Welcome from the Chairs

Welcome to MSR 2013, the Tenth International Working Conference on Mining Software Repositories, held on May 18-19 in San Francisco, California, and co-located with the 35th ACM/IEEE International Conference on Software Engineering (ICSE 2013).

We are happy to report that the community of MSR researchers continues to grow, while at the same time the research being performed is maturing. This year marks the 10th edition of MSR. Since 2004 MSR has grown from a one-day workshop with 38 submissions to a two-day conference with 164 submissions combined across all submission categories. To account for the growing industrial interest in smart data and data science, the technical track introduced an additional submission category of practice papers, which report experiences of applying mining repository algorithms in an industry/open source organization context.

Compared to last year, we saw a further increase in the total number of submissions to the technical track: 86 full paper submissions (10 pages) and 22 short paper submissions (4 pages) for a total of 109. Full paper submissions increased by 22 (86 vs. 64) while short paper submissions remained constant with respect to MSR 2012.

Of the 86 full paper submissions, 31 were ultimately accepted, for a full paper acceptance rate of 38%. For short papers, of the 22 submissions 5 were accepted, for a short paper acceptance rate of 23%. Following last year’s conference, MSR 2013 accepted or rejected papers at their submitted length, and did not permit downgrading of a paper from full paper to short paper. Each paper received three reviews by members of the program committee. After we received all reviews, an online discussion process allowed variance among reviews to be understood, and in many cases, consensus to emerge. The program committee co-chairs then personally read all of the reviews and commentary prior to making final decisions.

It is important to note that we did not fix any quota on the number of papers to be accepted. For this reason, since the quality of papers was really high this year, MSR foresees a quite high number of contributions. In keeping with MSR’s roots as a working conference, each session includes significant time for discussion, while will keep the presentations short, and having the main purpose of soliciting the audience curiosity to read the papers, and to stimulate the forthcoming discussion. MSR 2013 foresees a wide variety of sessions on topics covering various research areas, including social mining, issue triaging and localization, text mining, bug and effort prediction. There will also be a retrospective session about the ten past years of MSR.

We are delighted to have two great keynote speakers this year.

Prof. Gail C. Murphy, from University of British Columbia, Canada, will present a talk entitled "What is Software Development Productivity, Anyway?". This talk will focus on the way software productivity has been characterized in software engineering and how software repository can support such a characterization.

Brian Doll, from GitHub, will present a talk entitled "Striking Gold: Building Software Better, Together". Brian will tell us how, by exploring GitHub, we can answer questions about the success of software projects and of software development teams.

For the first time, this year MSR will have a data showcase track. Rather than describing research achievements, data papers describe datasets curated by their authors and made available to others. Such papers provide description of the data, including its source; methodology used to gather it; description of
the schema used to store it, and any limitations and/or challenges of this data set. The data showcase track received 27 submissions, out of which 15 were accepted (56%).

We are pleased to continue the tradition of the MSR challenge track, where researchers from across the community apply their mining techniques to a common problem. This year the challenge is about StackOverflow, a very successful Questions and Answers web service. Being an important venue for sharing knowledge on software development, Stack Overflow has recently raised the interest of many researchers including those in the mining software repository area. This year the challenge provides participants with the latest official StackOverflow dump, with the aim of attracting the interest of more researchers on this still largely unexplored source of data. Submissions were invited to report their findings in a 4-page submission. Challenge track submissions almost doubled with respect to last year (29 vs. 15), showing an increasing interest on this track and the provided data. The program committee was composed of experienced Ph.D. students and postdocs. Each submission was reviewed by three PC members, and 12 (41%) challenge submissions were accepted and will be presented at the conference.

We are also delighted to announce that a selection of the best papers from this year’s conference will be invited to submit an extended version for consideration in a special issue of the Springer journal Empirical Software Engineering. The invitees will be announced during the opening session of MSR 2013.

Finally, we would like to thank the committee for the research, data and challenge track, and invited reviewers for their diligence in reviewing and discussing the submissions. The high quality of reviews is a hallmark of the Mining Software Repositories conference. We thank Microsoft Research for providing the prizes for the challenge. We thank Julius Davies for his speedy website and social media updates. We thank Dirk Beyer and his team at Conference Publishing Solutions for organizing the proceedings. We thank David Notkin and the ICSE 2013 organization team for their help with co-locating MSR with ICSE and local arrangements. We thank the MSR steering committee for their advice in organizing the conference. And finally, we thank the authors for their submissions and the MSR community for your enthusiasm.

Welcome to San Francisco, part of the California Gold Rush, 1849. Let’s mine some gold!
Enjoy the conference!

Thomas Zimmermann, Microsoft Research, USA
MSR 2013 General Chair

Massimiliano Di Penta, University of Sannio, Italy
Sunghun Kim, Hong Kong University of Science and Technology, China
MSR 2013 Program Co-Chairs

Daniel M. German, University of Victoria, Canada
MSR 2013 Chief of Data

Alberto Bacchelli, University of Lugano, Switzerland
MSR 2013 Challenge Chair
MSR 2013 Team

Organizing Committee Members
General Chair
Thomas Zimmermann, Microsoft Research, USA

Program Co-Chairs
Massimiliano Di Penta, University of Sannio, Italy
Sunghun Kim, Hong Kong University of Science and Technology, China

Chief of Data
Daniel M. German, University of Victoria, Canada

Challenge Chair
Alberto Bacchelli, University of Lugano, Switzerland

Web Chair
Julius Davies, University of British Columbia, Canada

Steering Committee Members
Ahmed E. Hassan, Queen’s University, Canada
Audris Mockus, Avaya Labs Research, USA
Richard Holt, University of Waterloo, Canada
Katsuro Inoue, University of Osaka, Japan
Sunghun Kim, Hong Kong University of Science and Technology, China
Stephan Diehl, University of Trier, Germany
Massimiliano Di Penta, University of Sannio, Italy
Harald Gall, University of Zurich, Switzerland
Michele Lanza, University of Lugano, Switzerland
Mike W. Godfrey, University of Waterloo, Canada
Jim Whitehead, University of California, Santa Cruz, USA
Tao Xie, North Carolina State University, USA
Thomas Zimmermann, Microsoft Research, USA
Program Committee Members (Research Track)

Mithun Acharya, ABB Research, USA
Bram Adams, École Polytechnique de Montréal, Canada
Giuliano Antoniol, École Polytechnique de Montréal, Canada
Christian Bird, Microsoft Research, USA
Raymond Buse, Google Inc., USA
Gerardo Canfora, University of Sannio, Italy
Andy Chou, Coverity Inc., USA
Prem Devanbu, University of California, Davis, USA
Stephan Diehl, University Trier, Germany
Vladimir Filkov, University of California, Davis, USA
Thomas Fritz, University of Zurich, Switzerland
Harald Gall, University of Zurich, Switzerland
Daniel German, University of Victoria, Canada
Yossi Gill, Technion, Israel
Mike W. Godfrey, University of Waterloo, Canada
Mark Harman, University College London, UK
Ahmed E. Hassan, Queen’s University, Canada
Israel Herraiz, Universidad Politécnica de Madrid, Spain
Emily Hill, Montclair State University, USA
Abram Hindle, University of Alberta, Canada
Reid Holmes, University of Waterloo, Canada
Katsuro Inoue, University of Osaka, Japan
Foutse Khomh, Queen’s University, Canada
Michele Lanza, University of Lugano, Switzerland
David Lo, Singapore Management University, Singapore
Audris Mockus, Avaya Labs Research, USA
David Morgenthaler, Google Inc., USA
Thomas Ostrand, Rutgers University, USA
John Penix, Google Inc., USA
Martin Pinzger, Alpen-Adria-Universität, Klagenfurt, Austria
Denys Poshlyvanyk, The College of William and Mary, USA
Romain Robbes, University of Chile, Chile
Peter Rotella, Cisco Systems, USA
Chanchal K. Roy, University of Saskatchewan, Canada
Emad Shihab, Rochester Institute of Technology, USA
Lin Tan, University of Waterloo, Canada
Suresh Thummalapenta, IBM Research, India
Jim Whitehead, University of California, Santa Cruz, USA
Laurie Williams, North Carolina State University, USA
Tao Xie, North Carolina State University, USA
Andreas Zeller, Saarland University, Germany
Dongmei Zhang, Microsoft Research, China
Hongyu Zhang, Tsinghua University, China
Program Committee Members (Data Showcase Track)
Julius Davies, University of British Columbia, Canada
Bogdan Dit, College of William and Mary, USA
Latifa Guerrouj, École Polytechnique de Montréal, Canada
Israel Herraz, Universidad Politécnica de Madrid, Spain
Abram Hindle, University of Alberta, Canada
Zhen Ming (Jack) Jiang, Queen’s University, Canada
Charles D. Knutson, Brigham Young University, USA
Jonathan L. Krein, Brigham Young University, USA
Peter Rigby, Concordia University, Canada
Gregorio Robles, Universidad Rey Juan Carlos, Spain
Emad Shihab, Rochester Institute of Technology, USA
Yasutaka Kamei, Kyushu University, Japan
Annie Ying, McGill University, Canada

Program Committee Members (Challenge Track)
Gabriele Bavota, University of Sannio, Italy
Olga Baysal, University of Waterloo, Canada
Oscar Callau, University of Chile, Chile
Julius Davies, University of British Columbia, Canada
Emanuel Giger, University of Zurich, Switzerland
Georgios Gousios, Delft University of Technology, Netherlands
Anja Guzzi, Delft University of Technology, Netherlands
Sonia Haiduc, Wayne State University, USA
Chris Parnin, Georgia Institute of Technology, USA
Shaun Phillips, University of Calgary, Canada
Sarah Rastkar, University of British Columbia, Canada
Stephen Thomas, Queen’s University, Canada
Christoph Treude, McGill University, Canada
Additional Reviewers

Venera Arnaoudova
Vinay Augustine
Fabian Beck
Nicolas Bettenburg
Benjamin Biegel
Andreas Bollin
Fenghong Chen
Peter Chen
Yingnong Dang
Patrick Francis
Shi Han
Laura Inozemtseva
Yujuan Jiang
Iman Keivanloo
Katja Keivanloo
Hammad Khalid
Jason King
Oleksii Kononenko
Mario Linares Vásquez
Jian-Guang Lou
Rainer Lutz
John Majikes
Andr Meyer
Ahmed Moamen
Manishankar Mondal
Patrick Morrison
Sebastian Mueller

Thanh Nguyen
Rahul Pandita
Daryl Posnett
Xiao Qu
Maria Riaz
Daniele Romano
Caitlin Sadowski
Raluka Sauciuc
Roger Scott
Kiran Shakya
John Slankas
Will Snipes
Zéphyrin Soh
Mark Syer
Parastou Tourani
Ruowen Wang
Wei Wang
Jochen Wuttke
Xusheng Xiao
Jialiang Xie
Qi Xuan
Wei Yang
Yuan Yao
Shamima Yeasmin
Xin Yi
Minghui Zhou
Lijie Zou