



MSR 2005

Mining Software Repositories

MSR 2005

International Workshop on Mining Software Repositories
(17th May 2005, co-located with ICSE 2005)

["http://msr.uwaterloo.ca"](http://msr.uwaterloo.ca)

Organizers

Ahmed E. Hassan

(aehassa@plg.uwaterloo.ca)

Richard C. Holt

(holt@plg.uwaterloo.ca)

School of Computer Science
University of Waterloo
Ontario, Canada

Stephan Diehl

(diehl@cs.uni-sb.de)

Department of Computer Science
Catholic University Eichstätt
Germany

Program Committee

Alexander Dekhtyar (U. of Kentucky, USA)
Premkumar T. Devanbu (UC Davis, USA)
Stephen G. Eick (SSS Research, Inc.)
Harald Gall (U. of Zurich, Switzerland)
Les Gasser (U. of Illinois, UC, USA)
Daniel German (U. of Victoria, Canada)
Jane Hayes (U. of Kentucky, USA)
Katsuro Inoue (Osaka U., Japan)
Philip Johnson (U. of Hawaii, USA)
Tim Lethbridge (U. of Ottawa, Canada)
Gail Murphy (U. of British Columbia, Canada)
Audris Mockus (Avaya Labs Research, USA)
Thomas J. Ostrand (AT&T Research, USA)
Dewayne Perry (U. of Texas, USA)
Jelber Sayyad Shirabad (U. of Ottawa, Canada)
Andreas Zeller (Saarland U., Germany)

Important Dates

Intent to submit:

11th February 2005

Deadline for submission:

21st February 2005

Paper Notification:

21st March 2005

Final papers due:

4th April 2005

Workshop date:

17th May 2005

Location

Co-located with ICSE 2005



Saint Louis, Missouri, USA

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 beginning to recognize the potential benefit 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, to support predictions about software development, and to plan various aspects of software projects.

The goal of this one-day workshop is to establish a community of researchers and practitioners who are working to recover and use the data stored in software repositories to further understanding of software development practices. We expect the presentations and discussions in this workshop to continue on a number of general themes and challenges, from the previous workshop (MSR 2004) held at ICSE 2004, such as:

- Engineering challenges related to the infrastructure and tools needed to recover useful data from these repositories
- Methods of integrating mined data from various historical sources
- Development and validation of approaches to visualize and present such data
- Use of recovered history for system understanding and analysis of change patterns
- Modeling of defects and software reliability using data from such repositories
- Uncovering of the social processes and interaction between the development community
- Discovery of techniques to facilitate software reuse

Since MSR 2004 was the first workshop to focus on the field of mining software repositories, it was exploratory in nature with a variety of results in different areas of the field. We expect MSR 2005 to retain an exploratory feel while at the same time promoting collaboration and systematic comparison of approaches and techniques. This mixture of exploratory work with systematic comparisons and collaborations should assist in maturing the field of mining software repositories and in directing it to take a central role in supporting software development practices.

Position papers should be at most 5 pages and may address issues along the general themes, including but not limited to the following:

- Approaches to study the quality of the mined data along with guidelines to ensure the quality of the recovered data
- Proposals for exchange formats, meta-models, and infrastructure tools to facilitate the sharing of extracted data and to encourage reuse and repeatability
- Models for social and development processes that occur in large software development projects
- Search techniques to assist developers in finding suitable components for reuse
- Techniques to model reliability and defect occurrences
- Analysis of change patterns to assist in future development
- Case studies on extracting data from repositories of large long lived projects
- Suggestions for benchmarks, consisting of large software repositories, to be shared among the community