MSR 2010

International Working Conference on Mining Software Repositories
(co-located with ICSE 2010)
2-3 May 2010, Cape Town, South Africa
http://msrconf.org/   twitter: @msrconf

General Chair
Audris Mockus — Avaya, USA

Program Co-chairs
Jim Whitehead —
UC Santa Cruz, USA

Thomas Zimmermann —
Microsoft Research, USA

Challenge Chair
Abram Hindle —
University of Waterloo, Canada

Program Committee
Giuliano Antoniol —
    École Polytechnique de Montréal, Canada
Andrew Begel — Microsoft Corp., USA
Christian Bird — UC Davis, USA
Li-Te Cheng — IBM Research, USA
Stephan Diehl — U of Trier, Germany
Massimiliano Di Penta — U of Sannio, Italy
Harald Gall — U of Zurich, Switzerland
Tudor Girba — U of Bern, Switzerland
Mike Godfrey — U of Waterloo, Canada
Jesus Gonzalez-Barahona — U Rey Juan Carlos
Ahmed Hassan — Queen's U, Canada
Reid Holmes — U of Washington, USA
Katsuro Inoue — U of Osaka, Japan
Huzefa Kagdi —
    Missouri U of Science and Technology, USA
Miryung Kim — U of Texas Austin, USA
Sung Kim —
    Hong Kong U of Science and Technology, China
Michele Lanza — U of Lugano, Switzerland
Andrian Marcus — Wayne State U, USA
Ken-ichi Matsumoto —
    Nara Institute of Science and Technology, Japan
Tim Menzies — West Virginia U, USA
Nachiappan Nagappan — Microsoft Corp., USA
Martin Pinzger — Delft Technical U, Netherlands
Rahul Premraj — Vrije U Amsterdam, Netherlands
Lori Pollock — U of Delaware, USA
Bill Pugh — U of Maryland, USA
Martin Robillard — McGill U, Canada
Gregorio Robles — U Rey Juan Carlos, Spain
Anita Sarma — U of Nebraska, USA
Tao Xie — North Caroline State U, USA
Westley Weimer — U of Virginia, USA
Laurie Williams — North Carolina State U, USA
Andy Zaidman — Delft Technical U, Netherlands
Andreas Zeller — Saarland U, Germany

Co-located with ICSE 2010

Call for Papers

Software repositories such as source control systems, archived communications between project personnel, and defect tracking systems are used to help manage the progress of software projects. Software practitioners and researchers are recognizing the benefits of mining this information to support the maintenance of software systems, improve software design/reuse, and empirically validate novel ideas and techniques. Research is now proceeding to uncover the ways in which mining these repositories can help to understand software development and software evolution, to support predictions about software development, and to exploit this knowledge concretely in planning future development.

The goal of this two-day working conference is to advance the science and practice of software engineering via the analysis of data stored in software repositories. We solicit **short papers** (4 pages) and **research papers** (10 pages). Short papers should discuss controversial issues in the field, or describe interesting or thought provoking ideas that are not yet fully developed. Accepted short papers will present their ideas in poster form during a poster session at the conference, and in a short lightning talk. Full research papers are expected to describe new research results, and have a higher degree of technical rigor than short papers. Accepted full papers will present their ideas in a research talk at the conference. A selection of the best research papers will be invited for consideration in a **special issue of the Springer journal Empirical Software Engineering**.

Papers may address issues along the general themes, including but not limited to the following:

- Analysis of software ecosystems and mining of repositories across multiple projects
- Models for social and development processes that occur in large software projects
- Prediction of future software qualities via analysis of software repositories
- Models of software project evolution based on historical repository data
- Characterization, classification, and prediction of software defects based on analysis of software repositories
- Techniques to model reliability and defect occurrences
- Search-based software engineering, including search techniques to assist developers in finding suitable components and code fragments for reuse, and software search engines
- Analysis of change patterns and trends to assist in future development
- Visualization techniques and models of mined data
- Techniques and tools for capturing new forms of data for storage in software repositories, such as effort data, fine-grained changes, and refactoring
- Approaches, applications, and tools for software repository mining
- Characterization of bias in mining and guidelines to ensure quality results
- Meta-models, exchange formats, and infrastructure tools to facilitate the sharing of extracted data and to encourage reuse and repeatability
- Case studies on extracting data from large long-lived and/or industrial projects
- Methods of integrating mined data from various historical sources

In the **MSR Challenge**, we invite researchers to demonstrate the usefulness of their tools on the FreeBSD distribution, Ultimate Debian Database, and the GNOME desktop suite. Discover interesting facts about these repositories for the *General Track* and report results as 4-page submissions. Accepted submissions will be included in the proceedings as challenge papers. For the *Prediction Track*, we challenge you to predict the newest bug number for the Debian project on April 30th, 2010. The winners of both tracks will receive an award. See the MSR homepage for more information about requirements and rules.

Important Dates
Research/short papers: January 11, 2010 (abstracts), January 14, 2010 (papers)
Mining challenge: February 6, 2010 (challenge papers), February 20, 2010 (predictions)
Author notification: February 19, 2010