



**CURRICULUM VITAE  
FOR  
ROY, Chanchal Kumar  
Department of Computer Science, University of Saskatchewan**

### ACADEMIC CREDENTIALS

- 1998 **B.Sc. Engg.**, Khulna University, Discipline of Computer Science and Engineering, Computer Science and Engineering, Bangladesh.
- 2004 **M.Sc.**, RWTH Aachen University of Technology, Dept. of Computer Science, Software Systems Engineering, Germany.
- 2009 **Ph.D.**, Queen's University, School of Computing, Software Engineering, Canada.

### EMPLOYMENT

- 2020-Present **Program Director**, NSERC CREATE on Software Analytics Research (SOAR), <https://soar.usask.ca/>
- 2019-Present **Professor**, Tenured, Computer Science, University of Saskatchewan, <https://clones.usask.ca/>
- 2009-Present **Co-Director**, Software Research Lab (SRLab), University of Saskatchewan, <https://srlab.usask.ca/>
- 2014-2019 Associate Professor, Tenured, Computer Science, University of Saskatchewan, Saskatchewan, Canada
- 2009-2014 Assistant Professor, Probationary, Computer Science, University of Saskatchewan, Saskatchewan, Canada
- 2005-2009 Teaching and Research Assistant, part-time, Queen's University, Kingston, Ontario, Canada
- 2005-2005 Teaching and Research Assistant, Dublin City University, Dublin, Ireland
- 2002-2005 Research Assistant/Hiwi, part-time, Chair of Communication Networks, RWTH Aachen, Germany
- 2002-2004 Research Assistant/Hiwi, part-time, Institute of Building Materials Research, RWTH Aachen, Germany
- 2003-2004 Computer Programmer, part-time, FiberVision GmbH, Würselen, Germany
- 1999-2002 Lecturer, Discipline of Computer Science and Engineering, Khulna University, Khulna, Bangladesh
- 1998-1999 Analyst Programmer, Southtech Limited, Dhaka, Bangladesh

### ASSOCIATE MEMBERSHIPS

- 2009-Present College of Graduate Studies and Research

### HONOURS (Medals, Fellowships, Prizes)

Major recognitions/awards	10
Best paper and best poster awards	12
Other awards, fellowships, scholarships, and major nominations	15
Total number of recognitions, fellowships and awards	36 (selected)

### MAJOR RECOGNITIONS/AWARDS

10. **Clones Lifetime Achievement Award**, *International Workshop on Software Clones (IWSC)*, October 2023 (Dr. Roy appeared to be only the second researcher to be honoured with this prestigious recognition internationally.)
9. **GSA Advising Excellence Award**, *Graduate Student Association (GSA), University of Saskatchewan, 2021*. This award is presented to a graduate student advisor who has shown excellent supervisory and mentoring skills as well as exceptional support for their advisees' ideas, scholarly work, and financial needs.
8. **10-year Most Influential Paper Award SANER 2021**, *the 28th International Conference on Software Analysis, Evolution, and Reengineering (SANER 2021)*. The *Most Influential Paper Award* recognizes papers from ten years ago that significantly impacted practice and theory of software engineering. This award was given for WCRE 2011 paper on "Sharif Uddin, Chanchal K. Roy, Kevin A. Schneider and Abram Hindle: *On the Effectiveness of Simhashing in Clone Detection on Large Scale Software System*".
7. **New Researcher Award**, *University of Saskatchewan* – given annually to at most two recipients for excellence in research across the university considering all the researchers who are within 10 years of their PhD, 2019.
6. **CS-Can/Info-Can Outstanding Young Computer Science Researcher Prizes**. This competition is among all the young computer science professors (within 10 years of PhD) in Canada.
5. **10-year Most Influential Paper Award (MIP)** at the *25th International Conference on Software Analysis, Evolution, and Reengineering (SANER'18)*, Campobasso, Italy for WCRE 2008 paper, "C. K. Roy, James R. Cordy: An empirical study of function clones in open-source software". An MIP award certificate with the following citation: "*For influencing future studies in software cloning through the development of an extensive benchmark dataset*"
4. **10-year Most Influential Paper Award (MIP)** at the *26th International Conference on Program Comprehension (ICPC 2018)* for ICPC 2008 paper, "C. K. Roy, James R. Cordy: NICAD: Accurate Detection of Near-Miss Intentional Clones Using Flexible Pretty-Printing and Code Normalization".
3. **New Scientist Research Award**, *College of Arts and Science, University of Saskatchewan, 2018*. This competition is among all the young Science Professors (within 10 years) of the College and only one was awarded in this category.
2. **Excellence in Supervision Award**, Department of Computer Science of University of Saskatchewan –given to a single professor by the students, 2016
1. **Best Reviewer Award**, *18<sup>th</sup> Working Conference on Reverse Engineering (WCRE 2011)* –given to a single program committee member for exceptional service in reviewing of the papers, 2011

## BEST PAPER AND BEST POSTER AWARDS

12. **IEEE TCSE Distinguished Paper Award**, J. Das, S. Mondal & C. K. Roy, "Why Do Developers Engage with ChatGPT in Issue-Tracker? Investigating Usage and Reliance on ChatGPT-Generated Code", at the 32nd IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2025), Montreal, Canada, March 2025.
11. **People's Choice Award**, Roy PR, Alam, AI, Al-omari, F., Roy, B., Roy, CK, & Schneider, KA at IWSC 202, "Unveiling the potential of large language models in generating semantic and cross-language clones" the 17th International Workshop on Software Clones (IWSC 2023), pp. 22-28, Bogota, Colombia
10. **People's Choice Award** with V. Bandi and C. Gutwin at IWSC 2020, 14th International Workshop on Software Clones, on "Clone Swarm: A Cloud Based Code-Clone Analysis Tool", 2020
9. **TCSE Distinguished Paper Award Nomination** with M. Rahman at ICSME 2018, the 34th International Conference on Software Maintenance and Evolution (ICSME 2018), pp. 473-484, Madrid, Spain, September 2018
8. **Best paper award** with J. Svajlenko at SEKE 2016, 28th International Conference on Software Engineering and Knowledge Engineering (SEKE 2016), 426-433, Redwood City, California, July 2016.
7. **Best Paper Award** with Kintab et al., CASCON 2014, the 2014 Conference of the Centre for Advanced Studies on Collaborative Research, Toronto, Canada, November 2014
6. **Best Paper Nomination Certificate Award**, ICSME 2014, 30th International Conference on Software Maintenance and Evolution, October 2014.
5. **Second Best Poster Award**, CSER 2014, Consortium for Software Engineering Research (CSER), Spring Meeting, April 2014, Edmonton, Canada.
4. **Best Paper Award**, with Mondal et al., ACM SAC 2012, 27th ACM Symposium on Applied Computing, 2012
3. **Best Poster Award**, with Khalid Billah, College of Arts and Science, University of Saskatchewan, 2011
2. **Best Paper Award**, 4<sup>th</sup> Mutation Workshop, 2009
1. **Best Poster paper Award**, Canadian Conference on Computer Science and Software Engineering, 2008

## OTHER AWARDS, FELLOWSHIPS, SCHOLARSHIPS AND MAJOR NOMINATIONS

15. **Arthur B. McDonald Fellowships nominee** (formerly E.W.R. Steacie Memorial Fellowships), University of Saskatchewan, 2024, Natural Sciences and Engineering Research Council of Canada (NSERC)
14. **Arthur B. McDonald Fellowships nominee** (formerly E.W.R. Steacie Memorial Fellowships), University of Saskatchewan, 2023, Natural Sciences and Engineering Research Council of Canada (NSERC)
13. **E.W.R. Steacie Memorial Fellowships nominee**, University of Saskatchewan, 2022, Natural Sciences and Engineering Research Council of Canada (NSERC)
12. **E.W.R. Steacie Memorial Fellowships nominee**, University of Saskatchewan, 2020, Natural Sciences and Engineering Research Council of Canada (NSERC)
11. **Distinguished Graduate Supervisor Award Nominee**, The College of Graduate and Postdoctoral Studies University of Saskatchewan, Canada, 2017
10. **NSERC Doctoral Prize Nominee**, Queen's University, 2009
9. **IEEE Kingston Research Excellence Honourable Mention Award**—given to a single recipient, 2009
8. **Ph.D. Research Achievement Award**, School of Computing, Queen's University—given to a single recipient, 2009
7. **NSERC Postdoctoral Fellowships** (declined for joining as faculty member at U of S), 2009
6. **NSERC Postgraduate Scholarships for PhD study**, Queen's University, 2007 to 2009
5. **Ontario Graduate Scholarship (OGS)**, Queen's University, (declined for taking up NSERC PGS-D), 2007
4. **Ontario Graduate Scholarship in Science and Technology (OGST)**, Queen's University, 2006 to 2007
3. **Queen's Discretionary Conference Award**, Queen's University, 2006 to 2009
2. **Queen's Graduate Award**, Queen's University, 2005 to 2009
1. 1993-2009: 6 graduate scholarships and awards from funding institutions not reported.

**HIGHLY QUALIFIED PERSONEL (HQP) SUPERVISION****HQP Supervision at a glance**

	Currently		In the past (Excluding the current students)		Total
	Supervising	Co-Supervising	Supervised	Co-supervised	
Undergraduate	1	0	24	3	28
Master's	6	4	10	17	37
Doctoral	4	7	8	4	23
Postdoctoral	0	2	0	6	8
Others	1	0	3	14	18
<b>Total</b>	<b>12</b>	<b>12</b>	<b>45</b>	<b>44</b>	<b>114</b>

\*As Program Director and Principal Investigator (PI) of the industry-stream NSERC CREATE program in software analytics research (SOAR), Dr. Roy also provides (or has provided) funding and administers (or has administered) their different SOAR program requirements for over 30 other graduate students of his colleagues at the department of computer science at University of Saskatchewan, University of Alberta, University of British Columbia and University of Calgary.

**UNDERGRADUATE SUPERVISION**

Name	Type of HQP Training and Status	Years Supervised or Co-supervised	Title of Project or Thesis	Present Position, publications and awards (if any)
Mittal, Kartik	Bachelor (In Progress)	Supervised 2024-2025	Software bug report analysis and visualization	
Rahman, Shahriar	Undergraduate (Completed)	Co-supervised 2018 - 2018	Software Migration	Software Developer, Diitalk Communications Inc.
Mittal, Kartik	Bachelors (Completed)	Supervised 2024 – 2024	Building recommender systems for Scientific workflows	
Kowshik, Mazumdar	Bachelors (Completed)	Supervised 2024 – 2024	Exploring large language models for software engineering	University of Toronto
Hardy, Ryder	Bachelors (Completed)	Supervised 2024 – 2024	Exploring large language models for building software benchmarks	
Sajjani, Neha	Bachelors (Completed)	Supervised 2020 – 2020	Analysis of branch policies in continuous integration (CI)	MSc Student, Oregon State University
Senick, Matthew	Bachelors (Completed)	Supervised 2019 – 2019	A framework for cloning in the cloud	Business Process Data Scientist, StandardAero
Dayo-Olaide, Olaoluwa	Bachelors (Completed)	Supervised 2023 – 2023	Validation of software bug reports and generative clones	
Cai, Jayse	Bachelors (Completed)	Supervised 2023 – 2024	Deep learning-based clone detection	Business Intelligence Analyst, Cameco
Ayshi, Dipika	Bachelors (Completed)	Supervised 2023 – 2023	Validation of GPT generated clones	
Ukil, Souvik	Bachelors (Completed)	Supervised 2023 – 2023	Benchmarking experiments of software clones	
Shvets, Mykyta	Bachelors (Completed)	Co-supervised 2022 – 2023	Visualizing large networks with magnetic force graph layouts	
Ukil, Shouvik	Bachelors (Completed)	Supervised 2022 – 2022	Studying backporting in software	
Paterson, Jesse	Bachelors (Completed)	Supervised 2022 – 2022	Managing clones in the cloud	
Khan, Asif	Bachelors (Completed)	Supervised 2020 – 2020	Scientific workflow and Data management portal	Research Assistant, University of Saskatchewan
Loyinmi, Victor	Bachelors (Completed)	Supervised 2020 – 2020	Renovation of hydrological modelling systems	
Zhang, Xuesong	Bachelors (Completed)	Supervised 2019 – 2019	Large scale clone visualization	Computer Vision Engineer at Nullmax.AI

Name	Type of HQP Training and Status	Years Supervised or Co-supervised	Title of Project or Thesis	Present Position, publications and awards (if any)
Wang, Sean	Bachelors (Completed)	Co-supervised 2018 – 2018	CloneWorld: A fine-grained clone visualization tool	Software Developer, University of Saskatchewan
Li, Yukun	Bachelors (Completed)	Co-supervised 2018 – 2018	Scientific workflow and Data management portal	
Sen Gupta, Paromita	Bachelors (Completed)	Co-supervised 2018 – 2018	Meta-data repository	Software Developer, SkillShark Athlete Evaluations
Iqra, Mashrafi	Bachelors (Completed)	Supervised 2018 – 2018	Software bugs and clones: a case study	Data Analyst, Suncor
Saha, Bishal	Bachelors (Completed)	Supervised 2018 – 2018	Renovating legacy systems	Design Engineer, Microchip Technology Inc.
Neufeld, Parker	Bachelors (Completed)	Co-supervised 2017 – 2017	Distributed processing with SciWorCS	Chief Executive Officer and Design Lead at Loot Fox Games Inc.
Saha, Bishal	Bachelors (Completed)	Co-supervised 2017 – 2017	Renovating legacy systems	SoC Verification Engineer at Microchip Technology Inc.
Svajlenko, Jeffrey	Bachelors (Completed)	Supervised 2010 – 2011	Benchmarking of clones	Applied Scientist, Amazon Web Services (AWS)
Sutcliffe, Steve	Bachelors (Completed)	Supervised 2010 – 2011	Use case development for buggy code	Assistant Professor (Teaching), University of Calgary
Hume, Teresa	Bachelors (Completed)	Supervised 2010 – 2011	Use case development for buggy code	Tech Lead at Nc Smart Call
Gryschuk, Shawn	Bachelors (Completed)	Supervised 2010 – 2011	Use case development for buggy code	Technical Lead at Vendasta

### MSC STUDENT SUPERVISION

(J:n, B:n, C:n, P:n where n=1, 2, 3... are number of journals, book chapters, conference papers and posters or significant presentations of the students.)

Name	Type of HQP Training and Status	Years Supervised or Co-supervised	Title of Project or Thesis	Present Position, publications and awards (if any)
Amin, Sadman	Master's (In Progress)	Supervised 2023 - 2025	Software migration analysis	Siemens Saskatoon (to join)
Wang, Sean	Master's (Completed)	Co-supervised 2019 - 2022	Large scale visualization of hydrological modelling	Software Engineer at a University, J:0, C:2,
Awal, Abdul	Master's (Completed)	Co-supervised 2019 - 2021	Explainable Software Analytics	PhD Student, USask, P:2
Rahman, Shamimur	Master's (Completed)	Supervised 2019 - 2021	Explainable Software Analytics	PhD Student, USask C:1, P:2
Ahmed, Zonayed	Master's (Completed)	Co-supervised 2018 - 2021	Visualizing water modelling systems	Software Developer, Vendasta Inc. C:1, P:2
Nadim, Md.	Master's (Completed)	Supervised 2018 - 2020	Software changes and clones	PhD student, USask, Dean's Scholarships, C:2, P:2
Mondal, Saikat	Master's (Completed)	Supervised 2018 - 2020	Improving developers' productivity	PhD student, USask, Research Excellence Awards (MSc), Graduate Thesis Award (USask, MSc), Geddes awards (MSc), C:2, P:3
Saifullah, CM	Master's (Completed)	Supervised 2017 - 2019	Learning APIs through Mining Code Snippet Examples	Research Scientist, University of Saskatchewan, C:2, P:2
Khodabandehi Hamid	Master's (Completed)	Co-supervised 2018 - 2021	Automated testing for legacy systems migration	Software Quality Assurance Engineer at Fortinet, C:1, P:2
Golpayegani, Reza	Master's (In Progress)	Supervised 2023-2025	Large language models for software bug analysis	C:1, P:1
Chakroborti, Debashish	Master's (Completed)	Co-supervised 2016 - 2018	Provenance management for Big data systems	Ph.D. Student, USask, J:1, C:2, P:2
Ferdous, Rayhan	Master's (Completed)	Co-supervised 2016 - 2018	Workflow Provenance: from Modeling to Reporting	Big Data Platform Developer, LotLinx Inc. B:1, P:2

Name	Type of HQP Training and Status	Years Supervised or Co-supervised	Title of Project or Thesis	Present Position, publications and awards (if any)
Mostaeen, Golam	Master's (Completed)	Supervised 2016 - 2018	Towards Collaborative Scientific Workflow Management System	Software Developer, Gameloft, Montreal, Research Excellence Awards (MSc), Dean's Scholarships, J:2, C:3, B:1, P:3
Kar, Tonny	Master's (Completed)	Supervised 2015 - 2017	Differential Privacy for Big Databases	PhD student, USAs, C:2, P:2
Mondal, Amit	Master's (Completed)	Co-supervised 2015 - 2017	Towards a Reference Architecture for Big Data Frameworks	Ph.D. Candidate, USask, C:1, P:3
Rahman, Masud	Master's (Completed)	Supervised 2012 - 2014	Exploiting Context in Dealing with Programming Errors	Ph.D. Candidate, USask, C:3, P:3
Yeasmin, Shamima	Master's (Completed)	Co-supervised 2012 - 2014	Interactive Visualization of Software Bug Reports	Ph.D. Candidate, USask, C:2, P:3
Uddin, Sharif	Master's (Completed)	Co-supervised 2010 - 2014	Dealing with clones in software: a practical approach	Applications Analyst, ISM Canada, C:2, P:2
Maha, Lamyea	Master's (In Progress)	Supervised 2024 -	Code review analysis	C:1, P:2
Bekele, Yeabsira	Master's (In Progress)	Supervised 2024 -	Software documentation smell analysis	P:1
Ahmed, Tasnova	Master's (In Progress)	Co-supervised 2024 -	Evaluating effectiveness of LLMs for software engineering tasks?	
Nowshin, Priata	Master's (In Progress)	Co-Supervised 2024 -	Context-aware validation of Stack Overflow answers using LLMs	
Muttakin, Zoarder Al	Master's (In Progress)	Supervised 2024 -	Code review automation using LLMs	C:1, P:2
Ahmed, Mumtahina	Master's (In Progress)	Supervised 2023 -	Automated test case generation	P:1
Bappon, Suborno	Master's (In Progress)	Co-Supervised 2023 -	Automated generation of software comments using LLMs	C:2, P2
Das, Joy	Master's (In Progress)	Supervised 2023 -	Automated program repair	C:2, P2
Alam, Ajmain	Master's (In Progress)	Co-Supervised 2023 - 2024	Building semantic and cross-language clone benchmarks	
Svajlenko, Jeffrey	Master's (Transferred to PhD)	Supervised 2011-2013	Supervisor, Similarity Benchmarking	PhD Student, USask, C:2, P:2
Pinku, Subroto	Master's (Completed)	Co-Supervised 2022 - 2024	Deep Clone Detection	Software Developer, Siemens Saskatoon, C:3, P:2
Jarin, Tasnim	Master's (Completed)	Co-Supervised 2022 - 2024	Code smells and bugs	Telus Digital
Alam, Mehjabin	Master's (Withdrawn)	Co-Supervised 2018 - 2019	Reverse engineering legacy systems	P:1
Yasir, Muhammad	Master's (Visiting Completed)	Co-supervised 2018 - 2018	Scientific workflow and Data management portal	Programmer Analyst, SaskREALTORS®
Kintab, Ghadeer	Master's (Completed)	Co-supervised 2010 - 2013	Recommendation Systems	IT PMO at Saudi Airlines, C:1
Billah, Khaled	Master's (Completed)	Supervised 2010 - 2013	Defects Detection and analysis	Technical Lead at Calia Advanced Technologies, Best Poster Award, P:2
Khan, Mohammad	Master's (Completed)	Co-supervised 2010 - 2013	Program Understanding	Developer at ESTI Consulting Service, C:2, P:2
Mondal, Manishankar	Master's (Completed)	Co-supervised 2011 - 2013	Clone Stability	PhD Candidate, USask, C:2, P:2
Saha, Ripon	Master's (Completed)	Co-supervised 2010 - 2012	Clone Genealogies	Research Scientist at Meta, C:8, P:5
Asaduzzaman, Muhammad	Master's (Completed)	Co-supervised 2010 - 2012	Clone Visualization	PhD Candidate, USask, C:2, P:2

**DOCTORAL STUDENT SUPERVISION**

Name	Type of HQP Training and Status	Years Supervised or Co-supervised	Title of Project or Thesis	Present Position, publications and awards (if any)
Awal, Abdul	Doctoral (In Progress)	Supervised 2022 -	Explainable clone detection and analytics	J:1, P:3
Roy, Saumendu	Doctoral (In Progress)	Co-supervised 2021 -	Explainable Software Analytics	Best Poste Award (Runner-up), C:1, P:2
Nadim, Md.	Doctoral (In Progress)	Supervised 2020 -	Quantum machine learning in software bug analysis	J:3, C:3, P:3
Saikat Mondal	Doctoral (In Progress)	Supervised 2020 -	Reproducibility of rejected edits in Stack Overflow	Geddes awards (PhD), Citizenship Award (PhD), GSA Research Excellence in STEM Award, Best Papear Award, Carl McCrosky Innovation Sch.,J:2, C:9, P:7
Chakroborti, Debasish	Doctoral (Completed)	Co-supervised 2019 - 2025	Backporting analysis in large software	Lecturer, SaskPoly, Saskatchewan Innovation and Opportunity Scholarship, Saskatchewan Innovation and Opportunity Schol, Carl McCrosky Innovation Scholarship for Graduate Studies, Best Poster award, Exceptional Student Leadership Award, J:2, C:4, P:5
Mondal, Amit	Doctoral (Completed)	Co-supervised 2018 - 2023	Software Architectural Change Detection	Assistant Professor, Khulna University, Research Excellence Award (PhD), Best Poster Award, J:2, C:4, P:4
Al-Omari, Farouq	Doctoral (Completed)	Supervised 2014 - 2021	Detection and Analysis of Semantic Software Clones	Postdoctoral Fellow, USask, C:6, P:3
Islam, Judith	Doctoral (Completed)	Supervised 2015 - 2020	Large scale software analytics for understanding cloned bugs	Assistant Professor, Iowa State University, Dean's Scholarships, J:1, C:5, P:3
Fernandes da Silva, Rodrigo	Doctoral (Visiting Completed)	Supervised 2018 - 2019	Recommending Comprehensive Solutions for Programming Tasks	Postdoctoral Fellow, Federal University of Uberlândia, J:1, C:1, P:1
Rahman, Masud	Doctoral (Completed)	Supervised 2014 - 2019	Supporting Source Code Search with Software Analytics	Assistant Professor, Dalhousie University, Governor General's Gold Medal, Doctoral Thesis Award (USask), Research Excellence Awards (PhD), Geddes awards (PhD), Saskatchewan Innovation Award, Dean's Scholarships, J:5, C:25, P:10
Asaduzzaman, Muhammad	Doctoral (Completed)	Co-supervised 2012 - 2018	Simple and Efficient Code Recommendation Systems	Assistant Professor, University of Windsor, J2: C:17, P:8
Svajlenko, Jeffrey	Doctoral (Completed)	Supervised 2014 - 2018	Large scale clone detection and benchmarking	Applied Software Engineer, Amazon Inc., NSERC PGS-D fellowships, NSERC Postdoctoral fellowship, Best paper award, J:4, C:15, P:10
Rahman, Md Saidur	Doctoral (Completed)	Supervised 2011 - 2018	Change Impact Analysis of Code Clones	Senior Data Scientist (Research and Development), Desjardins, J1:C:4, P:4
Mondal, Manishankar	Doctoral (Completed)	Co-supervised 2013 - 2017	Fine-grain Analysis of Software Evolutionary Coupling	Assistant Professor, Khulna University, Dean's Scholarship, J:7, C:25, P:15
Zibran, Minhaz	Doctoral (Completed)	Supervised 2009 - 2014	Management Aspects of Software Clone Detection and Analysis	Associate Professor, Idaho State Universit, NSERC PGS-D fellowships, J:2, C:9, P:7
Dhali,	Doctoral	Supervised	Ranking of refactoring	P:1

Name	Type of HQP Training and Status	Years Supervised or Co-supervised	Title of Project or Thesis	Present Position, publications and awards (if any)
Aditi	(In Progress)	2024 -	candidates using technical debt analysis	
Barman, Mala	Doctoral (In Progress)	Co-supervised 2024 -	Software quality assurance using large language models	
Ranjbaran, Golshid	Doctoral (In Progress)	Co-supervised 2024 -	Feature interpretability for machine learning models	J:1, C:2, P:1
Rahman, Shamimur	Doctoral (In Progress)	Co-supervised 2022 -	Code review analytics	Geddes awards (PhD), SOAR Distinguished Research Award J:1, C:3, P:3
Hossain, Mohammad	Doctoral (In Progress)	Co-Supervised 2019 -	Building scientific workflows for large scale analytics	Delayed due to medical reasons but now submitted the final thesis, J:2, C:3, P:3
Kar, Tonny	Doctoral (Withdrawn)	Supervised 2018 – 2020	Bug localization	Faculty Member, American International University-Bangladesh, Dean's Scholarships, C:2, P:2
Wang, Pengcheng	Doctoral (visiting) (Completed)	Supervised 2018 – 2018	Machine learning based approach to localize buggy clones	C:3
Wazed, Kawser	Doctoral (In Progress)	Co-supervised 2017 -	Cross-language similarity analytics	Delayed due to personal reasons, Dean's Scholarships, J:1, C:4, P:2
Corraya, Sonia	Doctoral (Withdrawn)	Supervised 2015 – 2016	Software bug detection and analytics	Assistant Professor, Jagannath University
Yeasmin, Shamima	Doctoral (In Progress)	Co-supervised 2014 -	Visualization and Analysis of software bugs	Delayed due to medical reasons and parental leaves, J:1, C:4. P:2

## POSTDOCTORAL SUPERVISION

Name	Type of HQP Training and Status	Years Supervised or Co-supervised	Title of Project or Thesis	Present Position, publications and awards (if any)
Dr. Sristy Nath	Post-Doctoral (In progress)	Co-supervised 2025 -	Software release engineering	
Dr. Almarini, Nuri	Post-Doctoral (In progress)	Co-supervised 2023 -	Social software analytics	
Dr. Mondal, Manishankar	Postdoctoral (Completed)	Co-supervised 2018 - 2020	Software maintenance and evolution research	Associate Professor, Khulna University
Dr. A-omari, Farouq	Post-Doctoral (Completed)	Co-supervised 2023 - 2024	Scientific workflow management systems	Assistant Professor, Thompson Rivers University (Starting January 2025)
Dr. Mandal, Ashis	Post-Doctoral (Completed)	Co-supervised 2023 – 2024	Quantum Software Engineering	Professor, Hajee Mohammad Danesh Science and Technology
Dr. Rahman, Tajmilur	Post-Doctoral (Completed)	Co-supervised 2023 - 2024	Empirical software engineering	Assistant Professor, Gannon University
Dr. Jahan, Munima	Post-Doctoral (Completed)	Co-supervised 2023 – 2023	Software Release Engineering with workflows	Assistant Professor, Thompson Rivers University
Dr. Fatahi, Somayeh	Post-Doctoral (Completed)	Co-supervised 2022 – 2023	Software Analysis with Generative AI	

**OTHER SUPERVISION**

Name	Type of HQP Training and Status	Years Supervised or Co-supervised	Title of Project or Thesis	Present Position, publications and awards (if any)
C M Khaled Saifullah	Data Scientist (Completed)	Supervised 2019 - 2021	P2IRC Data Management and Repository	Software Engineer, Microsoft Inc.
Mondal, Manishankar	Res. Associate (Completed)	Co-supervised 2017 - 2018	Software Clone and Coupling Analysis	Assistant Professor, Khulna University
Dr. Choudhury, Dipankar	Res. Associate (Completed)	Co-supervised 2015 - 2016	Exploring Human Aspects of Software Clones	
Hossain, Mohammad	Data Scientist (Completed)	Co-supervised 2019 - 2023	P2IRC cloud to support large scale analytics	
Mazumdar, Souvik	Research Assistant (In Progress)	Supervised 2024 -	Software analytics using large language models	
Dr. Hassan, Mahdi Mohammad	Research Associate (Completed)	Co-Supervised 2023 – 2023	Model driven development and testing	Assistant Professor, University of Prince Edward Island
Khan, Asif	Research Assistant (Completed)	Supervised 2023 – 2024	Clone detection using quantum algorithms	
Dr. Kaur, Jasleen	Research Associate (Completed)	Co-Supervised 2023 – 2023	Software linguistic analytics	Sessional Lecturer – University of Saskatchewan
Dr. Al-omari, Farouq	Research Associate (completed)	Supervised 2022 – 2023	Software clone management	Research Associate, University of Saskatchewan
Alam, Tanvir	Research Associate (Completed)	Supervised 2022 – 2022	Large scale clone data analytics	Software Developer, Morning Star LTD
Islam, E.	Contract Programmer (Completed)	Co-Supervised 2022 – 2023	Scientific workflow and Data management portal	
Islam, S.	Contract Programmer (Completed)	Co-Supervised 2022 – 2023	Scientific workflow and Data management portal	
Dr. Azeem, Syed	Research Associated (Completed)	Supervised 2022 – 2022	Human studies in software engineering	Sessional Lecturer – University of Saskatchewan
Moniruzzaman, Md.	Contract Developer (Completed)	Co-Supervised 2021 – 2022	Scientific workflow and Data management portal	
Qamar, Hassan	Programmer (Completed)	Supervised 2019 – 2020	Cloned buggy code management	Software Developer, Push Interactions Inc.
Islam, Moksedul	Contract Programmer (Completed)	Co-Supervised 2020 – 2022	Scientific workflow and Data management portal	
Mohammad, Sultan	Contract Programmer (Completed)	Co-Supervised 2019 – 2021	Scientific workflow and Data management portal	
Hasan, Mredul	Contract Programmer (Completed)	Co-Supervised 2019 – 2021	Scientific workflow and Data management portal	



**PUBLICATIONS****Impact Chart**

	<i>h</i> -index	Citations
Google Scholar	53 (i10-index 176)	13,628 (7,207 since 2020)

**Publication Chart**

Edited Books	1
Book Chapters	8
Papers in Peer-Reviewed Journals	61
Invited Papers	4
Papers in Peer-Reviewed Archival Conference/Workshop Proceedings	185
Technical Reports	5
Invited Lectures and Presentations	30

Many of my publications are in conferences, rather than journals, which is customary in Computer Science. Conference papers in my field are substantial in length and detail, are accessible through digital libraries, and are fully refereed by 3-6 reviewers. I mostly publish in top-tier journals and conferences. The acceptance rate of the conferences is also very competitive. In the publications below, my name and the names of my HQPs are marked in boldface.

**EDITED BOOKS**

1. Katsuro Inoue, **C. K. Roy** (2021): Code Clone Analysis. Springer Singapore 2021, ISBN 978-981-16-1926-7.

**PAPERS IN REFEREED JOURNALS**

61. **A. K. Mandal, M. Nadim, C. K. Roy**, B. Roy, K. A. Schneider, "Quantum Software Engineering and Potential of Quantum Computing in Software Engineering Research: A Review", *Automated Software Engineering*, 32, 27 (2025). <https://doi.org/10.1007/s10515-025-00493-w>.
60. **M. Nadim, M. Hassan, A. K. Mandal**, C. K. Roy, B. Roy and K. A. Schneider, "Comparative analysis of quantum and classical support vector classifiers for software bug prediction: an exploratory study", *Quantum Machine Intelligence*, 7, 32 (2025). <https://doi.org/10.1007/s42484-025-00236-w>.
59. **G. Ranjbaran**, D.R. Recupero, C.K. Roy, K.A. Schneider, "C-SHAP: A Hybrid Method for Fast and Efficient Interpretability", *Applied Sciences*, 23pp., 2025, 15, 672. <https://doi.org/10.3390/app15020672>
58. **A. K. Mondal, M. Hossain, C. K. Roy**, B. Roy, K. A. Schneider, "FSECAM: A contextual thematic approach for linking feature to multi-level software architectural components", *Journal of Systems and Software*, 219: 112245 (2025)
57. **S. Fatahi**, J. Vassileva, **C. K. Roy**, "Comparing emotions in ChatGPT answers and human answers to the coding questions on Stack Overflow", *Frontiers Artif. Intell.* 7:1-11, (2024)
56. **M. A. Awal, C. K. Roy**: EvaluateXAI, "A framework to evaluate the reliability and consistency of rule-based XAI techniques for software analytics tasks", *Journal of Systems and Software*, 217: 112159 (2024)
55. **M.S. Rahman**, Z. Codabux, **C. K. Roy**, "Do Words Have Power? Understanding and Fostering Civility in Code Review Discussion", *Journal Proceedings of the ACM on Software Engineering 1(issue FSE)* 73:1632-1655 (2024)
54. **D. Chakroborti**, K. A. Schneider, **C. K. Roy**, "ReBack: recommending backports in social coding environments", *Automated Software Engineering* 31(1): 18 (2024)
53. M. Zakeri-Nasrabadi, S. Parsa, M. Ramezani, **C.K. Roy**, M. Ekhtiarzadeh, "A systematic literature review on source code similarity measurement and clone detection: Techniques, applications, and challenges", *Journal of Systems and Software* 204: 111796 (2023)
52. M. M. Rahman and **C.K. Roy**, "A Systematic Review of Automated Query Reformulations in Source Code Search", *ACM Transactions on Software Engineering and Methodology (TOSEM)*, 81 pages, (Accepted on May 24, 2023) <https://doi.org/10.1145/3607179>
51. **S. Mondal**, G. Uddin, and **C. K. Roy**, "Automatic Prediction of Rejected Edits in Stack Overflow", *Empirical Software Engineering (EMSE)*, 28 9 (2023):1-46, <https://doi.org/10.1007/s10664-022-10242-2>
50. **A. K. Mondal**, K. A. Schneider, B. Roy, **C. K. Roy**, "A survey of software architectural change detection and categorization techniques", *Journal of Systems & Software* 194(111505): 1-28, 2022
49. **D. Chakroborti**, S. Nath, K. A. Schneider, and **C. K. Roy**, "Release conventions of open source software: An exploratory study", *Journal of Software: Evolution and Process* 05: 1-22, 2022.
48. **Md Nadim**, D. Mondal, **C. K. Roy**, "Leveraging Structural Properties of Source Code Graphs for Just-In-Time Bug Prediction", *Automated Software Engineering* 29, 27:1-30, 2022.
47. G. Uddin, Y. G. Guéhéneuc F. Khomh, and **C. K. Roy**, "An Empirical Study of the Effectiveness of an Ensemble of Stand-alone Sentiment Detection Tools for Software Engineering Datasets", *ACM Transactions on Software Engineering and Methodology* 31(3), Article 48: 1-38, July 2022.
46. **S. Mondal**, M. M. Rahman, **C. K. Roy**, and K. Schneider, "The Reproducibility of Programming-Related Issues in Stack Overflow Questions", *Empirical Software Engineering Journal (EMSE)* 27, 62:1-52, 2022.
45. **Md. Nadim, M. Mondal, C. K. Roy**, K. A. Schneider, "Evaluating the Performance of Clone Detection Tools in Detecting Cloned Co-change Candidates", *Journal of Systems and Software* 187:111229: 1-19, 2022.

44. **M. M. Rahman**, F. Khomh, **S. Yeasmin**, **C. K. Roy**, "The forgotten role of search queries in IR-based bug localization: an empirical study", *Empirical Software Engineering* 26(6): 116, 2021.
43. G. Uddin, F. Khomh, **C. K. Roy**, "Automatic API Usage Scenario Documentation from Technical Q&A Sites", *ACM Transactions on Software Engineering and Methodology* 30(3): 1-45, 2021.
42. **M. Mondal**, B. Roy, **C. K. Roy**, K. A. Schneider, "ID-correspondence: a measure for detecting evolutionary coupling", *Empirical Software Engineering* 26(1): 5, 2021.
41. **J. Svajlenko** and **C. K. Roy**, The Mutation and Injection Framework: Evaluating Clone Detection Tools with Mutation Analysis, *Transactions on Software Engineering*, 47(5): 1060-1087, 2021.
40. **R. F. Silva**, **M. M. Rahman**, C. E. de Carvalho Dantas, **C. K. Roy**, F. Khomh, Marcelo de Almeida Maia, "Improved retrieval of programming solutions with code examples using a multi-featured score", *Journal of Systems and Software* 181 (111063):1-14, 2021.
39. M. Wu, **P. Wang**, K. Yin, H. Cheng, Y. Xu, **C. K. Roy**, "LVMapper: A Large-Variance Clone Detector Using Sequencing Alignment Approach", *IEEE Access* 8: 27986-27997, 2020.
38. **R. Gomes da Silva**, **C. K. Roy**, M. M. Rahman, K. A. Schneider, K. Paixão, C. E. de Carvalho Dantas, M. de Almeida Maia, "CROKAGE: Effective Solution Recommendation for Programming Tasks by Leveraging Crowd Knowledge", *Empirical Software Engineering*, 25(6): 4707—4758, 2020.
37. **G. Mostaen**, B. Roy, **C. K. Roy**, K. A. Schneider, J. Svajlenko, "A Machine Learning Based Framework for Code Clone Validation", *Journal of Systems and Software*, 169 (110686):1-19, 2020.
36. **M. M. Hossain**, B. Roy, **C. K. Roy**, K. A. Schneider, "VizSciFlow: A Visually Guided Scripting Framework for Supporting Complex Scientific Data Analysis", *Journal Proceedings of the ACM on Human-Computer Interaction (EICS 2020)*, 4(74):1-37, 2020.
35. **M. Mondal**, **C. K. Roy**, K. A. Schneider, "A Survey on Clone Refactoring and Tracking", *Journal of Systems and Software*, 159(110429): 1-27, 2020.
34. **K. W. Nafi**, B. Roy, **C. K. Roy**, K. A. Schneider, "A Universal Cross Language Software Similarity Detector for Open-source Software Categorization", *Journal of Systems and Software*, 162(110491): 1-19, 2020.
33. M. Ahasanuzzaman, **M. Asaduzzaman**, **C. K. Roy**, K. A. Schneider, "CAPS: a supervised technique for classifying Stack Overflow posts concerning API issues", *Empirical Software Engineering* 25(2): 1493-1532, 2020.
32. M. Wu, **P. Wang**, K. Yin, H. Cheng, Y. Xu, **C. K. Roy**, "LVMapper: A Large-Variance Clone Detector Using Sequencing Alignment Approach", *IEEE Access* 8: 27986-27997, 2020.
31. G. Uddin, F. Khomh, **C. K. Roy**, "Mining API usage scenarios from stack overflow", *Information and Software Technology* 122(106277):1-16, 2020.
30. **M. Mondal**, B. Roy, **C. K. Roy**, and K. A. Schneider, "An Empirical Study on Bug Propagation through Code Cloning", *Journal of Systems and Software*, 158(110407): 1-18, 2019.
29. **G. Mostaen**, B. Roy, **C. K. Roy**, and K. A. Schneider, "Designing for Real-Time Groupware Systems to Support Complex Scientific Data Analysis", *Journal Proceedings of the ACM on Human-Computer Interaction – EICS Volume 3 (9)*:1-28, 2019.
28. **M. M. Rahman**, **C. K. Roy** and David Lo, "Automatic Query Reformulation for Code Search using Crowdsourced Knowledge", *Empirical Software Engineering (EMSE)*, 24(4): 1869-1924, 2019.
27. D. Mondal, **M. Mondal**, **C. K. Roy**, K. A. Schneider, **S. Wang** and **Y. Li**, "Clone-World: A visual analytic system for large scale software clones", *Journal of Visual Informatics*, 3(1): 18-26, 2019.
26. **M. Mondal**, **C. K. Roy**, and K. A. Schneider, "Bug-proneness and Late Propagation Tendency of Code Clones: A Comparative Study on Different Clone Types", *Journal of Systems and Software*, Volume 144:41-59, 2018.
25. **M. Mondal**, **M. S. Rahman**, **C. K. Roy**, K. A. Schneider, "Is Cloned Code Really Stable?", *Empirical Software Engineering (EMSE)*, Volume 23, Issue 2, pp 693–770 (April 2018).
24. **J. F. Islam**, **M. Mondal**, **C. K. Roy**, and K. A. Schneider, "Comparing Software Bugs in Clone and Non-clone Code: An Empirical Study", *International Journal of Software Engineering and Knowledge Engineering*, Vol. 27, No. 09n10: 1507-1527, 2017.
23. **D. Choudhury**, M. Vrbka, A. Bin Mamat, I. Stavness, **C. K. Roy**, R. Mootanah, I. Krupka, "The Impact of Surface and Geometry on Coefficient of Friction of Artificial Hip Joints", *Journal of the Mechanical Behavior of Biomedical Materials*, Volume 72:192-199, 2017.
22. **J. Svajlenko** and **C. K. Roy**, "A Machine Learning Based Approach for Evaluating Clone Detection Tools for a Generalized and Accurate Precision", *International Journal of Software Engineering and Knowledge Engineering* 26(9-10): 1399-1430, 2016.
21. **M. Asaduzzaman**, **C. K. Roy**, K. A. Schneider and D. Hou, "A Simple, Efficient, Context Sensitive Approach for Code Completion", *Journal of Software: Evolution and Process*, 28(7): 512-541, 2016.
20. **D. Choudhury**, J. M. Lackner, L. Major, T. Morita, Y. Sawae, A. Mamat, I. Stavness, **C. K. Roy**, and I. Krupka, "Improved Wear Resistance of Functional Diamond like Carbon Coated Ti-6Al-4V Alloys in an Edge Loading Condition" *Journal of the Mechanical Behavior of Biomedical Materials*, 04/2016; 59, 2016, DOI: 10.1016/j.jmbbm.2016.04.004.
19. **M. Mondal**, **C. K. Roy**, and K. A. Schneider, "A comparative study on the intensity and harmfulness of late propagation in near-miss code clones", *Software Quality Journal* 24(4): 883-915, 2016.
18. **J. Svajlenko**, I. Keivanloo, and **C. K. Roy**, "Big-data clone detection using classical detectors: an exploratory study", *Journal of Software: Evolution and Process (June 2015)*, Volume 7(6), pp. 430-464, 2015
17. I. Keivanloo, **C. K. Roy** and J. Rilling, "SeByte: Scalable Clone and Similarity Search for Bytecode", *Science of Computer Programming (December 2015)*, Volume 95, Part 4, pp. 426-444, 2015.

16. **Mondal, C. K. Roy** and K. A. Schneider, "An Insight into the Dispersion of Changes in Cloned and Non-cloned Code: A Genealogy Based Empirical Study", *Science of Computer Programming (December 2014)*, Volume 95, Part 4, pp. 445–468, 2014.
15. **M. F. Zibran, R. K. Saha, C. K. Roy**, and K. A. Schneider, "Genealogical Insights into the Facts and Fictions of Clone Removal", *ACM Applied Computing Review*, 13 (4): 30 - 42, 2013.
14. J.R. Cordy and **C. K. Roy**, "Tuning Research Tools for Scalability and Performance: The NICAD Experience", *Science of Computer Programming*, 79,1 (January 2014), pp. 158-171, 2012
13. **M. F. Zibran** and **C. K. Roy**, "Conflict-aware Optimal Scheduling of Code Clone Refactoring", *IET Software*, Volume 7(3), June 2013, pp. 167-186, 2013
12. **M. Mondal, C. K. Roy**, and K.A. Schneider, "An Empirical Study on Clone Stability", *ACM SIGAPP Applied Computing Review (ACR)*, Volume 12, Issue 3, pp. 20-36, 2012.
11. H. Malik, A.S. Malik and **C. K. Roy**, 2011. "A Methodology to Optimize Query in Wireless Sensor Networks Using Historical Data", *Journal of Ambient Intelligence and Humanized Computing*, 2:227–238, 2011.
10. **C. K. Roy** and J.R. Cordy, 2010. "Near-miss Function Clones in Open Source Software: An Empirical Study", *Journal of Software: Evolution and Process*, 22:3 (April 2010), pp. 165-189.
9. **C. K. Roy**, J.R. Cordy and R. Koschke, 2009. "Comparison and Evaluation of Code Clone Detection Techniques and Tools: A Qualitative Approach", *Science of Computer Programming*, 74 (2009) 470-495, 2009.
8. B. Roy, M. Einhaus and **C. K. Roy**, 2009. "A Framework for Development and Evaluation of a Dynamic Subchannel Allocation Scheme in an OFDMA System", *Journal of Supercomputing*, 47(2):198-227, 2009.
7. **C. K. Roy**, B. Roy and M.G. Uddin, 2008. "Bayesian Approaches to Modeling Genetic Regulatory Networks: A Literature Review", *Queen's Health Science Journal* 9(1): 23-25, 2008.
6. **C. K. Roy** and J.R. Cordy, 2006. "Evaluating the Evolution of Small Scale Open Source Software Systems", *Research in Computing Science* 23:123-136, 2006.
5. A. Rahman, **C. K. Roy**, K.S. Islam and A. Islam, 2003. "Structural Operational Semantics of Concurrency in Java", *Khulna University Studies* 4(2): 306-312, 2003.
4. K.S. Islam, **C. K. Roy** and A. Rahman, 2002. "Ontology Based Directory Enabled Network Design Using Java Naming and Directory Interface", *Khulna University Studies* 4(1): 603-610, 2002.
3. K.S. Islam, **C. K. Roy** and A. Rahman, 2002. "An Agent Marketplace Supporting Negotiation for Business to Business", *Khulna University Studies* 3(2): 565-571, 2002.
2. **C. K. Roy**, M.M. Assaduzzaman, A. Rahman and D. Kamal, 2002. "Computer Education in Khulna City: An Empirical Study", *Khulna University Studies* 3(2): 505-509, 2002.
1. **C. K. Roy**, M.M. Assaduzzaman, R. Haque and B. Roy, 2001. "Abstract Data Types and Objects for Developing Component Based Software", *Khulna University Studies* 3(1): 413-420, 2001.

#### INVITED PAPERS/ABSTRACTS IN PUBLISHED CONFERENCE PROCEEDINGS

4. **C. K. Roy** and James. R. Cordy, "Benchmarks for software clone detection: a ten-year retrospective", in *Proceedings of the 25th International Conference on Software Analysis, Evolution, and Reengineering (SANER'18)*, pp. 26-37, Campobasso, Italy, March 2018.
3. **C. K. Roy**, "On the detection and management of code clones: The state of the art", (Keynote talk), *the 12th International Workshop on Software Clones (IWSC 2018)*, pp. 1, Campobasso, Italy, March 2018.
2. **C. K. Roy** and James. R. Cordy, "Adventures in NICAD: a ten-year retrospective", Most Influential Paper Award talk for ICPC 2008 paper at the *26th International Conference on Program Comprehension (ICPC 2018)*, pp. 19.
1. **C. K. Roy**, **M. F. Zibran**, and R. Koschke, "The Vision of Software Clone Management: Past, Present and Future (keynote paper)", In *Proceeding of the IEEE CSMR-18/WCRE-21 Software Evolution Week (SEW'14)*, pp. 18--33, Antwerp, Belgium, February 2014.

#### PAPERS IN FULLY REFEREED CONFERENCE/WORKSHOP PROCEEDINGS

##### Accepted (fully refereed, archival):

185. **M. A. Awal**, M. Rochan and **C. K. Roy**, "Large Language Models as Robust Data Generators in Software Analytics: Are We There Yet?" in *Proceedings of the AI Models /Data track at the 29th International Conference on Evaluation and Assessment in Software Engineering (EASE 2025)*, 12pp, Istanbul, Turkey, June 2025. (Accepted)
184. **S. Roy, S. Mondal, B. Roy** and **C. K. Roy**, "From Questions to Insights: Exploring XAI Challenges Reported on Stack Overflow Questions", in *Proceedings of the 29th International Conference on Evaluation and Assessment in Software Engineering (EASE 2025)*, 11pp, Istanbul, Turkey, June 2025. (Accepted)
183. **M.M. Hossain**, B.Roy, **C. K. Roy**, K.A. Schneider , "A Collaborative Framework for Cross-Domain Scientific Experiments for Society 5.0", in *Proceedings of the Software Engineering in Society (SEIS) track of the 47th IEEE/ACM International Conference on Software Engineering (ICSE 2025)*, 12 pp., Ottawa, Ontario, Canada, April 2025 (Accepted)
182. **M. Nadim, M. Hassan, A. K. Mandal** and **C. K. Roy**, "Quantum vs. Classical Machine Learning Algorithms for Software Defect Prediction: Challenges and Opportunities", in *Proceedings of the 6th Quantum Software Engineering (Q-SE) workshop at the 47th IEEE/ACM International Conference on Software Engineering (ICSE 2025)*, 12 pp., Ottawa, Ontario, Canada, April 2025 (Accepted)
181. **S. Rahman**, Z. Codabux, **C. K. Roy**, "Investigating the Understandability of Review Comments on Code Change Requests", in *Proceedings of the 22nd International Conference on Mining Software Repositories (MSR 2025)*, 12pp., Ottawa, Canada, April 2025.

- 180 **S. Yeasmin, C. K. Roy, K. A. Schneider, M. M. Rahman, K. Mittal, R. Hardy**, "Towards Enhancing IR-based Bug Localization Leveraging Texts and Multimedia from Bug Reports", in *Proceedings of the 33rd IEEE/ACM International Conference on Program Comprehension (ICPC 2025) Early Research Achievements*, 5pp., Ottawa, Canada, April 2025. (Accepted)

**Published (fully refereed, archival):**

179. **J. K. Das, S. Mondal, C. K. Roy**, "Why Do Developers Engage with ChatGPT in Issue-Tracker? Investigating Usage and Reliance on ChatGPT-Generated Code", in *Proceedings of the 32nd International Conference on Software Analysis, Evolution and Reengineering (SANER 2025)*, 12 pp., Montréal, Québec, Canada, March 2025.
178. **G. Ranjbaran, B. Roy, C.K. Roy**, "Exploring Sentiments in Stack Overflow Score and Discussion: A Dual Approach with Machine Learning Models and Expert Evaluation", in *Proceedings of the 15th International Conference on Cloud Computing, Data Science & Engineering (Confluence-2025)*, 8pp., Noida, India, January 2025.
177. **A. K. Mandal, M. Nadim, C. K. Roy, B. Roy and K. A. Schneider**, "Evaluating the Performance of a D-Wave Quantum Annealing System for Feature Subset Selection in Software Defect Prediction," in *Proceedings of the 2024 IEEE International Conference on Quantum Computing and Engineering (QCE 2024)*, pp. 103-108, Montreal, QC, Canada, September 2024.
176. **S. N. Pinku, D. Mondal, C. K. Roy**, "On the Use of Deep Learning Models for Semantic Clone Detection", in *Proceedings of the 40<sup>th</sup> International Conference on Software Maintenance and Evolution (ICSME 2024)*, pp. 12, Flagstaff, United States, October 2024.
175. **M. Jahan, M.M. Hassan, R. Golpayegani, G. Ranjbaran, C. K. Roy, B. Roy, K. A. Schneider**, "Automated Derivation of UML Sequence Diagrams from User Stories: Unleashing the Power of Generative AI vs. a Rule-Based Approach", in *Proceedings of the ACM / IEEE 27th International Conference on Model Driven Engineering Languages and Systems (MODELS 2024)*, pp 138-148, Linz, Austria, September 2024
174. **S. Mondal, M.M. Rahman, C. K. Roy**, "Can We Identify Stack Overflow Questions Requiring Code Snippets? Investigating the Cause & Effect of Missing Code Snippets", in *Proceedings of the 31st International Conference on Software Analysis, Evolution and Reengineering (SANER 2024)*, pp. 764-775, Rovaniemi, Finland, March 2024.
173. **P. R. Roy, A. I. Alam, F. Al-omari, B. Roy, C. K. Roy and K. A. Schneider**, "Unveiling the Potential of Large Language Models in Generating Semantic and Cross-Language Clones," in *Proceedings of the 2023 17th International Workshop on Software Clones (IWSC)*, Bogotá, Colombia, 2023, pp. 22-28, doi: 10.1109/IWSC60764.2023.00011
172. **S. Mondal, S. D. Bappon and C. K. Roy**, "Enhancing User Interaction in ChatGPT: Characterizing and Consolidating Multiple Prompts for Issue Resolution", in *Proceedings of the 21st International Conference on Mining Software Repositories (MSR 2024)*, pp. 222-226, Lisbon, Portugal, April 2024.
171. **J. K. Das, S. Mondal, and C. K. Roy**, "Investigating the Utility of ChatGPT in the Issue Tracking System: An Exploratory Study", in *Proceedings of the 21st International Conference on Mining Software Repositories (MSR 2024)*, pp. 217 - 221, Lisbon, Portugal, April 2024
170. **D. Chakroborti, C. K. Roy, K. A. Schneider**, "A Study of Backporting Code in Open-Source Software for Characterizing Changesets", in *Proceedings of the 46th International Conference on Software Engineering Companion (ICSE 2024 Companion)*, pp. 296-297, Lisbon, Portugal, April 2024.
169. **S. N. Pinku, D. Mondal, and C. K. Roy**, "TransClone: A Language Agnostic Code Clone Detector", in *Proceedings of the 17th International Workshop on Software Clones (IWSC 2023)*, Bogotá, Colombia, 2023, pp. 29-32
168. **M.S. Rahman, D. Mondal, Z. Codabux, C.K. Roy**, "Integrating Visual Aids to Enhance the Code Reviewer Selection Process", in *Proceedings of the 39th IEEE International Conference on Software Maintenance and Evolution (ICSME 2023)*, 12 pages, (Accepted on June 24, 2023)
167. **A.I. Alam, P.R. Roy, F. Al-omari, C. K. Roy, B. Roy, K.A. Schneider**, "GPTCloneBench: A comprehensive benchmark of semantic clones and cross-language clones using GPT-3 model and SemanticCloneBench", in *Proceedings of the 39th IEEE International Conference on Software Maintenance and Evolution (ICSME 2023)*, 13 pages (Accepted on June 24, 2023)
166. **M.M. Hossain, B. Roy, C. K. Roy, K.A. Schneider**, "Extensibility Challenges of Scientific Workflow Management Systems", in *Proceedings of the International Conference on Human-Computer Interaction (6) 2023: 51-70* (Accepted on December 13, 2022)
165. **S. Mondal, M.M. Rahman, C.K. Roy**, "Do Subjectivity and Objectivity Always Agree? A Case Study with Stack Overflow Questions", in *Proceedings of the 20th International Conference on Mining Software Repositories (MSR 2023)*, pp. 389-401, Melbourne, Australia, May 2023.
164. **S.N. Pinku, D. Mondal, C.K. Roy**, "Pathways to Leverage Transcompiler based Data Augmentation for Cross-Language Clone Detection", in *Proceedings of the 31st IEEE/ACM International Conference on Program Comprehension (ICPC 2023)*, pp. 169-180, Melbourne, Australia, May 2023.
163. **J. Tasnim, D. Chakroborti, C.K. Roy, Kevin A. Schneider**, "How does quality deviate in stable releases by backporting?", in *Proceedings of the 2023 IEEE/ACM 45th International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER)*, pp. 140-145, Melbourne, Australia, 2023.
162. **M. M. Hossain, B. Roy, C. K. Roy, K. A. Schneider**, "A Domain-Specific Provenance Query Composition Environment for Scientific Workflows", in *Proceedings of the 17th Workshop on Workflows in Support of Large-Scale Science (WORKS22)*, 8pp., Dallas, November 2022.

161. **M. S. Rahman**, C. K. Roy, "An Insight into the Reusability of Stack Overflow Code Fragments in Mobile Applications", in *Proceedings of the 16th International Workshop on Software Clones (IWSC 2022)*, pp. 69-75, Limassol, Cyprus, October 2022.
160. J. Svajlenko, **C. K. Roy**, "BigCloneBench: A Retrospective and Roadmap", in *Proceedings of the 16th International Workshop on Software Clones (IWSC 2022)*, pp. 8-9, Limassol, Cyprus, October 2022.
159. **D. Chakroborti**, K. A. Schneider and **C. K. Roy**, "Backports: Change Types, Challenges and Strategies," in *Proceedings of the IEEE/ACM 30th International Conference on Program Comprehension (ICPC)*, 2022, pp. 636-647, doi: 10.1145/3524610.3527920.
158. **S. Wang**, D. Mondal, S. Sadri, **C. K. Roy**, J. S. Famiglietti, and K. A. Schneider, "SET-STAT-MAP: Extending Parallel Sets for Visualizing Mixed Data", in *Proceedings of the 15th IEEE PacificVis symposium*, pp. 151-160, 2022.
157. **K. W. Nafi**, M. Asaduzzaman, B. Roy, **C. K. Roy** and K. A. Schneider, "Mining Software Information Sites to Recommend Cross-Language Analogical Libraries", in *Proceedings of the 29th IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2022)*, pp. 913-924, March 2022.
156. **S. Mondal**, G. Uddin and **C. K. Roy**, "Rollback Edit Inconsistencies in Developer Forum", *IEEE 18th International Conference on Mining Software Repositories (MSR 2021)*, pp. 380-391, Madrid, Spain, May 2021.
155. **S. Mondal**, C M K. Saifullah, A. Bhattacharjee, M. M. Rahman, and **C. K. Roy**, "Early Detection and Guidelines to Improve Unanswered Questions on Stack Overflow", *ACM 14th Innovation in Software Engineering Conference (ISEC 2021)*, pp. 9:1-9:11, Bhubaneswar, India, February 2021
154. **M. Mondal**, **C. K. Roy**, B. Roy, K. A. Schneider: FLeCCS: A Technique for Suggesting Fragment-Level Similar Co-change Candidates. in *Proceedings of the 29th ACM/IEEE International Conference on Program Comprehension (ICPC 2021)*, pages, 160-171, 2021.
153. **C. M. K. Saifullah**, M. Asaduzzaman, **C. K. Roy**, "COSTER: A Tool for Finding Fully Qualified Names of API Elements in Online Code Snippets", in *Proceedings of the 43rd ACM/IEEE International Conference on Software Engineering (ICSE 2021 Companion Volume)*, pp. 73-76, May 2021.
152. **A. K. Mondal**, **C. K. Roy**, K. A. Schneider, B. Roy, and S. S. Nath. "Semantic Slicing of Architectural Change Commits: Towards Semantic Design Review." In *Proceedings of the 15th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, pp. 1-6, October 2021.
151. **H. Khodabandehloo**, B. Roy, **M. Mondal**, **C. K. Roy**, K. A. Schneider: A Testing Approach While Re-Engineering Legacy Systems: An Industrial Case Study, in *Proceedings of the 27th edition of the IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2021)*, pages 600-604.
150. **Z. Ahmed**, M. Beyene, D. Mondal, **C. K. Roy**, C. Dutchyn, K. A. Schneider, "ContourDiff: Revealing Differential Trends in Spatiotemporal Data", in *Proceedings of the 25th International Conference Information Visualisation (IV)*, 2021, pp. 35-41
149. **M. Mondal**, **C. K. Roy**, K. A. Schneider, "A Fine-Grained Analysis on the Inconsistent Changes in Code Clones", in *Proceeding of the 36th IEEE International Conference on Software Maintenance and Evolution (ICSME 2020)*, pp. 220-231, Adelaide, Australia, September 2020.
148. **S. Mondal**, G. Uddin and **C. K. Roy**, "Automatic Identification of Rollback Edit with Reasons in Stack Overflow Q&A Site". In *Proceeding of the 36th IEEE International Conference on Software Maintenance and Evolution (ICSME 2020 RR Track)*, pp. 856-856, Adelaide, Australia, September 2020.
147. **M. Mondal**, B. Roy, **C. K. Roy**, K. A. Schneider, "Investigating Near-Miss Micro-Clones in Evolving Software", in *Proceedings of the 28th ACM/IEEE International Conference on Program Comprehension (ICPC 2020)*, pp. 208-218, May 2020.
146. **A. Bhattacharjee**, S. S. Nath, S. Zhou, **D. Chakroborti**, B. Roy, **C. K. Roy**, K. A. Schneider: (2020), "An Exploratory Study to Find Motives Behind Cross-platform Forks from Software Heritage Dataset", in *Proceeding of the 17th International Conference on Mining Software Repositories (MSR 2020)*, Mining Challenge Track., pp. 11-15, May 2020.
145. Venkat Bandi, **C. K. Roy**, Carl Gutwin, "Clone Swarm: A Cloud Based Code-Clone Analysis Tool", in *Proceedings of the 14th International Workshop on Software Clones (IWSC 2020)*, February 2020, London, Ontario, Canada, pp. 52-56.
144. **C. M. K. Saifullah**, M. Asaduzzaman, **C. K. Roy**, "Exploring Type Inference Techniques of Dynamically Typed Languages", in *Proceedings of the 27th edition of the IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2020)*, London, Ontario, February 2020, pp. 70-80. (**AR (AR): 23.11%**)
143. **M. Mondal**, B. Roy, **C. K. Roy**, K. A. Schneider, "Associating Code Clones with Association Rules for Change Impact Analysis", in *Proceedings of the 27th edition of the IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2020)*, London, Ontario, February 2020, pp. 93-103. (**AR: 23.11%**)
142. **F. Al-Omari**, **C. K. Roy**, T. Chen, "SemanticCloneBench: A Semantic Code Clone Benchmark using Crowd-Source Knowledge", in *Proceedings of the 14th International Workshop on Software Clones (IWSC 2020)*, February 2020, London, Ontario, Canada, pp. 57-63.
141. **M. Mondal**, B. Roy, **C. K. Roy**, K. A. Schneider, "HistoRank: History-Based Ranking of Co-change Candidates", in *Proceedings of the 27th edition of the IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2020)*, London, Ontario, February 2020, pp. 240-250. (**AR: 23.11%**)
140. G. Li, Yijian Wu, **C. K. Roy**, J. Sun, X. Peng, N. Zhan, B. Hu, J. Ma, "SAGA: Efficient and Large-Scale Detection of Near-Miss Clones with GPU Acceleration", in *Proceedings of the 27th edition of the IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2020)*, London, Ontario, February 2020, 272-283. (**AR: 23.11%**)

139. **M. Nadim, M. Mondal, C. K. Roy**, "Evaluating Performance of Clone Detection Tools in Detecting Cloned Co-change Candidates", in *Proceedings of the 14th International Workshop on Software Clones (IWSC 2020)*, February 2020, London, Ontario, Canada, pp. 15-21
138. **C M K Saifullah, M. Asaduzzaman and C. K. Roy** "Learning from examples to Find Fully Qualified Names of API Elements in Code Snippets", in *Proceedings of the 34th IEEE/ACM International Conference on Automated Software Engineering (ASE 2019)*, San Diego, California, United States, November 2019, pp. 243-254. (AR: 22%)
137. **A. Mondal, B. Roy, C. K. Roy, K. A. Schneider**, "Automatic Components Separation of Obfuscated Android Applications: An Empirical Study of Design Based Features", in *Proceedings of the 34th IEEE/ACM International Conference on Automated Software Engineering Workshop (ASEW 2019)*, San Diego, United States, pp. 23-28.
136. **K. W. Nafi, T. Kar, B. Roy, C. K. Roy, and K. A. Schneider**, "CLCDSA: Cross Language Code Clone Detection using Syntactical Features and API Documentation", in *Proceedings of the 34th IEEE/ACM International Conference on Automated Software Engineering (ASE 2019)*, San Diego, California, United States, November 2019, pp. 1026-1037. (AR: 22%)
135. **G. Mostaeen, J. Svajlenko, B. Roy, C. K. Roy, K. A. Schneider**, "CloneCognition: machine learning based code clone validation tool", in *Proceeding of The 27th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2019)*, pp. 1105-1109
134. **M. Mondal, B. Roy, C. K. Roy, K. A. Schneider**, "Ranking Co-change Candidates of Micro-Clones", in *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering (CASCON)*, Toronto, Canada, November 2019, pp. 244-253. (AR: 30%)
133. **M. Mondal, B. Roy, C. K. Roy, K. A. Schneider**, "Investigating the Relationship between Evolutionary Coupling and Software Bug-proneness", in *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering (CASCON 2019)*, Toronto, Canada, 2019, pp. 173-182. (AR: 30%)
132. **M. Mondal, B. Roy, C. K. Roy, K. A. Schneider**, "Investigating Context Adaptation Bugs in Code Clones", in *Proceedings of the 35th International Conference on Software Maintenance and Evolution (ICSME)*, Cleveland, USA, September 2019, pp. 157-168. (AR: 22.96%)
131. **S. Mondal, M. M. Rahman and C. K. Roy**, "Can Issues Reported at Stack Overflow Questions be Reproduced? An Exploratory Study", In *Proceeding of the 16th International Conference on Mining Software Repositories (MSR 2019)*, pp. 479-489, Montreal, Canada, May 2019. (AR: 24.5%)
130. **G. Uddin, F. Khomh and C. K. Roy**, "Towards Crowd-Sourced API Documentation", In *Proceedings of the 41st ACM/IEEE International Conference on Software Engineering (ICSE 2019, poster track)*, pp. 310-311, Montreal, Canada, May 2019.
129. **R.F. Gomes Da Silva, C. K. Roy, M. M. Rahman, KA Schneider, K. Paixão and M. Maia**, "Recommending Comprehensive Solutions for Programming Tasks by Mining Crowd Knowledge", In *Proceedings of The 27th International Conference on Program Comprehension (ICPC 2019)*, pp. 358-368, Montreal, Canada, May 2019. (AR: 30.1%)
128. **J.F. Islam, M Mondal, CK. Roy and KA Schneider**, "Comparing Bug Replication in Regular and Micro Code Clones", In *Proceedings of The 27th International Conference on Program Comprehension (ICPC 2019)*, pp. 81-92, Montreal, Canada, May 2019. (AR: 30.1%)
127. **D. Mondal, M. Mondal, C. K. Roy, K. A. Schneider, S. Wang and Y. Li**, "Towards Visualizing Large Scale Evolving Clones", In *Proceedings of the 41st ACM/IEEE International Conference on Software Engineering (ICSE 2019, poster track)*, pp. 302-303, Montreal, Canada, May 2019.
126. **J. F. Islam, M. Mondal, C. K. Roy**, "A Comparative Study of Software Bugs in Micro-clones and Regular Code Clones", In *Proceedings of the 25th IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, Hangzhou, China, February 2019, pp. 73-83. (AR: 30.4%)
125. **D. Chakraborti, M. Mondal, B. Roy, C. K. Roy, K. A. Schneider**, "Optimized Storing of Workflow Outputs through Mining Association Rules", In *Proceedings of the IEEE BigData 2018*, Seattle, USA, pp. 508-515, December 2018.
124. **M. M. Rahman and C. K. Roy**, "Improving IR-Based Bug Localization with Context-Aware Query Reformulation", In *Proceeding of The 26th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2018)*, pp. 621--632, Florida, USA, November 2018. (AR 19%)
123. **M. M. Rahman and C. K. Roy**, "Effective Reformulation of Query for Code Search using Crowdsourced Knowledge and Extra-Large Data Analytics", In *Proceeding of The 34th International Conference on Software Maintenance and Evolution (ICSME 2018)*, pp. 473-484, Madrid, Spain, September, 2018 [TCSE Distinguished Paper Award Nomination] (AR 21%)
122. **M. M. Rahman and C. K. Roy**, "NLP2API: Query Reformulation for Code Search using Crowdsourced Knowledge and Extra-Large Data Analytics", In *Proceeding of The 34th International Conference on Software Maintenance and Evolution (Artifact Track)* (ICSME 2018), pp. 714, Madrid, Spain, September 2018.
121. **K. W. Nafi, B. Roy, C. K. Roy, and K. A. Schneider**, "CroLSim: Cross Language Software Similarity Detector Using API Documentation", In *Proceedings of the 18th IEEE International Working Conference on Source Code Analysis and Manipulation (SCAM 2018)*, pp. 139-148, Madrid, Spain, September 2018. (AR: 25%)
120. **G. Mostaeen, J. Svajlenko, B. Roy, C. K. Roy, and K. A. Schneider**, "On the Use of Machine Learning Techniques Towards the Design of Automatic Code Clone Validation Tools", In *Proceedings of the 18th IEEE International Working Conference on Source Code Analysis and Manipulation (SCAM 2018)*, pp. 155-164, Madrid, Spain, September 2018. (AR: 25%)
119. **M. A. Islam, M. Islam, M. Mondal, B. Roy, C. K. Roy, K. A. Schneider**, "Detecting Evolutionary Coupling Using Transitive Association Rules", In *Proceedings of the 18th IEEE International Working Conference on Source Code Analysis and Manipulation (SCAM 2018)*, pp. 113-122, Madrid, Spain, September 2018. (AR: 25%)

118. **G. Mostaeen, B. Roy, C. K. Roy**, and K. A. Schneider, "Fine-Grained Attribute Level Locking Scheme for Collaborative Scientific Workflow Development", In *Proc. of IEEE Conference on Service Computing (SCC'18)*, pp. 273-277, San Francisco, USA, July 2018.
117. **P. Wang, J. Svajlenko**, Y. Wu, Y. Xu and **C. K. Roy**, "CCAligner: a token based large-gap clone detector", In *Proceeding of The 40th International Conference on Software Engineering (ICSE 2018)*, pp. 1066-1077, Gothenburg, Sweden, May 2018 (**AR 20.9%**).
116. **M. M. Rahman** and **C. K. Roy**, "Poster: Improving Bug Localization with Report Quality Dynamics and Query Reformulation", In *Proceeding of The 40th International Conference on Software Engineering: Companion Proceedings (ICSE 2018, poster)*, pp. 348-349, Gothenburg, Sweden, May 2018.
115. **J. Svajlenko** and **C. K. Roy**, "Poster: Fast, Scalable and User-Guided Clone Detection", In *Proceeding of The 40th International Conference on Software Engineering: Companion Proceedings (ICSE 2018, poster)*, pp. 352-353, Gothenburg, Sweden, May 2018.
114. M. Ahasanuzzaman, **M. Asaduzzaman, C. K. Roy** and K. A. Schneider, "Classifying Stack Overflow Posts On API Issues", In *Proceedings of the 25th IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER 2018)*, pp. 244-254, Campobasso, Italy, March 2018 (**AR 27%**).
113. **M. Mondal, C. K. Roy** and K. A. Schneider, "Micro-Clones in Evolving Software", In *Proceedings of the 25th IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER 2018)*, pp. 50-60, Campobasso, Italy, March 2018 (**AR 27%**).
112. **M. M. Rahman** and **C. K. Roy**, "Improved Query Reformulation for Concept Location using CodeRank and Document Structures", In *Proceeding of The 32nd International Conference on Automated Software Engineering (ASE 2017)*, pp. 428-439, Urbana-Champaign, Illinois, USA, October 2017. (**AR 21%**)
111. **M. Asaduzzaman, C. K. Roy**, K. A. Schneider and D. Hou, "FEMIR: a tool for recommending framework extension examples", In *Proceeding of the Tool Demonstration Track of the 32nd International Conference on Automated Software Engineering (ASE 2017)*, pp. 967-972, Urbana-Champaign, USA, October 2017.
110. **M. Asaduzzaman, C. K. Roy**, K. A. Schneider and D. Hou, "Recommending Framework Extension Examples", In *Proceedings of the 33rd IEEE International Conference on Software Maintenance and Evolution (ICSME)*, pp. 456-466, Shanghai, China, September 2017. (**AR 27.7%**).
109. **M. Mondal, C. K. Roy** and K. A. Schneider, "Bug Propagation through Code Cloning: An Empirical Study", In *Proceedings of the 33rd IEEE International Conference on Software Maintenance and Evolution (ICSME)*, pp. 227-237, Shanghai, China, September 2017. (**AR 27.7%**).
108. **Md Saidur Rahman** and **C. K. Roy**, "On the Relationships between Stability and Bug-proneness of Code Clones: An Empirical Study", In *Proceedings of the 17th IEEE International Working Conference on Source Code Analysis and Manipulation (SCAM 2017)*, pp. 131-140, Shanghai, China, September 2017.
107. **J. F. Islam, M. Mondal, C. K. Roy**, K. A. Schneider, "A Comparative Study of Software Bugs in Clone and Non-Clone Code", In *Proceedings of the 29th International Conference on Software Engineering and Knowledge Engineering (SEKE 2017)*, pp. 436-443, Pittsburgh, USA, July 2017.
106. **J. Svajlenko** and **C. K. Roy**, "Fast and Flexible Large-Scale Clone Detection with CloneWorks", In *Proceeding of the Poster Track of the 39th International Conference on Software Engineering (ICSE 2017 Companion)*, pp. 27-30, Buenos Aires, Argentina, May, 2017.
105. **J. Svajlenko** and **C. K. Roy**, "CloneWorks: a fast and flexible large-scale near-miss clone detection tool", In *Proceeding of the Tool Demonstration Track of the 39th International Conference on Software Engineering (ICSE 2017 Companion)*, pp. 177-179, Buenos Aires, Argentina, May 2017.
104. **M. Mondal, C. K. Roy**, K. A. Schneider, "Identifying Code Clones Having High Possibilities of Containing Bugs", In *Proceedings of the 25th International Conference on Program Comprehension (ICPC 2017)*, pp. 99-109, Buenos Aires, Argentina, May 2017.
103. **M. M. Rahman** and **C. K. Roy** and R.G. Kula, "Predicting Usefulness of Code Review Comments using Textual Features and Developer Experience", In *Proceeding of The 14th International Conference on Mining Software Repositories (MSR 2017)*, pp. 215-226, Buenos Aires, Argentina, May 2017 (**AR 30.60%**)
102. **M. M. Rahman** and **C. K. Roy** and David Lo, "RACK: Code Search in the IDE using Crowdsourced Knowledge", In *Proceeding of the 39th International Conference on Software Engineering (ICSE 2017 Companion)*, pp. 51-54, Buenos Aires, Argentina, May 2017 (**AR 31.58%**)
101. **M. M. Rahman** and **C. K. Roy**, "Impact of Continuous Integration on Code Reviews", In *Proceeding of the Mining Challenge Track of the 14th International Conference on Mining Software Repositories (2017)*, pp. 499-502, Buenos Aires, Argentina, May 2017.
100. B. Roy, **A. K. Mondal, C. K. Roy**, K. A. Schneider, K. Wazed, "Towards a Reference Architecture for Cloud-Based Plant Genotyping and Phenotyping Analysis Frameworks", in *Proceedings of the 2017 International Conference on Software Architecture (ICSA 2017)*, pp. 41-50, Gothenburg, Sweden, April 2017. (**AR 22%**)
99. **M. Mondal, C. K. Roy**, K. A. Schneider, "Does cloned code increase maintenance effort?", In *Proceedings of the 11th International Workshop on Software Clones (IWSC 2017)*, pp. 38-44, Klagenfurt, Austria, February 2017.
98. **M. M. Rahman** and **C. K. Roy**, "STRICT: Information Retrieval Based Search Term Identification for Concept Location", In *Proceeding of The 24th IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER 2017)*, pp. 79--90, Klagenfurt, Austria, February 2017 (**AR 24%**).
97. **M. Mondal, C. K. Roy**, K. A. Schneider, "An exploratory study on change suggestions for methods using clone detection", in *Proceedings of the 2016 Conference of the Centre for Advanced Studies on Collaborative Research (CASCON 2016)*, pp. 85-95, Toronto, Canada, November 2016. (**AR 31.7%**)
96. **J. Svajlenko**, and **C. K. Roy**. "BigCloneEval: A Clone Detection Tool Evaluation Framework with BigCloneBench", in *Proceedings of the Tool Demonstration Track of the Tools Demos Track of the 32nd*

- International Conference on Software Maintenance and Evolution (ICSME 2016)*, pp. 596-600, Raleigh, North Carolina, USA, October 2016.
95. **M. M. Rahman** and **C. K. Roy**, "QUICKAR: Automatic Query Reformulation for Concept Location Using Crowdsourced Knowledge", In *Proceeding of the New Ideas Track of the 31st IEEE/ACM International Conference on Automated Software Engineering (ASE 2016)*, pp. 220--225, Singapore, September 2016.
94. **M. M. Rahman, C. K. Roy**, J. Redl, and J. Collins, "CORRECT: Code Reviewer Recommendation at GitHub for Vendasta Technologies", In *Proceeding of the Tool Demonstration Track of the 31st IEEE/ACM International Conference on Automated Software Engineering (ASE 2016)*, pp. 792--797, Singapore, September 2016.
93. **J. Svajlenko, C. K. Roy**, "Efficiently Measuring an Accurate and Generalized Clone Detection Precision using Clone Clustering", In *Proceedings of the 28th International Conference on Software Engineering and Knowledge Engineering (SEKE 2016)*, 426-433, Redwood City, California, July 2016. (Best Paper Award)
92. **Amit K. Mondal, M. M. Rahman** and **C. K. Roy**, "Embedded Emotion-based Classification of Stack Overflow Questions Towards the Question Quality Prediction", In *Proc. of The 28th International Conference on Software Engineering & Knowledge Engineering (SEKE 2016)*, pp. 521-526, California, July 2016.
91. H. Sajnani, V. Saini, **J. Svajlenko, C. K. Roy**, and C. V. Lopes, "SourcererCC: Scaling Code Clone Detection to Big-Code", In *Proceedings of the 38th International Conference on Software Engineering (ICSE 2016)*, pp. 1157-1168, Austin, TX, May 2016. (AR 19%).
90. **M. M. Rahman, C. K. Roy**, and Jason Collins, "CORRECT: Code Reviewer Recommendation in GitHub Based on Cross-Project and Technology Experience", In *Proceeding of The 38th International Conference on Software Engineering: Companion volume (ICSE 2016, SEIP)*, pp. 222—231, Austin Texas, USA, May 2016. (AR 26%).
89. M. Ahasanuzzaman, **M. Asaduzzaman, C. K. Roy**, K. A. Schneider: Mining duplicate questions in stack overflow", In *Proceedings of the 13th International Conference on Mining Software Repositories (MSR 2016)*, pp. 402-412, Austin, TX, USA, May 2016. (AR 27%).
88. **M. Asaduzzaman**, M. Ahasanuzzaman, **C. K. Roy** and K. A. Schneider, "How Developers Use Exception Handling in Java?", In *Proceedings of 13th International Conference on Mining Software Repositories (MSR 2016)*, pp. 516-519, Austin, TX, USA, May 2016. (Mining Challenge paper).
87. **Farouq Al-omari** and **C. K. Roy**, "Is Code Cloning in Games Really Different?", In *Proceedings of the Software Engineering Track of the 31st ACM Symposium on Applied Computing (ACM SAC 2016)*, Pisa, Italy, April 2016, pp. 1512-1519.
86. **M. M. Rahman, C. K. Roy** and David Lo, "RACK: Automatic API Recommendation using Crowdsourced Knowledge", In *Proceeding of The 23rd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER 2016)*, pp. 349-359, Osaka, Japan, March 2016.
85. **J. F. Islam, M. Mondal, C. K. Roy**, "Bug Replication in Code Clones: An Empirical Study", in *Proceedings of the 23rd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, pp. 68-78, Osaka, Japan, March 2016. IEEE.
84. **M. Mondal, C. K. Roy**, K. A. Schneider, "An Empirical Study on Ranking Change Recommendations Retrieved using Code Similarity", In *Proceedings of the 10th International Workshop on Software Clones (IWSC)*, pp. 44-50, Osaka, Japan, March 2016.
83. **S. Yeasmin, C. K. Roy** and K. A. Schneider, "How Should We Read and Analyze Bug Reports: An Interactive Visualization using Extractive Summaries and Topic Evolution", In *Proceedings of the 2015 Conference of the Centre for Advanced Studies on Collaborative Research (CASCON 2015)*, pp. 171-180, Toronto, Canada, November 2015. (AR 29.5%)
82. **M. M. Rahman** and **C. K. Roy**, "Recommending Relevant Sections from a Webpage about Programming Errors and Exceptions", In *Proceedings of the 2015 Conference of the Centre for Advanced Studies on Collaborative Research (CASCON 2015)*, pp. 181-190, Toronto, Canada, November 2015. (AR 29.5%)
81. **M. Mondal, C. K. Roy** and K. A. Schneider, "An Empirical Study on Change Recommendation", In *Proceedings of the 2015 Conference of the Centre for Advanced Studies on Collaborative Research (CASCON 2015)*, pp. 141-150, Toronto, Canada, November 2015. (AR 29.5%)
80. **S. Uddin, C. K. Roy** and K. A. Schneider, "Towards Convenient Management of Software Clone Codes in Practice: An Integrated Approach", In *Proceedings of the 2015 Conference of the Centre for Advanced Studies on Collaborative Research (CASCON 2015)*, pp. 211-220, Toronto, Canada, November 2015. (AR 29.5%)
79. **M. M. Rahman, C. K. Roy**, Iman Keivanloo, "Recommending insightful comments for source code using crowdsourced knowledge", In *Proceedings of the 15th IEEE Intl. Working Conference on Software Code Analysis and Manipulation (SCAM 2015)*, pp. 81-90, Bremen, Germany, September 2015. (AR 29.5%)
78. **M. S. Uddin, V. Gaur, C. Gutwin, C. K. Roy**, "On the comprehension of code clone visualizations: A controlled study using eye tracking", In *Proceedings of the 15th IEEE International Working Conference on Software Code Analysis and Manipulation (SCAM 2015)*, pp. 161-170, Bremen, Germany, September 2015.
77. **M. Mondal, C. K. Roy** and K. A. Schneider, "A Comparative Study on the Bug-proneness of Different Types of Code Clones", In *Proceedings of the 31st International Conference on Software Maintenance and Evolution (ICSME 2015)*, pp. 91-100, Bremen, Germany, October 2015. (AR 22%)
76. **J. Svajlenko** and **C. K. Roy**, "Evaluating Clone Detection Tools with BigCloneBench", In *Proceedings of the 31st International Conference on Software Maintenance and Evolution (ICSME2015)*, pp. 131-140, Bremen, Germany, October 2015. (AR 22%)
75. **M. Asaduzzaman, C. K. Roy**, Samiul Monir, and K. A. Schneider, "Exploring API Method Parameter Recommendations", In *Proceedings of the 31st International Conference on Software Maintenance and Evolution (ICSME2015)*, pp. 271-280, Bremen, Germany, October 2015. (AR 22%)



74. **M. Asaduzzaman, C. K. Roy, K. A. Schneider**, "PARC: Recommending API methods parameters", In *Proceedings of the Tool Demonstration Track of the 31st International Conference on Software Maintenance and Evolution (ICSME 2015)*, pp. 330-332, Bremen, Germany, October 2015.
73. **M. M. Rahman and C. K. Roy**, "An Insight into the Unresolved Questions at Stack Overflow", In *Proceeding of the Mining Challenge Track of the 12th Working Conference on Mining Software Repositories (MSR 2015)*, pp. 426-429, Florence, Italy, May 2015.
72. **M. M. Rahman and C. K. Roy**, "TextRank Based Search Term Identification for Software Change Tasks", In *Proceeding of The 22nd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER 2015, ERA)*, pp. 540-544, Montreal, Canada, March 2015.
71. **M. Mondal, C. K. Roy and K. A. Schneider**, "SPCP-Miner: A Tool for Mining Code Clones that are Important for Refactoring or Tracking", In *Proceedings of the Tool Demonstration Track of the 22nd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, 2015, pp. 482--486, Montreal, Canada, March 2015.
70. **G. Kintab, G. McCalla and C. K. Roy**, "Recommending Software Experts Using Code Similarity and Social Heuristics", In *Proceedings of the 2014 Conference of the Centre for Advanced Studies on Collaborative Research (CASCON 2014)*, pp. 4-18, Toronto, November 2014. (Best Paper Award).
69. **M. Mondal, C. K. Roy and K. A. Schneider**, "Late Propagation in Near-Miss Clones: An Empirical Study", In *Proceedings of the 8th International Workshop on Software Clones (IWSC 2014)*, published in the Journal of Electronic Communications of the EASST, Volume 63: Software Clones, 17 pp., 2014.
68. **M. Khan, C. K. Roy and K. A. Schneider**, "Active Clones: Source Code Clones at Runtime", In *Proceedings of the 8th International Workshop on Software Clones (IWSC 2014)*, published in the Journal of Electronic Communications of the EASST, Volume 63: Software Clones, 18 pp., 2014.
67. **M. S. Rahman and C. K. Roy**, "A Change-Type Based Empirical Study on the Stability of Cloned Code", In *Proceedings of the 14th IEEE International Working Conference on Software Code Analysis and Manipulation (SCAM 2014)*, pp. 31-40, Victoria, Canada, September 2014. (AR 31.7%)
66. **M. Mondal, C. K. Roy and K. A. Schneider**, "Automatic Identification of Important Clones for Refactoring and Tracking", In *Proceedings of the 14th IEEE International Working Conference on Software Code Analysis and Manipulation (SCAM 2014)*, pp. 11-20, Victoria, Canada, September 2014. (AR 31.7%)
65. **M. M. Rahman and C. K. Roy**, "On the Use of Context in Recommending Exception Handling Code Examples", In *Proceedings of the 14th IEEE International Working Conference on Software Code Analysis and Manipulation (SCAM 2014)*, pp. 285-294, Victoria, Canada, September 2014. (AR 31.7%)
64. **J. Svajlenko and C. K. Roy**, "Evaluating Modern Clone Detection Tools", In *Proceedings of the 30th International Conference on Software Maintenance and Evolution (ICSME 2014)*, pp. 321-330, Victoria, Canada, September 2014. (AR 19%)
63. **M. Asaduzzaman, C. K. Roy, K. A. Schneider and D. Hou**, "CSCC: Simple, Efficient, Context Sensitive Code Completion", In *Proceedings of the 30th International Conference on Software Maintenance and Evolution (ICSME 2014)*, pp. 71-80, Victoria, Canada, September 2014. (Best Paper Nomination Award). (AR 19%)
62. **M. Mondal, C. K. Roy and K. A. Schneider**, "A Fine-Grained Analysis on the Evolutionary Coupling of Cloned Code", In *Proceedings of the 30th International Conference on Software Maintenance and Evolution (ICSME 2014)*, pp. 51-60, Victoria, Canada, September 2014. (AR 19%)
61. **J. Svajlenko, J. F. Islam, I. Keivanloo, C. K. Roy and M. M. Mia**, "Towards a Big Data Curated Benchmark of Inter-Project Code Clones", In *Proceedings of the Early Research Achievements track of the 30th International Conference on Software Maintenance and Evolution (ICSME 2014)*, pp. 476-480, Victoria, Canada, September 2014.
60. **S. Yeasmin, C. K. Roy and K. A. Schneider**, "Interactive Visualization of Bug Reports using Topic Evolution and Extractive Summaries", In *Proceedings of the Early Research Achievements track of the 30th International Conference on Software Maintenance and Evolution (ICSME 2014)*, pp. 421-425, Victoria, Canada, September 2014.
59. **M. M. Rahman and C. K. Roy**, "SurfClipse: Context-Aware Meta Search in the IDE", In *Proceedings of the Tool Demonstrations track of the 30th International Conference on Software Maintenance and Evolution (ICSME 2014)*, pp. 617-620, Victoria, Canada, September 2014.
58. **M. Asaduzzaman, C. K. Roy, K. A. Schneider and D. Hou**, "Context-sensitive Code Completion Tool for Better API Usability", In *Proceedings of the Tool Demonstrations track of the 30th International Conference on Software Maintenance and Evolution (ICSME 2014)*, pp. 621-624, Victoria, Canada, September 2014.
57. **M. Mondal, C. K. Roy and K. A. Schneider**, "Prediction and Ranking of Co-change Candidates for Clones", In *Proceedings of the 11th Working Conference on Mining Software Repositories (MSR 2014)*, pp. 32-41, Hyderabad, India, May 2014.
56. **M. M. Rahman and C. K. Roy**, 2014, "An Insight into the Pull Requests of GitHub", In *Proceedings of the Mining Challenge Track of the 11th Working Conference on Mining Software Repositories (MSR 2014)*, pp. 364-367, Hyderabad, India, May 2014.
55. **Y. Chen, I. Keivanloo, and C. K. Roy**, "Near-miss Software Clones in Open Source Games: An Empirical Study", In *Proceedings of the 27th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE 2014)*, 7 pp., Toronto, Canada, May 2014.
54. **M. Mondal, C. K. Roy and K. A. Schneider**, "Automatic Ranking of Clones for Refactoring through Mining Association Rules", In *Proceeding of the IEEE CSMR-18/WCRE-21 Software Evolution Week (SEW'14)*, pp. 114--123, Antwerp, Belgium, February 2014. (AR 31%)
53. **M. M. Rahman, S. Yeasmin and C. K. Roy**, "Towards a Context-Aware Meta Search Engine for IDE-Based Recommendation about Programming Errors and Exceptions", In *Proceeding of the IEEE CSMR-18/WCRE-21 Software Evolution Week (SEW'14)*, pp. 94--203, Antwerp, Belgium, February 2014. (AR 31%)

52. **M. Mondal, C. K. Roy** and K. A. Schneider, 2014, "Improving the Detection Accuracy of Evolutionary Coupling by Measuring Change Correspondence", In *Proceeding of the ERA track of the IEEE CSMR-18/WCRE-21 Software Evolution Week (SEW'14)*, pp. 358--362, Antwerp, Belgium, February 2014.
51. **M. M. Rahman, S. Yeasmin, C. K. Roy**, "An IDE-Based Context-Aware Meta Search Engine", In *Proceedings of Early Research Achievement (ERA) Track of the 20th Working Conference on Reverse Engineering (WCRE 2013)*, Koblenz, Germany, October 2013.
50. **M. Asaduzzaman, C. K. Roy**, K. A. Schneider and M. Di Penta, "LHDiff: Tracking Source Code Lines To Support Software Maintenance Activities", In *Proceedings of the Tool Demonstration Track of the 29th IEEE International Conference on Software Maintenance (ICSM'13)*, pp. 484-487, Eindhoven, the Netherlands, September 2013.
49. **R. K. Saha, C. K. Roy**, and K. A. Schneider, "gCad: A Near-Miss Clone Genealogy Extractor to Support Clone Evolution Analysis", In *Proceedings of the Tool Demonstration Track of the 29th IEEE International Conference on Software Maintenance (ICSM'13)*, pp. 488--491, Eindhoven, the Netherlands, September 2013.
48. **J. Svajlenko, C. K. Roy** and S. Duszynski, "ForkSim: Generating Software Forks for Evaluating Cross-Project Similarity Analysis Tools", In *Proceedings of the Tool Paper track of the 13th IEEE International Working Conference on Source Code Analysis and Manipulation (SCAM 2013)*, 37--42, Eindhoven, the Netherlands, September 2013.
47. **M. Asaduzzaman, C. K. Roy**, K. A. Schneider and M. Di Penta, "LHDiff: A Language-Independent Hybrid Approach for Tracking Source Code Lines", In *Proceedings of the 29th IEEE International Conference on Software Maintenance (ICSM'13)*, pp. 230--239, Eindhoven, the Netherlands, September 2013. **(AR 22%)**
46. **M.S. Rahman**, A. Aryani, **C. K. Roy**, and F. Perin, "On the Relationships between Domain-Based Coupling and Code Clones: An Exploratory Study", In *Proceedings of the New Ideas and Emerging Results Track of the 35th International Conference on Software Engineering (ICSE 2013)*, San Francisco, CA, May 2013, pp. 1265-1268.
45. **R.K. Saha, C. K. Roy**, K.A. Schneider, and D. E. Perry, "Understanding the Evolution of Type-3 Clones: An Exploratory Study", In *Proceedings of the 10th Working Conference on Mining Software Repositories (MSR 2013)*, San Francisco, CA, May 2013, pp. 139-148.
44. **M. Asaduzzaman**, A.S. Mashiyat, **C. K. Roy**, and K.A. Schneider, "Answering Questions about Unanswered Questions of Stack Overflow", In *Proceedings of the Mining Challenge Track of the 10th Working Conference on Mining Software Repositories (MSR 2013)*, San Francisco, May 2013, pp. 97-100.
43. **M. Mondal, C. K. Roy**, and K. A. Schneider, "Insight into a Method Co-change Pattern to Identify Highly Coupled Methods: An Empirical Study", In *Proceedings of the 21st IEEE International Conference on Program Comprehension (ICPC 2013)*, San Francisco, CA, May 2013, pp. 103-112.
42. **M.S. Uddin, C. K. Roy**, and K.A. Schneider, "SimCad: An Extensible and Faster Clone Detection Tool for Large Scale Software Systems", In *Proceedings of the Tool Demonstration Track of the 21st IEEE International Conference on Program Comprehension (ICPC 2013)*, San Francisco, CA, May 2013, pp. 236-238.
41. **M. Mondal, C. K. Roy**, and K.A. Schneider, "Improving the Detection Accuracy of Evolutionary Coupling", In *Proceedings of the Early Research Achievements Track of the 21st IEEE International Conference on Program Comprehension (ICPC 2013)*, San Francisco, CA, May 2013, pp. 223-226.
40. **J.T. Svajlenko**, I. Keivanloo, and **C. K. Roy**, "Scaling Classical Clone Detection Tools for Ultra-Large Datasets: An Exploratory Study", In *Proceedings of the ICSE 7th International Workshop on Software Clones (IWSC 2013)*, San Francisco, CA, May 2013, pp. 16-22.
39. **J. Svajlenko, C. K. Roy**, and J.R. Cordy, "A Mutation Analysis Based Benchmarking Framework for Clone Detectors", In *Proceedings of Tool Papers Track of the ICSE 7th International Workshop on Software Clones (IWSC 2013)*, San Francisco, CA, May 2013, pp. 8-9.
38. T. Muhammad, **Minhaz F. Zibran**, Y. Yamamoto, and **C. K. Roy**, "Near-Miss Clone Patterns in Web Applications: An Empirical Study with Industrial Systems", In *Proceedings of the 2013 Canadian Conference on Electrical and Computer Engineering (CCECE 2013)*, Regina, Canada, May 2013, pp. 1-6.
37. **M.F. Zibran, R.K. Saha, C. K. Roy**, and K.A. Schneider, "Evaluating the Conventional Wisdom in Clone Removal: A Genealogy-based Empirical Study", In *Proceedings of the Software Engineering track of the 28th ACM Symposium On Applied Computing (ACM SAC 2013)*, Coimbra, Portugal, March 2013, pp. 1223-1230
36. **M. Mondal, C. K. Roy**, and K. A. Schneider, "Connectivity of Co-changed Method Groups: A Case Study on Open Source Systems", In *Proceedings of the 2012 Conference of the Centre for Advanced Studies on Collaborative Research (CASCON 2012)*, Toronto, Canada, November 2012, pp. 205-219.
35. **F. Al-Omari**, I. Keivanloo, **C. K. Roy** and J. Rilling, "Detecting Clones across Microsoft .NET Programming Languages", In *Proceedings of the 19th Working Conference on Reverse Engineering (WCRE 2012)*, Kingston, Canada, October 2012, pp. 405-414.
34. **I. Keivanloo, C. K. Roy** and J. Rilling, "SeByte: A Semantic Clone Detection Tool for Intermediate Languages", In *Proceedings of the Tool Demonstration Track of the 20th IEEE International Conference on Program Comprehension (ICPC 2012)*, Passau, Bavaria, Germany, June 2012, pp. 247-249.
33. **M. Asaduzzaman**, M. Bullock, **C. K. Roy** and K.A. Schneider, "Bug Introducing Changes: A Study with Android", In *Proceedings of the Mining challenge track of the 9th Working Conference on Mining Software Repositories (MSR 2012)*, Zurich, Switzerland, June 2012, pp. 116-119.
32. **M. Mondal, C. K. Roy**, and K.A. Schneider, "Dispersion of Changes in Cloned and Non-cloned Code", In *Proceedings of the ICSE 6th International Workshop on Software Clones (IWSC 2012)*, Zurich, Switzerland, June 2012, pp. 29-35.
31. **I. Keivanloo, C. K. Roy**, and J. Rilling, "Java Bytecode Clone Detection via Relaxation on Code Fingerprint and Semantic Web Reasoning", In *Proceedings of the ICSE 6th International Workshop on Software Clones (IWSC 2012)*, Zurich, Switzerland, June 2012, pp. 36-42.

30. **I. Keivanloo, C. K. Roy, J. Rilling, and P. Charland**, "Shuffling and Randomization for Scalable Source Code Clone Detection", In *Proceedings of the ICSE 6th International Workshop on Software Clones (IWSC 2012)*, Zurich, Switzerland, June 2012, pp. 82-83.
29. **M. Mondal, C. K. Roy, M.S. Rahman, R.K. Saha, J. Krinke and K.A. Schneider**, "Comparative Stability of Cloned and Non-cloned Code: An Empirical Study", In *Proceedings of the Software Engineering Track of the 27th ACM Symposium on Applied Computing (ACM SAC 2012)*, Riva del Garda, Trento, Italy, March 2012, pp. 1227-1234.
28. **M.F. Zibran and C. K. Roy**, "IDE-based Real-time Focused Search for Near-miss Clones", In *Proceedings of the Software Engineering Track of the 27th ACM Symposium on Applied Computing (ACM SAC 2012)*, Riva del Garda, Trento, Italy, March 2012, pp. 1235-1242.
27. **M.S. Uddin, and C. K. Roy, K.A. Schneider and A. Hindle**, "On the Effectiveness of Simhash for Detecting Near-Miss Clones in Large Scale Software Systems", In *Proceedings of the 18th IEEE Working Conference on Reverse Engineering (WCRE 2011)*, IEEE Press, Lero, Limerick, Ireland, October 2011, pp. 13-22.
26. **R.K. Saha, C. K. Roy and K.A. Schneider**, "An Automatic Framework for Extracting and Classifying Near-Miss Clone Genealogies", In *Proceedings of the 27th IEEE International Conference on Software Maintenance (ICSM 2011)*, IEEE Press, Williamsburg, Virginia, USA, September 2011, pp. 293-302.
25. **M.F. Zibran and C.K Roy**, "A Constraint Programming Approach to Conflict-aware Optimal Scheduling of Prioritized Code Clone Refactoring", In *Proceedings of the 11th IEEE International Working Conference on Source Code Analysis and Manipulation (SCAM 2011)*, IEEE Press, Williamsburg, VA, USA, September 2011, pp. 105-114.
24. **M. Mondal, M.S. Rahman, R.K. Saha, C. K. Roy, J. Krinke and K.A. Schneider**, "An Empirical Study of the Impacts of Clones in Software Maintenance", In *Proceedings of the Student Research Symposium Track of the 19th International Conference on Program Comprehension (ICPC 2011)*, IEEE Press, Kingston, Canada, June 2011, pp. 242-245
23. **J.R. Cordy and C. K. Roy**, "The NiCad Clone Detector", In *Proceedings of the Tool Demo Track of the 19th International Conference on Program Comprehension (ICPC 2011)*, IEEE Press, Kingston, Canada, June 2011, pp. 219-220.
22. **M.F. Zibran and C.K Roy**, "Conflict-aware Optimal Scheduling of Code Clone Refactoring: A Constraint Programming Approach", In *Proceedings of the Student Research Symposium Track of the 19th International Conference on Program Comprehension (ICPC 2011)*, IEEE Press, Kingston, Canada, June 2011, pp. 266-269.
21. **J.R. Cordy and C. K. Roy**, "DebCheck: Efficient Checking for Open Source Clones in Software Systems", In *Proceedings of the Tool Demo Track of the 19th International Conference on Program Comprehension (ICPC 2011)*, IEEE Press, Kingston, Canada, June 2011, pp. 217-218.
20. **M.F. Zibran and C. K. Roy**, "Towards Flexible Code Clone Detection, Management, and Refactoring in IDE", In *Proceedings of the ICSE 5th International Workshop on Software Clones (IWSC 2011)*, ACM Press, Waikiki, Hawaii, USA, May 2011, pp. 75-76. (Position Paper).
19. **R.K. Saha, C. K. Roy and K.A. Schneider**, "Visualizing the Evolution of Code Clones", In *Proceedings of ICSE 5th International Workshop on Software Clones (IWSC 2011)*, ACM Press, Waikiki, Hawaii, USA, May 2011, pp. 71-72. (Position Paper).
18. **M. Asaduzzaman, C. K. Roy and K.A. Schneider**, "VisCad: Flexible Code Clone Analysis Support For NiCad", In *Proceedings of the Tool Demo Track of the ICSE 5th International Workshop on Software Clones (IWSC 2011)*, ACM Press, Waikiki, Hawaii, USA, May 2011, pp. 77-78.
17. **M.F. Zibran, R.K. Saha, M. Asaduzzaman and C. K. Roy**, "Analyzing and Forecasting Near-miss Clones in Evolving Software: An Empirical Study", In *Proceedings of the 16th IEEE International Conference on Engineering of Complex Computer Systems (ICECCS 2011)*, IEEE Press, Las Vegas, USA, April 2011, pp. 295-304.
16. **R.K. Saha, M. Asaduzzaman, M.F. Zibran, C. K. Roy and K. A. Schneider**, "Evaluating Code Clone Genealogies at Release level: An Empirical Study" In *Proceedings of the 10th IEEE International Conference on Source Code Analysis and Manipulation (SCAM 2010)*, IEEE Press, Timisoara, Romania, September 2010, pp. 87-96.
15. **M.F. Zibran, F.Z. Eishita and C. K. Roy**, "Useful, but usable? Factors Affecting the Usability of APIs", In *Proceedings of the 18th IEEE Working Conference on Reverse Engineering (WCRE 2011)*, IEEE Press, Lero, Limerick, Ireland, October 2011, pp. 151-155.
14. **C. K. Roy**, "Detection and Analysis of Near-Miss Software Clones" In *Proceedings of the Doctoral Symposium Track of the 25th IEEE International Conference on Software Maintenance (ICSM 2009)*, IEEE Press, Edmonton, Canada, September 2009, pp. 447-450.
13. **C. K. Roy and J.R. Cordy**, "Are Scripting Languages Really Different?", In *Proceedings of the ICSE 4th International Workshop on Software Clones (IWSC 2010)*, ACM Press, Cape Town, South Africa, May 2010, pp. 17-24.
12. **C. K. Roy and J.R. Cordy**, "A Mutation / Injection-based Automatic Framework for Evaluating Code Clone Detection Tools", in *Proceedings of the 4th International Workshop on Mutation Analysis (Mutation 2009)*, IEEE Press, Denver, Colorado, USA, April 2009, pp. 157-166. (Received the Best Paper Award).
11. **C. K. Roy and J.R. Cordy**, "An Empirical Evaluation of Function Clones in Open Source Software", in *Proceedings of the 15th Working Conference on Reverse Engineering (WCRE 2008)*, pp. 81-90, IEEE Press, Antwerp, Belgium, October 2008.
10. **C. K. Roy and J.R. Cordy**, "NICAD: Accurate Detection of Near-Miss Intentional Clones Using Flexible Pretty-Printing and Code Normalization", in *Proceedings of the 16th IEEE International Conference on Program Comprehension (ICPC 2008)*, pp. 172-181, IEEE Press, Amsterdam, The Netherlands, June 2008.
9. **C. K. Roy and J.R. Cordy**, "Scenario-based Comparison of Clone Detection Techniques", in *Proceedings of the 16th IEEE International Conference on Program Comprehension (ICPC 2008)*, pp.153-162, IEEE Press, Amsterdam, The Netherlands, June 2008. (Invited for special journal issue as one of the selected best papers).

8. **C. K. Roy** and J.R. Cordy, "Towards a Mutation-Based Automatic Framework for Evaluating Code Clone Detection Tools", in *Proceedings of the Poster Paper Track of the Canadian Conference on Computer Science and Software Engineering (C3S2E 2008)*, ACM Press, Montreal, Canada, May 2008, pp. 137-140.
7. **C. K. Roy**, M.G. Uddin, B. Roy and T.R. Dean, "Evaluating Aspect Mining Techniques: A Case Study", in *Proceedings of the 15th IEEE International Conference on Program Comprehension (ICPC 2007)*, IEEE Press, Banff, Canada, June 2007, pp. 167-176.
6. B. Roy, **C. K. Roy** and M. Einhaus, "Interference Aware Dynamic Subchannel Allocation in a Multi-cellular OFDMA System based on Traffic Situation", in *Proceeding of the 4th International Symposium in Parallel and Distributed Processing and Applications (ISPA2006)*, LNCS Springer Press, Sorrento, Italy, December 2006, pp. 341-352.
5. **Farouq Al-omari** and **C. K. Roy**, "Is Code Cloning in Games Really Different?", In *Proceedings of the Software Engineering Track of the 31st ACM Symposium on Applied Computing (ACM SAC 2016)*, Pisa, Italy, April 2016, pp. 1512-1519.
4. **C. K. Roy**, T. Noll, B. Roy and J.R. Cordy, "Towards Automatic Verification of Erlang Programs by pi-Calculus Translation", in *Proceedings of the ACM SIGPLAN 2006 5th Erlang Workshop (Erlang 2006)*, ACM Press, Portland, Oregon, September 2006, pp. 38-49.
3. T. Noll and **C.K Roy**, "Modeling Erlang in the Pi-Calculus", in *Proceedings of the ACM SIGPLAN 2005 4th Erlang Workshop (Erlang 2005)*, ACM Press, Tallinn, Estonia, September 2005, pp. 72-77.
2. M.Z. Hoque, **C. K. Roy**, N.Y. Ali and M.R. Khan, "Largest Subset of Disjoint Products in ESOP", in *Proceedings of the 5th International Conference on Computer and Information Technology (ICCIT 2002)*, Dhaka, Bangladesh, December 2002, pp. 305-311.
1. **C. K. Roy**, M. Masud, M.M. Asaduzzaman and H.H. Babu, "A Modification of Huffman Header", in *Proceedings of the 4th International Conference on Computer and Information Technology (ICCIT 2001)*, Dhaka, Bangladesh, December 2001, pp. 409- 414.

#### CHAPTERS IN BOOKS, EXPOSITORY AND REVIEW ARTICLES

8. **M.M. Hossain**, B. Roy, **C. Roy**, K.A. Schneider, "Reproducibility Challenges of External Computational Experiments in Scientific Workflow Management Systems", in: *Kurosu, M., Hashizume, A. (eds) Human-Computer Interaction. HCII 2024. Lecture Notes in Computer Science*, vol 14687 (189-207). Springer, Cham. [https://doi.org/10.1007/978-3-031-60441-6\\_13](https://doi.org/10.1007/978-3-031-60441-6_13)
7. **Amit Kumar Mondal**, Banani Roy, **Chanchal K. Roy**, and Kevin A. Schneider, Banani Roy. Large Scale Image Registration Utilizing Data-Tunneling in the MapReduce Cluster. In *Lecture Notes on Data Engineering and Communications Technologies*, vol 95. 2022.
6. **Manishankar Mondal**, **Chanchal K. Roy**, James R. Cordy: NiCad: A Modern Clone Detector. *Code Clone Analysis 2021*: 45-50
5. **Manishankar Mondal**, **Chanchal K. Roy**, Kevin A. Schneider: A Summary on the Stability of Code Clones and Current Research Trends. *Code Clone Analysis 2021*: 169-180
4. Hitesh Sajjani, Vaibhav Saini, **Chanchal K. Roy**, Cristina V. Lopes: SourcererCC: Scalable and Accurate Clone Detection. *Code Clone Analysis 2021*: 51-62
3. Jeffrey Svajlenko, **Chanchal K. Roy**: BigCloneBench. *Code Clone Analysis*, 2021: 93-105
2. **R. Ferdous**, B. Roy, **C. K. Roy** and K. A. Schneider, "Workflow Provenance for Big Data: From Modelling to Reporting", Alhaji R, Moshirpour M, Far B. In: Alhaji R., Moshirpour M., Far B. (eds) *Data Management and Analysis. Studies in Big Data*. (65): pp. 185-200.
1. **D. Chakroborti**, B. Roy, **A. K. Mondal**, **G. Mostaeen**, R. Deters, **C. K. Roy** and K. A. Schneider, "A Data Management Scheme for Micro-Level Modular Computation-intensive Programs in Big Data Platforms", Alhaji R, Moshirpour M, Far B. In: Alhaji R., Moshirpour M., Far B. (eds) *Data Management and Analysis. Studies in Big Data*. (65): pp. 135-153.

#### TECHNICAL REPORTS RELEVANT TO ACADEMIC FIELD (NON-REFERRED)

5. **J. Svajlenko** and **C. K. Roy**, "A Survey on the Evaluation of Clone Detection Performance and Benchmarking", arXiv:2006.15682 [cs.SE], 2020, 109 pages.
4. **I. Keivanloo**, **C. K. Roy**, and J. Rilling, 2014, "Towards Source Code Clone Search via Information Retrieval", Technical Report 2014-02, Department of Computer Science. 21. University of Saskatchewan
3. **M. Rahman**, **C. K. Roy**, I. Keivanloo, 2013, "Subjective Evaluation of Software Quality Using Crowdsourcing Knowledge: An Exploratory Study", Technical Report 2013-1, Department of Computer Science, University of Saskatchewan, 10 pp, 2013.
2. **M. F. Zibran** and **C. K. Roy**, 2012. "The Road to Software Clone Management: A Survey", Technical Report 2012-03, Department of Computer Science, The University of Saskatchewan, Canada, February 2012, 62 pp.
1. **C. K. Roy** and J.R. Cordy, 2007. "A Survey on Software Clone Detection Research", Technical Report 2007-541, School of Computing, Queen's University, September 2007, 115 pp.

## INVITED LECTURES (OUTSIDE THE U OF S) AND INVITED CONFERENCE PRESENTATIONS (Selected)

30. **C. K. Roy**, The roles of large language models in clone detection, analysis and benchmarking, *the 15th International Conference on Cloud Computing, Data Science & Engineering (Confluence-2025)*, Noida, India, January 2025 (Keynote talk)
29. **C. K. Roy**, AI-based clone analytics, *Advances in Computing Technology (A-FACT) on Artificial Intelligence, Data Analytics, Machine Learning, Cloud Computing, Internet of Things and Cyber Security*, Amity University Uttar Pradesh, Noida, May 2024 (Expert speaker)
28. **C. K. Roy**, AI-enabled Software Clone detection, Analysis and Benchmarking, *the 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence-2024)*, Noida, India, January 2024 (Keynote talk)
27. **C. K. Roy**, An approach to reproducible, repeatable and scalable software clone analytics, *the 2nd International Conference on Big Data, IoT and Machine Learning (BIM 2023)*, Dhaka, Bangladesh, September 2023 (keynote talk)
26. **C. K. Roy**, SciClone: A Scientific Workflow based approach to reproducible, repeatable and scalable software clone analytics, *the 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence-2023)*, Noida, India, January 2023 (Keynote talk)
25. **C.K. Roy**, "Clone Detection and Benchmarking in Big Code", *the 3rd International Conference on Sustainable Technologies for Industry 4.0 (STI 2021)*, Dhaka, Bangladesh, December 2021 (keynote talk).
24. **C.K. Roy**, "A Machine Learning Based Framework for validating code clones from Big Code", *the International Conference on Big Data, IoT and Machine Learning (BIM 2021)*, Chittagong, Bangladesh, September 2021 (keynote talk).
23. **C.K Roy**, "Clone Detection Big Code: Techniques, benchmarks and applications", *Software Engineering Research in India (SERI 2020)*, July 2020.
22. **C. K. Roy**, "Large scale and collaborative data management and analytics for advanced agriculture research", *GIFS-BARC Technology Centre Workshop*, Dhaka, Bangladesh February 2020.
21. **C. K. Roy**, "Towards a next generation clone management system", *Khulna University*, Khulna, Bangladesh, January 2020.
20. **C. K. Roy**, "Data Management and Analytics in Digitizing Agriculture in Bangladesh", *Workshop on Digitizing Agriculture in Bangladesh*, Khulna University, Khulna, Bangladesh, January 2020.
19. **C. K. Roy**, "5-Minute Introduction to Research", *Celebration of ASE 2019*, 2019.
18. **C. K. Roy**, "On large scale buggy clone detection and management", *Department of Computer Science and Software Engineering, Concordia University*, Montreal, Canada, June 2019.
17. **C. K. Roy**, "Visualization of large scale evolving clones for clone management", *Polytechnique Montreal*, Montreal, Canada, June 2019.
16. **C. K. Roy**, "State of the art in clone analysis in Big Code: Past, Present and Future", *Department of Computer Science, Columbia University*, New York, USA, April 2019.
15. **C. K. Roy**, "Big Data clone detection and their applications", *Department of Computer Science, Rochester Institute of Technology*, Rochester, USA, April 2019.
14. **C. K. Roy** and James. R. Cordy, "Adventures in NICAD: a ten-year retrospective", *Most Influential Paper Award talk for ICPC 2008 paper at the 26th International Conference on Program Comprehension (ICPC 2018)*.
13. **C. K. Roy** and James. R. Cordy, "Benchmarks for software clone detection: a ten-year retrospective", *Most Influential Paper Award talk for WCRE 2008 paper at the 25th International Conference on Software Analysis, Evolution, and Reengineering (SANER'18)*.
12. **C. K. Roy**, "On the detection and management of code clones: The state of the art", (Keynote talk), *the 12th International Workshop on Software Clones (IWSC 2018)*, Campobasso, Italy, March 2018.
11. **C. K. Roy**, "Code Reviewer Recommendation at GitHub", *Discipline of computer Science and Engineering, Khulna University*, December 2017
10. **C. K. Roy**, "Research methodologies and automation in Software Engineering: an introduction with peer code review", *Department of Computer Science & Engineering, Khulna University of Engineering & Technology (KUET) Khulna-9203*, Bangladesh.
9. **C. K. Roy**, "Towards Automated Software technologies for Sustainable Humanity: Peer Code Review and Big Data Cloud Technologies for Food Security (Keynote talk)", *the fifth IEEE Region 10 Humanitarian Technology Conference (R10HTC)*, December 2017, Dhaka.
8. **C. K. Roy**, "Towards Automated Supports for Code Reviews using Reviewer Recommendation and Review Quality Modelling", *the 56th CREST Open Workshop - Code Review and Continuous Inspection/Integration*, University College London, November 2017.
7. **C. K. Roy**, M. F. Zibran, and R. Koschke, "The Vision of Software Clone Management: Past, Present and Future (keynote paper)", *Vision keynote presented at the IEEE CSMR-18/WCRE-21 Software Evolution Week (SEW'14)*, pp. 18--33, Antwerp, Belgium, February 2014. (Vision keynote speakers).
6. **C. K. Roy** and Ahmed Abdel Moamen, "Exploring Development Practices of Android Mobile Apps from Different Categories", *the Eight International Workshop on Software Quality and Maintainability (SQM 2014)*, Antwerp, Belgium, February 2014.
5. **C.R. Roy**, "Toward a Comprehensive Clone Management System", *the Consortium for Software Engineering Research (CSER) Spring 2012 Meeting*, University of Victoria, May 2012.
4. **C.R. Roy**, "A Benchmarking Framework for Evaluating Clone Detection Tools", *Departmental Seminar*, in the Faculty of Science (Computer Science) at the University of Ontario Institute of Technology (UOIT), November 2011.

3. **C.R. Roy**, “Code Clone Detection and Management: Past, Present and the Future”, *the Consortium for Software Engineering Research (CSER) Spring 2011 Meeting co-located with ICPC’11*, Queen’s University at Kingston, June 2011. (Keynote Speaker)
2. **C.R. Roy**, “Evaluating Software Clone Detection Tools: A Mutation-Based Approach”, *the 9th CREST Open Workshop Code Provenance and Clone Detection*, Department of Computer Science, University College London, November 2010.
1. **C. K. Roy**, “Near-miss Function Clones in Open Source Software: An Empirical Study”, *CSER Workshop on Mining Software Repositories*, Montréal, Canada, April 2009.

#### RESEARCH GRANT AND CONTRACT INFORMATION (Selected)

Principal Investigator/Co-Applicant/Co-Investigator	Over \$6M
-----------------------------------------------------	-----------

#### Over the last four years:

24. **C. K. Roy**, Principal Investigator, *SciClone: A Scientific Workflow based approach for studying the change, similarity and redundancy in Software*, NSERC Discovery grant, \$295,000 over five years, 2023-28.
23. **C. K. Roy**, Principal Investigator (co-applicants, K. Schneider, J. Pomeroy), *Actionable software analytics and data management & analytics for global water futures and global institute of food security*, Living Skies Postdoctoral Fellowship (PDF) Program, \$360,000 over two years, 2022-2024.
22. **C. K. Roy**, Principal Investigator (co-applicants, K. Schneider, D. Mondal), *Software Analytics Research Lab*, \$875,000 (\$350K from CFI), John R. Evans Leaders Fund (including the matching fund).
21. **C. K. Roy**, Principal Investigator, (with seven more co-applicants), *SOAR: Software Analytics Research. CREATE Industrial Stream*, \$1,650,000 over six years, September 2020 — August 2027.
20. **C. K. Roy**, Academic Supervisor for intern Dylan Froh, *Archiving Groups and Athletes - Dylan Froh*, Mitacs Business Strategy Internship (BSI), \$15,000, May 2024 to December 2024.
19. **C. K. Roy**, Principal Investigator (co-applicants, K. Schneider, D. Mondal), *Software Analytics Research Lab*, CFI Infrastructure Operating Fund (IOF), \$78,750, December 2023 to present.
18. **C. K. Roy**, Principal Investigator, College of Arts and Science (A&S) - *Match to CREATE SOAR: Software Analytics Research*, \$60,000 over six year, September 2020 — August 2026.
17. **C. K. Roy**, Principal Investigator, Office of the Vice President of Research (OVPR)- *Match to CREATE SOAR: Software Analytics Research*, \$100,000 over six year, September 2020 — August 2026.
16. **C. K. Roy**, Principal Investigator, College of Graduate and Postdoctoral Studies (CGPS) - *Match CREATE SOAR: Software Analytics Research*, \$100,000 over six year, September 2020 — August 2026.
15. **C. K. Roy**, Principal Investigator, Global Institute for Food Security (GIFS) – *Match to CREATE SOAR: Software Analytics Research*, \$390,000 over six year, September 2020 — August 2026.
14. **C.K. Roy**, Co-Investigator (PI-Kevin Schneider for the sub-grant), *Global Water Future, Transformative Technologies for Canadian Water Futures: Big Data Platform and “Smart” Watersheds*, NSERC Canada First Research Excellence Fund (CFREF), over \$500,000 for April 1, 2018 to February 28, 2025, Total grant: 37M, 2018-2024, PI: John Pomeroy for the primary grant.

#### In the past:

13. **C. K. Roy**, Principal Investigator, *A framework for studying and predicting bug propagation through code cloning for mobile apps*, \$25,000, September 2019 — August 2020. NSERC Engage grant with Push Interactions.
12. **C. K. Roy**, Co-Principal Investigator, *Maintenance Project: Data Management Portal and Repository* (Project within CFREF project Designing Crops for Global Food Security), \$902,520 over three years, April 2019—March 2023.
11. **C. K. Roy**, Principal Investigator, *Change, Similarity and Redundancy in Software Systems*, \$120,000, Discovery Accelerator Supplements Grant, April 1, 2016 – March 31, 2019, \$40,000/year.
10. **C. K. Roy**, Co-Principal Investigator, *Plant Phenotyping and Imaging Research Centre (P2IRC)*, Co-Lead of SubTheme 3.1 P2IRC Cloud: Big Data Analytics for Crop Phenomics, NSERC Canada First Research Excellence Fund, \$791,823 (\$263,941/year, 2016-2019; Total grant: \$37M, P.I. Maurice Moloney).
9. **C. K. Roy**, Principal Investigator, *Code Reviewer Recommendation Based on Cross-Project & Technology Experience*, NSERC Engage Grant with Vendasta Technologies, Mar 1, 2016 to Aug 31, 2016, \$25,000.00
8. **C. K. Roy**, Principal Investigator, *A Simple, Efficient and Robust Code Completion System for ESTI Development Environment*, NSERC Engage Grant with ESTI Consulting Services, May 1, 2016 to Oct 31, 2016, \$25,000.00
7. **C. K. Roy**, Principal Investigator, *Change, Similarity and Redundancy in Software Systems*, \$215,000, NSERC Discovery Grants, April 1, 2015 – March 31, 2022, \$43,000/year. (extended for two years because of serving in the NSERC Evaluation group 1507)
6. **C. K. Roy**, Principal Investigator, New Faculty Graduate Support Program Fund, \$66,000, College of Graduate Studies and Department of Computer Science, September 1, 2010 – August 31, 2013, \$22,000/year.
5. **C. K. Roy**, Principal Investigator, *Change, Similarity and Redundancy in Software Systems*, NSERC DG Early Career Researcher (ECR) Supplement, \$20,000.00, April 1, 2011 – March 31, 2015, \$5,000.00/year
4. **C. K. Roy**, Principal Investigator, *Change, Similarity and Redundancy in Software Systems*, \$75,000, NSERC Discovery Grants, April 1, 2010 – March 31, 2015, \$15,000/year.
3. **C. K. Roy**, Principal Investigator, *Research on Software Maintenance and Evolution: Continuous Research Fund*, \$20,000, U of S College of Arts and Science Start-up Fund (Operational), July 1 2009 –to date.
2. **C. K. Roy**, Principal Investigator, *Establishing a Software Research Lab*, \$30,000, U of S and College of Arts and Science

Capital Start-up Fund, July 1, 2009 – June 30, 2012.

1. **C. K. Roy**, Principal Investigator, *Establishing a Software Research Lab*, \$5,000, U of S VP Academic Start-up Fund (Operational), July 1, 2009 – June 30, 2012.

## CONFERENCE AND JOURNAL PARTICIPATION

Journal Editorial Boards/Guest Editors	4 editions
Organizing Chairs of Conferences	19 conferences editions
Conference Session Chairs	23 conferences editions
Program Committee Members of Conferences	70 conference editions
Reviewer for Conferences	350+ papers reviewed for over 70 conference editions
Reviewer for Journals	Reviewed over 130 journal articles for over 15 journals

## JOURNAL EDITORIAL BOARDS

5. Editorial Board Member, Journal of Software: Evolution and Process, 2024-present
4. Guest co-editor, Journal of Systems and Software, ISEC 2022 special issue
3. Editorial Board Member, Journal of Systems and Software, 2020-23
2. Guest co-editor, Empirical Software Engineering, special issue for ICPC 2018
1. Guest co-editor, Journal of Software: Evolution and Process, Special Issue for ICPC 2014.

## CONFERENCE/WORKSHOP ORGANIZATION

19. **Steering Committee member**, Innovations in Software Engineering Conference (ISEC), 2022 to date
18. **Program Co-Chair**, the 15<sup>th</sup> Innovations in Software Engineering Conference (ISEC 2022)
17. **Organizing/Steering Committee Member**, International Workshop on Software Clones (IWSC)
16. **General Chair**, SCAM 2019, 19th IEEE International Working Conference on Source Code Analysis and Manipulation (SCAM), Cleveland, Ohio, USA.
15. **Steering Committee Member**, IEEE/ACM International Conference on Program Comprehension (ICPC) for three years, 2018-2021
14. **Program Co-Chair**, ICPC 2018, the 26th IEEE/ACM International Conference on Program Comprehension.
13. **NIER track Co-Chair**, ICSME 2017, the New Ideas and Emerging Research Track of the 33rd IEEE International Conference on Software Maintenance and Evolution (ICSME), Shanghai, China, 12pp., September 20-22.
12. **General Chair**, IWSC 2015, 9th International Workshop on Software Clones, March 2015.
11. **General Chair**, ICPC 2014, 22nd IEEE International Conference on Program Comprehension, June 2014.
10. **Finance Chair**, ICSM 2013, 29<sup>th</sup> IEEE International Conference on Software Maintenance, Eindhoven, The Netherlands, September.
9. **Program co-chair**, IWSC 2012, the ICSE 6th International Workshop on Software Clones Switzerland, June.
8. **Tool Track Co-Chair**, ICSM 2012, 28<sup>th</sup> IEEE International Conference on Software Maintenance, Italy, September.
7. **Poster Chair**, ICPC 2012, 20th IEEE International Conference on Program Comprehension, Germany, June.
6. **Tool Track Chair**, SCAM 2012, 11th IEEE International Working Conference on Source Code Analysis and Manipulation, Riva del Garda, Trento, Italy, September.
5. **Tool Track Co-chair**, WCRE 2012, the 19<sup>th</sup> Working Conference on Reverse Engineering, October.
4. **Organizing co-chair**, CSER 2011 (Fall), The Consortium for Software Engineering Research (CSER) Fall 2011.
3. **Organizing Chair**, CSER 2011 Spring Workshop on the Future Trends of Detection, Evolution, Management and Applications of Code Clones, Kingston, On, Canada, June.
2. **Finance Chair**, ICPC 2011, 19th IEEE International Conference on Program Comprehension, Canada, June.
1. **Publicity Chair**, WCRE 2009, 16th Working Conference on Reverse Engineering, Lille, France, October.

## CONFERENCE SESSION CHAIRS

23. **MSR 2023**, the 20th International Conference on Mining Software Repositories (MSR 2023), Melbourne, Australia, May 2023
22. **ICPC 2023**, the 31st IEEE/ACM International Conference on Program Comprehension (ICPC 2023), Melbourne, Australia, May 2023
21. **ICSE 2020**, The 42nd International Conference on Software Engineering, July 2020 (held virtually based in South Korea).
20. **IWSC 2020**, 14th International Workshop on Software Clones, London, Ontario, Canada, February 2020.
19. **ASE 2019**, the 34th IEEE/ACM International Conference on Automated Software Engineering, San Diego, California, United States, November, 2019.
18. **MSR 2019**, the 16th International Conference on Mining Software Repositories, Montreal, May 2019.

17. **SCAM 2019**, the 19th IEEE International Working Conference on Source Code Analysis and Manipulation, Cleveland, United States, September 2019.
16. **ICPC 2019**, the 27th International Conference on Program Comprehension, Montreal, Canada, May 2019.
15. **ICSE 2018**, the 40th International Conference on Software Engineering (ICSE 2018), Gothenburg, Sweden.
14. **SANER 2018**, The 25th International Conference on Software Analysis, Evolution, and Reengineering, Italy.
13. **ICSME 2017**, International Conference on Software Maintenance and Evolution (ICSME), Shanghai, China, 2017.
12. **MSR 2017**, The 14th International Conference on Mining Software Repositories, Buenos Aires, Argentina, May 2017
11. **CSER 2014 Spring Meeting**, Consortium for Software Engineering Research (CSER), Canada April 2014.
10. **MSR 2014**, the 11th Working Conference on Mining Software Repositories (MSR 2014), Hyderabad, May 2014.
9. **CSMR-18/WCRE-21 2014**, the joint 18th European Conference on Software Maintenance and Reengineering / 21st Working Conference on Reverse Engineering, Antwerp, Belgium, February 2014
8. **IWSC 2014**, the 8th International Workshop on Software Clones (IWSC 2014), Antwerp, Belgium, February 2014.
7. **ICSM 2012**, the 28<sup>th</sup> IEEE International Conference on Software Maintenance, Italy, September 2012.
6. **SCAM 2012**, the 12th IEEE International Working Conference on Source Code Analysis and Manipulation,
5. **ICPC 2012**, the 20th IEEE International Conference on Program Comprehension (ICPC 2012), Passau, June 2012.
4. **IWSC 2012**, the ICSE 6th International Workshop on Software Clones (IWSC 2012), Zurich, June 2012.
3. **SCAM 2011**, 11th IEEE International Working Conference on Source Code Analysis and Manipulation, 2011.
2. **WCRE 2011**, 18<sup>th</sup> Working Conference on Reverse Engineering, Lero, Limerick, Ireland, October 2011.
1. **ICSM 2010**, 26th IEEE International Conference on Software Maintenance, Timișoara, Romania, September 2010.

### MEMBER OF THE PROGRAM COMMITTEE FOR CONFERENCES AND WORKSHOPS OF OVER 70 EDITIONS

Total number of papers of average length 10-12 pages reviewed	350+
---------------------------------------------------------------	------

70. **ASE 2024**, the 39th IEEE/ACM International Conference on Automated Software Engineering (ASE 2024), January 2024 to October 2024 (Technical Research Track)
69. **ASE 2024**, the 39th IEEE/ACM International Conference on Automated Software Engineering (ASE 2024), February 2024 to October 2024 (ASE workshop RENE track)
68. **ICSE 2024**, the IEEE/ACM 46th International Conference on Software Engineering Companion (ICSE 2024), July 2023 to April 2024 (Software Engineering in Society)
67. **SANER 2024**, the 31st International Conference on Software Analysis, Evolution and Reengineering (SANER 2024), August 2023 to March 2024 (Technical Research Track)
66. **ICSME 2024**, the 40th IEEE International Conference on Software Maintenance and Evolution (ICSME), November 2023 to September 2024 (Technical Research Track)
65. **CASCON 2024**, the 34th International Conference on Collaborative Advances in Software and COmputiNg (CASCON 2024), February 2024 to November 2024 (Technical Research Track)
64. **ICSME 2023**, the 39th IEEE International Conference on Software Maintenance and Evolution (ICSME) October 2023 (New Ideas and Emerging Research track)
63. **ICSME 2023**, the 39th IEEE International Conference on Software Maintenance and Evolution (ICSME) October 2023 (Technical research track)
62. **SANER 2023**, the 30th IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), Macao SAR, China, March 2023
61. **ICCIT 2022**, 25th International Conference on Computer and Information Technology (ICCIT), December 2022.
60. **ICSME 2022** (Research track), the 38<sup>th</sup> IEEE International Conference on Software Maintenance and Evolution (ICSME), September 2022.
59. **ICSME 2020** (Research track), the 36th IEEE International Conference on Software Maintenance and Evolution (ICSME), September 2020 (will happen virtually).
58. **ICSME 2020** (New Ideas track), the 36th IEEE International Conference on Software Maintenance and Evolution (ICSME), September 2020 (will happen virtually).
57. **ICSE 2020**, The Demonstration track of the 42nd International Conference on Software Engineering, 2020.
56. **CASCON 2020**, the 30th Annual International Conference on Computer Science and Software Engineering (CASCON), Toronto, Canada, 2019.
55. **TENSYM 2020**, IEEE Region 10 Symposium (TENSYMP) Dhaka, April 2020
54. **SANER 2020**, 27th edition of the IEEE International Conference on Software Analysis, Evolution and Reengineering will be held at Western University in London, Ontario, February 2020.
53. **ASE 2019**, the 34th IEEE/ACM International Conference on Automated Software Engineering (ASE 2019), San Diego, USA, November 2019.
52. **ASE 2019**, the Industry Showcase of the 34th IEEE/ACM International Conference on Automated Software Engineering (ASE 2019), San Diego, USA, November 2019.
51. **ICPC 2019** (Replication Track), the 27th International Conference on Program Comprehension, Montreal, Canada, May 2019.
50. **ICPC 2019** (Research), the 27th International Conference on Program Comprehension, Montreal, Canada, May 2019.
49. **MSR 2019**, the 16th International Conference on Mining Software Repositories, Montreal, Canada, May 2019.
48. **SANER 2019**, the 26th IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER 2019), Hangzhou, China, February 2019.
47. **IWSC 2019**, the 13th International Workshop on Software Clones, Hangzhou, China, February 2019.
46. **ICSME 2018**, the 34th IEEE International Conference on Software Maintenance and Evolution (ICSME 2018).



45. **MSR 2018**, the 15th International Conference on Mining Software Repositories (MSR 2018).
44. **ICSE 2018**, the 40th International Conference on Software Engineering (ICSE 2018).
43. **IWSC 2018**, the 12th International Workshop on Software Clones (IWSC 2018).
42. **ICSME 2017**, the 33rd IEEE International Conference on Software Maintenance and Evolution (ICSME 2017).
41. **MSR 2017**, the 14th International Conference on Mining Software Repositories (MSR 2017).
40. **IWSC 2017**, the 11th International Workshop on Software Clones (IWSC 2017).
39. **PIC 2016**, the 2016 International Conference on Progress in Informatics and Computing (PIC-2016).
38. **SCAM 2016**, the 16th IEEE International Working Conference on Source Code Analysis and Manipulation (SCAM 2016), Raleigh, NC, U.S.A, October 2016.
37. **IWSC 2016**, the 10th International Workshop on Software Clones (IWSC 2016), Osaka, Japan, March 2016.
36. **SANER 2016**, the 23rd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER 2016), pp. 349-359, Osaka, Japan, March 2016.
35. **SCAM 2015**, the 15th IEEE International Working Conference on Source Code Analysis and Manipulation (SCAM 2015), Bremen, Germany, September 2015
34. **MSR 2015**, The 12th Working Conference on Mining Software Repositories (MSR 2015), Florence, Italy, May 2015.
33. **ICPC 2015**, the 23rd IEEE International Conference on Program Comprehension (ICPC 2015), Florence, Italy, May 2015.
32. **IWSC 2015**, the 9th International Workshop on Software Clones (IWSC 2015), Montreal, March 2015.
31. **SCAM 2014**, The 14th IEEE International Working Conference on Source Code Analysis and Manipulation, Victoria, British Columbia, Canada, September 2014.
30. **CASCON 2014**, the 24th Annual International Conference of the Centre for Advanced Studies on Collaborative Research, Toronto, Canada, November 2014.
29. **ICPC 2014**, the 22nd IEEE International Conference on Program Comprehension (ICPC 2014), Hyderabad, India, May 2014.
28. **MSR 2014**, the 11th Working Conference on Mining Software Repositories (MSR 2014), Hyderabad, India, May 2014.
27. **CCECE 2014**, the 26th Annual IEEE Canadian Conference on Electrical and Computer Engineering, Regina, Canada, May 2013.
26. **ICSOFT 2014**, 9th International Conference on Software Engineering and Applications, Vienna, Austria, August 2014.
25. **PIC 2014**, International Conference on Progress in Informatics and Computing, Shanghai, China, May 2014.
24. **IWSC 2014**, the 8th International Workshop on Software Clones (IWSC 2014), Antwerp, Belgium, February 2014.
23. **CSMR-18/WCRE-21 2014**, the joint 18th European Conference on Software Maintenance and Reengineering / 21st Working Conference on Reverse Engineering, Antwerp, Belgium, February 2014.
22. **ICSM 2013**, the 29<sup>th</sup> IEEE International Conference on Software Maintenance, Eindhoven, The Netherlands, September 2013.
21. **WCRE 2013**, the 20<sup>th</sup> Working Conference on Reverse Engineering, Koblenz, Germany, October 2013.
20. **CASCON 2013**, the 23rd IBM Centre for Advanced Studies International Conference on Computer Science and Software Engineering, Toronto, Canada, November 2013.
19. **SCAM 2013**, the 13th IEEE International Working Conference on Source Code Analysis and Manipulation, Eindhoven, The Netherlands, September 2012.
18. **IWSC 2013**, the ICSE 7th International Workshop on Software Clones, San Francisco, California, May 2013.
17. **MSR 2013**, the 10th Working Conference on Mining Software Repositories, San Francisco, California, May 2013
16. **ICPC 2013**, the 21st IEEE International Conference on Program Comprehension, San Francisco, California, May 2013.
15. **CCECE 2013**, the 26<sup>th</sup> Annual Canadian Conference on Electrical and Computer Engineering, Regina, Canada, May 2013.
14. **ICSM 2012**, the 28th IEEE International Conference on Software Maintenance, Riva del Garda, Italy, September 2012.
13. **SCAM 2012**, the 12th IEEE International Working Conference on Source Code Analysis and Manipulation, Riva del Garda, Italy September 2012.
12. **WCRE 2012**, the 19th Working Conference on Reverse Engineering, Kingston, Canada, October 2012.
11. **MSR 2012**, the 9th Working Conference on Mining Software Repositories, Zurich, Switzerland, June 2012.
10. **CASCON 2012**, 22nd IBM Centre for Advanced Studies International Conference on Computer Science and Software Engineering, Toronto, Canada, November 2012.
9. **ICSE 2012**, Poster and Informal Tool Demonstrations Track of the 34<sup>th</sup> International Conference on Software Engineering, Zurich, Switzerland, June 2012.
8. **ICSM 2011**, the 27<sup>th</sup> IEEE International Conference on Software Maintenance, Williamsburg, VA, USA, September 2011.
7. **SCAM 2011**, the 11th IEEE International Working Conference on Source Code Analysis and Manipulation, Williamsburg, VA, USA, September 2011.
6. **WCRE 2011**, the 18<sup>th</sup> Working Conference on Reverse Engineering, Lero, Limerick, Ireland, October 2011.
5. **IWDSC 2011**, the 5th International Workshop on the Detection of Software Clones, 2011, co-located with ICSE'11, Hawaii, USA, May 2011.
4. **WCRE 2010**, the 17th Working Conference on Reverse Engineering, Boston, USA, October 2010.
3. **ICSM 2010**, the 26th IEEE International Conference on Software Maintenance, Timișoara, Romania, September 2010.
2. **IWDSC 2010**, the 4th International Workshop on the Detection of Software Clones, 2010, co-located with ICSE'10, Cape Town, South Africa.
1. **IWDSC 2009**, the 3<sup>rd</sup> International Workshop on the Detection of Software Clones, 2009, co-located with CSMR 2009

**REVIEWER FOR JOURNALS**

Number of articles reviewed since 2009 for over 24 journals (avg. length 20-30 pages)	130+
---------------------------------------------------------------------------------------	------

24. Automated Software Engineering
23. IEEE Transactions on Reliability
22. Empirical Software Engineering
21. IEEE Transactions on Software Engineering
20. IEEE Access
19. ACM Transactions on Software Engineering and Methodology
18. ACM Computing Survey
17. Journal of Systems and Software
16. IET Software
15. Software: Practice and Experience
14. Journal of Engineering Research
13. Symmetry
12. Computer Science Review
11. Journal of Computer Science and Technology
10. Software Quality Journal
9. Journal of Software: Evolution and Process
8. British Journal of Mathematics & Computer Science
7. Journal of Software Maintenance and Evolution: Research and Practice
6. Science of Computer Programming
5. International Journal of Innovative Computing and Applications
4. Computers & Mathematics with Applications
3. Journal of Supercomputing
2. Higher-Order and Symbolic Computation
1. Information and Software Technology

**GRANTS, PHD EXTERNAL EXAMINER, AND TENURE & PROMOTION REVIEWS**

Grants reviewed from NSERC, Mitacs, and provincial grants since 2009	100+
PhD External Examiner	7
Tenure and Promotion reviews	10+

4. Committee Member, NSERC Discovery Horizons (DH) program, 2022-23
3. Section Chair and Member, NSERC Evaluation Group Member (1507) for Discovery Grants 2020 to 2021 (couldn't finish the term due to critical illness)
2. Acting Section Chair and Member, NSERC Evaluation Group Member (1507) for Discovery Grants, 2019 to 2020
1. NSERC Evaluation Group Member (1507) for Discovery Grants, 2018 to 2021

**OTHER ACTIVITIES**

I am the Director of the Industry Stream inter-University NSERC CREATE program on Software Analytics Research (SOAR), Co-Director of the Software Research Lab (SRLab) and have/had been Co-Leads of a couple of P2IRC projects. I closely worked with the Government of Saskatchewan in helping to make it easier for graduate students working in our province to obtain citizenship and help build Saskatchewan. I have been also working as a coordinator for a Memorandum of Understanding to promote collaboration between University of Saskatchewan and Khulna University, Bangladesh over the decade.

5. **Program Director, NSERC CREATE in Software Analytics Research (SOAR), September 2020 to August 2027**

SOAR (<https://soar.usask.ca>) is an industry-stream and inter-University graduate program with over 15 industry partners and four Universities (University of Saskatchewan, University of Calgary/York University, University of Alberta, and University of British Columbia). SOAR trainees analyze massive quantities of software and usage data to construct software that is more secure, sustainable, scalable, and cost effective. The project's industrial partners include Microsoft, Siemens, Nutrien, Calian, and Vendasta who significant cash contributions for internship stipends and offer diverse perspectives and opportunities for over 100 students and post-doctoral fellows. As Director, I am responsible for dealing with all the financial management of the students and post-doctoral fellows including managing all the related administrative activities. Furthermore, as Program Director, I need to ensure that all the SOAR HQP meet the program requirements. SOAR trainees need to gain both professional and technical skills of different types. Professional skills include Professional skills development programs available at partner universities, Industry/Mitacs workshops and Entrepreneurship training and technical skills include yearly symposium, SOAR foundation courses, industry workshops, and seminars featuring industry, academic or other prominent speakers. Mobility is another major part of the SOAR program where SOAR trainees participate both national and international mobility programs. I am responsible for overseeing all these program requirements of the SOAR CREATE program.

4. **Promoting research via media presence**, July 2018 to Present

I have also given a number of media interviews over the last few years. I was interviewed and then featured in Saskatoon's newspaper of record, the Star Phoenix (June 17, 2019). I was also interviewed by Prothom Alo (Chhutihir Diney, Feb 22, 2020), the largest daily in Bangladesh with 7.6 million daily readers (print + online), in both the Bangla and English versions of the news. I was also featured in Education News Canada (June 3, 2020) on our SOAR CREATE program. I have been interviewed multiple times by the Communication office at USask and featured multiple times in the campus newspaper. I was also interviewed by two local radios, Voice of Saskatoon and CJWW radio.

3. **Co-Lead, Data Management Portal and Repository**, NSERC Canada First Research Excellence Fund (CFREF), Plant Phenotyping and Imaging Research Centre (P2IRC), University of Saskatchewan, Global Institute for Food Security (GIFS), April 2019 to March 2023

I have been serving as the Co-Lead of the Data Management Portal and Repository platform of an NSERC CFREF on Plant Phenotyping and Imaging (<https://p2irc.usask.ca>) with \$902,520 funding and was responsible for hiring of students, postdocs, research associates and staffs and supervising them (total funding \$37M, 2016-2023, P.I. Maurice Moloney). The purpose of this theme has been to build a data management portal and repository integrated with a cloud-based data analytics system that we have been also building.

2. **Co-Lead, Theme 3.1 P2IRC Cloud: Big Data Analytics for Crop Phenomics**, NSERC Canada First Research Excellence Fund (CFREF), Plant Phenotyping and Imaging Research Centre (P2IRC), University of Saskatchewan, Global Institute for Food Security (GIFS), April 2016 to March 2019

I served as the Co-Lead of the Theme 3.1 P2IRC Cloud: Big Data Analytics for Crop Phenomics of an NSERC CFREF on Plant Phenotyping and Imaging (<https://p2irc.usask.ca>) with \$791,823 and was responsible for hiring of students, postdocs, research associates and staffs and supervising them (total funding \$37M, 2016-2023, P.I. Maurice Moloney). The purpose of this theme has been to build a cloud framework for conducting large scale data analytics in the cloud, especially with crop phenomics data.

3. **Canadian Citizenship for our Graduate Students**

Because of my initiative, our graduates (also University of Regina graduates) could apply for Canadian landed immigration status under Saskatchewan Immigrant Nominee Program (SINP) and can ultimately get Canadian citizenship. Previously, our graduates were not eligible to apply for SINP or Canadian immigration unless they had job offers or enough Canadian experience. I proposed a new immigration rule for our graduates to the then honourable minister, Rob Norris, Minister of Advanced Education, Employment and Immigration, and after a lot of discussion and revisions, on November 16, 2010, he announced the new "Master's and PhD Graduate Stream" to SINP, under which graduates from both the University of Saskatchewan and University of Regina were eligible to apply for Canadian immigration. We were able to attract high quality international students who might ultimately settle in Saskatchewan. We have already seen the results. Even within the Department of Computer Science, we have had a large pool of talented international students. This program has now been replaced with some other related programs.

1. **MOU for External Collaboration**

I have established collaboration between the University of Saskatchewan and Khulna University (KU) Bangladesh. A Memorandum of Understanding (MOU) has also been signed between the two Universities (between our Vice-President Research Dr. Karan Chad and KU Vice-Chancellor, Prof. Dr. Md. Saifuddin Shah) on March 13, 2011. Honourable Premier Brad Wall witnessed the signing ceremony during his visit to Bangladesh. The primary purpose of this MOU is to attract high quality students as we have been doing and promote collaboration. We have further established another MOU with Global Institute for Food Security (GIFS) and Khulna University for which I am the primary coordinator.

**MOST SIGNIFICANT CONTRIBUTIONS TO RESEARCH AND/OR PRACTICAL APPLICATIONS**

(References are from the CV itself above with journal articles indexed as J, conferences as C, and book chapters as B)

5. **Codifying software clone research.** Software bugs spread by developers who copy or "clone" segments of programming codes are costly, amounting to well in excess of a trillion dollars annually and even resulting in potentially fatal accidents in some applications [C132]. I have made significant contributions to advancing code clone research by reviewing, codifying, and documenting the range of research in the code clone community, and working to set the research on a firm foundation. A code clone is a pair of code fragments, within or between similar software systems. By 2006, software clones were thought to be a dead area of study. However, my 2007 review [T1] of clone detection techniques demonstrated that while important work had been done in the area, many more challenges remained, both in the detection and management of code clones.

Since 2008, that technical report [T1] has received 1,141 citations, the scenario-based comparison of the tools and techniques from it has received 94 citations [C9] while the extended journal version comparing the tools and techniques derived from this review has received 1,376 citations [J9] totalling over 2,600 citations on this review article. This has been one of the top two most cited papers in this field and has been the most cited paper on any topic in the journal, *Science of Computer Programming* for a number of years, since its publication in 2009. My work sparked renewed interest in clones in the software engineering community; hundreds of new papers have been published over the past decade on the topic, and the international workshop on software clones (IWSC) was resumed in 2009. I was appointed program co-chair in 2012 for the workshop, General Chair in 2015, and a keynote speaker in 2018. I have also been on the program committee every year since 2009 and am now a steering committee member of the workshop. In recognition of my expertise and contributions, I was recognized with the *Clones Lifetime Achievement Award* from IWSC in October 2023, becoming only the second person to receive this award. Given my expertise in software clones and their management, I was also invited as one of the two keynote vision speakers at the joint International Conferences of Software Maintenance, Reengineering, and Reverse Engineering, (CSMR-WCRE) 2014 [IP1] (now known as Software Analysis, Evolution and Reengineering, or SANER). This vision keynote not only demonstrates the recognition of my expertise in the area but shows that software cloning has been rekindled as an important topic in software maintenance and evolution. Cloning has become an established topic for all major software engineering conferences and journals, as well for industry; for example, Microsoft Corporation has published cloning papers and included clone detection and management features in Microsoft Visual Studio 2012

4. **Next generation scalable clone detection systems.** A challenge I identified was near-miss clone detection, that is clones that differ significantly in their syntactic constructs but still carry similarity in the context of use. Over the past several years, I have greatly advanced state-of-the-art research in this area. Using novel source transformation techniques, we showed that it is possible to build scalable near-miss clone (copy/pasted code fragments with significant changes in the pasted fragments) detection tools, such as NiCad [C10, J14, C23] (over 1,075 citations altogether, hundreds of downloads) with high precision and recall [C10, C12, C64, C76, J22, J41]. NiCad [A] has been widely used and cited in the community and has been instrumental in my research lab. We received the *10-year Most Influential Paper (MIP) award* from the 26th IEEE International Conference on Program Comprehension (ICPC 2018) for the NiCad paper that we published at ICPC 2008. We further aimed for large scale detection of near-miss clones, a problem that not only deals with the vague definition of clones but also its inherent timing complexity. My contributions to such large-scale clone detection have been highly influential as well. Our first attempt using Google's simhashing technology [B] received another *10-year Most Influential Paper (MIP) award* from the 28th International Conference on Software Analysis, Evolution, and Reengineering (SANER 2021). Using BigCloneBench (contribution 3 below and [C61, C96]) as the vehicle, I continue to push the field into new areas in large scale clone detection. With my HQP and collaborators, we built SourcererCC [C91] which has been also widely used and cited over 670 times, which we then also scaled further with CloneWorks [C105, C106] based on partitioned partial indexes technology. We also built large-gapped clone detectors, CCAAligner [C117] and LVMapper [J39] with our collaborators in China as well. Recently, we also published work on the scalable near-miss clone detection tool, SAGA that is capable of processing one billion lines of code. It represents the cutting edge of clone detection research [C140].

I also made novel contributions in finding cross-language similar software applications (CroLSim) [C121] and finding cross-language code clones (CLCDSA)[C136], contributions that appeared at ASE 2019, one of the top three conferences in general software engineering, and in the major *Journal of Systems and Software* [J34]. While there are novel approaches to detect and categorize similar software applications developed in the same programming language, we cannot detect similarity in applications written in different programming languages. Cross-language software similarity detection is inherently more challenging due to variations in language, application structures, support libraries used, and naming conventions. But we overcame these challenges by establishing relationships among the software libraries and Application Programming Interface (API) methods of those diverse systems with various programming languages. We also extended the approach for detecting cross-language clone detection, one of the open challenges in clone research, and an essential component for multilingual software management. I have pioneered conducting large scale empirical studies on clones in software evolution and their relation to different properties, such as clones and bugs [J24, C107, C126], and predicting, tracking, and refactoring of clones [C66, C71]. These studies, which pioneered our work on clone management, have been widely cited.

[A] C. K. Roy and J.R. Cordy, "NICAD: Accurate Detection of Near-Miss Intentional Clones Using Flexible Pretty-Printing and Code Normalization", *Proc. ICPC 2008*, June 2008, pp. 172-181. (*MIP award winner at ICPC 2018*)

[B] M.S. Uddin, and C. K. Roy, K.A. Schneider and A. Hindle, "On the Effectiveness of Simhash for Detecting Near-Miss Clones in Large Scale Software Systems", *Proc. of the 18th IEEE Working Conf. on Reverse Engineering (WCRE 2011)*, October 2011, pp. 13-22. (*MIP award winner at SANER 2021*)

3. **Tackling the clone oracling/benchmarking problem.** My contributions to large scale comparative evaluation and benchmarking of clone detection tools have been highly influential as well with a third *10-year Most Influential Paper*

**(MIP) award** from 25th International Conference on Software Analysis, Evolution, and Reengineering (SANER 2018) with the award citation, “*For influencing future studies in software cloning through the development of an extensive benchmark dataset*” for our seminal work on clone benchmark creation [C]. Benchmarks are essential to objectively evaluating and comparing code clone detection tools, but previous approaches all had severe weaknesses. To address this, I tackled the longstanding oracling problem (e.g., how to measure recall and precision) [Baker, TSE’07] in the cloning community in two innovative ways. First, we proposed a mutation-based framework that can automatically evaluate and compare the clone detection tools at a finer granularity of clone types by artificially generating thousands of different clones using a taxonomy of developers’ editing activities during software development [C12, C39, J41]. Second, we created a large-scale, real-world benchmark, BigCloneBench [C61, C96], and tools to use the benchmark. This benchmark now has more than 8.9 million manually validated clone pairs from a repository of 25,000 Java projects and 365 million lines of code. It’s the largest benchmark ever created in software clone detection, making it an ideal tool for evaluating Big Data clone detectors. This benchmark can also be used as a testbed for evaluating the performance of emerging Big Data frameworks in similar contexts. My former PhD student Dr. J. Svajlenko, whose PhD work included BigCloneBench, is now continuing his work in this context at Amazon Inc. Another co-author of the benchmark, Dr. I. Keivanloo also continues their work at Amazon Inc surrounding this context. Contributions from these works appeared at the major software engineering conferences and journals including IEEE Transactions on Software Engineering, the top journals in general software engineering including the **10-year Most Influential Paper (MIP) award** from SANER 2018.

Since BigCloneBench was released in 2014, it has been widely used for evaluating modern clone detection tools. Furthermore, we observe that its dataset has been used in training innovative machine-learning based approaches to semantic clone (Type-4) detection and there are over 380 citations of this benchmark [D] although it was not specifically designed for this purpose, and may not be optimal. To support the community, we thus built SemanticCloneBench [C142], which has semantic clones of four different languages. We exploited Stack Overflow questions and their answers and manual validation to build this benchmark. However, SemanticCloneBench does not have enough semantic clones for each programming language either, and it was not feasible to build a large benchmark. We thus introduced GPTCloneBench [C167] exploiting large language models where we have 37,149 true semantic clone pairs, 19,288 false semantic pairs, and 20,770 cross-language clones across four major programming languages. GPTCloneBench is a significant contribution to software engineering, especially in the realm of code clone detection and related problems in code similarity analysis. Its development not only demonstrates the possibilities of AI in software analysis but also sets a new standard in the field, encouraging further innovation and research.

We are also the first group that proposed a cloud-based machine learning framework for validating code clone detection tools using machine learning approaches [C135]. Since code clones often impact the maintainability of a software system, over 200 code clone detection techniques and tools have been proposed and studied over the last decade. However, these tools are not always perfect and often produce irrelevant results. Manually validating clones is highly time-consuming and error-prone, particularly for large-scale clone detection. Furthermore, one might only be interested in looking at code clones for a certain usage scenario (e.g., clones that may contain bugs). To deal with this pressing issue, we proposed a machine learning-based cloud framework for predicting user code clone validation patterns. The proposed method works on top of any code clone detection tools for classifying the reported clones as per user preferences. We published a conference paper [C120] and a journal [J37] including a demonstration at FSE [C135], one of the top three software engineering conferences.

[C] C. K. Roy and J.R. Cordy, “An Empirical Evaluation of Function Clones in Open-Source Software”, *Proc. WCRE 2008*, October 2008, pp. 81-90. **(MIP award winner at SANER 2018)**

2. **Recommendation systems in Software Engineering.** Another core aspect of my research program is to build recommendation systems to help software engineers in the development and maintenance of software that is sustainable, secure, reliable, scalable, efficient, evolvable, and cost-effective. Towards this goal, my group has made ground-breaking advances in several related areas such as context-sensitive code completion [J21, C110, C111], query reformulation for concept location [J28, C95, C112, C124], bug localization [J44, C128], recommending programming solutions [J40], code reviews and code reviewer recommendations [J55, C103 C168, ], code search [C102, C123], and many more. Contributions from these works appeared in the top conferences (ICSE, ASE, FSE, ICSME, SANER and so on) in Software Engineering, as well as in the top journals (e.g., TSE, JSS, and EMSE). Our work on a new way to search Stack Overflow (SO) (the CROKAGE tool) [J38] has been featured in SO blogs and subsequently featured in other tech news publications such as TechRepublic, SD Times, and ACM Tech News. Our work on IR-based bug localization [J44] is being adopted by a California based start-up, and a Mitacs Accelerate International has been awarded to one of my students with them. The reviewer recommendation tool, CORRECT [C90, C94] developed as part of an Engage grant with a local software service

company has also drawn attention from them, and they have committed for large Alliance grants. HQP working in these areas also won noteworthy awards, such as M. M. Rahman receiving the Governor General's Gold Medal (2019) and also the Doctoral Thesis Award in Physical and Engineering Sciences category at USask. Another recent MSc student, S. Mondal, also received the Graduate Thesis Award and GSA Research Excellence Award at USask. Several others have won a number of research awards at the department level, such as Research Excellence Awards and Geddes Awards.

1. **P2IRC Cloud: A cloud based collaborative framework for supporting data intensive discoveries for Plant Phenotyping and Genotyping.** As co-lead and Investigator of the Big Data Analytics group and Core Computer Science Team of two large Canada First Research Excellence Funds (CFREFs) projects on Food Security and Water Security respectively, I contributed significantly to designing and developing a cloud platform to support their large-scale data analytics needs. We developed scientific workflow-based frameworks such as VizSciFlow [J36] where multidisciplinary scientists could compose workflow pipelines of their analytical experiments using simple drag-and-drop programming elements and perform large-scale analysis tasks based on the composed workflows. We also published in several top-tier publications findings from this work, such as a reference architecture for big data platforms [C100] which could be used by others in designing their big data analytics frameworks, as well as a collaborative version of the analytical framework, SciWorkCS [J29], where multidisciplinary scientists could compose workflow pipelines for large scale analytics collaboratively, both in real-time or asynchronously. We developed a programming model for provenance management and visualization [B2, B1, C125] which provides support for reusability of the composed workflows, and improved monitoring of errors in composing the workflows [C125, B1]. Furthermore, this framework is now being used as a foundation for building P2IRC Data Management Portal and Repository (<http://scidatamanager.usask.ca>) to ensure the preservation, retention, and sharing of P2IRC data and metadata during the research project and beyond. To support this goal, we built a federated repository and a central public-facing portal, incorporating well-defined data and metadata formats, standards, and protocols based on best practices.