The 4th International Workshop on Mining Software Repositories
Co-located with ICSE 2007, May 19-20, 2007 - Minneapolis, USA
Workshop web site: http://msr.uwaterloo.ca/ — Contact: msr2007@swag.uwaterloo.ca

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 two-day workshop is to strengthen the 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 at MSR 2007 in Minneapolis to continue on a number of general themes and challenges from the previous workshops held at ICSE 2006 in China, ICSE 2005 in the US, and ICSE 2004 in Europe.

We solicit short position papers (4 pages) and research papers (8 pages). Short papers will be expected to discuss controversial issues in the field, or describe interesting or thoughtprovoking ideas that are not yet fully developed, while full papers will be expected to describe new research results, and have a higher degree of technical rigor than short papers. The papers must be in ICSE format. We plan to invite the best papers of MSR in a revised and extended version for publication in an international journal.

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

- Approaches, applications, and tools for software repository mining
- Quality aspects and guidelines to ensure quality results in mining
- Meta-models, exchange formats, 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 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
- Visualization techniques and models of mined data

In addition to the MSR paper track we invite researchers to participate in the MSR Challenge and demonstrate the usefulness of their mining tools. The main task will be to find interesting insights by analyzing the software repositories of Eclipse and Firefox (Challenge #1: Scale). Both systems are large in size, several years mature, and provide lots of input for mining tools. Experts and developers of the investigated projects will give feedback on the submissions (Challenge #2: Convince) and authors of selected submissions will be invited to report their findings at the workshop. More information about the requirements and rules for the MSR Challenge are available on the MSR homepage.