked defects (in JIRA and Bugzilla).
  - Performed Functional & Regression Testing for the user interface to validate functionalities of different software components.
  - Performed Scoping and estimated the timeline of the project, considering the efforts and number of resources available for a particular production release.

- 2009-2012 **Project Engineer**, *Wipro Technologies, Ltd*, Bangalore, India,  
Domain: Online stock/equity trading for the client Charles Schwab.
- Developed the automation suite for regression testing (in Perl) and debugging of Perl/Shell scripts to run batch processes in UNIX platform.
  - Created and executed test cases, identified and logged defects/bugs using Clear Quest.
  - Performed Sanity testing for new builds & version controlling using Clear-Case.
- Domain: Multi Functional Peripheral for the client Hewlett-Packard.
- Created test plans and test sessions in QC, performed testing on builds - both in development as well as delivery phase using in-house testing tools HP DSS, HP Web Jet Admin.
  - Found critical defects through ad-hoc/ exploratory testing, reviewed and fixed bugs in the source code.
  - Performed build qualification with Smoke & Sanity testing.

---

## Honors & Awards

- **Frank A. Pehrson Graduate Student Award for Outstanding Computer Science Research** - Awarded by Department of Computer and Information Science, University of Delaware, a monetary award to one or more CIS graduate students in recognition of outstanding performance and future potential in the field, 2021.
- **Lauri Pfeffer Shinn Memorial Award** - Awarded by Department of Computer and Information Science, University of Delaware, a monetary award to one or more women (undergrad or graduate) in recognition of a contribution to the CIS department and/or in recognition of academic success, 2016.
- **NSF Student Travel Support** - Awarded by the U.S. National Science Foundation to financially support selected students in attending the 41st International Conference on Software Engineering, 2019.
- **ACM-SIGSOFT CAPS Travel Award** - Awarded by SIGSOFT, ACM to help defray costs of attending SIGSOFT sponsored conferences and workshops, 2017.
- **Professional Development Award** - Awarded by University of Delaware to help selected graduate students participate in significant professional conferences pertaining to their field of study, 2017. Renounced to use it for future event.
- **WIE Conference Travel Scholarship** - Awarded by Women in Engineering Association, University of Delaware, to help selected female graduate students participate in significant professional conferences, 2017.
- **GHC Sponsorship** - Awarded by University of Delaware to attend virtual Grace Hopper Celebration of Women in Computing, 2020.
- **CRA-W Travel Scholarship** - Awarded by Computing Research Association Women to attend Grad Cohort Workshop, 2017 and 2015.
- **GHC Travel Scholarship** - Awarded by University of Delaware to attend Grace Hopper Celebration of Women in Computing as a mentor to the undergraduate students, 2015.

---

## Professional Memberships & Affiliations

- Member, Association for Computing Machinery, Special Interest Group on Software Engineering (ACM-SIGSOFT)
- Member, Association for Computing Machinery, Women (ACM-W)
- Member, Association for Computing Machinery (ACM)

---

## References

Available on request.