MSR 2017 General Chair Report

Jesus M. Gonzalez-Barahona

Buenos Aires (Argentina), May 20-21 2017

Abstract

This document intends to report on the preparation and celebration of the International Conference on Mining Software Repositories 2017. Its intention is twofold: to serve as a summary of the activities related to the conference, and the conference itself, and to be useful for future organizers, so that they may be aware of some details that could be interesting to know.

Contents

Chapter 1

Celebration

MSR 2017 (International Conference on Mining Software Repositories) was held in Buenos Aires, Argentina, during May 20th and 21st 2017. The Call for Papers received 210 submissions in all the categories (research papers, mining challenge, data showcase), of which 65 were accepted for presentation. For review, several committees collaborated: 69 persons in the Program Committee, 19 persons in the Mining Challenge Committee, 14 persons in the Data Showcase Committee. The review process for the research track was double blind (authors of papers were not known to reviewers during the review process, and reviewers were not known to authors). This year, 143 persons registered for attending the conference, of which 139 attended the conference.

MSR 2017 has a program composed of 8 sessions, of which five were parallel (in two different rooms), and three were plenary. The conference started with an opening session with an introduction, the presentation of the MSR 2017 Most Influential Paper award, and the first general discussion session. During the rest of the first day, we had two parallel sessions for presenting research papers, and one plenary session for presenting Mining Challenge papers. During the second day we started with a plenary session which included the keynote, by Diomidis Spinellis, and the second discussion session. Then we had two parallel sessions for presenting research papers and the contributions to the Data Showcase. The conference finished with a plenary with included the awards ceremony, the third discussion session, and the presentation of MSR 2018.

1.1 Contributions

All papers were presented in place, except for one, which was presented with a video. For this, authors provided a reasonable rationale for not being able of presenting in place.

1.2 Registration

A total of 143 persons registered for MSR. Of those, 4 didn't show at the conference. The details of the registration were as follows:

- Registered as IEEE or ACM members: 69

- Registered as non IEEE or ACEM members: 11

- Registered as Argentina nationals: 1

- Registered as students, non IEEE or ACM members: 16

- Registered as students, IEEE or ACM members: 46

1.3 Venue

MSR took place in Auditorios Puerto Madero of Pontificia Universidad Catolica Argentina, in Buenos Aires. More specifically at the Building San Jose - Auditoriums UCA (Edificio San José - Auditorios UCA), Av. Alicia M. de Justo 1600. We used rooms "Auditorio 1" (for plenaries and some parallel sessions) and "Cine" (for some parallel sessions).

1.4 Sessions

First day (Saturday May 20th):

9:00-10:30 Plenary:

- Opening (15 min.)
- 2007 Most Influential Paper. Announcement of the awarded paper, and presentation by Tom Zimmerman, one of its authors. (30 min. including Q&A).
- Discussion 1 (45 min.)

11:00-12:30 Research presentations (two parallel sessions)

14:00-15:30 Research presentations (two parallel sessions)

16:00-17:30 Plenary: Data Challenge presentations

19:00-21:00 MSR Dinner

Second day (Sunday May 21th):

9:00-10:30 Plenary:

- Keynote (55 min. including Q&A).
- Discussion 1 (35 min.)

11:00-12:30 Research presentations (two parallel sessions)

14:00-15:30 Data Showcase presentations and Research presentations (two parallel sessions)

16:00-17:30 Plenary:

- Awards Ceremony (15 min).
- Foundational Contributions Presentation (40 min. including Q&A)
- Discussion 3 (30 min.)
- MSR 2018 Presentation (3 min.)
- Closing (2 min.)

18:00-21:00 Post-MSR Social Event

1.4.1 Awards

The following awards were announced during the conference:

- MSR 2007 Most Influential Paper

  Awarded: Cathrin Weiß, Rahul Premraj, Thomas Zimmermann, Andreas Zeller for "How Long Will It Take to Fix This Bug?"

- MSR Foundational Contribution Award

  Awarded: Tim Menzies

- MSR Early Career Achievement Award

  Awarded: Abram Hindle

- SIGSOFT Distinguished Papers

  Awarded: Luca Pascarella and Alberto Bacchelli, for "Classifying code comments in Java open-source software systems"

  Awarded: Mohammad Gharehyazie, Baishakhi Ray, and Vladimir Filkov, for "Some From Here, Some From There: Cross-Project Code Reuse in GitHub"

- MSR 2017 Mining Challenge Best Paper Award

  Awarded: Marcel Rebouças, Renato Oliveira Dos Santos, Gustavo Pinto, and Fernando Castor, for "Best Mining Challenge Paper,How Does Contributors' Involvement Influence the Build Status of an Open-Source Software Project?"

- MSR 2017 Mining Challenge Best Presentation Award

  Awarded: Rodrigo Souza and Bruno Silva, for "Sentiment analysis of Travis CI builds"

- MSR 2017 Data Showcase Best Contribution Award

  Awarded: Jeroen F.H. Noten, Josh G.M. Mengerink and Alexander Serebrenik, for "A Data Set of OCL Expressions on GitHub"

1.4.2 Keynote

The keynote was: "Half a Century of of Unix Hist Preservation, and Lessons Lear", by Diomidis Spinellis

Diomidis Spinellis is professor in the Department of Management Science and Technology of the Athens University of Economics and Business. He serves as Editor in Chief of IEEE Software, director of the Information Systems Technology Laboratory (ISTLab) and its Software Engineering and Security group (SENSE). Diomidis is author of most than 40 papers with more than 40 citations each (according to Google Scholar), in several fields related to computer science and engineering. He is also author and editor of several well known books, and software programs. More information about him can be obtained from his web site[1].

---

[1] http://www.spinellis.gr

1.4.3   Discussions

The program included three general discussions, intended to be an open forum on topics of interest for the MSR community. They were scheduled in plenary sessions, so that all attendees could participate. All discussions had a moderator, and usually a short introducing presentation.

- Discussion 1: "Research Infrastructure for MSR: Needs, Wants, and Where we Are". Moderator: Lori Pollock.

- Discussion 2: "Double-blind reviewing: lessons learned and paths ahead". Presentation by Lin Tan and Abram Hindle. Moderator: Tom Zimmerman

- Discussion 3: "Go deeper, not wider". Moderator: Mike Godfrey.

1.4.4   Research Sessions

We had the following Research Sessions (in parenthesis, session chairs):

- Mobile (Gregorio Robles): 6 papers

- Dependencies (Shane McIntosh): 6 papers

- Modelling and prediction (Yasutaka Kamei): 6 papers

- NLP and code review (Christian Bird): 7 papers

- Clones and edits (Chanchal Roy): 6 papers

- Continuous integration and build (Diomidis Spinellis): 6 papers

- Testing and bugs (John Wittern): 6 papers

1.4.5   Mining Challenge

The mining challenge started with the open call for datasets (the "Open Challenge"), which received four contributions. Of these, the Program Chairs, in coordination with the General Chair, selected the dataset that was to be used in the challenge: "TravisTorrent", submitted by Moritz Beller, Georgios Gousios and Andy Zaidman.

The authors of the dataset became the Mining Challenge Chairs, and selected the Mining Challenge Committee, that reviewed the 29 contributions received. Of these, they selected 14 for presentation during a plenary session in MSR.

The Mining Challenge Chairs also awarded the best Mining Challenge paper and the best Mining Challenge presentation.

1.4.6   Data Showcase

The Data Showcase Chair selected the Data Showcase Committee, which reviewed the 22 contributions received. Of these, they selected 9 for presentation during an specific parallel session in MSR.

The Data Showcase Chair also awarded the best Data Showcase contribution.

1.4.7 Unofficial BoF

This year MSR experimented with allowing for unofficial BoFs (Birds of a Feather Sessions) at the end of the regular sessions, on Saturday. One BoF was proposed, on "Tools for Mining Software Repositories". About 30 people attended to it. Some tools were presented, and a lively discussion happened on the matter.

1.5 Social events

We had two social events:

- MSR Dinner, during the evening of Saturday. It was at El Mirasol at Puerto Madero, in Av. Alicia Moreau de Justo 202, Buenos Aires.

- Post-MSR Event, held at a Buller Pub Downtown, in Calle Paraguay 428.

  Both were attended by most MSR attendees.

1.6 Steering Committee Meeting

During lunch time, on Sunday, we had the annual meeting of the MSR Steering Committee with the following agenda:

- MSR 2017 feedback

- Double-blind discussion

- MSR 2018 update

- MSR 2019

- Mining challenge

- Mining awards

- CORE ranking

- Publication options

- Relationship with the open source community

- Any other issue.

1.7 Website

The MSR maintained the http://2017.msrconf.org/ website, with the same structure, and templates than the websites of past editions. The website was updated continuously during all the year. It included information about:

- Call for contributions

- Program and schedule

- Organizing Team, including Organizing Committee, Program Committee, and Influential Paper Committee

- MSR Awards Section, including Awards Committee

- Mining Challenge Section, including call for contributions and Mining Challenge Committee

- Data Showcase Section, including call for contributions and Data Showcase Committee

The program included links to all the papers that were openly published as preprints.

1.8 Social media and publicity

Social media activity focused on Twitter and Facebook. The Publicity Chair ensured that all main milestones previous to the conference (call for papers, deadlines, publication of the program, availability of preprints, etc.) were broadcast. She also engaged in conversations with interested persons. All activities during the conference, including pictures of all presentations, and links to all papers with available preprints were also broadcast. The hashtag used was was "MSR17".

1.9 Budget

Most of the budget is covered by ICSE. That includes getting the payment from all registered persons, and carrying with all the costs, including the venue, proceedings, and the conference dinner.

The agreement with them had the following budgetary implications for MSR:

- ICSE invoices all people registering for MSR, according to the ICSE prices for co-located events.

- ICSE acts as financial representative for MSR, taking charge of all accounting and reimbursements.

- ICSE provides free of charge to MSR all venue facilities as a co-located event, including room space, supporting personnel and volunteers.

- ICSE provides free registration for the keynote speaker, the PC co-chairs and the general chair.

- ICSE pays for the MSR dinner for all registered people, and supports MSR in the selection of the venue for it.

- ICSE provides USD 500 for MSR expenses.

However, there is some amount that still had to be managed by the MSR organization. The details of this amount are as follows.

Income:

Item	Amount (USD)
Contribution by ICSE	500
Sponsorship by Microsoft	2,500
Total	3,000

Expenses:

Item	Amount (ARS)	Amount (USD)
Gadgets for keynote, awards and recognitions	1,750	109.50
Gadgets for attendees (1)	6.120	382.97
Gadgets for attendees (2)		1,578
Certificates		261.10
Paper copies (program)	800	50.06

In addition, source{d} contributed in kind, sponsoring the Post–MSR Social Event, on the evening of the second day.

Chapter 2

Preparation

When I received the invitation to chair MSR 2017, I had little idea of what a MSR General Chair does. With the awesome help of the Chair of the MSR Steering Committee (Tom Zimmerman), the MSR 2016 General Chair (Miryung Kim), the MSR Steering Committee at large, and of course the MSR 2017 Organizing Committee, I was learning over time, in which resulting to be a amazing process. I enjoined it, but at some points I also decided that maybe the transfer of knowledge could be improved a bit if some notes on the preparation of MSR were available. This is my attempt to start such notes.

2.1 Negotiation with ICSE

The negotiation with ICSE was done in collaboration between the General Chair and the Chair of the Steering Committee, mainly with Domenico Bianculli, the Co-located Events Chair of ICSE 2017, via email. First exchanges of messages date from October 2015, but most of the negotiation took place during April-July 2016. The contract was signed on July 2016.

Some specific details that were discussed with ICSE:

Deadlines

The call for papers for MSR specified some deadlines which a bit later than those for most of the other ICSE co-located events. We had negotiated in advance about those dates, checking with ICSE during the negotiation conversations the latest deadlines possible with the publisher of the proceedings, which in this case were:

- March 15th 2017: Proceedings chairs send to the Publisher the list of accepted papers, authors and emails

- March 19th 2017: Proceedings chairs send to the Publisher the table of contents, the message of the chairs, and the steering and program committees

- March 30th 2017: Camera Ready of colocated events.

These deadlines, which constrain the MSR deadlines, were known in early January 2016.

Space requirements

Space requirements (number and size of rooms needed for MSR) were settled during May 2016. To do that, it is important to know the degree of parallelism (how many sessions will be happening in parallel), and to have an estimation of people attending.

We needed to reserve a room for the Steering Committee meeting as well.

Submissions by chairs

ICSE does not allow the Program Chairs or the General Chair to submit papers to their own conference. This also affects their students.

Free registrations

ICSE allowed for free registration for Program Chairs, General Chair, and Keynote Speaker. The rules for how to register in these cases were known some days in advance of the early registration deadline.

Timetable

ICSE informed us of the timetable, that was common to all co-located events, during February 2017. This timetable was needed to produce the MSR schedule.

Signs / banners

ICSE was in charge of producing signs to reach the MSR rooms, and banners in the rooms. For that, they asked for appropriate logos during March 2017.

Registration

ICSE took charge of all activities related to the registration of the attendees, according to the ICSE procedures for co-located events. ICSE produced, when requested, numbers and lists of people registered. The first numbers came in late March.

No-shows policy

ICSE has the policy of removing from the Proceedings those papers that were not presented at the conference. Program Chairs were requested to provide a list of such papers, if any.

2.2 Organizing Committee

2.3 Website

The website was maintained during all the year, since the termination of MSR 2016, by the Web Chair, with the help mainly of the Program Chairs. During May-July 2016, the website for the new edition was built from the templates of past years. Major events in its maintenance were:

- June 30th 2016: first complete version of the website was up.

- July 11th 2016: published the Open Call for Mining Challenge Proposals

- August 31th: published the Call for Papers for the Mining Challenge

-

2.4 Social media and publicity

Activity in social media (mainly Facebook and Twitter) was maintained during all the year before the conference, and specially during the conference itself by the Publicity Chair. Previous to the conference, these were the main activities before the conference:

- Presentation of the persons in the organizing committee

- Announcement and reminders of deadlines

- Invitations to submit and to register

- Results of the review processes

- Publication of all papers accepted (title and authors), including preprint when available.

- Travel tips

During the conference:

- Picture, title and authors of all presentations

- Local tips

- Last-minute information (dinner venue, events happening during the conference)

Social media helped to collect the list of links to preprints, and showed to be a very efficient way of doing this.

2.5 Call for papers

The call for papers was produced by the Program Chairs. The main dates (for all calls and related actions) were:

- Call for mining challenge proposals: July 20th

- Mining challenge proposals deadline: August 1st

- Mining challenge proposals notifications sent: August 6th

- Call for papers: December 1st

- Abstracts due: February 3rd

- Papers due: February 10th

- Notifications sent: March 15th

- Camera ready: March 30th

- Call for nominations of distinguished papers and invites to the journal special issue: April 26th

- Call for session chairs: May 2nd

2.6 Mining Challenge

The Mining Challenge has its own open call, which was produced by the Program Chairs. The main dates related to the Mining Challenge were:

- Open Call for Datasets for the Mining Challenge: July 11th 2017. Important notice: this call could not be made public through the SEWORLD mailing list because they only accept one call for contributions message per conference.

- Challenge PC Formed: August 31st 2016

- Challenge Call for Contributions: August 31st 2016

- Challenge Dataset made available: August 31st, 2016

- Challenge Papers Deadline: Feb 20th, 2017

- Challenge notifications sent: March 10th

- Challenge camera ready: March 30th

Almost all tasks related to the Mining Challenge, including the selection of the Mining Challenge Committee, were carried on by their chairs, once they had been selected. Before that, those tasks were dealt with mainly by the Program Chairs.

The Mining Challenge Chair coordinated with the Proceedings Chair the details related to the production of this part of the MSR Proceedings.

The Mining Challenge Chair chaired the Mining Challenge session during the conference.

2.7 Data Showcase

Most of the tasks related to the Data Showcase, including the selection of the Data Showcase Committee, were carried on by its chair. The deadlines of the showcase were:

- Data showcase papers due: February 20th

- Data showcase notifications sent: March 8th

- Data showcase camera ready: March 17th

The Data Showcase Chair coordinated with the Proceedings Chair the details related to the production of this part of the MSR Proceedings.

The Data Showcase Chair chaired the Data Showcase session during the conference.

2.8 Keynote

The MSR keynote presentation was "Half Century of Unix: History, Preservation, and Lessons Learned", delivered by Diomidis Spinellis.

2.9 Review process

The review process was conducted completely in Easychair, with no in-person meeting. All papers were reviewed following a double-blind process (for the first time in MSR). All authors who submitted an abstract were sent a reminder about this new process before the full paper submission deadline. PC Chairs checked the submitted PDF files to identify potential violations, and gave the relevant authors time to remove their identities from their PDF files before paper bidding started.

The review process started with a bidding period, followed by the reviews, and then a discussion period for deciding the final accepted papers. Following tradition, papers were accepted or rejected, not permitting downgrading a paper from a full paper to a short paper. 158 submissions of research papers (121 full and 37 short) were received. Out of the 121 full paper submissions, 37 papers were accepted, and out of the 37 short paper submissions, six were accepted. These are acceptance rates of 30.6% for full papers and 16.2% for short papers. Of the accepted papers, two were submitted in the practice experience category, four in the reusable tools category, and four additional papers in both. Authors of accepted papers come from 23 different countries.

2.10 Preprints

Authors were asked to send links to the preprints of their papers, when they published such preprints, to let people attending have the papers in advance. Preprints for 40 papers were received, and linked from the program in the website.

2.11 Proceedings

ICSE required the nomination of a Proceedings Chair, to deal with everything related to the MSR Proceedings. One of the Program Chairs, Abram Hindle, volunteered to do the job.

The work of the Proceedings Chair started in March 2017, acknowledging the deadlines that were defined in the contract, and being introduced to the ICSE Proceedings Chair, Juan Pablo Galeotti.

2.12 Schedule

Once the review process was completed, the program chairs prepared a draft of a program, with papers organized in sessions, and all the other activities as well. After some iterations, this program became final. Some key points in those iterations were: duration of the talks, degree of parallelism, special needs of some speakers, avoidance of parallel presentations by same presenters, room for all non-presentation activities, relevance of the presentations selected as a part of the Mining Challenge and the Data Showcase.

Once the program was ready, some coordination was needed to:

- Update the website with all the information.

- Ensure that all people participating in the non-presentation activities (keynote, awards, etc.) knew when they were expected to participate.

- Broadcast about it via social media.

- Produce the version of the program in paper.

Once ICSE defined the rooms that were allocated for MSR, the program and schedule were annotated with them.

2.13 Sponsoring

This year Microsoft sponsored the MSR with USD 2,500. The sponsorship was negotiated by the Chair of the Steering Committee, and was announced during May 2017. The financial details of the sponsorship were dealt with by ICSE.

Source{d} sponsored in kind, taking charge of organizing the post-MSR social event.

2.14 Social events

We had two social events:

- The MSR dinner, during the evening of Saturday. This event was included in the agreement with ICSE. They took charge of finding a convenient venue, after some discussion with the General Chair about the preferences of the conference (location, kind of venue, etc.). ICSE also arranged all details with the venue, and provided a volunteer that helped with the organization. The MSR dinner was at El Mirasol at Puerto Madero, Av. Alicia Moreau de Justo 202, Buenos Aires.

- The post-MSR event, held at a pub nearby. The sponsoring company took charge of organizing everything, so the only activity needed from the organizing committee was coordination with them and ensuring that the event was properly announced.

2.15 Gadgets

Some gadgets were prepared as souvenirs for people attending the conference:

- Stickers and pins, ordered by the Chair of the Steering Committee. These have become traditional over the years.

- Mates with the MSR logo, ordered by the General Chair.

2.16 Other issues before the conference

- Call for session chairs. Some weeks before the conference, program chairs assembled the team of session chairs for all the sessions.

- Certificates. Certificates for all awards were prepared by the general chair. Some of them would be printed during the conference, because the corresponding names were not known (this was the case of the best presentation award for the Mining Challenge). Certificates for all the members of the organizing committee were produced by the Chair of the Steering Committee.

- Program in paper. The Publicity Chair prepared the sheets that composed the program that was later delivered in paper to attendees.

- Message to all people attending the conference, with practical information.

2.17 During the conference

The days immediately before the conference, and the conference itself, are very intense. During these days, the organizing team needs to track last-minute changes, in some cases solving upcoming issues, coordinate about them with affected parties, and communicate to the MSR community when needed. Some examples of issues that arose during these days (some predicted, some unpredictable) are: final updates to the website, communication in social media, production of the certificates, authors that could not come and prepared a video for their presentation, late deliver of some swags to attendees, lack of power plugs in the rooms, problems with the projector in one room, final arrangements to the presentation by the organizing committee, changes in place and time of the post-MSR social event, last-minute changes of the paper version of the program, etc.

However, preparation was in general good, and most issues were already planned or sorted out thanks to the dedication of the members of the organizing committee.

The most relevant tasks that the team dealt with were:

- Meeting with the local organization for last minute issues, and familiarization with the ways to contact them when needed.

- Visit to the venue where MSR took place, the day before, to discuss any last-minute issue, and to check for every possible detail.

- Production of the paper version of the program for attendees.

- Buying locally some small swags for awarded people, for the keynote speaker.

- Retrieving gadgets for attendees, that were ordered to be delivered in place, or bringing them to the venue, if they were acquired in advance.

- Preparation of the presentation of MSR, jointly by the program co-chairs and the general chair, to be delivered during the inaugural session.

- Support to session chairs, in case it was needed.

- Preparation of the closing session, which this year included the awards session.

- Broadcasting in social media the sessions, including references to all papers presented, links to open access papers, practical advice, pictures of the event, etc.

- Twitter contest, which mainly consisted in keeping track of retweets and favorites.

- Steering Committee meeting was organized by the Chair of the Steering Committee, with all the logistics being provided by ICSE.

2.18 After the conference

After the conference, some tasks were still performed:

- Preparation of the conference report.

- Reimbursement of all expenses.

- Help to the next Organizing Committee.

2.19 Responsibilities

During the preparation and execution of the conference, we didn't have a very detailed split of responsibilities, but the coordination worked, in my opinion, very well. From my point of view, and in general terms, the responsibilities were as follows:

- Program co-Chairs: Everything related to the production of the final research program, and how it was presented. This included (the list is by no means exhaustive): selecting the Program Committee, preparing the call for papers, configuring EasyChair, communicating with authors, coordinating the review process (including ensuring that double blind rules were preserved), announcing the results of the review process for the main track, coordinating with the chairs of the Mining Challenge and Data Showcase about their calls, review processes and announcement of results, coordination with the Publicity Chair and Web

Chair about all aspects related to the above, selecting chairs for the different sessions of the conference, selection of the discussion topics. One of the program co-chairs also took charge of the position of proceedings chair, which is used by ICSE as the contact point for everything related to the production of the proceedings.

- Chair of the Steering Committee: General advice on any aspects that the rest of the committee asks, and negotiations with ICSE. One of the key aspects of the Chair of the Steering Committee is to pass to a new organizing team the contacts and some of the knowledge about the organization of past editions. This includes providing information and advice about how specific issues were handled by past organizing committees, and send some heads-up when deadlines are approaching, or in general when important stuff is missed.

- Steering Committee: Selection of program chairs, general chair, and support and opinions on ideas proposed about new activities, or changes to traditional activities. They were also proactive, proposing ideas about changes or improvements.

- General Chair: Overall responsibility on all aspects of the conference not covered by somebody else. This includes the participation with the chair of the Steering Committee on the negotiations with ICSE; being available for comments and feedback on all the activities related to the selection and working of the Program Committee; being available for some specific committees, such as that selecting the Mining Challenge dataset and chairs, or the MSR Awards Committee; selection, in collaboration with the program co-chairs, of the remaining members of the organizing committee; preparation of certificates and small swags to awardees and guest speakers; coordination with ICSE before and during the conference, including the ICSE local team, etc.

Some responsibilities were taken together by the Program co-Chairs and the General Chair, usually with the advice of the Chair of the Steering Committee (the list is not exhaustive): selection of the Organizing Committee, dates for calls, management of the Open Mining Challenge, selection of keynote speaker, structure and schedule of the conference sessions.

All the team was very proactive, and helped as much as possible in issues outside their direct area of responsibility. In particular, the Program co-Chairs were in constant communication with the General Chair, and consulted with me many of their decisions.

Appendix A

Organzing team

A.1   Organizing committee

- General Chair: Jesus M. Gonzalez-Barahona, Universidad Rey Juan Carlos

- Program Co-Chairs:

  - Lin Tan, University of Waterloo
  - Abram Hindle, University of Alberta

- Data Showcase Chair: Megan Squire, Elon University

- Mining Challenge Co-Chairs:

  - Moritz Beller, TU Delft
  - Georgios Gousios, TU Delft
  - Andy Zaidman, TU Delft

- Web Chair: Thibaud Lutellier, University of Waterloo

- Publicity Chair: Gema Rodríguez-Pérez, Universidad Rey Juan Carlos

- Most Influential Paper Award Co-Chairs:

  - Michele Lanza, Università della Svizzera italiana - USI
  - Harald Gall, University of Zurich

- Free, Open Source Software Chair: Stefano Zacchiroli, University Paris Diderot and Inria

- MSR Awards Co-Chairs:

  - Massimiliano Di Penta, University of Sannio
  - Ahmed E. Hassan, Queen's University

A.2   Steering Committee

- Christian Bird, Microsoft Research
- Prem Devanbu, University of California Davis
- Massimiliano Di Penta, University of Sannio
- Jesus M. Gonzalez-Barahona, Universidad Rey Juan Carlos
- Ahmed E. Hassan, Queen's University
- Emily Hill, Drew University
- Abram Hindle, University of Alberta
- Yasutaka Kamei, Kyushu University
- Miryung Kim, University of California Los Angeles
- Sunghun Kim, Hong Kong University of Science and Technology
- Gail Murphy, University of British Columbia
- Martin Pinzger, University of Klagenfurt
- Romain Robbes, University of Chile
- Lin Tan, University of Waterloo
- Andy Zaidman, TU Delft
- Andreas Zeller, Saarland University
- Thomas Zimmermann, Microsoft Research

A.3   Program committee

- Bram Adams, MCIS, Polytechnique Montréal
- Giuliano Antoniol, Ecole Polytechnique de Montréal
- Earl Barr, University College London
- Gabriele Bavota, Università della Svizzera italiana (USI)
- Christian Bird, Microsoft Research
- Raymond Buse, Google
- Hoa Khanh Dam, University of Wollongong
- Prem Devanbu, UC Davis
- Massimiliano Di Penta, University of Sannio
- Robert Dyer, Bowling Green State University

- Fernando Figueira Filho, Federal University of Rio Grande do Norte
- Vladimir Filkov, University of California Davis
- Thomas Fritz, University of Zurich
- Daniel German, University of Victoria
- Ahmed E. Hassan, Queen's University
- Felienne Hermans, Delft University of Technology
- Israel Herraiz, Amadeus IT Group
- Yoshiki Higo, Osaka University
- Emily Hill, Drew University
- Reid Holmes, University of British Columbia
- Akinori Ihara, Nara Institute of Science and Technology
- Yasutaka Kamei, Kyushu University
- Foutse Khomh, DGIGL, École Polytechnique de Montréal
- Sunghun Kim, The University of Hong Kong Science and Technology
- Nicholas A. Kraft, ABB Corporate Research
- Jens Krinke, University College London
- Lucas Layman, Fraunhofer CESE
- Alina Lazar, Youngstown State University
- Zhongpeng Lin, Microsoft
- Xuanzhe Liu, Peking University
- David Lo, Singapore Management University
- Walid Maalej, University of Hamburg
- Shane Mcintosh, McGill University
- Tim Menzies, North Carolina State University
- Audris Mockus, The University of Tennessee
- Sarah Nadi, University of Alberta
- Meiyappan Nagappan, University of Waterloo
- Jaechang Nam, University of Waterloo
- Hoan Nguyen, Iowa State University
- Tien Nguyen, University of Texas at Dallas

- Rocco Oliveto, STAKE Lab - University of Molise
- Chris Parnin, North Carolina State University
- Dewayne E Perry, The University of Texas at Austin
- Lori Pollock, University of Delaware
- Denys Poshyvanyk, College of William and Mary
- Baishakhi Ray, University of Virginia
- Peter Rigby, Concordia University
- Romain Robbes, University of Chile
- Gregorio Robles, Universidad Rey Juan Carlos
- Chanchal K. Roy, University of Saskatchewan
- Barbara Russo, Free University of Bolzano/Bozen
- Anita Sarma, Oregon State University
- Bonita Sharif, Youngstown State University
- Vibha Sinha, IBM Watson
- Diomidis Spinellis, Athens University of Economics and Business
- Megan Squire, Elon University
- Bogdan Vasilescu, Carnegie Mellon University
- Patrick Wagstrom, Capital One
- Jim Whitehead, University of California, Santa Cruz
- Laurie Williams, North Carolina State University
- Erik Wittern, IBM Research
- Xin Xia, University of British Columbia
- Annie T.T. Ying, IBM Watson
- Stefano Zacchiroli, University Paris Diderot and Inria
- Andy Zaidman, TU Delft
- Andreas Zeller, Saarland University
- Hongyu Zhang, The University of Newcastle
- Minghui Zhou, Peking University
- Ying Zou, Queen's University

A.4 Mining challenge program committee

- Sven Aman, TU Darmstadt
- Maurício Aniche, Delft University of Technology
- Titus Barik, North Carolina State University
- Jonathan Bell, Columbia University
- Stefanie Beyer, University of Klagenfurt
- Cor-Paul Bezemer, Queen's University
- Caius Brindescu, Oregon State University
- Jürgen Cito, University of Zurich
- Roberta De Souza Coelho, PUC-Rio
- Joseph Hejderup, Delft University of Technology
- Michael Hilton, Oregon State University
- Maria Kechagia, Athens University of Economics and Business
- Anna Nagy, Travis CI
- Sebastian Proksch, TU Darmstadt
- Ayushi Rastogi, Indraprastha Institute of Information Technology, Delhi
- Mohammed Sayagh, Ecole Polytechnique Montreal
- Gerald Schermann, University of Zurich
- Patanamon Thongtanunam, Nara Institute of Science and Technology
- Bogdan Vasilescu, Carnegie Mellon University
- Yue Yu, National University of Defense Technology

A.5 Data showcase program committee

- Giuseppe Destefanis, Brunel University
- Neil Ernst, Software Engineering Institute
- Daniel Izquierdo-Cortazar, Bitergia
- Zhen Ming Jack Jiang, York University
- Raula Gaikovina Kula, Osaka University
- Senthil Mani, IBM Research
- Tien Nguyen, University of Texas at Dallas

- Hoan Nguyen, Iowa State University

- Marco Ortu, University of Cagliari

- Fabio Palomba, Delft University of Technology

- Germán Poo-Caamaño, University of Victoria

- Sebastian Proksch, TU Darmstadt

- Gregorio Robles, Universidad Rey Juan Carlos

- Norihiro Yoshida, Nagoya University

A.6   Additional reviewers

- Akond Rahman

- Alaaeddin Swidan

- Andre Meyer

- Boyang Li

- Bushra Aloraini

- Carlos Gabriel Gavidia Calderon

- Chaiyong Ragkhitwetsagul

- Christopher Theisen

- Christopher Vendome

- Cody Watson

- Cor-Paul Bezemer

- DongGyun Han

- Ehsan Noei

- Eirini Kalliamvakou

- Fang-Hsiang Su

- Feng Zhang

- Ferdian Thung

- Fiorella Zampetti

- Katja Kevic

- Kevin Moran

- Maria Kechagia

- Mariam Ei Mezouar
- Marios Fragkoulis
- Mauricio Aniche
- Md. Masudur Rahman
- Michele Tufano
- Mijung Kim
- Ozgur Kafali
- Patrick Morrison
- Pavneet Singh Kochhar
- Pradeep Venkatesh
- Profir-Petru Partachi
- Rubén Saborido Infantes
- Sadika Amreen
- Safwat Hassan
- Santanu Kumar Dash
- Sarah Elder
- Shaowei Wang
- Simone Scalabrino
- Tapajit Dey
- Tse-Hsun Peter Chen
- Tushar Sharma
- Victor Kwan
- Wai Ting Cheung
- Xin Yang
- Xuan Lu
- Yonghui Huang
- Yu Zhao
- Yuan Tian
- Yun Zhang
- Yuxing Ma
- Zheng Gao
- Zijad Kurtanović

Appendix B

Message from the chairs in the proceedings

Welcome to MSR 2017, the Fourteenth International Conference on Mining Software Repositories, held on May 20 and 21 in Buenos Aires, Argentina, co-located with the 39th International Conference on Software Engineering (ICSE 2017).

When the first MSR was held, 13 years ago, it was difficult to foresee that it was to evolve into a large, mature, and reputable conference, meeting point for a then emergent, but today fully established, research community. With more than 200 submissions in all our submission categories, MSR now includes categories for both reusable tools (new this year) and practical experiences in the main research track, along with the now already classical Mining Challenge and Data Showcase tracks. As any other research community with some history, and some future, this year we have started two new recognitions: the Early Career Achievement Award, and the Foundational Contribution Award.

MSR 2017 used double-blind reviewing, the very first time at MSR. We decided so this year partially because of the biases introduced[1] when reviewers know the authors, and the many requests from people from the MSR community. In addition to reducing biases, we would like to remain inviting and friendly. Since this is a learning experience for everyone, we have tried our best to remind authors of this new process. For example, we have sent all authors who have submitted an abstract a reminder about this new process before the full paper submission deadline. We have checked the submitted PDF files to identify potential violations as much as we could and have given the relevant authors time to remove their identities from their PDF files before paper bidding starts.

For this edition of MSR we introduced a new category of research papers: reusable tools. With this new category, we intend to promote and recognize the creation and use of tools that are designed and built not only for a specific research project, but for the MSR community as a whole. Reviewed, along with the papers in the practice experiences category, together with other research papers, they ensure that we take into account these two fundamental areas for our community.

This year, we received a record number of 158 submissions of research papers (121 full and 37 short). Out of the 121 full paper submissions, we accepted 37 papers, and out of the 37 short paper submissions, we accepted six. These are acceptance rates of 30.6% for full papers and 16.2% for short papers. Of the accepted papers, two were submitted in the practice experience category,

---

[1] http://www.cs.princeton.edu/~dpw/popl/15/dbr-faq.html#studies and https://arxiv.org/abs/1702.00502

four in the reusable tools category, and four additional papers in both. Authors of accepted papers come from 23 different countries. We did not fix any quota on the number of papers to be accepted. Rather, the acceptance decisions were made based on the program committee's discussions on the content and quality of each paper individually. Following tradition, MSR 2017 accepted or rejected papers at their submitted length, and did not permit downgrading a paper from a full paper to a short paper.

Acknowledging our roots as a working conference, the schedule has allocated significant time for discussion, trying a new model for encouraging participation. MSR 2017 has included plenary sessions for general discussions, and discussion time in all sessions (plenary and parallel).

For the fifth year in a row, MSR 2017 has a Data Showcase, in which papers present datasets curated by their authors, and made available to the community at large. These papers provide a description of the dataset, how it was put together (including the methodology to gather it), its schema, how it can be used, and its limitations/challenges. This year we received 22 submissions, of which eight (36

We also continue the tradition of the Mining Challenge, where researchers from across the community apply their mining techniques to a common problem. For the second year, we issued an open call for challenges, which attracted four competitive proposals. From those, we then selected the one we considered the most appealing to be the Mining Challenge for 2016: TravisTorrent, a dataset synthesized from Travis CI and GitHub. The Mining Challenge 29 submissions, out of which 14 (48

As it has already became a tradition, a selection of the best research papers, and for the third year also data showcase papers, 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 2017. As we did already for two years, MSR 2017 will also recognize papers with ACM SIGSOFT Distinguished Paper Awards.

Following the practice of the last editions, MSR 2017 presents a Most Influential Paper award honoring a paper from MSR 2007, that showed to have a high impact on the MSR community. The award has been selected by a committee consisting of the organisers of MSR 2007, who came to an unanimous decision. The award goes to Cathrin Weiss, Rahul Premraj, Thomas Zimmermann, and Andreas Zeller for their paper "How Long will it Take to Fix This Bug?". The paper in question not only has had by far the most impact in purely academic terms (as in: number of citations), but also paved the way for the concept of "actionable" MSR research, i.e., the conception and implementation of an approach which can be reused "out of the box".

This year, for the first time, the MSR community recognizes outstanding contributions in this field, by establishing two series of awards: the MSR Early Career Achievement Award, and the MSR Foundational Contribution Award. The MSR Early Career Achievement Award recognizes outstanding junior researchers who provided outstanding contributions in the area of mining software repositories. The MSR Foundational Contribution Award recognizes individuals, or groups of individuals, having produced fundamental contributions in the field of mining software repositories.

This year the MSR Early Career Achievement Award goes to Abram Hindle from University of Alberta, Canada, to recognize the rigor, fearlessness, and breadth of research related to Mining Software Repositories, and for establishing a new area of research related to "green mining".

The MSR Foundational Contribution Award goes to Tim Menzies, from North Carolina State University, United State, for pioneering work in defect prediction, and for setting widely-adopted standards and frameworks for statistical work in software engineering, including the creation of the PROMISE reusable data repository.

MSR is prepared by a large team. We thank all the members of the Program Committee (69 members this year), and invited reviewers, the chairs of the Mining Challenge (Moritz Beller,